Maths - Bio

Introduction à la topologie

François Bienvenu

Séances des 12 et 19 février et du 5 mars 2015

Remarques préliminaires

Le titre de ce document est un peu ambitieux. J'aurais sans doute dû lui préférer quelque chose comme *Aperçu de quelques notions de topologie* : en effet, seuls un petit nombre de concepts seront abordés ici. Pire, puisque nous sommes des biologistes intéressés avant tout par les mathématiques appliquées, il ne s'agira pas forcément des concepts les plus fondamentaux d'un point de vue mathématique !

Le but de ce document est donc uniquement de donner un avant goût de ce que peut être la topologie, tout en restant très simple d'accès ; les personnes que sa lecture aura intéressées seront invitées à consulter un véritable cours d'introduction au sujet. Je n'ai pas de recommandation particulière, mais le net en regorge.

Vous pouvez me contacter pour me signaler des erreurs dans ce document, me faire des suggestions ou me poser des questions à l'adresse :

`francois -point- bienvenu -arobase- free -point- fr`

Bonne lecture, et n'oubliez pas de chercher les exercices proposés !

1 Introduction

Topologie. L'étude (*logos*) du lieu (*topos*). D'accord, mais qu'est-ce que ça veut dire ?

Quand on consulte des ouvrages de vulgarisation sur la topologie (par exemple, l'excellent *Topologicon* de Jean-Pierre Petit [Petit, 1999]), on nous parle de tasses qui sont équivalentes à des donuts, de propriétés qui sont vraies sur la sphère mais pas sur le tore [1]... Bref, on nous parle d'objets "identiques à une déformation continue près". Eh bien ce n'est pas ce dont on va parler ici.

En fait, ces résultats appartiennent à des domaines comme la topologie algébrique ou la topologie différentielle, tandis que nous allons parler de *topologie générale*. Cette discipline a pour but de généraliser des notions naturelles, comme la continuité, en étudiant les espaces munis d'une structure que nous appellerons bientôt leur topologie. Le cadre théorique qui en résulte est très abstrait, et on peut donc se demander quel est l'intérêt pour des biologistes comme nous de s'y intéresser. Outre l'aspect

1. Voir, par exemple, le fameux théorème de la boule chevelue...

divertissant que certains pourront y voir, la réponse à cette question est : l'analyse fonctionnelle. Explications.

Jusqu'à maintenant (du moins pour ceux qui sortent de prépa bio), vous avez principalement rencontré des fonctions réelles d'une variable réelle, i.e $f : \mathbb{R} \to \mathbb{R}$. Vous avez aussi vu des fonctions de plusieurs variables, que ce soit en probabilités avec la densité d'un couple de variables aléatoires réelles ou encore avec les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$ étudiées dans les chapitres d'algèbre linéaire. Dans tous ces cas, la situation était relativement simple car on voit bien la "tête" des espaces étudiées : $\mathbb{R}$ est une droite, $\mathbb{R}^2$ est un plan, etc. Bref, une fonction est continue si on peut tracer son graphe sans lever le stylo, la dérivée de la fonction en un point est la pente de la tangente à son graphe en ce point, et on arrive facilement à formaliser tout ça...

Mais, en tant que scientifique, on est souvent amené à considérer des fonctions dont les espaces de départ et d'arrivée sont bien moins sympas que $\mathbb{R}^n$. En voici un exemple : vous avez rencontré beaucoup d'équations différentielles en physique, et notamment des équations comme l'équation de la chaleur :

$$\frac{\partial T}{\partial t} = D\frac{\partial^2 T}{\partial x^2}\,.$$

Il s'agit d'un exemple d'équation aux dérivées partielles. Pour trouver la distribution de température à l'équilibre (c'est-à-dire "résoudre" cette équation pour $t = +\infty$), il vous avait fallu imposer des conditions sur les bords du domaine étudié – par exemple, fixer la température aux deux extrémités d'une barre. Si on veut vraiment modéliser la répartition de la température dans l'espace au cours du temps, on doit travailler dans $\mathbb{R}^3$ et l'équation s'écrit alors

$$\begin{cases} \partial_t T = D\Delta T & \text{sur } \Omega \subset \mathbb{R}^3 \\ T(\,\cdot\,, t) = f & \text{sur } \partial\Omega, \forall t \quad \text{(condition de bord)} \\ T(\,\cdot\,, 0) = g & \text{sur } \Omega \qquad\quad\;\; \text{(condition initiale)} \end{cases}$$

où $\partial_t T$ est une notation simplifiée pour $\frac{\partial T}{\partial t}$, $\Delta T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}$ (Laplacien de T), Ω est le domaine considéré et $\partial\Omega$ est son bord (les deux extrémités de la barre dans le cas précédent).

Peu importe ici l'interprétation physique de cette équation, ce qui nous intéresse est qu'il s'agit d'un problème très concret qui peut s'écrire sous la forme :

$$\begin{cases} P(T) = 0 & \text{sur } \Omega \\ T = f & \text{sur } \partial\Omega \\ T_{|t=0} = g & \text{sur } \Omega \end{cases}$$

où $P : T \mapsto \partial_t T - D\Delta T$ est un *opérateur*, c'est-à-dire une fonction qui prend une fonction comme argument et en renvoie une autre. Attention ici au léger abus de notation : 0 ne désigne pas le réel 0 mais l'application de Ω dans $\mathbb{R}$ qui à $x \in \Omega$ associe 0. De même, $T = f$ et $T_{|t=0} = g$ sont des égalités entre fonctions.

On en arrive enfin au point important : *des questions comme l'existence et l'unicité d'une solution vont pouvoir être traitées en raisonnant directement sur l'opérateur P*. De plus, dans le cas général, les réponses à ces questions vont dépendre à la fois de l'espace de fonctions dans lequel on choisit de chercher des solutions (suis-je intéressé par des solutions $\mathcal{C}^\infty$, ou bien simplement continues ? Est-ce qu'une solution discontinue aurait un sens ?) et des fonctions f et g paramétrant le problème.

Bref, ce qu'il faut retenir c'est que, quand on étudie des équations aux dérivées partielles, les variables comme les paramètres sont des fonctions ! Il est donc crucial de pouvoir généraliser à des espaces de fonctions les propriétés de $\mathbb{R}$ qui nous avaient permis de définir des notions comme la continuité.

L'intérêt de l'analyse fonctionnelle ne se limite pas à l'étude des équations différentielles, mais il faut bien reconnaître que c'en est *la* motivation majeure. On peut également mentionner le fait que l'analyse fonctionnelle joue un rôle important en théorie des probabilités : en effet, l'objet central des probabilités est la variable aléatoire. Or une variable aléatoire n'est rien d'autre qu'une fonction ! [2]

---

2. On peut faire une remarque à ce propos, que nous ne développerons pas davantage dans ce document : en analyse fonctionnelle, on cherche à définir des notions comme la continuité. La bonne structure pour faire cela est la topologie. En probabilités, ce qu'on demande avant tout aux variables aléatoires est d'être *intégrables*. La structure de base dont on doit munir l'espace est alors la *tribu* (σ-algèbre). Topologie et tribu sont donc des structures assez différentes, bien que leurs axiomes aient une ressemblance superficielle. Néanmoins, la notion d'intégrabilité permet de définir des normes (normes L^p) – or une norme induit une topologie ! Ainsi, des espaces comme L^2 (fonctions de carré intégrable), dont l'étude est au cœur de l'analyse fonctionnelle, jouent également un rôle central en probabilités...

Maintenant que nous sommes convaincus de l'intérêt de l'analyse fonctionnelle, nous allons pouvoir commencer par le début et nous intéresser à ces structures qu'on appelle des topologie. Loin de les introduire de la façon la plus élégante qui soit et d'en étudier toutes les propriétés, notre objectif sera simplement d'avoir une démarche aussi naturelle que possible et d'introduire du vocabulaire et un petit nombre de notions indispensables à la lecture d'articles de biologie mathématique.

2 Concepts de base sur les espaces métriques

2.1 Notion de distance

Puisque notre but est de généraliser la notion de continuité, et qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite continue en a lorsque $f(x) \xrightarrow[x \to a]{} f(a)$, commençons par écrire le fait que f admet ℓ pour limite en a :

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \text{ t.q.} \quad |x - a| < \delta \implies |f(x) - \ell| < \varepsilon\,.$$

Ce que cela veut dire, c'est qu'on peut se rapprocher autant que l'on veut de ℓ en se rapprochant de a : quelle que soit la distance ε que l'on se fixe, on est sûr qu'en étant assez proche de a, la distance entre $f(a)$ et ℓ lui sera inférieure. Dans cette définition, c'est la valeur absolue de la différence qui est utilisée pour exprimer la distance – et en effet, cela correspond bien à la distance “physique” sur la droite.

Une idée naturelle pour généraliser la notion de limite est donc de réécrire cette définition pour n'importe quelle notion de distance. On dira donc qu'une fonction $f : E \to F$ admet ℓ pour limite en a lorsque :

$$\forall \varepsilon>0, \quad \exists \delta>0 \text{ t.q.} \quad d_E(x,a) < \delta \implies d_F(f(x),\ell) < \varepsilon\,,$$

où d_E et d_F sont des fonctions qu'on se donne qui rendent compte de la distance entre deux éléments de E et de F, respectivement.

Néanmoins, il va de soi que toute fonction d ne fera pas l'affaire ; il faut que les fonctions considérées correspondent bien à l'idée intuitive qu'on se fait d'une distance pour que cette définition ait un intérêt. Pour préciser cela, il est nécessaire de définir la notion de distance de façon mathématique ; ceci se fait en identifiant les propriétés que nous jugeons caractéristiques de cette notion, puis en en faisant des axiomes. On aboutit à la définition suivante :

Une *distance* sur E est une application $d : E \times E \to \mathbb{R}$ vérifiant les propriétés suivantes :

1. $d(x, y) \geq 0$
2. $d(x, y) = d(y, x)$ *(symétrie)*
3. $d(x, y) = 0 \iff x = y$ *(séparation)*
4. $d(x, z) \leq d(x, y) + d(y, z)$ *(inégalité triangulaire)*

(on a omis les quantificateurs ; il est clair qu'il s'agit de "pour tout" à chaque fois)

Il est facile de voir que la notion habituelle de distance satisfait ces propriétés, mais il n'est pas évident de voir qu'il s'agit des propriétés qui caractérisent la notion. Néanmoins, il semble que c'est bien la notion de distance répondant à ces caractéristiques qui est intéressante en mathématiques.

Exercice 1 :

1. *Montrer que $d : (x, y) \mapsto |x - y|$ définit bien une distance sur $\mathbb{R}$.* *($\star$)*

2. *Montrer que*

$$d : (f, g) \mapsto \max_{x \in [a,b]} |f(x) - g(x)|$$

définit une distance sur l'espace $\mathcal{C}^0([a, b])$ *des fonctions continues de* $[a, b]$ *dans* $\mathbb{R}$. *($\star$)*

3. *Montrer que*

$$d : (\mathbf{x}, \mathbf{y}) \mapsto \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} ,$$

où $p \in [1, +\infty[$ *, définit une distance distance sur* $\mathbb{R}^n$*. Qu'obtient-on lorsque* $p \to \infty$ *? ($\star\star\star$)*

Une remarque de vocabulaire : un ensemble E muni d'une distance est appelé un *espace métrique*.

Enfin, notons qu'on peut définir de façon naturelle la distance d'un point $x \in E$ à un ensemble $A \subset E$. Il suffit de poser :

$$d(x, A) = \inf_{y \in A} d(x, y)$$

Il est évident qu'on a alors $x \in A \implies d(x, A) = 0$. Mais **attention**, la réciproque est fausse ! Par exemple, sur $\mathbb{R}$ muni de la distance usuelle (i.e. d :

$(x, y) \mapsto |x-y|)$, $d\big(1,]0,1[\big) = 0$ mais $1 \notin]0,1[$. Nous reviendrons sur cela...

2.2 Un peu de structure (et beaucoup de vocabulaire)

Maintenant que nous avons muni l'espace E d'une fonction permettant de définir la notion de limite (et donc de continuité), nous allons étudier la structure qu'elle engendre sur l'espace. Pour cela, commençons par définir la notion de boule ouverte :

On appelle *boule ouverte de rayon r centrée en x* l'ensemble :

$$B_r(x) = \{y \in E \mid d(x, y) < r\}$$

Attention au fait qu'il s'agit ici d'une inégalité *stricte*, car c'est le point crucial. Par la suite, nous serons amenés à parler également de *boule fermée de rayon r centrée en x* :

$$\overline{B}_r(x) = \{y \in E \mid d(x, y) \leq r\}$$

(la notation $\overline{B}$, qui n'est qu'une notation pour l'instant, sera justifiée par la suite).

Exercice 2 : *Représenter les boules ouvertes de rayon 1 centrées en 0 associées aux distances sur* $\mathbb{R}^2$ *:*

$$d_1 : \big((x_1, y_1), (x_2, y_2)\big) \mapsto |x_1 - x_2| + |y_1 - y_2|$$
$$d_2 : \big((x_1, y_1), (x_2, y_2)\big) \mapsto |x_1 - x_2|^2 + |y_1 - y_2|^2$$
$$d_\infty : \big((x_1, y_1), (x_2, y_2)\big) \mapsto \max\left(|x_1 - x_2|, |y_1 - y_2|\right)$$

On appelle *voisinage* de x tout ensemble V tel que

$$\exists r > 0 \text{ t.q. } B_r(x) \subset V$$

La notion de voisinage est importante, car elle permet de formuler certaines propriétés de façon concise (on va en voir un exemple juste après). De plus, ce concept permet à lui seul une construction relativement intuitive de la topologie [3]. Néanmoins, nous ne serons pas amenés à manipuler beaucoup de voisinages car nous raisonnerons directement en termes d'*ouverts* :

3. Dans ce document, les topologies seront définies en termes d'ouverts ; mais nous aurions pu faire la même construction à partir de la notion de voisinage. Cette façon de procéder aurait été plus intuitive, mais aussi plus lourde...

On appelle *ouvert* tout ensemble qui est voisinage de chacun de ses points ; on peut reformuler cela de la façon suivante : un ensemble O est ouvert si et seulement si $\forall x \in O$, $\exists r > 0$ tel que $B_r(x) \subset O$.

Le concept compagnon de celui d'ouvert est celui de *fermé* : un ensemble est dit fermé lorsque son complémentaire est un ouvert. Pour l'instant, c'est un peu mystérieux, mais tout cela va devenir bien plus concret lorsque nous manipulerons ces concepts. Mais avant cela, il nous faut introduire un peu de vocabulaire supplémentaire :

- On appelle *point limite* d'un sous ensemble A de E tout point x tel que toute boule centrée en x intersecte à la fois A et son complémentaire, A^{c}. L'ensemble des points limites de A est noté ∂A et appelé la *frontière* de A :

$$\partial A = \{x \in E \mid \forall r > 0, \exists y \in A \cap B_r(x) \text{ et } \exists z \in A^{\mathsf{c}} \cap B_r(x)\}$$

  En fait, cette définition est hyper intuitive puisqu'elle veut tout simplement dire que la frontière de A correspond à l'ensemble des points qui sont à distance nulle de A et de son complémentaire – ça correspond bien à l'idée qu'on a d'une frontière.

- On appelle *point intérieur* d'un sous ensemble A de E tout point qui est le centre d'une boule incluse dans A. L'intérieur de A est noté $\mathring{A}$:

$$\mathring{A} = \{x \in A \mid \exists r > 0 \text{ t.q. } B_r(x) \subset A\}$$

  Là encore, cette définition est intuitive : un point est à l'intérieur de A si on peut l'isoler de A^{c} par une "petite boule de A". De plus, il est clair d'après cette définition qu'un ouvert est son propre intérieur : $A = \mathring{A}$ si et seulement si A est ouvert.

- On appelle *point d'adhérence* (ou point de contact) d'un sous ensemble A de E tout point x tel que toute boule centrée en x intersecte A. L'ensemble des points d'adhérence de A est appelé l'adhérence de A (ou sa fermeture) et est notée $\overline{A}$:

$$\overline{A} = \{x \in E \mid \forall r > 0, \, \exists y \in A \cap B_r(x)\}$$

  La différence entre cette définition et celle de point limite est qu'on n'impose plus d'être à distance nulle du complémentaire de A, ce qui fait qu'on va aussi prendre en compte les points intérieurs de A.

Donnons tout de suite quelques exemples très simples : plaçons-nous sur $\mathbb{R}$ muni de la distance usuelle, et considérons les cas : $A =]0,1[$, $A = [0,1]$ et $A =]0,1]$. On pourra vérifier que dans ces trois cas on a :

$$\mathring{A} =]0,1[\,, \quad \overline{A} = [0,1]\,, \quad \text{et} \quad \partial A = \{0\} \cup \{1\}.$$

Les petits exercices suivants permettent de s'assurer qu'on a bien compris toutes ces définitions avant de passer à la suite.

Exercice 3 :

1. *Montrer que $\overline{A}$ est l'ensemble des points à distance nulle de A.* $(\star)$

2. *Montrer que $\forall A \subset E$, $E = \mathring{A} \sqcup \partial A \sqcup \left(\overline{A}\right)^{\mathsf{c}}$ (où $\sqcup$ désigne l'union disjointe), et que $\mathring{A} = \left(\overline{A^{\mathsf{c}}}\right)^{\mathsf{c}}$. Remarquer que cela implique notamment :*

   (a) $\overline{A} = \mathring{A} \sqcup \partial A$

   (b) $\overline{A^{\mathsf{c}}} = \partial A \sqcup \left(\overline{A}\right)^{\mathsf{c}}$

   $(\star\star)$

Maintenant qu'on a un peu d'intuition sur ce vocabulaire, on va préciser les notions d'ouverts et de fermés, qui sont vraiment à la base de la topologie. La notion d'ouvert est déjà plutôt naturelle : ils s'agit d'ensembles égaux à leur intérieur. Ces ensembles seront agréables à manipuler car, en gros, on pourra raisonner de la même façon en chacun de leurs points ; par exemple, pas besoin de se demander "oui mais si on ne peut pas trouver de voisinage de ce point sur lequel la fonction soit définie ?" – lorsqu'une fonction est définie sur un ouvert, elle est automatiquement définie sur un voisinage de chacun des points de cet ouvert. Bref, quel que soit le point $x \in]0,1[$, je peux dire "je considère un l'intervalle $]x-\varepsilon, x+\varepsilon[$..." sans avoir à me soucier de ce qui se passe en 0 ou en 1...

En revanche, à ce stade, la notion de fermé reste assez mystérieuse... voire piégeuse ! En effet, on pourrait avoir l'impression que puisqu'un fermé est le complémentaire d'un ouvert, un ensemble est soit ouvert, soit fermé. C'est une erreur de raisonnement – et en effet, cette assertion est complètement **fausse** !

Exercice 4 :

1. *Commencer par vérifier que les intervalles ouverts sont des ouverts et que les intervalles fermés sont des fermés (pour $\mathbb{R}$ muni de la distance usuelle).* ($\star$)

2. *Trouver des exemples d'ensembles qui ne sont ni ouverts, ni fermés.* ($\star$)

3. *Trouver des exemples d'ensembles qui sont à la fois ouverts et fermés.* ($\star\star$)

(penser à préciser la distance considérée).

On vient de préciser un peu les relations entre ouverts et fermés. Pour cerner encore un peu mieux ces ensembles, on va maintenant voir comment ils se comportent vis à vis des opérations habituelles sur les ensembles.

Exercice 5 :

1. *Montrer qu'une réunion finie d'ouverts / fermés le reste. Même question pour une intersection finie.* ($\star$)

2. *Qu'en est-il des intersections / réunions dénombrables ? Quelconques ?* $(\star\star)$

On commence à mieux voir ce que sont les ouverts et les fermés... Mais on ne comprend toujours pas bien ce que représentent les fermés, et surtout quel est leur intérêt. L'exercice suivant a pour but de pallier cela.

Exercice 6 :

1. *Montrer que $A = \overline{A}$ ssi A est fermé.* $(\star)$
2. *Montrer que A est ouvert si et seulement si toute suite d'éléments de A qui converge a sa limite dans A.* $(\star\star)$

De même qu'il découlait immédiatement des définitions que $\mathring{A}$ est le plus grand ouvert contenant A, la propriété $A = \overline{A} \Leftrightarrow A$ fermé permet de montrer que $\overline{A}$ est le plus petit fermé contenant A (exercice).

Enfin, la deuxième propriété ("un ensemble est fermé si et seulement si la limite de toute suite convergente d'éléments de cet ensemble lui appartient") donne une nouvelle caractérisation des fermés qui s'avérera souvent plus utile que la définition un peu indirecte de "complémentaire d'un ouvert". S'il y a une propriété à retenir sur les fermés, c'est sans doute celle-là ; nous en verrons l'intérêt par la suite.

2.3 Densité

Nous allons maintenant aborder un concept dont nous avons déjà parlé cette année au GT : celui de densité. Un ensemble A est dit *dense* dans B si et seulement si :

$$B \subset \overline{A}\,.$$

Ce que cela veut dire, c'est tout simplement que tout point de B est à distance nulle de A (i.e. peut être approché autant que souhaité par un point de A). Au passage, une remarque triviale : cette relation n'est pas symétrique. En effet, ce n'est pas parce que tout point de A est à distance nulle de B que tout point de B est à distance nulle de A...

L'exemple le plus naturel est sans doute celui de la densité de $\mathbb{Q}$ dans $\mathbb{R}$: en effet, $\mathbb{Q}$ est dense dans $\mathbb{R}$ puisqu'il est connu que tout réel peut être approché autant que souhaité par un rationnel ; ce fait est acquis dès qu'on admet que tout réel peut être représenté – au passage, de façon non unique [4] – par la séquence infinie de ses décimales ; ainsi, tout réel x peut être approché autant que souhaité au moyen d'un terme de la suite

$$x_n = \frac{\lfloor x \times 10^n \rfloor}{10^n}$$

(où $\lfloor \cdot \rfloor$ désigne la partie entière – en gros c'est juste une écriture pour dire qu'on approche π par 3 puis 3,1 puis 3,14 etc). En effet, on vérifie facilement que

$$x - x_n = \frac{x \times 10^n - \lfloor x \times 10^n \rfloor}{10^n} \leq 10^{-n}\,.$$

Pour donner des exemples un peu plus "costauds" : les matrices inversibles sont denses dans l'espace des matrices (pour ceux qui savent ce qu'est le déterminant : en perturbant très légèrement une entrée d'une matrice on peut s'assurer que son déterminant ne s'annule pas et donc qu'elle est inversible) ; de même, les fonctions polynomiales sur $[a, b]$ sont denses dans l'ensemble des fonctions continues sur $[a, b]$ (d'après le théorème d'approximation de Weierstrass – il suffit alors d'utiliser la distance déjà rencontrée dans l'exercice 1 : $d(f, g) = \max_{[a,b]} |f - g|$).

4. C'est le fameux 0.99999... = 1

Introduisons maintenant le concept de nulle part densité : un ensemble A est dit *nulle part dense* dans B s'il n'est dense dans aucun ouvert non vide de B. La première question qui vient à l'esprit quand on voit cette définition est : "pourquoi a-t-on besoin d'une définition si tordue et ne peut-on pas se contenter de la notion de 'non dense' ?". Un élément de réponse est que, dans bien des situations, le concept de "non dense" est trop faible pour être satisfaisant. Par exemple, ni $\mathbb{Q} \setminus [0, 1]$ ni $\mathbb{N}$ ne sont denses dans $\mathbb{R}$. Pourtant, on voit bien que ces deux ensembles de "remplissent" pas $\mathbb{R}$ de la même manière du tout... Il y aura donc des situations ou une notion permettant de donner un sens à "la plupart des points de B ne peuvent pas être approchés par des points de A" sera plus intéressante que la non-densité, qui ne correspond qu'à "il existe un point de B ne pouvant pas être approché par des points de A".

Donnons maintenant un exemple d'ensemble nulle part dense : l'ensemble des matrices non-diagonalisables dans $\mathbb{C}$ est nulle part dense dans l'ensemble des matrices complexes. On peut de plus remarquer que dire ceci est plus fort que le fait de dire que les matrices diagonalisables dans $\mathbb{C}$ sont denses dans l'ensemble des matrices complexes [5].

Quelle est l'utilité des notions que nous venons d'introduire ? Une première réponse est qu'elles permettent d'avoir une idée de la pertinence de certaines hypothèses qu'on pourrait être amenés à faire. "J'ai fait l'hypothèse que ma matrice était diagonalisable dans $\mathbb{C}$, est-ce une grosse restriction ?" Ça dépend de ce qu'on veut faire exactement, mais si ses entrées proviennent de données expérimentales, alors quelle que soit la précision utilisée pour les mesurer cette hypothèse aurait aussi bien pu être vraie (i.e. il existe des matrices diagonalisables dans la "région de confiance" associée à la précision de mesure) donc il y a des chances pour qu'elle ne soit pas très restrictive...

Néanmoins, il faut faire attention : la densité est une notion d'*approchabilité*, qui n'a rien à voir avec la notion de *probabilité de choix*. Ni le fait que l'ensemble des matrices diagonalisables soit dense dans l'ensemble des matrices complexes, ni même le fait que l'ensemble des matrices non-diagonalisables

5. Attention, ces résultats sont valables pour les matrices complexes uniquement : les matrices diagonalisables dans $\mathbb{R}$ ne sont pas denses dans l'espaces des matrices réelles !

y soit nulle part dense ne permettent de conclure qu'une matrice choisie au hasard sera diagonalisable ! La bonne notion pour étudier cette question est la *négligeabilité*, qui est une notion de théorie de la mesure[6]. Rappelons qu'un ensemble A est dit négligeable pour la mesure μ lorsque $\mu(A) = 0$. Intuitivement, cela signifie qu'il est de volume nul – ou encore, lorsque μ est une mesure de probabilité, que la probabilité qu'un point choisi selon μ tombe dans A est nulle. Lorsqu'on ne connait pas a priori la mesure à utiliser, un bon choix par défaut est la mesure de Lebesgue, qui correspond à la mesure de volume usuelle.

Or non seulement on rencontre fréquemment des ensembles denses de mesure de Lebesgue nulle (exemple : $\mathbb{Q}$ dans $\mathbb{R}$), mais il existe aussi des ensembles nulle part denses de mesure de Lebesgue non-nulle !

Exercice 7 : *Trouver un exemple d'ensemble nulle part dense de mesure de Lebesgue strictement positive (c'est-à-dire, si l'on travaille sur* $\mathbb{R}$*, de longueur non-nulle).* $(\star\star\star)$

---

6. On retrouve l'opposition (proximité, continuité, topologie générale, topologie) *vs* (volume, intégrabilité, théorie de la mesure, tribu) déjà mentionnée dans la note 2.

Le principal intérêt de la notion de densité est en fait le raisonnement par *prolongement*. Le principe en est très simple : on montre une propriété en la montrant d'abord pour un ensemble "simple" (typiquement : dénombrable, ou composé d'éléments particulièrement agréables à manipuler, etc) puis en la prolongeant à un ensemble moins sympathique à l'aide d'un théorème de prolongement du type de celui qui suit :

Théorème de prolongement par continuité : Si deux applications continues coïncident sur une partie dense de E, alors elles coïncident sur E tout entier.

Exercice 8 : *Prouver ce théorème.* $(\star\star)$

Remarque : ce théorème est assez simple à montrer en "fonçant dans le tas", mais il existe aussi une démonstration très élégante...

C'est ce théorème que nous avions utilisé lors des séances de Guillaume Jeanne autour de la transformée de Laplace pour montrer que toute fonction vérifiant $f(a+b) = f(a) \times f(b)$ est une exponentielle : nous avions commencé par montrer la propriété sur $\mathbb{N}$, puis sur $\mathbb{Z}$, puis sur $\mathbb{Q}$ à l'aide de manipulations algébriques ; puis nous l'avions étendue à $\mathbb{R}$ en utilisant la continuité de f et la densité de $\mathbb{Q}$ dans $\mathbb{R}$.

3 Interlude : oublier un instant les distances...

Si on prend un peu de recul, on se rend compte que, bien qu'on ait basé toute notre construction sur la notion de distance, on l'a assez peu manipulée directement ; on a quasiment tout le temps raisonné en termes de boules, sans vraiment se soucier de leur rayon... On peut donc se demander s'il serait possible d'oublier complètement la notion de distance et de tout construire en termes d'un ensemble prédéfini jouant le rôle de nos boules. Mais avant de faire ça, on va commencer par se demander si cela présenterait un quelconque intérêt.

3.1 Pour quoi faire ?

La première question à se poser c'est :

Peut-on toujours munir un ensemble d'une distance ?

Et bien la réponse est... oui ! Je me suis bien gardé de l'introduire plus tôt, mais vous pourrez vérifier que la fonction suivante, appelée *distance triviale*, est bien une distance :

$$d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{sinon} \end{cases}$$

Mais alors, si on peut *toujours* munir l'ensemble sur lequel on travaille d'une distance, pourquoi pourrait-on avoir envie de définir des choses indépendamment de la notion de distance ?

En fait, la réponse à cette question est très importante. Bien sûr, il y a le désir de définir les choses de la manière la plus générale possible, pour “en extraire l'essence”... Mais il n'y a pas que ça ! En fait, il y a des situations dans lesquelles la “bonne” notion de convergence n'est associée à aucune distance. Nous allons en voir des exemples – mais avant ça, il nous faut définir le cadre général dans lequel travailler.

3.2 Espaces topologiques généraux

On appelle *espace topologique* un couple $(E, \mathcal{T})$ où $\mathcal{T}$ est un ensemble de sous-ensembles de E vérifiant :

- $\varnothing \in \mathcal{T}$ et $E \in \mathcal{T}$.
- $\mathcal{T}$ est stable par union quelconque.
- $\mathcal{T}$ est stable par intersection finie.

$\mathcal{T}$ est la *topologie* dont on munit E, et ses éléments sont appelés les *ouverts* de E.

On peut déjà remarquer que cette définition est compatible avec les résultats de l'exercice 5 : une intersection finie d'ouverts est un ouvert, et toute union d'ouvert en est également un. À titre d'exemples, on pourra vérifier que les ensembles suivants sont bien des topologies sur E :

- $\{\varnothing, E\}$ **(topologie grossière)**
- $\mathcal{P}(E)$, l'ensemble des parties de E **(topologie discrète)**
- $\{\varnothing, \{0\}, \{-1, 1\}, \{-1, 0, 1\}\}$ **(topologie GT sur $\{-1, 0, 1\}$)**

Étant donnée une famille d'ensembles (A_i), on peut également définir la *topologie engendrée* par cette famille comme la plus petite[7] topologie contenant tous les A_i. Cette définition un peu abstraite peut être rendue un peu plus "concrète" par la procédure de construction suivante :

1. Ajouter $\varnothing$ et E à $\mathcal{T}$.

2. Ajouter toutes les intersections finies d'éléments de la famille (A_i) à $\mathcal{T}$.

3. Ajouter à $\mathcal{T}$ toutes les unions (quelconques) de ses éléments.

L'étape 2 permet d'obtenir la stabilité par intersection finie, et l'étape 3 la stabilité par union, tout en conservant la stabilité par intersection finie. Il y a tout de même une petite subtilité : on ne peut pas intervertir les étapes 2 et 3 de cette construction car on pourrait perdre la stabilité par union !

Ainsi, lorsque l'espace est muni d'une distance, on peut parler de la topologie engendrée par les boules

7. Au sens de l'inclusion, c'est-à-dire que les éléments de cette topologie se retrouvent dans toute autre topologie contenant tous les A_i.

associées à cette distance. Cette topologie est appelée *topologie induite par la distance d*. On pourrait vérifier que, en plus d'être les unions quelconques d'intersections de boules, les ouverts de cette topologie sont les ensembles O tels que :

$$\forall x \in O,\ \exists r > 0 \text{ t.q. } B_r(x) \subset O\,,$$

ce qui justifie la définition que nous avions donnée dans la section 2.

Une fois qu'on s'est donné une topologie, toutes les notions vues précédemment (voisinage, fermé, intérieur, frontière, adhérence, convergence, continuité...) se généralisent de façon évidente en remplaçant "il existe une boule centrée en x telle que..." par "il existe un ouvert contenant x tel que...". Par exemple, le fait que la fonction f soit continue en x s'écrit :

$$\forall O_{f(x)},\ \exists O_x \text{ t.q. } \forall y \in O_x,\ f(y) \in O_{f(x)}$$

(où la notation générique O_a désigne un ouvert contenant a). On pourra vérifier en exercice que la continuité de f peut également s'exprimer de la façon suivante :

f est continue ssi l'image réciproque par f d'un ouvert est un ouvert.

3.2.1 Digression : topologies et tribus

Ceux qui ont déjà rencontré la notion de *tribu* (σ-algèbre) auront peut-être remarqué une certaine ressemblance dans le formalisme. En effet, les axiomes d'une tribu $\mathcal{F}$ sont :

- $\mathcal{F}$ n'est pas vide.
- $\mathcal{F}$ est stable par passage au complémentaire.
- $\mathcal{F}$ est stable par union dénombrable.

Mais il ne faut surtout pas confondre les deux notions, qui sont fondamentalement différentes ! Comme nous l'avons déjà mentionné plusieurs fois, topologie et tribus sont des objets construits dans des buts très différents. Pour se convaincre de la différence entre les deux, on peut remarquer que le comportement vis-à-vis du passage au complémentaire est une différence importante : les seuls ensembles à la fois ouverts et fermés de $\mathbb{R}$ muni de la topologie usuelle sont $\varnothing$ et $\mathbb{R}$, et donc la topologie usuelle de $\mathbb{R}$ n'est absolument pas stable par passage au complémentaire[8].

8. Point culture : un espace topologique $(E, \mathcal{T})$ dont les seuls ensembles à la fois ouverts et fermés sont $\varnothing$ et E est dit *connexe*. Nous n'aurons malheureusement pas le temps d'aborder cette notion, mais je vous invite à y jeter un coup d'œil.

Néanmoins, malgré cela il existe un lien fort entre topologies et tribus, notamment car (1) l'intégrabilité permet de définir une notion de distance – et donc une topologie – et surtout (2) une tribu particulièrement intéressante est la tribu borélienne, qui est définie comme la tribu engendrée par les ouverts d'un espace topologique...

3.2.2 Que perd-on ?

Un grand nombre des propriétés que nous avons montrées restent valides dans des espaces topologiques généraux. Par exemple, les relations entre intérieur, adhérence, et frontière vues dans l'exercice 3 sont vraies dans le cas général. Il en est de même de la propriété $A = \overline{A}$ ssi A fermé vue dans l'exercice 6.

En revanche, ce n'est pas le cas de la caractérisation séquentielle des fermés vue dans l'exercice 6 l'implication A fermé $\implies$ toute suite d'éléments de A qui converge a sa limite dans A reste vraie (et justifie d'ailleurs le terme de "fermé" : "on ne peut pas en sortir") mais la réciproque est fausse sur un espace topologique non métrique ! De même, le lemme 2 et le théorème de prolongement par continuité de l'exercice 8 ne sont pas valables dans le cas général. En effet, les arguments écrit en bleu ne fonctionnent plus : $x \neq y \not\Longrightarrow \exists O_x$ et O_y tels que $O_x \cap O_y = \varnothing$ – comme illustré par les éléments (-1) et 1 de l'espace muni de la topologie GT décrite plus haut ; de même, la caractérisation séquentielle de la continuité n'est valable que dans les espaces métriques.

Rien qu'avec ces exemples, on peut voir qu'en oubliant la notion de distance on a perdu des propriétés liées à des caractérisations séquentielles. Or ces propriétés sont bien pratiques, car il est souvent agréable de manipuler des suites ! C'est une des raisons pour lesquelles les espaces métriques sont "sympathiques".

Pour rentrer un peu plus dans certains détails et introduire du vocabulaire, on a perdu la propriété de séparation [9]. Un espace topologique est dit *séparé* (ou de Hausdorff) lorsque :

$$x \neq y \implies \exists O_x, \exists O_y \text{ t.q. } O_x \cap O_y = \varnothing$$

---

9. La propriété de séparation présentée est la propriété de séparation T_2 (ou de Hausdorff). Il existe de **nombreuses** notions de séparation (je vous invite à consulter l'article Wikipédia `https://fr.wikipedia.org/wiki/Axiome_de_séparation_(topologie)` pour vous en convaincre !) mais on se limitera à celle-ci...

Les espaces métriques sont séparés : en effet, en choisissant $r \leq d(x, y)/2$, les boules $B_r(x)$ et $B_r(y)$ fournissent les ouverts disjoints contenant x et y, respectivement. En revanche, il existe des espaces séparés non métrisables.

Un des principaux intérêts des espaces séparés est qu'on y a l'unicité de la limite. En effet, une suite (x_n) ne peut pas admettre deux limites ℓ_1 et ℓ_2 différentes car cela voudrait dire qu'il existe un certain rang à partir duquel tous les points de la suite seraient à la fois dans un ouvert O_1 contenant ℓ_1 et dans un ouvert O_2 contenant ℓ_2, avec O_1 et O_2 disjoints – ce qui est bien évidemment impossible. En revanche, rien ne garantit l'unicité de la limite dans un espace non-séparé ! Par exemple, considérons la suite de terme général $(-1)^n$ sur $\{-1, 0, 1\}$ muni de la topologie GT. On vérifie facilement que cette suite admet à la fois (-1) et 1 pour limite.

Un autre intérêt de ces espaces est justement le théorème de prolongement par continuité de l'exercice 8 – au passage, on remarque que seule la séparation de l'espace d'arrivée est nécessaire, et l'on comprend pourquoi j'ai présenté deux versions de la démonstration : la première car elle est plus générale, et la seconde, limitée aux espaces métriques (utilisation de la caractérisation séquentielle de la continuité), car elle me semble plus "visuelle".

3.2.3 Comparaisons de topologies

Lorsque j'ai parlé de "plus petite topologie contenant une famille d'ensemble", j'ai déjà dit de façon implicite qu'on pouvait comparer les topologies. En effet, la relation d'inclusion est bien une relation d'ordre pour l'ensemble des topologies définies sur un espace. Cet ordre admet même un élément minimal (la topologie grossière, dont les éléments appartiennent à toute topologie) et un élément maximal (la topologie discrète, qui contient les éléments de toute topologie), mais il n'est pas total : on ne peut pas toujours dire qu'une topologie est plus petite ou plus grande qu'une autre (par exemple, pour $x \neq y$, les topologies $\{\varnothing, \{x\}, E\}$ et $\{\varnothing, \{y\}, E\}$ ne sont pas comparables). Au passage, on préfère dire qu'une topologie $\mathcal{T}_1$ est *plus fine* que $\mathcal{T}_2$ lorsque $\mathcal{T}_2 \subset \mathcal{T}_1$ (cette dénomination est assez intuitive, car $\mathcal{T}_1$ contient alors plus d'ouverts et donc permet de mieux "distinguer" les éléments de E...).

Ici, ce ne sont pas les propriétés (un peu abstraites...) de cette relation d'ordre qui vont nous intéresser mais simplement le lien qu'on peut faire avec la "facilité" qu'il y a à converger dans une topologie donnée. Ce qu'il faut remarquer, c'est tout simplement que plus une topologie est fine, puis il est difficile de converger – et donc plus la notion de convergence associée est forte. En effet, plus une topologie est fine, plus il y a d'ouverts pour lesquels la propriété d'"appartenance à partir d'un certain rang" doit être vérifiée, et donc plus la convergence est exigeante. Par exemple, on remarque que toute suite est convergente pour la topologie grossière, tandis que seules les suites qui deviennent constantes à partir d'un certain rang le sont pour la topologie discrète (conséquence directe du fait que les singletons sont des ouverts).

De même, plus la topologie de son espace d'arrivée est fine, plus il est difficile pour une application d'être continue ; et inversement pour son espace de départ. Par exemple, munir l'espace d'arrivée de la topologie grossière rend tout application continue – de même que le fait de munir l'espace de départ de la topologie discrète.

3.3 Exemples d'espaces topologiques non-métrisables

Nous avons déjà vu un exemple de convergence qui ne peut pas être associée à une distance avec la topologie GT (on a vu que les topologies induites par une distance étaient séparées ; n'étant pas séparée, la topologie GT ne peut donc être induite par aucune distance). Mais il faut bien reconnaître que cet exemple n'est pas très intéressant... Nous allons donc maintenant donner de "vrais" exemples de situations où la bonne notion de convergence ne peut être associée à aucune distance.

La façon la plus simple de définir la convergence dans un espace fonctionnel est de dire qu'une suite de fonctions (f_n) converge vers f si et seulement si :

$$\forall x \in X,\ f_n(x) \xrightarrow[n\to\infty]{} f(x)\,.$$

On parle de *convergence simple* (ou *convergence ponctuelle*). À cette notion de convergence est associée une topologie, appelée topologie de la convergence simple. Comme le montre l'exercice suivant, cette topologie n'est pas métrisable !

Exercice 9 : *On considère les ensembles :*

$$B_\varepsilon^x(f) = \left\{ g \in \mathbb{R}^E \ t.q. \ |g(x) - f(x)| < \varepsilon \right\}$$

et on note $\mathcal{T}$ la topologie engendrée par la famille $\left(B_\varepsilon^x(f)\right)_{(x,\varepsilon,f)}$ lorsque (x, ε, f) décrit $E \times \mathbb{R}_+^* \times \mathbb{R}^E$.

1. *Montrer que $\mathcal{T}$ est bien la topologie correspondant à la convergence simple.($\star$)*

Soient maintenant a et b deux réels distincts. Considérons l'ensemble :

$$A = \left\{ f \in \mathbb{R}^E \ t.q. \ f = a \text{ sauf en un nombre fini de points} \right\}$$

2. *Montrer que $f : x \mapsto b \in \overline{A}$. ($\star$)*
3. *Discuter si f est limite simple de fonctions de A. ($\star\star$)*
4. *Conclure. ($\star$)*

Deux autres exemples de topologies non-métrisables ayant une place centrale en analyse fonctionnelle sont donnés par la topologie de la *convergence faible* et la topologie de la *convergence faible-$*$*. Ces topologies interviennent lorsqu'on manipule les distributions, objets ayant pour but de généraliser les fonctions et dont l'usage est constant dans l'étude des équations aux dérivées partielles.

Nous avons déjà entendu parler de distributions cette année (voir les notes sur la méthode des moments pour les processus ponctuels spatialisés). Je ne vais pas m'aventurer à essayer d'écrire une introduction au sujet, mais simplement donner quelques éléments permettant de comprendre la suite : l'idée derrière la théorie des distributions est qu'il n'est pas toujours pertinent de caractériser la répartition d'une quantité par sa valeur en chaque point, car cela fait jouer un rôle trop important aux irrégularités (et pose des problèmes, par exemple pour définir la dérivée). Un bon exemple de cela est la température : comme elle est due à des chocs d'atomes, des problèmes se posent pour assigner une valeur à cette quantité en dessous d'une certaine échelle d'espace et de temps. En revanche, il est indéniable que lorsqu'on met un thermomètre à un endroit, celui-ci nous donne une valeur de la température – valeur qui varie en fonction de la position et de la forme du thermomètre utilisé. Il est donc assez naturel de vouloir caractériser la température par son action sur des appareils de mesure plutôt que par sa valeur en chaque point.

Une fois qu'on a cette idée, reste à construire toute la théorie mathématique associée. Peu importent ici les détails, mais il semble alors que la bonne façon de procéder soit de faire jouer le rôle des appareils de mesure à certaines fonctions appelées *fonctions test*, dont on note l'ensemble $\mathcal{D}$. Les distributions sont alors les formes linéaires continues sur $\mathcal{D}$, dont on note l'ensemble $\mathcal{D}'$. Une distribution T est donc caractérisée par son action $\langle T, \varphi \rangle \in \mathbb{R}$ sur toute fonction test φ. Une fois qu'on a posé cela, on définit naturellement une notion de convergence sur l'ensemble $\mathcal{D}$ des fonctions test et sur l'ensemble $\mathcal{D}'$ des distributions : une suite de fonctions (φ_n) *converge faiblement* vers φ si et seulement si :

$$\forall T \in \mathcal{D}',\ \langle T, \varphi_n \rangle \xrightarrow[n\to\infty]{} \langle T, \varphi \rangle\,.$$

De même, une suite de distributions (T_n) tend vers T si et seulement si :

$$\forall \varphi \in \mathcal{D},\ \langle T_n, \varphi \rangle \xrightarrow[n\to\infty]{} \langle T, \varphi \rangle\,.$$

Il s'agit de la *convergence faible-$*$*, ou *convergence au* *sens des distributions*[10].

Comme on l'a déjà dit, il s'avère que ces notions de convergence ne sont associées à aucune distance sur les espaces considérés. Nous ne développerons pas plus ici, mais comme il s'agit de notions importantes dont vous ré-entendrez sans doute parler, je pensais qu'il était intéressant de les mentionner.

---

10. En fait, la convergence faible-$*$ a un lien très fort avec une convergence que vous avez déjà rencontrée en probabilités : la convergence en loi. Pour s'en convaincre, il suffit de comparer la définition de la convergence pour des distributions associées à des mesures de Borel et la définition de la convergence en loi d'une suite de variables aléatoires : il n'y a que les espaces des fonctions test à considérer qui diffèrent ($\mathcal{C}_c^\infty$ vs $\mathcal{C}_b^0$). Attention néanmoins, car la convergence en loi, vue comme une convergence sur l'ensemble des lois de probabilité, est métrisable (en revanche, elle ne l'est évidemment pas pour l'espace des variables aléatoires : deux variables aléatoires différentes pouvant avoir la même loi, la topologie associée n'a aucune chance d'être séparée).

4 Plus de structure !

Avant de continuer, on peut faire un petit bilan de ce qu'on a fait jusqu'à maintenant : on a d'abord commencé par introduire la notion de distance, à partir de laquelle on a pu munir l'espace d'une structure permettant de définir des notions comme la convergence ou la continuité. Puis, en manipulant un peu cette structure, on s'est aperçus qu'on avait rarement besoin d'avoir recours à la notion de distance de façon directe. Cela nous a conduits à introduire une structure abstraite, très épurée mais retenant néanmoins l'essence des objets que nous avions vus précédemment : la topologie.

Néanmoins attention : on n'a pas tout conservé en se passant de distance ! On a notamment vu qu'on avait perdu une propriété importante appelée *séparabilité*, qui garantit entre autres l'unicité de la limite. Plus généralement, on a vu qu'on avait perdu des propriétés de *caractérisations séquentielles* (caractérisation séquentielle de la continuité, des fermés...) pourtant bien pratiques.

À ce stade, deux possibilités s'offrent à nous : on pourrait soit continuer à explorer cette "structure minimale" pour trouver de belles propriétés très abstraites... soit manipuler des structure un peu plus "lourdes", mais obtenir des résultats plus balèzes dont on voit tout de suite les applications. Vous l'avez deviné, c'est la deuxième possibilité qu'on choisira ici [11].

À partir de maintenant, on se placera donc sur des espaces métriques. Néanmoins, lorsque cela ne demande pas un gros effort, je rédigerai les raisonnements de la façon la plus générale possible, qu'on voie bien où le caractère métrique intervient.

4.1 Complétude

Pour l'instant, on ne dispose que d'une seule méthode pour montrer qu'une suite converge : prendre un candidat pour la limite, et vérifier que la suite satisfait la définition de la convergence. Cette méthode demande donc de connaître la limite ! Or, au cours de votre scolarité, vous avez sans doute remarqué que cela s'avère souvent trop dur – voire impossible – et qu'il est alors plus simple d'avoir recours à des méthodes un peu plus sophistiquées comme le fait de montrer le caractère monotone et borné

11. Eh oui, désolé mais c'est le GT maths-bio ici...

d'une suite ou encore le recours au théorème des gendarmes.

On va donc maintenant s'attacher à présenter une notion permettant de caractériser la convergence de manière indirecte : la *complétude*[12]. Il nous faut pour cela commencer par introduire les *suites de Cauchy*.

Suites de Cauchy : une suite (u_n) est dite de Cauchy ssi :

$$\forall \varepsilon > 0,\ \exists N \text{ t.q. } \forall m, n > N,\ d(u_m, u_n) < \varepsilon\,.$$

Autrement dit, les termes de la suite "se rapprochent uniformément les uns des autres" : quelle que soit la distance qu'on se fixe, on sait qu'à partir d'un certain rang tous les termes de la suite seront à moins de cette distance les uns des autres.

Attention au caractère "uniforme" de la définition ! Le rang N doit être tel que $\forall\, \boldsymbol{m, n > N}$, etc – et pas seulement pour deux termes consécutifs. Par exemple, la suite de terme général $u_n = \ln(n)$ n'est pas de Cauchy car bien qu'on ait $\ln(n+1) - \ln(n) \sim 1/n \to 0$, en revanche $\ln(2n) - \ln(n) = \ln(2) \not\to 0$.

12. Rien à voir avec le fameux théorème d'incomplétude de Gödel.

Une première propriété immédiate est que toute suite convergente est de Cauchy. En effet, puisqu'à partir d'un certain rang tous les termes sont contenus dans une boule de rayon ε, par l'inégalité triangulaire on a également qu'à partir de ce même rang tous les termes sont à une distance inférieure à 2ε les uns des autres. Mais attention, la réciproque est fausse : il existe, dans certains espaces, des suites de Cauchy qui ne sont pas convergentes ! Essayez d'en trouver un exemple avant d'étudier celui présenté dans l'exercice suivant.

Exercice 10 : *Plaçons-nous dans $\mathbb{Q}$ muni de la distance usuelle, et considérons la suite (u_n) définie par :*

$$\begin{cases} u_{n+1} = \frac{u_n}{2} + \frac{1}{u_n} \\ u_0 = \frac{3}{2} \text{ (par exemple)} \end{cases}$$

1. *Montrer que cette suite est de Cauchy.* $(\star\star)$

2. *Montrer que cette suite ne converge pas dans l'espace considéré.* $(\star\star)$

Ce constat du fait que toute suite de Cauchy ne converge pas (alors que, franchement, on aimerait bien) nous amène à introduire la notion d'*espace complet* :

Définition : Un espace métrique est dit complet si et seulement si toute suite de Cauchy y est convergente.

Une remarque, c'est que la complétude est une notion métrique et non topologique : deux distances peuvent être équivalentes d'un point de vue topologique (au sens qu'elles induisent la même topologie[13]) mais être associées à des notions de complétude différentes – c'est-à-dire qu'un espace peut être complet pour une distance sans l'être pour une autre.

On a vu avec l'exercice 10 que $\mathbb{Q}$ muni de la distance usuelle n'est pas complet. En revanche, $\mathbb{R}$ l'est... Seulement pour le montrer, encore faut-il avoir défini $\mathbb{R}$! Or il s'avère qu'une des constructions classiques de $\mathbb{R}$ procède justement de manière à assurer cette complétude.

---

13. Par exemple, les distances d_1, d_2, d_p et d_∞ présentées dans la section 1 sont toutes équivalentes dans $\mathbb{R}^n$.

4.1.1 Point culture : construction de $\mathbb{R}$

Je vais me contenter de donner les grandes idées de la construction, sans rentrer dans les détails techniques :

Après avoir construit $\mathbb{N}$ (par exemple à partir des axiomes de Peano) et l'avoir muni de l'addition, on peut construire $\mathbb{Z}$ de manière à ce que tout nombre y ait un inverse pour l'addition. On obtient alors un *groupe*. Une fois qu'on a fait ça, le fait de munir $\mathbb{Z}$ de la multiplication nous donne un *anneau*. Mais on aimerait que tout nombre ait aussi un inverse pour la multiplication, de manière à disposer d'un *corps*. On construit alors $\mathbb{Q}$ comme le corps des fractions de $\mathbb{Z}$. Il ne nous manque alors plus qu'une chose : la complétude.

Là, il faut faire un peu attention à ne pas se mordre la queue : en effet, on a définit la notion de suite de Cauchy à partir de la notion de distance. Or on avait définit une distance comme une fonction à valeurs dans $\mathbb{R}_+$! Donc on voit mal comment on pourrait parler de suite de Cauchy et de topologie usuelle sur $\mathbb{Q}$ à ce stade... En fait, c'est tout simple.

Il suffit de définir les suites de Cauchy ainsi :

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

et bien évidemment la convergence est définie de la manière suivante :

$$\exists \ell \in \mathbb{Q} \text{ t.q. } \forall \varepsilon \in \mathbb{Q}_+^*,\ \exists N \text{ t.q. } \forall n \geq N,\ |u_n - \ell| < \varepsilon\,.$$

On évite ainsi tout raisonnement circulaire. En partant de ces définitions, on peut montrer qu'il existe des suites de Cauchy qui ne convergent pas dans $\mathbb{Q}$. On décide alors de construire un ensemble, appelé $\mathbb{R}$, "étendant" $\mathbb{Q}$ de manière à ce que toute suite de Cauchy y soit convergente. Pour ça, on utilise une idée simple : on définit les réels comme les limites de suites de Cauchy de rationnels ! En fait, il y a quand même un petit problème : deux suites différentes peuvent avoir la même limite, donc un réel donné ne correspondrait pas à une seule suite... Pour résoudre ce problème, on *quotiente* l'espace des suites de Cauchy par la *relation d'équivalence* $\mathcal{R}$ suivante :

$$(u_n)\,\mathcal{R}\,(v_n) \quad \text{ssi} \quad |u_n - v_n| \xrightarrow[n\to\infty]{} 0$$

En gros, ça veut dire qu'on construit un nouvel espace dans lequel on identifie deux suites de Cauchy ayant la même limite. Les éléments de cet espace sont appelés des *classes d'équivalence*, et ce sont eux qui correspondent aux réels.

Une fois qu'on a fait ça, il ne nous reste plus qu'à munir l'ensemble qu'on vient de définir de l'addition et de la multiplication. Cela revient à munir l'ensemble des suites de Cauchy de ces opérations, ce qu'on fait de la façon habituelle, c'est-à-dire terme à terme : $(u + v)_n = u_n + v_n$ et $(u \times v)_n = u_n \times v_n$. En fait, il y a quand même quelques subtilités techniques ; notamment, il faut vérifier que les opérations définies sur les suites de Cauchy vont bien se comporter lors du quotientage ("Comment additionner deux classes d'équivalence, et est-ce que la classe d'équivalence de la somme de deux suites est la bien somme des classes d'équivalence de ces suites ?", etc). Il y a une théorie générale derrière tout ça qui permet de répondre à ces questions simplement. Bref, une fois qu'on a fait ça il ne reste plus qu'à vérifier que l'espace qu'on vient de définir est bien complet, ce qui se fait assez facilement.

J'invite ceux que ça intéresse à consulter l'article Wikipédia

https://fr.wikipedia.org/wiki/Construction_des_nombres_réels dans lequel les aspects plus techniques de cette construction sont détaillés.

Vocabulaire : Pour tout espace métrique E, il existe un unique espace $\widetilde{E}$ tel que (1) E est dense dans $\widetilde{E}$ et (2) $\widetilde{E}$ est complet. $\widetilde{E}$ est appelée le *complété* de E.

Par exemple, $\mathbb{R}$ est le complété de $\mathbb{Q}$.

On va maintenant pouvoir introduire un théorème extrêmement puissant, dont vous ré-entendrez sans doute parler : le théorème du point fixe de Banach.

4.1.2 Théorème du point fixe de Banach

Avant de parler du théorème lui-même, une petite remarque sur quelque chose qui m'amuse beaucoup : ce théorème doit son nom au mathématicien polonais Stefan Banach, qui l'énonça et le démontra en 1922. Mais en France, il est souvent appelé théorème du point fixe de Picard – du nom d'Émile Picard, mathématicien français. Dans la section suivante, on verra un théorème souvent appelé théorème de Borel-Lebesgue en France... Alors que le reste du monde préfère parler du théorème de Heine-Borel. De même, pendant ma scolarité on m'a parlé plusieurs fois de la loi de Descartes pour la réfraction, sans que j'entende jamais le nom de Snell... Bref, je ne suis pas en train de dire que nous sommes chauvins au point de changer systématiquement les noms des théorèmes en faveur de scientifiques français... Mais ça pourrait être une hypothèse amusante à tester ! Si vous êtes intéressé, n'hésitez pas à me contacter !

Application lipschitzienne Une application $\varphi : E \to E$ est dite *lipschitzienne* ssi

$$\exists k \geq 0 \text{ t.q. } \forall x, y,\ d\big(\varphi(x), \varphi(y)\big) \leq k\, d(x, y)\,.$$

Lorsque $k < 1$, φ est dite *contractante*.

Une première remarque, c'est qu'il découle directement de cette définition que toute application lipschitzienne est continue. Une deuxième remarque, c'est qu'on peut voir ça comme les fonctions dérivables dont la dérivée est bornée. En fait, il y a quelques cas pathologiques qui font que ce n'est pas tout à fait rigoureux de dire ça (comme $x \mapsto |x|$), mais globalement c'est ça l'idée.

Théorème du point fixe Soit E un espace métrique complet, et $\varphi : E \to E$ une application contractante. Alors φ admet un unique point fixe, i.e. $\exists! x^* \in E$ tel que $\varphi(x^*) = x^*$.

Exercice 11 : *Montrer ce théorème.* ($\star\star$)

Indication : on pourra introduire la suite définie, $\forall x_0 \in E$, par $x_{n+1} = \varphi(x_n)$.

On va faire quelque chose que je déteste : passer à la suite sans détailler d'exemple d'application du théorème qu'on vient d'énoncer. La raison pour laquelle je m'autorise à faire ça, c'est que ce théorème est d'utilisation tellement fréquente que je suis sûr que vous en réentendrez parler. En effet, on peut citer comme exemples d'applications :

- la preuve du **théorème de Cauchy-Lipschitz**, qui établit l'existence et l'unicité de la solution de certaines équations différentielles.
- la preuve du **théorème d'inversion locale**, qui donne des conditions suffisantes pour qu'une fonction soit inversible au voisinage d'un point (et qui est plus utile qu'il en a l'air).

- la **méthode de Newton**, qui permet d'approcher numériquement les zéros de certaines fonctions (et dont dérivent la plupart des méthodes modernes de résolution numérique d'équation).

En plus d'être assez variés, ces résultats sont tous extrêmement importants ; ceci illustre bien le rôle central que joue le théorème de Banach en mathématique. Donc, même si pour l'instant il peut sembler un peu abstrait, gardez le en tête et cela vous servira un jour !

4.2 Compacité

Lorsqu'on travaille sur $\mathbb{R}$, les intervalles fermés possèdent des propriétés intéressantes. Par exemple, on sait qu'une fonction continue sur un intervalle fermé y atteint ses bornes. Plus généralement, même si cela n'est pas forcément évident à voir, une propriété assez générale des intervalles fermés est qu'il est souvent possible d'y *transformer des propriétés locales en propriétés globales*. Par exemple, vous savez peut-être qu'une fonction continue sur un intervalle fermé I y est uniformément continue. Or la continuité sur I est bien une propriété locale, au sens qu'elle est définie de façon indépendante en tout point de l'intervalle :

$$\forall x{\in}I, \forall\varepsilon{>}0, \exists\boldsymbol{\delta_x{>}0} \text{ t.q. } \forall y{\in}I, |x{-}y|{<}\delta_x \Longrightarrow |f(x){-}f(y)|{<}\varepsilon\,,$$

tandis que dans le cas de la continuité uniforme, on impose que la continuité soit la même en tout point de l'intervalle I :

$$\forall\varepsilon{>}0, \exists\boldsymbol{\delta_I{>}0} \text{ t.q. } \forall x,y{\in}I, |x{-}y|{<}\delta_I \Longrightarrow |f(x){-}f(y)|{<}\varepsilon\,.$$

On a donc bien, par le théorème qui affirme qu'une fonction continue sur un intervalle fermé y est uniformément continue, un exemple de passage d'une propriété locale à une propriété globale rendu possible par le fait qu'on travaille sur un intervalle fermé borné.

Quelle est donc la propriété des intervalles fermés qui permet d'obtenir ces résultats ? Le caractère fermé semble important puisque les propriétés énoncées ne sont pas vraies, par exemple, sur des intervalles ouverts ($x \mapsto 1/x$ est continue sur $]0,1[$ mais elle n'y atteint pas ses bornes et n'y est pas uniformément continue) ; mais il n'est pas suffisant : en effet, $\mathbb{R}$ est fermé et pourtant une fonction continue n'y atteint pas forcément ses bornes et n'y est pas forcément uniformément continue. En revanche, le caractère “d’un seul tenant” (i.e. la connexité) n’est pas nécessaire : il est clair que les propriétés énoncées restent vraies dans le cas d’une union finie d’intervalles fermés.

En fait, cette propriété – qu’on appellera *compacité* – est assez dure à cerner... au point que l’identification des compacts est l’une des premières étapes de l’étude d’un espace topologique, et s’avère souvent pas évidente ! Autre illustration du caractère un peu complexe de la notion : historiquement, la compacité s’est construite par deux approches assez différentes qui n’ont été unifiées qu’en 1929. Mais les compacts sont tellement pratiques qu’on sera récompensés de nos efforts pour essayer de comprendre ce dont il s’agit.

Propriété de Borel-Lebesgue : *E vérifie la propriété de Borel-Lebesgue si de tout recouvrement de E par des ouverts, on peut extraire un recouvrement fini.*

Un *recouvrement* de E est une famille (quelconque ; pas forcément dénombrable) d’ensembles (A_x) telle que $E \subset \bigcup_x A_x$. Vous vous en doutez, un recouvrement par des ouverts est un recouvrement dont tous les ensembles A_x sont des ouverts, et un recouvrement fini est un recouvrement composé d'un nombre fini d'éléments.

$$E \subset \bigcup_{x \in X} O_x \implies \exists_{x_1,\ldots,x_n \in X} \text{ t.q. } E \subset O_{x_1} \cup \cdots \cup O_{x_n}$$

Un espace séparé E est dit *compact* si et seulement si il vérifie la propriété de Borel-Lebesgue. Une partie d'un espace topologique est dite compacte si elle est compacte en tant qu'espace topologique muni de la topologie induite par l'espace de départ.

Si c'est la première fois que vous voyez cette définition et que vous êtes normalement constitué, vous devez avoir du mal à faire le lien avec ce qu'on a raconté sur les intervalles fermés au début de cette section... Si ce n'est être carrément perdu. C'est normal, mais ça ne va pas durer car les propriétés suivantes vont nous aider à mieux cerner les compacts, en précisant un peu leurs relations avec les fermés :

Exercice 12 : *Montrer les propriétés suivantes :*

1. *Toute partie fermée d'un compact est compacte.* $(\star)$

2. Tout compact est fermé. (★★)

Il y a donc un lien fort entre compacts et fermés... Mais on ne peut pour l'instant pas en dire beaucoup plus. On va donc poursuivre notre caractérisation des compacts.

Propriété de Bolzano-Weierstrass : *A vérifie la propriété de Bolzano-Weierstrass si et seulement si de toute suite d'éléments de A il est possible d'extraire une sous-suite convergente dans A.*

Un ensemble vérifiant la propriété de Bolzano-Weierstrass est dit *séquentiellement compact*.

Là où les choses se compliquent, c'est qu'un espace compact n'est pas forcément séquentiellement compact [14]... Mais qu'un espace séquentiellement

14. On ne va pas détailler d'exemple car ça nous emmènerait beaucoup trop loin. Ceux que ça intéresserait peuvent se pencher sur $[0,1]^{[0,1]}$ muni de la topologie produit (qui est compact par le théorème de Tykhonov, mais n'est pas séquentiellement compact car la suite (f_n) de fonctions définie par "$\forall x \in [0,1]$, $f_n(x)$ est égal au $n^{\text{ème}}$ digit dans l'écriture binaire de x" n'admet pas de sous-suite convergente). Un autre exemple fréquemment mentionné est le compactifié de Stone-Čech de $\mathbb{N}$ – ne m'en demandez pas plus.

compact n'est pas forcément compact non plus[15] ! On ne voit donc pas trop ce qui justifie l'appellation "compacité séquentielle". En fait, elle vient du résultat suivant : dans un espace **métrique**, la compacité séquentielle est équivalente à la compacité. C'est ce qu'on se propose de montrer dans l'exercice suivant :

Exercice 13 : *Soit E un espace métrique.*

1. ***Borel-Lebesgue $\Longrightarrow$ Bolzano-Weierstrass***

   (a) *Montrer que si E vérifie la propriété de Borel-Lebesgue, alors l'intersection d'une suite décroissante de fermés non-vides est non-vide.*

   (b) *Remarquer que si une suite (x_n) admet une valeur d'adhérence x^*, i.e.*

   $$\exists x^* \ t.q.\ \forall \varepsilon > 0, \forall n, \exists p \geq n \ t.q.\ x_p \in B_\varepsilon(x^*)$$

   alors on peut extraire une sous-suite convergente de (x_n).

---

15. Pareil que pour le point précédent : ceux que ça intéressent peuvent se renseigner, par exemple, sur l'ensemble des ordinaux inférieurs au premier ordinal non-dénombrable muni de la topologie de l'ordre.

(c) *En considérant les ensembles*

$$A_n = \{x_k \ t.q. \ k \geq n\}$$

associés à une suite (x_n), *en déduire que la propriété de Borel-Lebesgue implique la propriété de Bolzano-Weierstrass.*

2. ***Bolzano-Weierstrass*** $\Longrightarrow$ ***Borel-Lebesgue***

   (a) *Montrer que si* E *vérifie la propriété de Bolzano-Weierstrass, alors* $\forall \varepsilon > 0$, *il existe un recouvrement fini de* E *par des boules de rayon* ε.

   (b) *Montrer que si* E *vérifie la propriété de Bolzano-Weierstrass, alors pour tout recouvrement* $(O_i)_{i \in I}$ *de* E,

$$\exists \alpha > 0 \ \textit{tel que} \ \forall x \in E, \exists i_x \in I \ \textit{tel que} \ B_\alpha(x) \subset O_{i_x}.$$

   (c) *En déduire que la propriété de Bolzano-Weierstrass implique la propriété de Borel-Lebesgue.*

Ainsi, dans un espace métrique, un compact est un fermé qui possède en plus la propriété d'être "petit", au sens qu'on finit par y tourner en rond. On va maintenant exploiter cette idée pour achever notre caractérisation des compacts en établissant un lien entre compacts et fermés bornés.

Exercice 14 :

1. *Montrer que, dans un espace métrique, un compact est borné (c'est-à-dire inclus dans une boule de rayon fini).* ($\star$)

2. *Montrer que les fermés bornés de* $\mathbb{R}^n$ *sont compacts.* ($\star\star$)

3. *Trouver un exemple de fermé borné non-compact.* ($\star$)

On vient donc de montrer que les compacts de $\mathbb{R}^n$ sont exactement les fermés bornés, un résultat connu sous le nom de *théorème de Heine-Borel* (ou théorème de Borel-Lebesgue ☺). En fait, ce théorème se généralise facilement à tout espace vectoriel normé de dimension finie.

Ce résultat pourrait donner l'impression qu'on s'est embêtés avec des définitions assez abstraites pour pas grand-chose... Rien n'est moins vrai : d'une part, rappelons que (du moins en mathématiques appliquées) la principale utilité de la topologie est de fournir une base à l'analyse fonctionnelle, et que les espaces de fonctions étudiés en pratique sont toujours de dimension infinie... Or, en dimension infinie, la boule fermée de rayon 1 n'est *jamais* compacte ! De plus, comme on aura l'occasion de le voir, même sur $\mathbb{R}^n$, certaines propriétés se démontrent beaucoup plus naturellement en utilisant la propriété de Borel-Lebesgue qu'en raisonnant en termes de "fermés bornés".

Maintenant que nous avons réussi à cerner un peu la notion de compacité, nous allons en étudier quelques propriétés – pour nous assurer qu'elle généralise bien les intervalles fermés de $\mathbb{R}$ comme nous le souhaitions.

Exercice 15 :

1. *Montrer que l'image d'un compact par une fonction continue est un compact.* ($\star$)

2. *En déduire que toute fonction à valeurs réelles atteint ses bornes sur un compact.* ($\star$)

3. *Sur des espaces métriques, il existe une notion de continuité plus exigente que la continuité simple ("l'image réciproque d'un ouvert est un ouvert") : la continuité uniforme. Une fonction* $f : E \to F$ *est dite* uniformément continue *ssi :*

$$\forall \varepsilon{>}0, \exists \delta{>}0 \ t.q. \ \forall x{,}y{\in}E, d_E(x,y){<}\delta \Longrightarrow d_F\Big(f(x), f(y)\Big){<}\varepsilon\,.$$

Montrer qu'une fonction continue sur un compact y est uniformément continue (théorème de Heine). $(\star\star)$

Finalement, on voit que notre définition des compacts nous a bien permis d'obtenir des objets assez généraux qui se réduisent bien aux objets qui nous intéressaient dans $\mathbb{R}^n$ tout en permettant de généraliser certaines propriétés à des espaces plus "abstraits". C'est donc un succès. ☺

Pour conclure sur les compacts, je dirais que même si ce sont des objets assez complexes, il ne faut pas en avoir peur. D'autant que le mot est souvent utilisé même lorsqu'on travaille dans des espaces comme $\mathbb{R}^n$, où il ne s'agit que de bêtes "fermés bornés" (mais *compac'*, ça claque). En plus, il faut bien voir que lorsqu'on manipule des compacts, dans bien des cas ce n'est pas parce qu'on a montré qu'il n'y aurait aucun sens à travailler sur d'autres ensembles, mais simplement pour faire en sorte que l'ensemble vérifie un certain nombre de propriétés désirables : en gros, généralement, on ne se retrouve pas sur un compact mais on choisit de s'y placer simplement parce que cela ne coûte pas cher et qu'on sait qu'on pourra extraire des sous-suites convergentes, que les fonctions continues atteindront leurs bornes, etc.

Un exemple de situation dans laquelle on entend souvent le mot "compact" : les *fonctions continues à support compact* (i.e. nulles en dehors d'un compact). Mais bien souvent, on ne se sert de cette propriété que pour dire que ces fonctions sont intégrables (et utiliser le fait que la continuité implique qu'elles s'annulent sur la frontière du compact – ce qui s'avère bien pratique puisque cela permet, par exemple, de ne pas se soucier des termes de bord lors d'une intégration par parties). Bref, il ne faut pas se laisser impressionner lorsqu'on vous parle de compacts car il y a des chances pour qu'il n'y ait en réalité rien de bien compliqué derrière.

5 Le mot de la fin

Il est temps de terminer cette petite excursion au pays des ouverts. J'espère que cela vous aura plus, et surtout que cela aura donné envie à ceux qui seraient arrivés jusqu'ici d'aller plus loin – que ce soit en poursuivant avec une introduction à l'analyse fonctionnelle ou en reprenant ce que nous avons fait de façon un peu plus poussée et rigoureuse : en effet, si la topologie était une recouverte d'une grosse pierre, nous n'aurions fait que la soulever juste assez pour avoir une idée du type de bestioles qui vivent en dessous (se dont on pourra se convaincre, par exemple, en jetant un coup d'œil au glossaire de topologie de Wikipédia).

Remerciements

Tout d'abord, je voudrais remercier les personnes venues assister aux trois séances du GT qui ont servi de base à ces notes. Je remercie également Camille Pouchol, qui m'a encouragé à les écrire en les relisant au fur et à mesure que je les rédigeais et avec qui j'ai eu des discussions constructives (notamment à propos des exemples de topologies non-métrisables), Guillaume Jeanne, qui a relu des corrections d'exercices, et Michel Bienvenu, qui m'a aidé à corriger l'orthographe.

Références

R. Palais. A simple proof of the Banach contraction principle. *Journal of Fixed Point Theory and Applications*, 2 :221–223, 2007.

J.P. Petit. *Le Topologicon*. Editions Belin, 1999. URL http://www.savoir-sans-frontieres.com/download/fre/topologicon.htm.

Exercices corrigés

Les étoiles sont une indication (complètement subjective) du niveau de difficulté des exercices :

- Les questions $\star$ ne nécessitent qu'une application directe des définitions – cela ne veut pas pour autant dire qu'elles sont *faciles* : il n'est pas du tout anormal de se retrouver complètement buggué face à raisonnement pourtant simple lorsqu'on est pas habitué à manipuler un formalisme donné....

- Les questions $\star\star$ reposent essentiellement sur une application des définitions, mais peuvent nécessiter de faire preuve d'un peu d'habilité calculatoire ou d'intuition pour orienter le développement.

- Le questions $\star\star\star$ nécessitent des raisonnements plus complexes, en plusieurs étapes, ou un peu de culture mathématique.

Exercice 1 : exemples de distances

1. *Montrer que* $d : (x, y) \mapsto |x - y|$ *définit bien une distance sur* $\mathbb{R}$. *(⋆)*

2. *Montrer que*

$$d : (f, g) \mapsto \max_{x \in [a,b]} |f(x) - g(x)|$$

définit une distance sur l'espace $\mathcal{C}^0([a, b])$ *des fonctions continues de* $[a, b]$ *dans* $\mathbb{R}$. *(⋆)*

3. *Montrer que*

$$d : (\mathbf{x}, \mathbf{y}) \mapsto \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p},$$

où $p \in [1, +\infty[$ *, définit une distance distance sur* $\mathbb{R}^n$. *Qu'obtient-on lorsque* $p \to \infty$ *? (⋆⋆⋆)*

1) La positivité, la séparation et la symétrie sont immédiates :

- La positivité découle du fait que $x \mapsto |x|$ est positive par définition.

- La séparation vient de $|x-y| = 0 \iff x-y = 0 \iff x = y$.
- La symétrie vient de $|x| = |-x|$ et $x - y = -(y - x)$.

L'inégalité triangulaire est légèrement plus délicate et nécessite une disjonction de cas : soient x, y et z trois réels. Supposons, sans perte de généralité, que $x \leq z$. On a donc $|x - z| = z - x$. Trois cas se présentent alors : (1) $y < x$; (2) $x \leq y \leq z$; (3) $z < y$. Commençons par le cas (2) : on a $x \leq y$ d'où $|x-y| = y-x$, et $|y-z| = z-y$. Ainsi, $|x-y|+|y-z| = y - x + z - y = z - x = |x - z|$. L'inégalité est donc vérifiée dans le cas (2). Dans le cas (3), on a maintenant $|x-y| = y-x \geq z-x = |x-z|$ (car $y > z$) et $|y-z| \geq 0$, et donc $|x-y|+|y-z| \geq |x-z|$. L'inégalité est donc encore vérifiée. Le cas (1) se traite de façon tout à fait similaire.

2) La positivité et la symétrie sont immédiates.

Pour ce qui est de la séparation, $f \neq g$ sur $[a, b]$ ssi $\exists x \in [a, b]$ tel que $f(x) \neq g(x)$. On a alors $|f(x) - g(x)| > 0$ et donc $\max_{[a,b]} |f - g| > 0$. Par contraposée, $\max_{[a,b]} |f-g| = 0 \implies f = g$ sur $[a, b]$.

La réciproque est immédiate.

L'inégalité triangulaire se déduit de la question (1) : $\max_{[a,b]} |f-g| = \max_{[a,b]} |f-h+h-g|$. Or comme

$$\forall x \in [a,b], |f(x)-h(x)+h(x)-g(x)| \leq |f(x)-h(x)|+|h(x)-g(x)|,$$

on a $\max_{[a,b]} |f-h+h-g| \leq \max_{[a,b]} (|f-h|+|h-g|) \leq \max_{[a,b]} |f-h| + \max_{[a,b]} |h-g|$, d'où $\max_{[a,b]} |f-g| \leq \max_{[a,b]} |f-h| + \max_{[a,b]} |h-g|$.

On vient donc de voir un exemple de distance entre fonctions. Y est associée une première notion de convergence dans des espaces fonctionnels, appelée *convergence uniforme*. Nous en verrons d'autres.

3) Cette question est assez compliquée, et surtout très calculatoire. Il y a deux raisons pour lesquelles j'ai choisi de l'inclure dans cette feuille : (1) ces distances (ou les normes L^p auxquelles elles sont associées) jouent un rôle fondamental en analyse fonctionnelle et surtout (2) la preuve va nous permettre de voir quelques inégalités classiques.

Tout d'abord, remarquons que ni la positivité, ni la symétrie, ni la séparation ne posent de problème.

Intéressons-nous donc à l'inégalité triangulaire. On doit montrer que :

$$\left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i - z_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |z_i - y_i|^p\right)^{1/p}.$$

En fait, il s'agit d'une inégalité connue appelée *inégalité de Minkowski* qu'on trouve le plus souvent sous la forme :

$$\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{1/p}.$$

Il suffit en effet de prendre $a_i = x_i - z_i$ et $b_i = z_i - y_i$. Ce qu'il nous faut, c'est donc montrer l'inégalité de Minkowski. On peut déjà remarquer que pour $p = 1$, cette inégalité est une conséquence de l'inégalité triangulaire que nous avons montrée à la question (1). On peut donc supposer $p > 1$. On utilise alors l'inégalité triangulaire et le fait que $(|a_i + b_i|)^p \leq |a_i||a_i + b_i|^{p-1} + |b_i||a_i + b_i|^{p-1}$ pour écrire :

$$\sum_{i=1}^{n} |a_i + b_i|^p \leq \sum_{i=1}^{n} |a_i||a_i + b_i|^{p-1} + \sum_{i=1}^{n} |b_i||a_i + b_i|^{p-1}$$

On va maintenant utiliser l'*inégalité de Hölder*, qui permet d'écrire que si $p \in]1, +\infty[$, alors en notant

$q = \frac{1}{1-1/p}$ le *conjugué de Hölder* de p (i.e. q tel que $\frac{1}{p} + \frac{1}{q} = 1$), on a pour tout $\mathbf{x}$ et $\mathbf{y}$ vecteurs de $\mathbb{R}^n$:

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q \right)^{1/q}$$

Au passage, on remarque que, pour $p = 2$, on a également $q = 2$ et donc que l'inégalité se réduit alors à la bien connue *inégalité de Cauchy–Schwarz*. L'inégalité de Hölder peut donc être vue comme une généralisation de l'inégalité de Cauchy–Schwarz.

On applique l'inégalité de Hölder aux deux sommes du membre de droite. Pour la première :

$$\sum_{i=1}^{n} |a_i||a_i+b_i|^{p-1} \leq \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} {\scriptstyle |a_i+b_i|^{(p-1)q}} \right)^{1/q}$$

Comme $(p-1)q = p$ (il suffit de multiplier des deux côtés par pq pour le voir), cela s'écrit :

$$\sum_{i=1}^{n} |a_i||a_i+b_i|^{p-1} \leq \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |a_i + b_i|^p \right)^{1/q}$$

On fait pareil avec l'autre et on met $\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/q}$

en facteur. On obtient :

$$\sum_{i=1}^{n} |a_i+b_i|^p \leq \left(\left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} + \left(\sum_{i=1}^{n} |b_i|^p \right)^{1/p} \right) \left(\sum_{i=1}^{n} |a_i+b_i|^p \right)$$

Pour conclure, il ne reste plus qu'à diviser les deux membres de cette inégalité par $\left(\sum_{i=1}^{n} |a_i + b_i|^p\right)^{1/q}$ et à utiliser le fait que $1 - 1/q = 1/p$.

Néanmoins, nous ne sommes pas au bout de nos peines car il nous reste à démontrer l'inégalité de Hölder. Cette inégalité repose sur l'inégalité suivante :

$$\forall p \in]1, +\infty[, \ \forall a, b > 0, \ ab \leq \frac{a^p}{p} + \frac{b^q}{q} \, ,$$

où q est toujours le conjugué de Hölder de p. La façon la plus simple de s'en convaincre est de passer par le logarithme : par concavité,

$$\log\left(\frac{a^p}{p}+\frac{b^q}{q}\right) \geq \frac{\log\left(a^p\right)}{p}+\frac{\log\left(b^q\right)}{q} = \log(a)+\log(b) = \log(ab)$$

D'où l'inégalité, par croissance de l'exponentielle.

Au passage, l'inégalité $\forall x, y > 0, \ \forall t \in [0, 1]$, $f(tx + (1-t)y) \geq tf(x) + (1-t)f(y)$ est la définition de la concavité de f, mais on peut montrer que si f est concave, alors elle vérifie une relation plus générale appelée *inégalité de Jensen* : soit f une fonction concave et soient (μ_i) positifs tels que $\sum_i \mu_i = 1$. Alors :

$$\sum_i \mu_i f(x_i) \leq f\left(\sum_i \mu_i x_i\right)$$

Cette inégalité, qui peut encore être généralisée en remplaçant la somme pondérée par l'intégrale par rapport à une mesure, est très utilisée en probabilités sous la forme :

$$\mathbb{E}\left(f(X)\right) \leq f\left(\mathbb{E}\left(X\right)\right) ,$$

pour tout fonction f concave (on a bien entendu l'inégalité dans l'autre sens pour les fonctions convexes).

Nous avons bientôt terminé. Pour conclure, nous allons maintenant appliquer l'inégalité $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ $(a, b > 0)$ terme à terme en sommant chacun des termes :

$$|x_i'| = \frac{|x_i|}{\left(\sum |x_i|^p\right)^{1/p}} \quad \text{et} \quad |y_i'| = \frac{|y_i|}{\left(\sum |y_i|^q\right)^{1/q}} .$$

On obtient :

$$\sum_i |x'_i y'_i| \leq \frac{1}{p} \sum_i |x'_i|^p + \frac{1}{q} \sum_i |y'_i|^q$$

Or $\sum_i |x'_i|^p = 1$ et $\sum_i |y'_i|^q = 1$. Comme on a également $\frac{1}{p} + \frac{1}{q} = 1$, le terme de droite est égal à 1 et l'inégalité devient :

$$\sum_i |x'_i y'_i| = \frac{\sum_i |x_i y_i|}{\left(\sum_i |x_i|^p\right)^{1/p} \left(\sum_i |y_i|^q\right)^{1/q}} \leq 1\,,$$

dont on déduit l'inégalité de Hölder.

On a enfin terminé notre preuve ! Finalement, les calculs ont peu d'intérêt, mais ils nous ont permis de voir les grandes inégalités classiques : Minkowski, Hölder, Cauchy-Schwarz et Jensen.

À noter que toutes ces inégalités se généralisent en remplaçant les sommes par des intégrales. On peut donc définir une distance entre fonctions L^p (i.e. f telle que $|f|^p$ est intégrable) par la formule suivante :

$$d(f,g) = \left(\int |f(x) - g(x)|^p \, dx \right)^{1/p}$$

Les espaces munis de cette distance sont d'une importance considérable.

Enfin, à noter que puisque :

$$a_1^p + \cdots + a_n^p \underset{p\to\infty}{\sim} \left(\max_i a_i\right)^p ,$$

on a

$$\left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p} \xrightarrow[p\to\infty]{} \max_i |x_i - y_i| ,$$

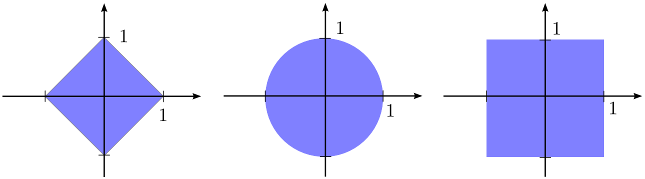
ce qui justifie la notation d_∞ que nous utiliserons à l'exercice suivant.

Exercice 2 : exemples de boules

Représenter les boules ouvertes de rayon 1 centrées en 0 associées aux distances sur $\mathbb{R}^2$ *:*

$$\begin{aligned}
d_1 &: \big((x_1, y_1), (x_2, y_2)\big) \mapsto |x_1 - x_2| + |y_1 - y_2| \\
d_2 &: \big((x_1, y_1), (x_2, y_2)\big) \mapsto |x_1 - x_2|^2 + |y_1 - y_2|^2 \\
d_\infty &: \big((x_1, y_1), (x_2, y_2)\big) \mapsto \max\left(|x_1 - x_2|, |y_1 - y_2|\right)
\end{aligned}$$

On se convainc facilement qu'il s'agit des ensembles suivants :



$d_1 < 1$ $\quad$ $d_2 < 1$ $\quad$ $d_\infty < 1$

Exercice 3 : intérieur, frontière et adhérence

1. *Montrer que* $\overline{A}$ *est l'ensemble des points à distance nulle de* A. *(*$\star$*)*

2. *Montrer que* $\forall A \subset E$, $E = \mathring{A} \sqcup \partial A \sqcup \left(\overline{A}\right)^{\mathsf{c}}$ *(où* $\sqcup$ *désigne l'union disjointe), et que* $\mathring{A} = \left(\overline{A^{\mathsf{c}}}\right)^{\mathsf{c}}$. *Remarquer que cela implique notamment :*

   (a) $\overline{A} = \mathring{A} \sqcup \partial A$

   (b) $\overline{A^{\mathsf{c}}} = \partial A \sqcup \left(\overline{A}\right)^{\mathsf{c}}$

   ($\star\star$*)*

1) $x \in \overline{A} \iff \forall r > 0,\ \exists y \in B_r(x) \cap A \iff \forall r > 0,\ d(x, A) \leq r \iff d(x, A) = 0.$

2) On va considérer un élément x de E et montrer qu'il appartient forcément soit à $\mathring{A}$, soit à ∂A, soit à $\left(\overline{A}\right)^{\mathsf{c}}$ en faisant une disjonction de cas. Tout d'abord, deux cas se présentent :

1. Soit $\exists B_r(x) \subset A$. Or ceci équivaut à $x \in \mathring{A}$.

2. Soit $\forall B_r(x)$, $B_r(x) \cap A^{\mathsf{c}} \neq \varnothing$. Dans ce cas, on distingue deux sous-cas :

   (a) Soit $\exists B_s(x) \subset A^{\mathsf{c}}$. Ceci équivaut au fait que $x \notin \overline{A}$, i.e. $x \in \left(\overline{A}\right)^{\mathsf{c}}$.

   (b) Soit $\forall B_s(x)$, $B_s(x) \cap A \neq \varnothing$. Couplé au fait que $\forall B_r(x)$, $B_r(x) \cap A^{\mathsf{c}} \neq \varnothing$, cela équivaut à $x \in \partial A$.

Pour montrer que $\mathring{A} = \left(\overline{A^{\mathsf{c}}}\right)^{\mathsf{c}}$, on peut montrer que $\left(\mathring{A}\right)^{\mathsf{c}} = \overline{A^{\mathsf{c}}}$. Or ceci est évident puisque, par définition, $x \notin \mathring{A} \iff \forall B_r(x),\ B_r(x) \cap A^{\mathsf{c}} \neq \varnothing \iff x \in \overline{A^{\mathsf{c}}}$.

Une fois qu'on dispose de ces relations, (a) et (b) se déduisent immédiatement du fait que $\forall X \subset E$, $E = X \sqcup X^{\mathsf{c}}$: il suffit alors de choisir $X = \overline{A}$ pour (a) et $X = \mathring{A}$ pour (b), en utilisant que $\mathring{A} = \left(\overline{A^{\mathsf{c}}}\right)^{\mathsf{c}}$.

Exercice 4 : liens entre ouverts et fermés

1. *Commencer par vérifier que les intervalles ouverts sont des ouverts et que les intervalles fermés sont des fermés (pour $\mathbb{R}$ muni de la distance usuelle).* $(\star)$

2. *Trouver des exemples d'ensembles qui ne sont ni ouverts, ni fermés.* $(\star)$

3. *Trouver des exemples d'ensembles qui sont à la fois ouverts et fermés.* $(\star\star)$

(penser à préciser la distance considérée).

1) Soit $x \in]a,b[$. Soit $0 < \varepsilon < \min(x-a, b-x)$. Il est clair que $B_\varepsilon(x) \subset]a,b[$, donc $]a,b[$ est ouvert.

Soit maintenant $x \in [a,b]^\mathsf{c}$. Le même procédé (avec $\varepsilon = \min(|x-a|, |x-b|)$) permet de montrer que $[a,b]^\mathsf{c}$ est ouvert, et donc que $[a,b]$ est fermé.

2) Plaçons-nous dans $\mathbb{R}$ muni de la distance usuelle, et considérons tout d'abord l'ensemble $]0,1]$. Cet ensemble n'est pas ouvert car $\forall r > 0$, $B_r(1) \not\subset]0,1]$. Il n'est pas non plus fermé car son complémentaire est $]-\infty, 0] \cup]1, +\infty[$, qui est pas un ouvert car il ne contient aucune boule centrée en 0.

Un exemple un peu plus amusant : on se place toujours dans $\mathbb{R}$ muni de la distance usuelle, et on considère maintenant $\mathbb{Q}$. Clairement, il ne s'agit pas d'un ouvert puisque toute boule centrée sur un rationnel contient un réel non-rationnel... Mais il n'est pas non plus fermé car de même, toute boule centrée sur un réel non-rationnel contient un rationnel ! Nous reviendrons sur cet exemple...

3) Soit $\mathbb{R}$ muni de la distance usuelle. On considère tout d'abord l'ensemble $\mathbb{R}$ tout entier. Clairement, pour tout point x, on peut trouver une boule centrée et x (et en plus, de rayon arbitrairement grand !) incluse dans $\mathbb{R}$. Donc $\mathbb{R}$ est ouvert. Son complémentaire dans $\mathbb{R}$ n'est autre que l'ensemble vide, $\varnothing$. Comme il est impossible de choisir un x dans $\varnothing$ pour invalider la propriété définissant un ouvert, on en déduit que $\varnothing$ est ouvert – et donc que $\mathbb{R}$ est fermé. De même, $\varnothing$ est à la fois ouvert et fermé. Plus généralement, lorsqu'on se donne un espace topologique E, E et $\varnothing$ sont toujours à la fois ouverts et fermés – nous reviendrons là-dessus dans la section 3 de ce document...

Exercice 5 : opérations sur les ouverts et les fermés

1. *Montrer qu'une réunion finie d'ouverts / fermés le reste. Même question pour une intersection finie.* ($\star$)

2. *Qu'en est-il des intersections / réunions dénombrables ? Quelconques ?* ($\star\star$)

1) On commence par montrer la propriété pour les ouverts.

Considérons le cas de la réunion : soient $O_1, \ldots, O_n$ des ouverts. Soit $x \in U = O_1 \cup \cdots \cup O_n$. Par définition, il existe i tel que $x \in O_i$. Puisque O_i est ouvert, il existe une boule $B_r(x) \subset O_i$. Mais comme $O_i \subset U$, on a également $B_r(x) \subset U$. On vient de montrer que $\forall x \in U$, $\exists B_r(x) \subset U$, i.e. U est ouvert.

Intéressons-nous maintenant à l'intersection I de $O_1, \ldots, O_n$: soit $x \in I$. Par définition, $\forall i \in \{1, \ldots, n\}$, $x \in O_i$. Les O_i étant des ouverts, $\exists\, r_1, \ldots, r_n$ strictement positifs tels que $\forall i$, $B_{r_i}(x) \subset$

O_i. Posons maintenant $r = \min_i r_i$. En tant que minimum d'un nombre fini de réels strictement positifs, $r > 0$. Considérons la boule de rayon r centrée en x. $\forall i, B_r(x) \subset B_{r_i}(x) \subset O_i$. Donc $B_r(x) \subset I = O_1 \cap \ldots \cap O_n$, ce qui termine la preuve.

On pourrait procéder exactement de la même manière pour les fermés, en re-faisant tout comme des bourrins... Mais on va être plus subtils et utiliser les lois de De Morgan [16] :

Rappel : les lois de De Morgan

Pour tout ensemble d'index I (dénombrable ou non),

$$\left(\bigcup_{i\in I} A_i\right)^{\mathsf{c}} = \bigcap_{i\in I} A_i^{\mathsf{c}}$$
$$\left(\bigcap_{i\in I} A_i\right)^{\mathsf{c}} = \bigcup_{i\in I} A_i^{\mathsf{c}}$$

Munis de cela, il suffit de remarquer que si $F_1, \ldots, F_n$ sont des fermés, $F_1^{\mathsf{c}}, \ldots, F_n^{\mathsf{c}}$ sont des ou-

16. Oui, c'est comme pour la formule de De Moivre, c'est bien *de De* Morgan et non de Morgan : il y a bel et bien deux "de" (ou, si vous préférez, "de De" ☺)...

verts. Par conséquent, leur union (resp. intersection) en sont également, d'après ce que nous venons de montrer. Il suffit alors d'appliquer les lois de De Morgan pour trouver que le complémentaire de l'intersection (resp. union) de $F_1, \ldots, F_n$ est ouverte, et donc que cet ensemble est fermé.

2) Commençons encore par nous intéresser au cas des ouverts. Le lecteur pourra vérifier que la démonstration faite dans le cas de l'union peut être transposée telle quelle au cas infini – y compris au cas infini non dénombrable.

En revanche, le cas de l'intersection pose problème ; l'étape problématique est la suivante : "$r = \min_i r_i$. En tant que minimum d'un nombre *fini* de réels strictement positifs, $r > 0$". Cette propriété n'est pas vraie dans le cas infini. Par exemple, les réels $(1/n)_{n\in\mathbb{N}^*}$ sont bien tous strictement positifs, et pourtant, $\inf_{n\in\mathbb{N}^*} 1/n = 0$. On peut donc penser au contre-exemple à la propriété suivant :

$$\bigcap_{n\in\mathbb{N}^*} B_{1/n}(0) = \{0\}$$

or $\{0\}$ n'est pas ouvert, bien que les $B_{1/n}(0)$ le soient.

Comme on l'avait fait précédemment, on peut utiliser les lois de De Morgan pour construire un contre-exemple dans le cas de l'union de fermés. Mais il existe un contre exemple encore plus évident : considérons $\mathbb{R}$ muni de la topologie usuelle. Clairement, $\forall x \in \mathbb{R}$, $\{x\}$ est un fermé. Considérons alors un ensemble $A \subset \mathbb{R}$ non fermé ; alors :

$$\bigcup_{x \in A} \{x\} = A$$

n'est pas un fermé...

Finalement, ce qu'il faut retenir c'est que les ouverts/fermés sont stables par union et intersection **sauf** pour l'intersection infinie dans le cas des ouverts et pour l'union infinie dans le cas des fermés. Nous reviendrons sur ce point...

Exercice 6 : caractérisation des fermés

1. *Montrer que $A = \overline{A}$ ssi A est fermé.* $(\star)$

2. *Montrer que A est ouvert si et seulement si toute suite d'éléments de A qui converge a sa limite dans A.* $(\star\star)$

Remarque : les raisonnements qui suivent sont élémentaires, mais peuvent s'avérer particulièrement "buggants". J'ai cherché à écrire des démonstrations les plus simples possibles, mais ça ne veut pas dire que ce sont les premières qui me sont venues à l'esprit ! Notamment, en cas de bug, ne surtout pas hésiter à retourner ces démonstrations pour les mettre sous une forme "par l'absurde" qui sera peut-être plus naturelle... Et surtout, penser à faire des dessins !

1) Commençons par remarquer qu'on a toujours $A \subset \overline{A}$: en effet, $\forall x \in A$, $\forall B_r(x)$, $x \in A \cap B_r(x)$ donc, par définition de $\overline{A}$, $x \in \overline{A}$.

Montrons maintenant que lorsque A est fermé, on a en plus l'inclusion réciproque : Soit $x \in \overline{A}$. Par définition, $\forall B_r(x)$, $B_r(x) \cap A \neq \varnothing$, i.e. $B_r(x) \not\subset A^\mathsf{c}$. Comme A est fermé, il est impossible que x appartienne à A^c car il cela contredirait l'existence d'une boule centrée en x incluse dans A^c, existence qui est assurée par le fait que A^c est ouvert. On en déduit donc que $x \in A$. D'où $\overline{A} \subset A$ lorsque A est fermé.

Montrons maintenant que $\overline{A} = A \implies A$ fermé : Soit $x \in A^\mathsf{c}$. Comme $\overline{A} = A$, on en déduit $x \notin \overline{A}$. Cela signifie qu'il existe une boule $B_r(x)$ centrée en x qui n'intersecte pas A, i.e. qui est incluse dans A^{c}. On vient donc de montrer que A^{c} est ouvert, c'est-à-dire que A est fermé.

On déduit de cette propriété que $\overline{A}$ est le plus petit fermé contenant A : d'une part $\overline{A}$ est bien un fermé contenant A. Maintenant, considérons un autre fermé B contenant A. $A \subset B \implies \overline{A} \subset \overline{B}$. Mais B étant fermé, $\overline{B} = B$ d'où $\overline{A} \subset B$.

2) Commençons par montrer que A fermé implique que toute suite d'éléments de A qui converge (au sens de la définition donnée dans ce document, bien entendu) a sa limite dans A : soit $(u_n) \in A^{\mathbb{N}}$ telle que $u_n \to \ell$, c'est-à-dire

$$\forall \delta,\, \exists N \text{ t.q. } \forall n \geq N, u_n \in B_\delta(\ell)\,.$$

Supposons maintenant que $\ell \in A^{\mathsf{c}}$. Puisque A est fermé, il existe donc $B_r(\ell) \subset A^{\mathsf{c}}$. D'après la convergence de (u_n), on pourrait trouver N tel que $\forall n \geq N$, $u_n \in B_r(\ell)$, i.e. $u_n \in A^{\mathsf{c}}$. Ceci est impossible, puisque $\forall n$, $u_n \in A$. On en déduit donc que $\ell \in A$.

Montrons maintenant la réciproque. Supposons que A ne soit pas fermé. Alors, il existe un élément

$x \in A^{\mathsf{c}}$ tel que $\forall B_r(x)$, $B_r(x) \not\subset A^{\mathsf{c}}$, i.e. $B_r(x) \cap A \neq \varnothing$. On va montrer qu'on peut alors construire une suite d'éléments de A admettant x pour limite : soit (ε_n) une suite de réels strictement positifs décroissant vers 0. $\forall \varepsilon_n$, $\exists y_n \in A \cap B_{\varepsilon_n}(x)$. Clairement, (y_n) converge vers x. Or il s'agit par construction d'une suite d'éléments de A. Par contraposée, on en déduit l'implication cherchée.

Exercice 7 : nulle part densité et négligeabilité

Trouver un exemple d'ensemble nulle part dense de mesure de Lebesgue strictement positive (c'est-à-dire, si l'on travaille sur $\mathbb{R}$, de longueur non-nulle). ($\star\star\star$)

On va construire un tel ensemble $A \subset [0, 1]$ à la manière de l'ensemble triadique de Cantor, c'est-à-dire au moyen d'un processus itératif de retrait d'intervalles. Mais on va retirer ces intervalles de manière à s'assurer de (1) conserver une mesure strictement positive et (2) d'obtenir un ensemble nulle part dense.

Commençons par nous préoccuper de la manière dont nous allons pouvoir conserver une mesure strictement positive : on sait que la longueur de l'ensemble final sera supérieure à

$$1 - \sum_{n=1}^{+\infty} \text{longueur enlevée à l'étape } n\,,$$

avec égalité si et seulement si il n'y a pas de chevauchement des intervalles retirés d'une étape sur l'autre. On voit donc que la manière la plus simple de s'assurer de conserver une longueur strictement positive est que la série des longueurs enlevées à chaque étape converge et soit strictement inférieure à 1. Personnellement, la série convergente la plus simple que je connaisse (si l'on excepte la série de terme général égal à zéro) est la série de terme général $1/2^n$. En sommant à partir de $n = 2$, cette somme vaut $1/2$. Bref,

$$\sum_{n=1}^{+\infty} \frac{1}{2^{n+1}} = \frac{1}{2} < 1\,,$$

donc une procédure retirant une longueur d'intervalle égale à $1/2^{n+1}$ à l'étape n semble prometteuse...

Se pose maintenant la question de savoir comment faire pour que l'ensemble A résultant soit nulle part dense. Essayons de voir ce qui pourrait faire qu'il ne le soit pas : il faudrait qu'il existe un ouvert dans lequel A serait dense. Cet ouvert contiendrait alors un intervalle $]a, b[$, dans lequel A serait également dense. Donc si on ne peut pas trouver de tel intervalle $]a, b[$, A ne peut être dense dans aucun ouvert et donc est bien nulle part dense. Il nous suffit donc de retirer des intervalles de manière à ce que tout intervalle $]a, b[$ se voie privé d'un intervalle I ouvert de longueur non nulle : ainsi, aucun point $x \in I$ ne pourra être approché autant que souhaité par des éléments de A. Maintenant, si lors de la construction on enlève des intervalles autour de chaque point d'un ensemble dense dans $[0, 1]$, on est sûr que tout intervalle $]a, b[$ contiendra un de ces points – et donc se verra bien privé d'un intervalle ouvert de longueur non nulle.

Il ne nous reste plus qu'à terminer la construction, en précisant quel ensemble de points considérer pour retirer les intervalles, et quelle longueur donner à ces intervalles.

Le premier ensemble dense dans $[0, 1]$ qui vient à l'esprit est $[0, 1] \cap \mathbb{Q}$, mais ce n'est pas le plus simple à manipuler. Notamment, même si on sait que $\mathbb{Q}$ est dénombrable et donc que ses éléments peuvent être indexés par des entiers naturels, ce n'est pas forcément évident à faire (vous pouvez essayer ; vous verrez que le fait que l'écriture sous forme de fractions ne soit pas unique (ex : $\frac{2}{4} = \frac{1}{2}$) introduit une petite difficulté. Et surtout, une fois qu'on l'a fait, l'ordre dans lequel on parcours les points de $[0,1] \cap \mathbb{Q}$ n'est pas forcément très intuitif géométriquement...

Heureusement, il existe des ensembles denses dans $[0,1]$ bien plus simples à manipuler, comme les fractions dyadiques, c'est-à-dire les rationnels de la forme :

$$\frac{a}{2^n},$$

avec $a \in \{1, \ldots, 2^n - 1\}$. Pour éviter de considérer plusieurs fois le même réel, on doit en plus imposer que a ne soit pas un multiple de 2. En fait cet ensemble n'est pas si étrange que cela : il s'agit tout simplement des fractions $\frac{1}{2}$ $(n = 1)$; $\frac{1}{4}$, $\frac{3}{4}$ $(n = 2)$, etc, dont la répartition dans $[0,1]$ est très naturelle. On montre facilement que cet ensemble est dense dans $[0,1]$ en procédant exactement comme lorsqu'on a montré que les rationnels étaient denses dans $\mathbb{R}$ (mais avec l'approximation $x_n = \lfloor x \times 2^n \rfloor / 2^n$ – de la même manière que ce qu'on avait fait consistait à tronquer l'écriture à la n-ième décimale, cela revient à tronquer l'écriture en base 2).

Maintenant qu'on a un bon candidat d'ensemble de points autour desquels retirer nos intervalles, ils ne nous reste plus qu'à préciser la longueur de ces intervalles de manière à ce que la somme de la série des longueurs retirées soit inférieure strictement à 1. On a vu qu'on aimerait retrancher une longueur d'au plus $1/2^{n+1}$ à l'étape n. Or à chaque étape n, on considère 2^{n-1} nouvelles fractions dyadiques. Si les intervalles qu'on retire ont tous la même longueur ℓ_n et qu'on veut retirer une longueur totale de $1/2^{n+1}$, on doit avoir $\ell_n = 1/2^{2n}$. Le plus simple est alors de retrancher les intervalles de la forme :

$$\left]\frac{a}{2^n} - \frac{1}{2^{2n+1}}, \frac{a}{2^n} + \frac{1}{2^{2n+1}}\right[.$$

Finalement, on pose donc :

$$A = [0,1] \setminus \left(\bigcup_{n=1}^{+\infty} \bigcup_{a \in I_n} \left]\frac{a}{2^n} - \frac{1}{2^{2n+1}}, \frac{a}{2^n} + \frac{1}{2^{2n+1}}\right[\right),$$

où $I_n = \{a \in \{1, \ldots, 2^n - 1\} \mid a \text{ impair}\}$.

Étant donnée la justification de cette construction, on vérifie facilement que A est bien nulle part dense dans $[0,1]$ et de mesure de Lebesgue supérieure à $1/2$ (strictement, à cause du chevauchement des intervalles retranchés d'une étape sur l'autre).

Ainsi, bien que la plupart des éléments de $[0,1]$ ne puissent pas être approchés par des éléments de A (par exemple, il n'y a aucun élément de A à distance inférieure à $1/8$ de $1/2$), si l'on choisit un réel de façon uniforme sur $[0,1]$, on a plus d'une chance sur deux de choisir un élément de A !

Bien entendu, cette construction n'a rien de "particulier". Notamment, on aurait pu choisir n'importe quelle autre série convergente de somme strictement inférieure à 1 et/ou n'importe quel autre ensemble dénombrable dense dans $[0,1]$ et ainsi construire des sous-ensembles de $[0,1]$ nulle part denses et de mesure de Lebesgue aussi proche de 1 que souhaité (mais strictement inférieure à 1, je pense – à vérifier !).

Exercice 8 : prolongement par continuité

Montrer que si deux applications continues coïncident sur une partie dense de E, alors elles coïncident sur E tout entier. ($\star\star$)

Remarque : ce théorème est assez simple à montrer en “fonçant dans le tas”, mais il existe aussi une démonstration très élégante...

Pour montrer ce résultat, on va commencer par montrer deux lemmes qui rendront la démonstration particulièrement élégante.

Lemme 1 : L'image réciproque d'un ouvert par une application continue est un ouvert.

En effet, soit A un ouvert et soit $x \in f^{-1}(A)$. Par définition, on a alors $f(x) \in A$. A étant ouvert, $\exists B_\varepsilon(f(x)) \subset A$. Mais comme f est continue, on a par définition que $\exists B_\delta(x)$ telle que $\forall y \in B_\delta(x)$, $f(y) \in B_\varepsilon(f(x))$ – et a fortiori, $f(y) \in A$. Finalement, on a montré que $\forall x \in f^{-1}(A)$, $\exists B_\delta(x) \subset f^{-1}(A)$, c'est-à-dire que $f^{-1}(A)$ est ouvert.

Ici, on n'a montré qu'une implication car c'est tout ce dont on a besoin ; mais on vérifie facilement qu'il s'agit d'une équivalence, qui peut donc être utilisée pour définir les applications continues. Le recours à ce lemme n'est donc pas vraiment une "astuce"...

Lemme 2 : Soient f et g deux applications continues de E dans F. Alors l'ensemble $\{f = g\} = \{x \in E \mid f(x) = g(x)\}$ sur lequel f et g coïncident est un fermé.

Soit $x \in \{f = g\}^{\mathsf{c}}$, i.e. $f(x) \neq g(x)$. Alors $\exists A$, B ouverts tels que $f(x) \in A$, $g(x) \in B$ et $A \cap B = \varnothing$ (il suffit de considérer $B_r(f(x))$ et $B_r(g(x))$, avec $r \leq d(x, y)/2$). Considérons $C = f^{-1}(A) \cap g^{-1}(A)$. C est ouvert, en tant qu'intersection finie d'ouverts (Exercice 5 + Lemme 1). De plus, $f(C) \subset A$ et $g(C) \subset B$ donc comme $A \cap B = \varnothing$, on en déduit que $\forall y \in C$, $f(y) \neq g(y)$, i.e. $C \subset \{f = g\}^{\mathsf{c}}$. On a bien exhibé un ouvert contenant x inclus dans $\{f = g\}^{\mathsf{c}}$, donc $\{f = g\}^{\mathsf{c}}$ est ouvert et $\{f = g\}$ est fermé.

J'ai rédigé la preuve de cette manière car, pour des raisons que j'expliquerai dans la section 3, j'avais envie d'insister sur un certain argument (devinez lequel). Mais il y a une autre façon de procéder, à laquelle il faut penser car elle est (à mon sens) plutôt plus simple :

Soit (x_n) une suite de points de $\{f = g\}$ qui converge vers une limite x. Puisque $\forall n$, $x_n \in \{f = g\}$, on a également $\forall n$, $f(x_n) = g(x_n)$. Par passage à la limite dans cette égalité, $\lim\limits_{+\infty} f(x_n) = \lim\limits_{+\infty} g(x_n)$. Mais f et g étant continues, par la caractérisation séquentielle de la continuité :

$$\lim_{+\infty} f(x_n) = f\Big(\lim_{+\infty} x_n\Big) = f(x) \quad \text{et} \quad \lim_{+\infty} g(x_n) = g\Big(\lim_{+\infty}$$

d'où $f(x) = g(x)$, c'est-à-dire $x \in \{f = g\}$. Ainsi, toute suite d'éléments de $\{f = g\}$ qui converge a sa limite dans $\{f = g\}$, et donc $\{f = g\}$ est fermé.

Bref, une fois qu'on dispose du Lemme 2, la preuve du théorème est immédiate, même si elle n'est pas forcément évidente à voir... Prenez le temps d'y réfléchir avant de poursuivre votre lecture !

$$\{f = g\} \text{ est fermé, donc } \{f = g\} = \overline{\{f = g\}}$$

$$\text{Or } \overline{\{f = g\}} = E, \text{ car } \{f = g\} \text{ est dense dans } E.$$

$$\text{D'où } \{f = g\} = E.$$

Exercice 9 : topologie de la convergence simple

Le but de cet exercice est de montrer que la topologie de la convergence simple n'est pas métrisable dans le cas général. On considère les ensembles :

$$B_\varepsilon^x(f) = \left\{ g \in \mathbb{R}^E \text{ } t.q. \text{ } |g(x) - f(x)| < \varepsilon \right\}$$

et on note $\mathcal{T}$ la topologie engendrée par la famille $\left(B_\varepsilon^x(f)\right)_{(x,\,\varepsilon,f)}$ lorsque $(x,\,\varepsilon, f)$ décrit $E \times \mathbb{R}_+^* \times \mathbb{R}^E$.

1. *Montrer que $\mathcal{T}$ est bien la topologie correspondant à la convergence simple.($\star$)*

Soient maintenant a et b deux réels distincts. Considérons l'ensemble :

$$A = \left\{ f \in \mathbb{R}^E \text{ } t.q. \text{ } f = a \text{ sauf en un nombre fini de poin}$$

2. *Montrer que $f : x \mapsto b \in \overline{A}$. ($\star$)*

3. *Discuter si f est limite simple de fonctions de A. ($\star\star$)*

4. *Conclure. ($\star$)*

1) On procède par double implication. Commençons par montrer que la convergence au sens de la topologie $\mathcal{T}$ engendrée par les $B_\varepsilon^x(f)$ implique la convergence simple : Soit $x \in E$. Soit $\varepsilon > 0$. Puisque $f_n \to \ell$ au sens de $\mathcal{T}$, cela signifie que $\forall O \in \mathcal{T}$ contenant ℓ, $\exists N$ tel que $n \geq N \implies f_n \in O$. En particulier, $B_\varepsilon^x(\ell)$ est un tel ouvert. Donc $\exists N$ tel que $n \geq N \implies f_n \in B_\varepsilon^x(\ell)$, i.e. $|f_n(x) - \ell(x)| < \varepsilon$. Ainsi, la convergence au sens de $\mathcal{T}$ implique la convergence simple.

Montrons maintenant la réciproque. Soit un ouvert O de $\mathcal{T}$ contenant ℓ. Par définition de $\mathcal{T}$, O s'écrit :

$$O = \bigcup_\alpha \bigcap_{i=1}^{n_\alpha} B_{\varepsilon_{\alpha,i}}^{x_{\alpha,i}}\left(g_{\alpha,i}\right),$$

et le fait que O contienne ℓ signifie que

$$\exists \alpha \text{ t.q. } \forall i,\, \ell \in B_{\varepsilon_{\alpha,i}}^{x_{\alpha,i}}\left(g_{\alpha,i}\right).$$

On s'intéresse alors à l'intersection correspondant à ce α (qu'on arrête d'indiquer pour alléger les notations) :

$$\forall i,\, \ell \in B_{\varepsilon_i}^{x_i}\left(g_i\right) \text{ i.e. } |\ell(x_i) - g_i(x_i)| = \varepsilon_i - \eta_i$$

Comme f_n converge simplement vers ℓ, $\exists N_i$ tel que $n \geq N_i \implies |f_n(x_i) - \ell(x_i)| < \eta_i$. Par inégalité triangulaire :

$$|f_n(x_i) - g_i(x_i)| \leq \underbrace{|f_n(x_i) - \ell(x_i)|}_{<\eta_i} + \underbrace{|\ell(x_i) - g_i(x_i)|}_{\varepsilon_i - \eta_i} < \varepsilon_i$$

Bref, $\exists N_i$ tel que $\forall n \geq N_i$, $f_n \in B_{\varepsilon_i}^{x_i}(g_i)$. Il suffit alors de prendre $N = \max_i N_i$ et on a bien $\forall n \geq N$, $\forall i$, $f_n \in B_{\varepsilon_i}^{x_i}(g_i)$ – et donc $f_n \in O$. Donc la convergence simple implique la convergence au sens de $\mathcal{T}$.

2) On veut montrer que tout ouvert contenant f contient également un élément de A. Comme à la question précédente, considérons donc un ouvert

$$O = \bigcap_{i=1}^{n} B_{\varepsilon_i}^{x_i}(g_i)$$

tel que $\forall i$, $|g_i(x_i) - b| < \varepsilon_i$. Considérons maintenant la fonction h valant a sur $E \setminus \{x_1, \ldots, x_n\}$ et b sur $\{x_1, \ldots, x_n\}$. Clairement, $h \in O$, et $h \in A$.

3) Si E est fini ou dénombrable, rien ne s'oppose à ce que f soit une limite d'éléments de A. En revanche, cela n'est pas possible lorsque E est infini non-dénombrable... En effet, soit :

$$X_n = \{x \in E \text{ t.q. } f_n(x) \neq a\}$$

Puisque $\forall n$, $f_n \in A$, on en déduit que $\forall n$, X_n est fini. Soit maintenant $X = \cup_{n=1}^{\infty} X_n$. Si f_n tend simplement vers f, alors $\forall x \in E$, $\exists N_x$ tel que $\forall n \geq N_x$, $f_n(x) \neq a$ et donc $X = E$. Or, en tant qu'union dénombrable d'ensembles finis, X est au plus dénombrable. Si E n'est pas dénombrable, on a donc une contradiction.

4) On a vu que dans un espace métrique, tout point de l'adhérence d'un ensemble est limite d'une suite d'éléments de cet ensemble (en fait on ne l'a pas vu de façon explicite donc si vous n'en êtes pas convaincus, n'hésitez pas à le re-démontrer). Or ici, avec f, on vient d'exhiber un élément qui appartient à $\overline{A}$ mais qui n'est limite d'aucune suite d'éléments de A. On ne peut donc avoir affaire à un espace métrique.

Exercice 10 : un exemple de suite de Cauchy non convergente

Plaçons nous dans $\mathbb{Q}$ muni de la distance usuelle, et considérons la suite (u_n) définie par :

$$\begin{cases} u_{n+1} = \frac{u_n}{2} + \frac{1}{u_n} \\ u_0 = \frac{3}{2} \ \textit{(par exemple)} \end{cases}$$

1. *Montrer que cette suite est de Cauchy.* ($\star\star$)
2. *Montrer que cette suite ne converge pas dans l'espace considéré.* ($\star\star$)

1) Pour montrer que cette suite est de Cauchy, on va montrer qu'elle converge dans $\mathbb{R}$: en effet, on a vu que toute suite convergente (pour une distance) est de Cauchy (pour cette même distance).

La façon la plus rapide de le faire c'est de remarquer que puisque $u_0 > 0$ et que $u_{n+1} = \frac{u_n}{2} + \frac{1}{u_n}$, on a par récurrence immédiate que $\forall n,\ u_n > 0$. Ensuite, on remarque que si (u_n) admet une limite, alors celle-ci vérifie $\ell = \frac{\ell}{2} + \frac{1}{\ell}$, c'est-à-dire $\ell^2 = 2$. Cela nous amène à nous intéresser à la quantité

$$u_{n+1}^2 - 2 = \left(\frac{u_n}{2} + \frac{1}{u_n}\right)^2 - 2 = \frac{u_n^2}{4} - 1 + \frac{1}{u_n^2} = \left(\frac{u_n^2 - 2}{2u_n}\right)^2$$

On en déduit que $\forall n,\ u_n^2 > 2$, c'est-à-dire (puisque $u_n > 0$) que $u_n > \sqrt{2}$. Or, on a également

$$u_{n+1} - u_n = -\frac{u_n}{2} + \frac{1}{u_n}\,.$$

Comme $\forall x > \sqrt{2},\ 1/x - x/2 < 0$, on en déduit que $\forall n,\ u_{n+1} - u_n < 0$ c'est-à-dire que (u_n) est décroissante. Décroissante et minorée, (u_n) converge dans $\mathbb{R}$ muni de la distance usuelle, i.e. $\exists \ell \in \mathbb{R}$ tel que $\forall \varepsilon > 0$, $\exists N$ tel que $\forall n > N$, $|u_n - \ell| < \varepsilon$. Comme on l'a fait dans le "cours", on en déduit par inégalité triangulaire que $\forall m, n > N$, $|u_m - u_n| < 2\varepsilon$, i.e. (u_n) est de Cauchy.

2) Il nous suffit de montrer que la limite ℓ de (u_n) n'appartient pas à $\mathbb{Q}$. Or a déjà vu que ℓ vérifie $\ell > 0$ et $\ell^2 = 2$, i.e. $\ell = \sqrt{2}$.

On sait bien que $\sqrt{2} \notin \mathbb{Q}$, mais on va quand même le montrer, histoire de faire les choses bien. La démonstration suivante de l'irrationalité de $\sqrt{2}$, que certains d'entre vous ont sans doute déjà vue au lycée, ne nécessite aucune connaissance particulière en arithmétique : par définition, les rationnels sont les nombres qui peuvent s'écrire sous la forme p/q avec p et q entiers (et q non nul bien sûr). Supposons alors que $\sqrt{2} = p/q$. Cette écriture n'est pas forcément unique ; on en choisit une telle que $\mathrm{PGCD}(p, q) = 1$. Une telle écriture existe, puisque que si $\mathrm{PGCD}(p, q) = k \neq 1$, on peut simplifier la fraction en divisant en haut et en bas par k et on obtient alors une écriture qui convient. Par définition de $\sqrt{2}$, on a $p^2/q^2 = 2$, i.e. $p^2 = 2q^2$. On en déduit que p est pair : en effet, si p était impair il s'écrirait $p = 2k+1$ et on aurait $p^2 = (2k+1)^2 = 4k^2+2k+1 = 2\,(2k^2+k) + 1$ qui serait impair. Donc $p = 2k$, i.e. $p^2 = 4k^2$. On peut alors simplifier l'égalité $p^2 = 2q^2$ qui devient $2k^2 = q^2$. On en déduit par le même raisonnement que q est également pair. Donc on a $\mathrm{PGCD}(p, q) \geq 2$, ce qui contredit $\mathrm{PGCD}(p, q) = 1$.

Si on admet l'unicité de l'écriture irréductible de tout rationnel, on peut donner une preuve encore plus rapide : on commence par supposer que $\sqrt{2}$ s'écrit p/q sous forme irréductible. On a alors $p^2/q^2 = 2$ et les deux membres de cette égalité sont sous forme irréductible ($2 = 2/1$). Par unicité de cette écriture, on en déduit $p^2 = 2$. Or il n'existe aucun entier naturel p vérifiant $p^2 = 2$, d'où la contradiction recherchée.

Pourquoi ai-je donné cet exercice, alors qu'on avait déjà vu un exemple de suite de rationnels tendant vers un irrationnel à la section 2.3 sur la densité ? D'une part, ça permet de réviser un peu les méthodes classiques d'étude de suites de réels, ce qui ne peut pas faire de mal ; d'autre part, il y a quand même une "astuce" que je trouve intéressante dans cet exercice : c'est de penser à montrer la convergence sur $\mathbb{R}$ pour montrer le caractère "de Cauchy" de la suite. La morale, c'est que ce n'est pas parce qu'on est n'est revenus aux fondements et qu'on a redéfini la notion de convergence qu'il faut oublier tout ce qu'on sait sur la convergence dans les cas usuels ! Enfin, cet exercice était surtout l'occasion de faire un point de culture générale :

- En rappelant la preuve classique de l'irrationalité de $\sqrt{2}$.

- En signalant qu'il s'agit d'exemples importants du point de vue de l'Histoire des maths : d'une part, le caractère irrationnel de $\sqrt{2}$ était connu des Grecs (cf notion d'incommensurabilité) et en a même tourmenté certains ; d'autre part, la suite présentée fournit une méthode de calcul de racine carrées appelée méthode de Héron, du nom de Héron d'Alexandrie qui l'a exposée clairement au I$^{\text{er}}$ siècle après J.-C. mais la méthode était probablement connue avant, voire aurait pu avoir été utilisée par les Babyloniens pour calculer l'approximation de $\sqrt{2}$ à 10^{-6} près de la tablette d'argile YBC 7289, tablette qui date de... $\sim$ 1700 avant J.-C. !

- En soulignant au passage le lien avec une méthode qu'on a mentionnée rapidement à l'oral : la méthode de Newton. Cette méthode permet, sous certaines conditions que je ne détaille pas, d'approcher les points d'annulation d'une fonction f au moyen de la suite :

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

On peut voir la suite qu'on a étudiée comme une application de la méthode de Newton à $f : x \mapsto x^2 - 2$.

Exercice 11 : preuve du théorème du point fixe

Soit E un espace métrique complet, et $\varphi : E \to E$ une application contractante. Montrer que φ admet un unique point fixe, i.e. $\exists! x^* \in E$ tel que $\varphi(x^*) = x^*$. (★★)

Indication : on pourra introduire la suite définie, $\forall x_0 \in E$, par $x_{n+1} = \varphi(x_n)$.

Je ne vais pas présenter ici la preuve classique, due à Banach (1922). Cette dernière présente l'avantage d'être assez "directe" (c'est-à-dire qu'elle ne requiert pas d'astuce), mais elle est légèrement calculatoire. À l'inverse, la preuve qui suit est très concise, mais... il fallait y penser ! De façon amusante, elle ne date que de 2007 et – pour ceux qui ne sauraient pas que ce type de journal existe – a été publiée dans le "Journal de la Théorie et des Applications du Point Fixe" [Palais, 2007].

Commençons par montrer un petit lemme :

Lemme :

$$\forall x, y \in E, \quad d(x, y) \leq \frac{d\big(\varphi(x),\, x\big) + d\big(\varphi(y),\, y\big)}{1-k}$$

En effet, par inégalité triangulaire,

$$d(x, y) \leq d(x, \varphi(x)) + d(\varphi(x), \varphi(y)) + d(\varphi(y), y)$$

Or puisque φ est k-contractante, $d(\varphi(x), \varphi(y)) \leq kd(x, y)$ – d'où le lemme.

Unicité : Grâce au lemme précédent, l'unicité du point fixe est immédiate. En effet, soient x^* et y^* deux points fixes, i.e. $\varphi(x^*) = x^*$ et $\varphi(y^*) = y^*$. En appliquant le lemme à x^* et y^* et en utilisant le fait que $d\big(\varphi(x^*),\, x^*\big) = d\big(\varphi(y^*),\, y^*\big) = 0$, on trouve que $d(x^*, y^*) \leq 0$, dont on déduit $d(x^*, y^*) = 0$, i.e. $x^* = y^*$.

Existence : Considérons la suite définie par :

$$x_{n+1} = \varphi(x_n),$$

et x_0 est un élément quelconque de E. On note φ^n la composée n fois de φ, de sorte que $x_n = \varphi^n(x_0)$.

On va montrer que la suite (x_n) est de Cauchy : pour tous rangs n et m,

$$\begin{aligned}
d\bigl(\varphi^n(x_0), \varphi^m(x_0)\bigr) &\leq \frac{d\bigl(\varphi(\varphi^n(x_0)),\, \varphi^n(x_0)\bigr) + d\bigl(\varphi(\varphi^m(x_0))}{1-k} \\
&\leq \frac{d\bigl(\varphi^n(\varphi(x_0)),\, \varphi^n(x_0)\bigr) + d\bigl(\varphi^m(\varphi(x_0))}{1-k} \\
&\leq \frac{k^n d\bigl(\varphi(x_0),\, x_0\bigr) + k^m d\bigl(\varphi(x_0),\, x_0\bigr)}{1-k} \\
&\leq \frac{k^n + k^m}{1-k}\, d\bigl(\varphi(x_0),\, x_0\bigr) \xrightarrow[n,m\to\infty]{} 0
\end{aligned}$$

Donc (x_n) est de Cauchy. E étant complet, elle admet une limite $x^* \in E$.

Montrons maintenant que x^* est un point fixe de φ : φ étant lipschitzienne, elle est continue. Or E est un espace métrique, d'où :

$$x^* = \lim_{n\to\infty} x_n = \underbrace{\lim_{n\to\infty} \varphi(x_{n-1}) = \varphi\left(\lim_{n\to\infty} x_{n-1}\right)}_{\text{caractéristation séquentielle de la continuité}} = \varphi(x^*)$$

Ce qui conclut la preuve.

Exercice 12 : compacts et fermés

Montrer les propriétés suivantes :

1. *Toute partie fermée d'un compact est compacte.* ($\star$)

2. *Tout compact est fermé.* ($\star\star$)

1) Soit A un fermé d'un compact E. Considérons un recouvrement (O_x) de A par des ouverts de la topologie induite sur A. On a :

$$A = \bigcup_x O_x .$$

Puisque A est fermé, A^c est ouvert. Or $A \cup A^\mathsf{c} = E$, d'où

$$E = A^\mathsf{c} \cup \left(\bigcup_x O_x \right) .$$

On a donc un recouvrement de E par des ouverts ; E étant compact, la propriété de Borel-Lebesgue nous permet d'affirmer l'existence d'un sous-recouvrement fini de E par ces ouverts. A^c fera forcément partie de ces ouverts (c'est le seul ouvert qui recouvre la partie

A^{c} de E, puisque les autres sont inclus dans A). On a donc :

$$E = A^{\mathsf{c}} \cup O_{x_1} \cup \cdots \cup O_{x_n}$$

d'où on déduit $A \subset O_{x_1} \cup \cdots \cup O_{x_n}$. On vient donc d'extraire un recouvrement fini du recouvrement (O_x) de A. On en déduit que A est compact.

2) *(je n'ai mis que deux étoiles à cette question car elle ne requiert aucune connaissance supplémentaire... Mais on s'est bien pris la tête dessus avec Jako !)*

Soit un compact A. Considérons $x \in A^{\mathsf{c}}$. Puisqu'on travaille sur un espace séparé, $\forall y \in A$, $\exists\, O_y$, O_x^y tels que $O_y \cap O_x^y = \varnothing$. Or, et c'est là l'"astuce" de cet exercice :

$$A \subset \bigcup_{y \in A} O_y$$

On a un recouvrement par des ouvert d'un compact ; on donc peut en extraire un recouvrement fini :

$$A \subset O_{y_1} \cup \cdots \cup O_{y_n} = O$$

Considérons maintenant l'intersection des ouverts $O_x^{y_n}$ associés aux ouverts O_{y_n} de ce recouvrement :

$$O_x = O_x^{y_1} \cap \cdots \cap O_x^{y_n}$$

On sait qu'il s'agit d'un ouvert puisqu'une intersection **finie** d'ouverts est un ouvert (cf question 1 de l'exercice 5). De plus, O_x contient bien x. Il ne nous reste donc plus qu'à montrer que $O_x \subset A^{\mathsf{c}}$ – et c'est bien le cas, puisque $O_x \cap O = \varnothing$ (ça se voit sur un dessin ; sinon, on peut dire qu'un élément de O_x appartient à chacun des $O_x^{y_i}$ et donc n'appartient à aucun des O_{y_i}. Il n'appartient donc pas à leur réunion, qui n'est autre que O) et que $A \subset O$.

On vient donc d'exhiber, pour un x quelconque de A^{c}, un ouvert de A^{c} contenant x. Donc A^{c} est ouvert, c'est-à-dire A est fermé.

Exercice 13 : compacité et compacité séquentielle

Le but de cet exercice est de montrer que, dans un espace métrique, la propriété de Borel-Lebesgue est équivalente à la propriété de Bolzano-Weierstrass, c'est-à-dire que la compacité séquentielle est équivalente à la compacité.

Soit E un espace métrique.

1. ***Borel-Lebesgue $\Longrightarrow$ Bolzano-Weierstrass***

   (a) *Montrer que si E vérifie la propriété de Borel-Lebesgue, alors l'intersection d'une suite décroissante de fermés non-vides est non-vide.*

   (b) *Remarquer que si une suite (x_n) admet une valeur d'adhérence x^*, i.e.*

   $$\exists x^* \text{ t.q. } \forall \varepsilon > 0, \forall n, \exists p \geq n \text{ t.q. } x_p \in B_\varepsilon(x^*)$$

   alors on peut extraire une sous-suite convergente de (x_n).

   (c) *En considérant les ensembles*

   $$A_n = \{x_k \text{ t.q. } k \geq n\}$$

   associés à une suite (x_n), en déduire que la propriété de Borel-Lebesgue implique la propriété de Bolzano-Weierstrass.

2. ***Bolzano-Weierstrass $\Longrightarrow$ Borel-Lebesgue***

   (a) *Montrer que si E vérifie la propriété de Bolzano-Weierstrass, alors $\forall \varepsilon > 0$, il existe un recouvrement fini de E par des boules de rayon ε.*

(b) *Montrer que si E vérifie la propriété de Bolzano-Weierstrass, alors pour tout recouvrement $(O_i)_{i\in I}$ de E,*

$$\exists\alpha > 0 \text{ tel que } \forall x \in E,\ \exists i_x \in I \text{ tel que } B_\alpha(x) \subset O_{i_x}.$$

(c) *En déduire que la propriété de Bolzano-Weierstrass implique la propriété de Borel-Lebesgue.*

1) (a) On va procéder par l'absurde : supposons qu'on ait une suite décroissante (F_n) de fermés non-vides de E telle que $\bigcap_n F_n = \varnothing$. On a alors

$$E = \left(\bigcap_n F_n\right)^{\mathsf{c}} = \bigcup_n F_n^{\mathsf{c}}.$$

Or, les F_n étant fermés, F_n^{c} sont des ouverts. On a donc un recouvrement de E par des ouverts. Puisque, par hypothèse, E vérifie la propriété de Borel-Lebesgue, on en déduit l'existence de $n_1, \ldots, n_p$ tels que

$$E = F_{n_1}^{\mathsf{c}} \cup \cdots \cup F_{n_p}^{\mathsf{c}}.$$

Mais (F_n) étant décroissante, (F_n^{c}) est croissante ($F_{n+1} \subset F_n \implies F_n^{\mathsf{c}} \subset F_{n+1}^{\mathsf{c}}$), donc $F_{n_1}^{\mathsf{c}} \cup \cdots \cup F_{n_p}^{\mathsf{c}} = F_{n_p}^{\mathsf{c}}$, d'où

$$E = F_{n_p}^{\mathsf{c}},$$

ce qui est impossible, puisque par hypothèse $F_{n_p} \neq \varnothing$. D'où la propriété.

(b) Il suffit, par exemple, de considérer la suite (ε_n) définie par $\varepsilon_n = 1/n$:

$$\forall n, \exists p_n \geq n \text{ t.q. } x_{p_n} \in B_{1/n}(x^*)$$

On peut alors extraire une suite en prenant, pour le rang n, le terme de (x_n) associé à un des entiers p_n vérifiant cette propriété – par exemple, le plus petit d'entre eux :

$$\forall n, \quad \tilde{x}_n := x_{\rho_n} \quad \text{où } \rho_n = \min\left\{p_n \text{ t.q. } x_{p_n} \in B_{1/n}(x^*)\right\}$$

On vérifie aisément que la suite ainsi définie est convergente, de limite x^*.

(c) Considérons une suite (x_n) d'éléments de E, et notons

$$A_n = \{x_k \text{ t.q. } k \geq n\} \; .$$

Clairement, ces ensembles sont non-vides et $A_{n+1} \subset A_n$ – d'où $\overline{A_{n+1}} \subset \overline{A_n}$. On a donc une suite décroissante de fermés non vides. D'après le résultat de la question (a),

$$\bigcap_n \overline{A_n} \neq \varnothing \, .$$

Ainsi, $\exists x^*$ tel que $\forall n$, $x^* \in \overline{A_n}$, c'est-à-dire, par définition de l'adhérence,

$$\exists x^* \text{ t.q. } \forall n, \quad \forall \varepsilon > 0,\ B_\varepsilon(x^*) \cap A_n \neq \varnothing$$

ou encore, par définition de A_n,

$$\exists x^* \text{ t.q. } \forall n,\ \forall \varepsilon > 0,\ \exists p \geq n \text{ t.q. } x_p \in B_\varepsilon(x^*)\,.$$

D'après la remarque de la question (b), on en déduit qu'on peut extraire une suite convergente de (x_n). On vient donc bien de montrer que E vérifie la propriété de Bolzano-Weierstrass.

2) (a) On procède par l'absurde : supposons qu'il existe $\varepsilon > 0$ tel qu'il n'existe aucun recouvrement fini de E par des boules de rayon ε. Dans ce cas, il est possible de définir une suite de la manière suivante : on choisit $x_0 \in E$. Puisque $B_\varepsilon(x_0)$ ne recouvre pas E, $B_\varepsilon(x_0)^\complement$ n'est pas vide. On peut donc y choisir x_1. Là encore, puisque $B_\varepsilon(x_0) \cup B_\varepsilon(x_1)$ ne recouvre pas E, on peut choisir x_2 dans $(B_\varepsilon(x_0) \cup B_\varepsilon(x_1))^\complement$, etc. Finalement, on construit une suite (x_n) telle que :

$$\forall n,\ x_{n+1} \notin \bigcup_{i=0}^{n} B_\varepsilon(x_i)\,.$$

On vient donc de construire une suite dont les termes sont tous à distance supérieure à ε les uns des autres – construction qui n'est rendue possible que parce que notre hypothèse de départ garantit qu'on pourra toujours trouver un point de E à distance supérieure à ε de tous les précédents.

Mais comme E vérifie la propriété de Bolzano-Weierstrass, on peut extraire une sous-suite convergente de (x_n), ce qui aboutit à une contradiction puisque les termes de cette suite extraite seront eux aussi à distance supérieure à ε les uns des autres ; ainsi, cette suite ne peut pas être de Cauchy, et donc ne peut pas être convergente.

(b) On procède par l'absurde : soit $(O_i)_{i\in I}$ un recouvrement de E par des ouverts. Supposons que $\forall\alpha > 0$, $\exists x \in E$ tel que $\forall i \in I$, $B_\alpha(x) \not\subset O_i$. En particulier, en prenant $\alpha = 1/n$,

$$\forall n,\ \exists x_n \in E \text{ tel que } \forall i \in I,\ B_{1/n}(x_n) \not\subset O_i$$

ce qui permet de définir une suite (x_n). Mais, comme E vérifie la propriété de Bolzano-Weierstrass, on peut en extraire une sous-suite convergente $(x_{\varphi(n)})$ (où $\varphi : \mathbb{N} \to \mathbb{N}$ est une fonction strictement croissante). Notons x^* sa limite. Puisque x^* appartient à

E qui est recouvert par la famille (O_i), $\exists j \in I$ tel que $x^* \in O_j$. O_j étant un ouvert, cela implique également l'existence d'un entier n tel que $B_{1/n}(x^*) \subset O_j$. Or, puisque $x_{\varphi(k)} \to x^*$, il existe un entier N tel que $B_{1/\varphi(N)}(x_{\varphi(N)}) \subset B_{1/n}(x^*)$. D'où

$$B_{1/\varphi(N)}(x_{\varphi(N)}) \subset O_j \, ,$$

ce qui contredit notre hypothèse de départ. D'où la propriété.

(c) Soit $(O_i)_{i\in I}$ un recouvrement de E par des ouverts. D'après la question (b), est associé à ce recouvrement un réel $\alpha > 0$ tel que

$$\forall x \in E, \exists i_x \in I \text{ tel que } B_\alpha(x) \in O_{i_x} \, .$$

Or d'après la question (a), il est possible de recouvrir E par un nombre fini de boules de rayon α :

$$E \subset B_\alpha(x_1) \cup \cdots \cup B_\alpha(x_n)$$

Mais on a justement choisi α de manière à ce que $\forall x_k$, $B_\alpha(x_k) \subset O_{i_{x_k}}$. D'où

$$E \subset O_{i_{x_1}} \cup \cdots \cup O_{i_{x_n}}$$

On vient donc bien d'extraire un recouvrement de E fini de $(O_i)_{i\in I}$. On en déduit que la propriété de Bolzano-Weierstrass implique la propriété de Borel-Lebesgue.

Exercice 14 : compacts et fermés bornés

1. *Montrer que, dans un espace métrique, un compact est borné (c'est-à-dire inclus dans une boule de rayon fini).* ($\star$)

2. *Montrer que les fermés bornés de $\mathbb{R}^n$ sont compacts.* ($\star\star$)

3. *Trouver un exemple de fermé borné non-compact.* ($\star$)

1) Puisqu'on est dans un espace métrique, on peut choisir d'utiliser la caractérisation de Bolzano-Weierstrass. Si un ensemble A n'est pas borné, cela signifie que quel que soit le centre et le rayon de la boule dans laquelle on tente de l'inclure, il existe des éléments de A qui sont hors de cette boule. On voit donc qu'on peut construire une suite une suite d'éléments de A dont les termes s'éloignent autant qu'on veut de tout point de A. On ne pourra donc pas extraire de sous-suite convergente de cette suite. Par contraposée, on en déduit qu'un ensemble vérifiant la propriété de Bolzano-Weierstrass est borné.

2) Il y a plusieurs façons de rédiger la réponse à cette question ; notamment, il ne serait pas difficile de généraliser notre raisonnement aux espaces vectoriels normés de dimension finie, mais cela demanderait de parler d'équivalence des normes, etc... Par souci de simplicité, on va donc se limiter à $\mathbb{R}^n$ muni de la topologie usuelle, et on dira un mot de la généralisation à la fin de la preuve.

On remarque que, puisque (1) un ensemble borné de $\mathbb{R}^n$ est inclus dans un pavé $[-a, a]^n$ et (2) on a montré que tout fermé inclus dans un compact est compact, il suffirait de montrer que tout pavé $[-a, a]^n$ est borné pour terminer la preuve...

Or c'est bien le cas, et cela se montre facilement à l'aide de la caractérisation de Bolzano-Weierstrass des compacts : en effet, soit une suite (x_n) d'éléments de $[-a, a]^n$. Subdivisons $[-a, a]^n$ en 2^n pavés de côté a (en le coupant en 2 selon chaque dimension). Au moins un de ces pavés contient une infinité de termes de (x_n). On ne conserve que les termes de (x_n) contenus dans ce pavé, et on choisit l'un d'eux pour être l'élément de rang 1 de notre suite extraite. En itérant ce processus sur le nouveau pavé, on obtient clairement une suite convergente. On vient donc de montrer que $[-a, a]^n$ est compact.

Digression

Une remarque au passage : on aurait pu procéder en montrant (1) que le segment $[-a, a]$ est compact (par le même procédé) et (2) que le produit cartésien de compacts est un compact, un résultat connu sous le nom de *théorème de Tykhonov*. L'unique raison pour laquelle je mentionne cette façon de faire, c'est de parler un peu de ce théorème. À première vue, il ne paye pas de mine ; d'autant qu'ici, on a besoin que de sa version finie (i.e. pour le cas d'un produit fini d'espaces compacts), qui n'est pas difficile à montrer (vous pouvez le faire en exercice)... Mais non seulement le théorème reste valide pour un produit quelconque d'espaces compacts, mais il prend alors une signification très profonde, au point qu'on dit parfois qu'il s'agit du théorème le plus important de la topologie générale ; en fait, on peut même montrer que le fait que le produit cartésien d'ensembles (pas forcément séparés) vérifiant la propriété de Borel-Lebesgue la vérifie lui-même est équivalent à l'axiome du choix [17].

17. Pour ceux qui n'en auraient jamais entendu parler, l'axiome du choix est une assertion assez innocente en apparence

Mais il ne faudrait pas croire que le fait que le théorème de Tykhonov reste valable dans le cas d'un produit quelconque d'ensembles compacts implique que la preuve précédente permet de généraliser le théorème de Heine-Borel à la dimension infinie ! En effet, on a omis de préciser les topologies considérées : les espaces de départ (ici, les segments $[-a, a]$), de même que l'espace produit (ici, le pavé $[-a, a]^n$), ont la topologie qu'on choisit de leur donner. Or le théorème de Tykhonov indique que le produit cartésien de compacts est compact *pour la topologie produit*[18]. Dans le cas de la dimension finie, on mon-

---

et dont on voit mal comment elle pourrait être fausse ou poser des problèmes : "*Étant donné un ensemble X d'ensembles non vides, il existe une fonction définie sur X qui à chaque ensemble de X associe un de ses éléments*". Cette assertion ne peut pas se démontrer, d'où son statut d'axiome. Certaines personnes n'aiment pas trop l'axiome du choix car il conduit à certains résultats contre-intuitifs comme le paradoxe de Banach-Tarski. Plus généralement, on lui reproche de permettre des preuves non-constructives (c'est-à-dire dans lesquelles on prouve l'existence d'objets sans les construire explicitement), ce qui va à l'encontre d'une certaine vision (dite constructiviste) des mathématiques.

18. Soient E_i des espaces topologiques. La topologie produit sur $E = \prod_i E_i$ est la topologie la moins fine rendant continues toutes les applications $p_i : E \to E_i$ qui à un élément de E associent sa composante selon E_i.

tre facilement que si les segments $[-a, a]^n$ sont munis de la topologie usuelle, alors la topologie produit associée est bien la topologie usuelle pour $[-a, a]^n$. Mais cela n'est vrai dans le cas d'un produit infini. Dans le cas infini non-dénombrable, la topologie produit n'est même pas métrisable (ce qu'on a déjà un peu vu via l'exercice 5 : en fait, la topologie de la convergence simple est une topologie produit...). Dans le cas dénombrable, elle est métrisable mais sauf erreur de ma part, le théorème de compacité de Riesz (cf infra) implique qu'elle n'est pas associée à une norme.

On va maintenant donner un contre-exemple en dimension infinie, et c'est sans doute ça qu'il faudra retenir plus que toute la discussion précédente.

---

3) Considérons un espace de dimension infinie – par exemple, l'espace des polynômes à coefficients réels, $\mathbb{R}[X]$. On sait qu'une base de cet espace est donnée par la famille des monômes $(X_n)_{n\in\mathbb{N}}$ et on peut s'en servir pour munir $\mathbb{R}[X]$ d'une distance – par exemple la distance associée à la norme L^∞ :

$$\forall P = \sum_{i=0}^{n} a_i X^i,\ \forall Q = \sum_{i=0}^{m} b_i X^i, \quad d_\infty(P, Q) = \max_i |a_i - b_i|$$

où les a_i (resp. b_i) sont nuls lorsque $i > \deg(P)$ (resp. $i > \deg(Q)$). Remarquons qu'il s'agit bien d'une fonction réelle, puisque tout polynôme a un degré fini. On se convainc ensuite aisément du fait qu'il s'agit d'une distance.

On munit $\mathbb{R}[X]$ de la topologie associée à cette distance, puis on considère la boule fermée centrée en 0 de rayon $1, \overline{B}_1(0)$. Clairement, cette boule est un fermé borné. Considérons alors la suite de monômes $(X^n)_{n\in\mathbb{N}}$. $\forall n \in \mathbb{N}$, $X^n \in \overline{B}_1(0)$ donc il s'agit d'une suite d'éléments de $\overline{B}_1(0)$. Mais on a aussi que $\forall n \neq m$, $d(X^n, X^m) = 1$ donc la suite (X^n) n'admet pas de sous-suite convergente (un telle sous-suite devrait être de Cauchy, ce qui est impossible quels que soient les termes qu'on choisit puisqu'ils sont tous à distance 1 les uns des autres). Donc $\overline{B}_1(0)$ ne vérifie pas la propriété de Bolzano-Weierstrass, et n'est donc pas compact.

En fait, ce qui se passe c'est que lorsque l'on dispose d'un nombre infini de dimensions, on peut rester dans un espace borné sans jamais repasser près d'un endroit où on est déjà passé : il suffit d'explorer à chaque fois une nouvelle dimension.

Bonus Avant de laisser cet exercice de côté, revenons rapidement sur un point : on a dit que la preuve de la question 2 se généralise facilement à un espace vectoriel réel normé de dimension finie. Voici comment : de même qu'on s'était ramenés à montrer que $[-a, a]^n$ était compact, on se ramène à montrer que la boule fermée de rayon a associée à la norme infinie, $\{x \in E \text{ t.q. } \|x\|_\infty \leq a\}$ est compacte (on peut le faire car, d'après l'équivalence des normes en dimension finie, on sait que toute boule est incluse dans une boule associée à la norme infinie). On utilise alors à nouveau un argument de type "di^n-chotomie" pour extraire une sous-suite convergente d'une suite quelconque, puis on conclut en utilisant une fois encore l'équivalence des normes en dimension finie : en dimension finie, toutes les normes induisent la même topologie – et donc la même notion de compacité.

Pour conclure, signalons que pour un espace vectoriel réel normé, il y a équivalence entre le fait d'être de dimension finie et le fait que la boule unité fermée soit compacte (il s'agit d'un théorème de Riesz).

Exercice 15 : compacts et continuité

1. *Montrer que l'image d'un compact par une fonction continue est un compact.* ($\star$)

2. *En déduire que toute fonction à valeurs réelles atteint ses bornes sur un compact.* ($\star$)

3. *Sur des espaces métriques, il existe une notion de continuité plus exigente que la continuité simple ("l'image réciproque d'un ouvert est un ouvert") : la continuité uniforme. Une fonction* $f : E \to F$ *est dite* uniformément continue *ssi :*

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ t.q. } \forall x, y \in E, d_E(x, y) < \delta \Longrightarrow d_F\Big(f(x), f(y)\Big) < \varepsilon \,.$$

   Montrer qu'une fonction continue sur un compact y est uniformément continue (théorème de Heine). ($\star\star$)

1) On suppose que l'espace d'arrivée de f est séparé.

Commençons par rappeler que l'image réciproque "commute avec l'union" : $f^{-1}\left(\bigcup \cdot\right) = \bigcup f^{-1}(\cdot)$. Il suffit de réécrire les définitions pour s'en assurer.

Soient alors K un compact et f une fonction continue sur K. Considérons un recouvrement (O_i) de $f(K)$ par des ouverts :

$$f(K) \subset \bigcup_i O_i$$

Par définition de $f^{-1}(K)$,

$$K \subset f^{-1}\left(\bigcup_i O_i\right) = \bigcup_i f^{-1}(O_i) \ .$$

Or, puisque f est continue, l'image réciproque d'un ouvert par f est un ouvert, par définition. D'où $\forall i$, $f^{-1}(O_i)$ est un ouvert. On est donc en présence d'un recouvrement de K par des ouverts ; K étant compact, il est possible d'en extraire un recouvrement fini :

$$K \subset f^{-1}(O_{i_1}) \cup \cdots \cup f^{-1}(O_{i_n}) = f^{-1}(O_{i_1} \cup \cdots \cup O_{i_n}) \ ,$$

c'est-à-dire

$$f(K) \subset O_{i_1} \cup \cdots \cup O_{i_n} \ .$$

On a donc extrait un recouvrement fini de $f(K)$ de (O_i). $f(K)$ est donc compact.

2) Si l'on recolle tout ce que l'on sait, ce théorème (parfois connu sous le nom de *théorème des bornes atteintes*) devient "évident" :

$\mathbb{R}$ est séparé et f est continue, donc, d'après la question précédente, $f(K)$ est un compact. C'est donc un fermé borné. Le caractère borné assure l'existence d'un infimum et d'un supremum finis. Le caractère fermé assure que ces bornes sont atteintes – une façon de le voir, c'est de remarquer que l'inf et le sup de $f(K)$ appartiennent à son adhérence (se montre facilement par l'absurde). Mais $f(K)$ étant fermé, $f(K) = \overline{f(K)}$.

3) Étant sur un espace métrique, on a le choix entre utiliser la propriété de Bolzano-Weierstrass ou celle de Borel-Lebesgue pour exploiter l'hypothèse de compacité. On n'a pas encore beaucoup manipulé Bolzano-Weierstrass donc on va choisir celle-là.

On procède par l'absurde. Supposons que $f : E \to F$ soit continue sur un compact K mais n'y soit pas uniformément continue. Alors

$$\exists \varepsilon>0 \text{ t.q. } \forall \delta>0, \exists x,y \in K \text{ t.q. } d_E(x,y)<\delta \text{ et } d_F\Big(f(x), f(y)\Big) \geq \varepsilon \,.$$

En particulier,

$$\exists \varepsilon > 0 \text{ t.q. } \forall n, \exists x_n, y_n \in K \text{ t.q. } d_E(x_n, y_n) < \frac{1}{n} \text{ et } d_F\Big(f(x_n), f(y_n)\Big) \geq \varepsilon \,.$$

Cela nous permet de construire deux suites d'éléments de K qui se rapprochent sans que ça ne soit le cas des images de leurs termes par f, qui restent à distance supérieure à ε les unes des autres.

K étant compact, on peut extraire de (x_n) une sous-suite convergente $(x_{\varphi(n)})$ (comme d'hab, $\varphi : \mathbb{N} \to \mathbb{N}$, strictement croissante). Notons $x^\star$ sa limite. Il est clair que $(y_{\varphi(n)})$ tend également vers $x^\star$, puisque $d_E\big(x_{\varphi(n)}, y_{\varphi(n)}\big) < 1/\varphi(n) \leq 1/n$. En revanche, $\big(f(y_{\varphi(n)})\big)$ ne peut pas tendre vers $f(x^\star)$ car

$$d_F\Big(f(x_{\varphi(n)}), f(y_{\varphi(n)})\Big) \leq d_F\Big(f(x_{\varphi(n)}), f(x^\star)\Big) + d_F\Big(f(x^\star), f(y_{\varphi(n)})\Big)$$

i.e.

$$d_F\Big(f(x^\star), f(y_{\varphi(n)})\Big) \geq \underbrace{d_F\Big(f(x_{\varphi(n)}), f(y_{\varphi(n)})\Big)}_{\geq\, \varepsilon} - \underbrace{d_F\Big(f(x_{\varphi(n)}), f(x^\star)\Big)}_{\to\, 0}$$

(le fait que $d_F\big(f(x_{\varphi(n)}), f(x^\star)\big) \to 0$ étant du à la continuité de f). Finalement, on a montré :

$$y_n \to x^\star \quad \text{et} \quad f(y_n) \not\to f(x^\star)$$

Donc f n'est pas séquentiellement continue. Or puisqu'on travaille sur des espaces métriques, cela signifie que f n'est pas continue. On a donc une contradiction.