UPS–L1-SFA–2015-2016 : Polycopié du module Mathématiques 4

Objectif. Ce module présente les notions fondamentales de l'analyse des suites numériques. Les modules de premier semestre ont servi à consolider les acquis du secondaire. Celui-ci vise à établir les bases permettant de poursuivre les études en Licence. Il met en avant les fondements (axiomes et définitions), les concepts et les démonstrations.

Mode d'emploi. Ce document contient les notes de cours et les exercices du module Math 4. Comme tout cours de Mathématique, il doit être lu avec un stylo et une feuille de papier pour vérifier pas à pas que toutes les assertions sont correctes.

Chaque section de ce cours se termine par des exercices non corrigés. Il est vivement recommandé de s'exercer à les résoudre : ce travail personnel permettra d'assimiler les notions introduites. On ne connaît pas d'autre méthode pour apprendre les mathématiques. Certains de ces exercices seront corrigés pendant les séances de TD, mais pas tous, faute de temps. N'hésitez pas à demander de l'aide aux enseignants si un exercice vous résiste ! Les exercices sont classés en trois catégories :

- Ceux dont le numéro n'est pas suivi d'étoile sont élémentaires. Ils permettent de se familiariser avec les notions introduites en cours.
- Les exercices (*) demandent un peu de réflexion, mais restent tout à fait faciles à traiter lorsque l'on a assimilé les techniques standard pour les résoudre.
- Les exercices (**) sont plus difficiles : ils demandent d'avoir bien compris les notions, ils sont parfois plus longs et certaines questions peuvent être dures. Il ne faut pas hésiter à y consacrer parfois plusieurs heures : c'est en essayant de résoudre ces exercices que l'on progressera en mathématiques.

Enfin, certaines parties du texte sont en petits caractères, comme ceci. Elles ne sont pas essentielles et l'on peut les sauter en première lecture. Cependant, les étudiants curieux auront avantage à les lire, quitte à ne pas tout comprendre. Elles donnent souvent un éclairage différent sur certains points du cours, ou introduisent des notions que l'on retrouvera plus tard dans d'autres cours de mathématiques.

Références et compléments. La Bibliothèque Universitaire propose plusieurs ouvrages qui couvrent le contenu de ce module. Citons entre autres :

- *Les mathématiques en Licence, Tome 1* par Élie Azoulay, Jean Avignant et Guy Auliac, Éditions ÉdiScience. Ce livre est très abordable.
- *Mathématiques. Tout-en-un pour la Licence. Niveau L1* sous la direction de Jean-Pierre Ramis et André Warusfel, Éditions Dunod. Ce livre est conseillé pour les nombreux compléments, éclaircissements et exercices supplémentaires qu'il contient.
- *Cours de Mathématiques du premier cycle* par Jacques Dixmier, Éditions Gauthier-Villars.
- *Mathématiques pour le DEUG, Analyse Première année* par François Liret et Dominique Martinais, Éditions Dunod.

Site web Vous trouverez une version électronique du polycopié, mais aussi les sujets de devoirs et d'examens donnés l'an passé, sur la page web :

`http://www.math.univ-toulouse.fr/~barthe/L1math4/`

Il est aussi possible d'accéder à cette page depuis le site Moodle de l'université.

Conventions. Nous noterons $\mathbb{N}$ l'ensemble des entiers naturels, $\mathbb{Q}$ l'ensemble des nombres rationnels, $\mathbb{R}$ l'ensemble des nombres réels, $\mathbb{C}$ l'ensemble des nombres complexes. Toutes ces notions (y compris la notion d'ensemble) seront détaillées plus bas.

Dans ce qui suit, les *mots en italiques* sont ceux que l'on est en train de définir. On emploie le symbole := lorsqu'une égalité sert à définir le membre gauche à partir du membre droit. Par exemple : On appelle *carré* du réel x le réel $x^2 := x.x$. On peut aussi introduire un terme sans définition complète et sans que sa connaissance soit exigible : on le mettra plutôt entre guillemets. Par exemple : On résume les propriétés de l'addition dans $\mathbb{R}$ en disant que $(\mathbb{R}, +)$ est un “groupe commutatif”.

Chapitre 1

Éléments de Logique et de théorie des ensembles

1.1 Logique

La logique est le langage des mathématiques. C'est une discipline à part entière, dont nous ne présentons ici que les bases élémentaires, qui permettront aux lecteurs de se familiariser avec les raisonnements mathématiques.

1.1.1 Énoncés

Un *énoncé* est une phrase ayant un sens mathématique précis, et qui peut être soit vrai, soit faux (mais jamais "entre les deux", ou "ni l'un ni l'autre"). Nous n'entrons pas dans la discussion de savoir quels types d'énoncés peuvent être considérés comme "mathématiques". Mais le lecteur verra au fur et à mesure des exemples ce que peut signifier un énoncé (ou une proposition).

Exemple d'énoncé : "$2 > 3$" (c'est un énoncé faux).

Un énoncé peut dépendre d'une variable. Par exemple $P(x)$: "$x > 3$". Alors $P(2)$ est faux mais $P(4)$ est vrai.

Il peut aussi dépendre de plusieurs variables. $P(x, y)$: "$x + y = 2$". Alors $P(1, 1)$ est vrai, mais $P(2, 1)$ est faux.

Quand on a un énoncé P, on a l'énoncé contraire, sa *négation* ($nonP$), encore noté $\neg P$ ou $\bar{P}$, qui est vrai lorsque P est faux et faux lorsque P est vrai.

Ainsi, si $P(x)$: "$x > 3$" alors $(nonP)(x)$: "$x \leq 3$" (et pas $x < 3$!!).

La négation de la négation d'un énoncé est l'énoncé de départ.

$$non(nonP) = P.$$

1.1.2 Quantificateurs

Il y en a deux. Ce sont les deux symboles $\forall$ (Quel que soit) et $\exists$ (Il existe).

Le symbole $\forall$ vient de l'allemand Alle, et $\exists$ de Existieren.

$\forall x \in E, P(x)$ se lit "Pour tout x dans l'ensemble E, $P(x)$ est vraie". C'est un nouvel énoncé fabriqué à partir de l'énoncé P. Nous n'avons pas encore vu ce qu'est un ensemble, ni ce que veut dire "x est dans l'ensemble E", mais nous espérons que la signification de cette phrase est bien claire pour tout le monde.

Par exemple, avec $P(x) : x > 3$, on voit que "$\forall x > 4, P(x)$" est un énoncé vrai, tandis que "$\forall x > 2, P(x)$" est un énoncé faux.

"$\exists x \in E, P(x)$" se lit "Il existe un élément x de l'ensemble E pour lequel $P(x)$ est vrai". On l'écrit aussi parfois "$\exists x \in E \mid P(x)$"

Ainsi, toujours avec $P(x) : x > 3$, "$\exists x > 2, P(x)$" est vrai tandis que "$\exists x < 2, P(x)$" est faux.

Dans ces deux derniers exemples, l'ensemble E est l'ensemble des nombres supérieurs strictement à 2 dans le premier cas, et strictement inférieurs à 2 dans le second. Ici, il faut faire attention : l'ensemble de nombres dont on parle peut être l'ensemble des entiers naturels, ou bien celui des nombres rationnels, ou celui des nombres réels, ou tout autre ensemble de nombres. Normalement, il faudrait l'indiquer dans l'énoncé, mais on suppose implicitement que cet ensemble de nombres est clair d'après le contexte : si l'on fait de l'arithmétique, ce sera l'ensemble des entiers, si on fait de l'analyse ce sera l'ensemble des nombres réels, etc... En tout cas, on suppose toujours dans un énoncé que l'ensemble E dont on parle est défini sans ambiguïté.

Avec ces quantificateurs, on peut donc construire de nouveaux énoncés, qui peuvent devenir de plus en plus compliqués. Par exemple "$\exists y > 2, \forall x > y, P(x)$", ou bien encore

$$Q(z) : \forall y > z, \exists x < y \mid P(x).$$

Une fois mise dans un quantificateur, une variable devient "muette" : c'est un nom arbitraire que nous lui avons donné. On peut changer son nom en un autre, **à condition de ne pas lui donner le nom d'une autre variable apparaissant ailleurs dans l'énoncé**. Rien ne nous oblige à l'appeler par une lettre : on aurait aussi bien pu l'appeler "Untel" : l'important est qu'il ne désigne personne en particulier : on n'aurait pas pu l'appeler "2", car "2" fait référence à un nombre bien défini.

Par exemple, "$\forall x \in E, P(x)$" et "$\forall a \in E, P(a)$", sont les mêmes énoncés.

Mais "$\exists y > 2, \forall x \geq y, P(x)$" n'est pas la même chose que "$\exists x > 2, \forall x \geq x, P(x)$".

La règle de **négation des propositions avec quantificateurs** est

$$non"\forall x \in E, P(x)" = "\exists x \in E, non P(x)",$$

et

$$non"\exists x \in E, P(x)" = "\forall x \in E, non P(x)".$$

Attention, une erreur fréquente (surtout dans les énoncés compliqués comme ceux qui apparaîtront à la fin de ce cours) est de penser que la négation de "$\forall x \in E, P(x)$" est "$\exists x \in E^c, P(x)$", où E^c est le complémentaire de E (voir prochaine section).

Ainsi, la négation de la proposition (fausse) "$\forall x > 2, x > 3$" est "$\exists x > 2, x \leq 3$" et non pas "$\exists x \leq 2, x \leq 3$".

Exercice 1. Quelle est la négation de l'énoncé $\exists x > a, \forall y < x, y > 4$?

On ne peut pas intervertir deux quantificateurs $\forall$ et $\exists$ successifs.

Ainsi, "$\forall x \in E,\ \exists y \in F,\ P(x, y)$" n'est pas la même chose que "$\exists y \in F, \forall x \in E, P(x, y)$" (mais est la même chose que "$\forall y \in E, \exists x \in F, P(y, x)$", en vertu de la règle de changement de nom).

Par exemple, "$\forall x > 3,\ \exists y > 2,\ y > x$" est vrai, mais "$\exists y > 2,\ \forall x > 3,\ y > x$" est faux.

Par contre, on peut intervertir "$\forall x \in E,\ \forall y \in F, P(x, y, z)$" et "$\forall y \in F, \forall x \in E, P(x, y, z)$" et de même "$\exists x \in E,\ \exists y \in F,$" et "$\exists y \in F, \exists x \in E,$". C'est-à-dire deux quantificateurs **successifs et de même type**.

Il faut faire attention ici à ce que les quantificateurs soient bien successifs. Ainsi, "$\forall x \in E, \exists y \in F,\ \forall z \in G,\ P(x,y,z)$" n'est pas la même chose que "$\forall z \in G, \exists y \in F,\ \forall x \in E,\ P(x,y,z)$". Par exemple, pour x, y, z dans l'ensemble à deux éléments $\{1,2\}$, la proposition

$$\forall x, \exists y, \forall z, x = y \text{ et } y \leq z+1$$

est vraie, alors que

$$\forall z, \exists y, \forall x, x = y \text{ et } y \leq z+1$$

est fausse. Ce qui est caché ici, c'est qu'on a intervertit un "quelque soit" et un "il existe".

Une autre situation où l'on peut intervertir $\forall$ et $\exists$ est la suivante. Si $\exists y, \forall x, P(x,y)$ est vrai, alors $\forall x, \exists y, P(x,y)$ est aussi vrai. Mais le contraire est faux (**et source d'un très grand nombre d'erreurs**) : dans le premier énoncé il existe un y tel que pour tout x, $P(x,y)$ est vrai ; dans le second, pour chaque x il existe un y avec cette propriété mais cette valeur de y peut varier avec x.

1.1.3 Connecteurs logiques

Conjonction et disjonction

À partir de deux énoncés P et Q, le connecteur de conjonction logique (*et*) et le connecteur de disjonction logique (*ou*) permettent d'en créer de nouveaux :

— L'énoncé $(P\, et\, Q)$, aussi noté $P \wedge Q$, est vrai si et seulement si à la fois P est vrai et Q est vrai.

— L'énoncé $(P\, ou\, Q)$, aussi noté $P \vee Q$, est vrai si et seulement si P est vrai ou Q est vrai. Attention, cela n'exclut pas que les deux énoncés puissent être vrais en même temps ! Ce n'est pas un ou "exclusif" comme dans "fromage ou dessert".

Par exemple, avec $P(x)$: "$x > 2$" et $Q(y)$: "$y < 3$", $(P\, ou\, Q)(x)$ est toujours vrai, tandis que $(P\, et\, Q)(x)$ n'est vrai que si $2 < x < 3$.

La négation échange le connecteurs *et* et *ou*, c'est à dire

$$non(P\, et\, Q) = (non P)\, ou\, (non Q), \qquad non(P\, ou\, Q) = (non P)\, et\, (non Q).$$

Les opérations *ou* et *et* sont "commutatives", c'est à dire

$$P\, et\, Q = Q\, et\, P, \qquad P\, ou\, Q = Q\, ou\, P.$$

Les opérations *ou* et *et* sont "associatives", c'est-à-dire

$$(P\, et\, Q)\, et\, R = P\, et\, (Q\, et\, R), \qquad (P\, ou\, Q)\, ou\, R = P\, ou\, (Q\, ou\, R).$$

C'est pourquoi on les note sans risquer d'erreur $P\, et\, Q\, et\, R$ et $P\, ou\, Q\, ou\, R$.

Les opérations *ou* et *et* sont "distributives l'une par rapport à l'autre" :

$$P\, et\, (Q\, ou\, R) = (P\, et\, Q)\, ou\, (P\, et\, R), \qquad P\, ou\, (Q\, et\, R) = (P\, ou\, Q)\, et\, (P\, ou\, R).$$

Enfin, quel que soit l'énoncé P, $P\, et\, (non P)$ est faux tandis que $P\, ou\, (non P)$ est vrai.

L'implication

C'est un énoncé de la forme "si P est vrai, alors Q l'est aussi". On le note $P \Longrightarrow Q$ et on le définit comme $(nonP)\, ou\, Q$. En d'autres termes, soit P est faux, soit P est vrai et Q est aussi vrai.

Attention : dans ce qui précède, P et Q peuvent dépendre d'un paramètre. Ainsi $(x > 2) \Longrightarrow (x \leq 3)$ est vrai pour $x \leq 3$ et faux pour $x > 3$, ce qui n'est bien sûr pas ce que l'on entend par là. **Lorsque P et Q dépendent du même paramètre x, alors on dit que $P \Longrightarrow Q$ lorsque**

$$\forall x,\ P(x) \Longrightarrow Q(x).$$

La *réciproque* de l'implication $P \Longrightarrow Q$ est l'implication $Q \Longrightarrow P$.

Deux énoncés sont *équivalents* (on note $P \iff Q$) si on a à la fois $P \Longrightarrow Q$ et $Q \Longrightarrow P$.

La *contraposée* de l'implication "$P \Longrightarrow Q$" est l'implication "$nonQ \Longrightarrow nonP$".

En fait

$$(nonQ \Longrightarrow nonP) \iff (P \Longrightarrow Q).$$

En effet,

$$(nonQ \Longrightarrow nonP) = (non(nonQ)\, ou\, nonP) = (Q\, ou\, nonP) = (P \Longrightarrow Q).$$

Attention, il ne faut pas confondre réciproque et contraposée. La contraposée est une "autre manière de dire la même chose", tandis que la réciproque est un énoncé différent de l'énoncé de départ.

L'implication est "transitive" Si $P \Longrightarrow Q$ et $Q \Longrightarrow R$ sont vraies, alors $P \Longrightarrow R$ est vraie. Cette transitivité peut se propager à plusieurs énoncés. Ainsi

$$P_1 \Longrightarrow P_2 \Longrightarrow P_3 \Longrightarrow \ldots \Longrightarrow P_n.$$

On peut ainsi obtenir des équivalences grâce au **"théorème tournant"**. Pour démontrer des équivalences entre énoncés, il suffit de montrer des implications en boucle. Si

$$P_1 \Longrightarrow P_2 \Longrightarrow P_3 \Longrightarrow \ldots \Longrightarrow P_n \Longrightarrow P_1,$$

alors

$$P_1 \iff P_2 \iff P_3 \iff \ldots \iff P_n.$$

L'unicité

On dit qu'un élément x dans un ensemble E ayant la propriété P est unique si deux éléments de E pour lesquels $P(x)$ est vraie sont nécessairement égaux. Ceci se traduit par

$$\forall x \in E, \forall y \in E,\ (P(x)\, et\, P(y)) \Longrightarrow x = y.$$

Ainsi, si $P(x) = (x \leq 2)\, et\, (x \geq 2)$ alors x est unique.

Attention, l'unicité n'implique pas l'existence. Ainsi, un x qui vérifie $(x > 2)\, et\, (x < 2)$ est unique (mais il n'existe pas !).

Lorsqu'il y a existence et unicité, on utilise souvent le symbole $\exists!$.

Ainsi $\exists! x, (x \leq 2)\, et\, (x \geq 2)$ (cet unique x est bien sûr $x = 2$).

1.1.4 Différents types de raisonnement

Le raisonnement direct ou syllogisme

Si $A \Longrightarrow B$ et que A est vrai, alors B est vrai. On l'utilise pour montrer que B est vrai.

Le raisonnement par équivalences

Il est parfois difficile de trouver un raisonnement direct pour établir une proposition A. On peut partir de la conclusion souhaitée et démontrer qu'elle est équivalente à des propositions plus simples, que l'on sait prouver directement.

Par exemple : si A est la propriété (pour les nombres réels) : $\forall x, \forall y, (x+y)^2 \leq 2x^2 + 2y^2$, en développant le carré et en regroupant les termes on peut écrire

$$(x+y)^2 \leq 2x^2+2y^2 \iff x^2+2xy+y^2 \leq 2x^2+2y^2 \iff 0 \leq x^2-2xy+y^2 \iff 0 \leq (x-y)^2.$$

Ce dernier énoncé est vrai (sur $\mathbb{R}$), donc A l'est aussi. En suivant les flèches de droite à gauche, on obtient une preuve directe de A.

Le raisonnement par contraposée

Pour démontrer l'implication $P \implies Q$, il est parfois plus facile d'établir sa contraposée $nonQ \implies nonP$ (nous avons vu que ces deux énoncés sont équivalents).

Par exemple : montrons (dans l'ensemble $\mathbb{N}$) que

$$(x > a \implies x > b) \implies a \geq b.$$

Ici $P =$ "$x > a \implies x > b$" et $Q =$ "$a \geq b$". Supposons $nonQ$, qui est $b > a$. Si l'on choisit $x = b$, alors pour cet x là , on a $x > a$ et non"$x > b$" : donc l'implication $x > a \implies x > b$ est fausse et on a $nonP$. Donc $nonQ \implies nonP$ et par conséquent $P \implies Q$.

Le raisonnement par l'absurde

On veut démontrer que A est vraie. On suppose que A est fausse et on cherche à en déduire une contradiction avec une autre propriété B dont on sait qu'elle est vraie. Plus formellement : on sait que B est vraie et on montre que $nonA \implies nonB$. Donc par contraposition $B \implies A$. Comme B est vraie, A l'est aussi.

Par exemple, démontrons que $\sqrt{2}$ n'est pas rationnel (propriété A). Par définition, $\sqrt{2}$ est un nombre positif x tel que $x^2 = 2$. Si x est rationnel, alors $x = \frac{p}{q}$ où p et q sont des entiers, $q \neq 0$. On a donc $x^2 = \frac{p^2}{q^2}$, d'où $p^2 = 2q^2$.

La propriété B sera la suivante : tout entier non nul s'écrit de façon unique $2^r n$, où n est impair (c'est une propriété bien connue en arithmétique). Supposons qu'il existe p et q non nuls tels que $p^2 = 2q^2$. On écrit $p = 2^r n$ et $q = 2^{r'} m$, avec n et m impairs. La relation $p^2 = 2q^2$ devient

$$2^{2r} n^2 = 2^{2r'+1} m^2.$$

Mais n^2 et m^2 sont impairs et $2r \neq 2r'+1$ puisqu'un nombre pair n'est jamais égal à un nombre impair. Donc on a obtenu deux écritures différentes comme "puissance de deux fois nombre impair". Ceci contredit la propriété B.

Le raisonnement par disjonction des cas

Si l'on sait que A est équivalent à $(A_1 \, ou \, A_2 \ldots \, ou \, A_n)$, et que $A_1 \implies B$, $A_2 \implies B$, $\ldots A_n \implies B$, alors $A \implies B$.

Par exemple pour montrer que $x > 0 \implies x^2 + \frac{1}{x} \geq 1$, il est commode de distinguer deux cas. Si $x \geq 1$ alors $x^2 \geq 1$ et on a bien $x^2 + \frac{1}{x} \geq 1$. Si $x \in (0,1]$ alors $1/x \geq 1$ et donc $x^2 + \frac{1}{x} \geq 1$.

Le raisonnement par récurrence

Il s'agit d'une propriété $P(n)$ qui dépend d'un nombre entier n et qu'on veut démontrer pour tout $n \geq n_0$.

On commence par démontrer $P(n_0)$ (c'est l'initialisation). On montre ensuite que pour tout $n \geq n_0$, $P(n) \Longrightarrow P(n+1)$ (on dit que P est héréditaire). Alors $P(n)$ sera vraie pour tous les entiers n à partir de n_0. Nous y reviendrons au chapitre suivant.

1.2 Ensembles

Le vocabulaire et les notations de la théorie des ensembles sont centraux en mathématiques. La théorie des ensembles "naïve" que nous décrivons ici laisse dans l'ombre des difficultés sérieuses (par exemple, on ne dit jamais explicitement ce qu'est un ensemble). Ce n'est pas par paresse, mais parce qu'une étude approfondie de ces notions nous entraînerait trop loin. Ce qui est important, ce sont les relations entre ensembles et éléments, ce qu'on a le droit de faire et ce qui est illicite, etc. Nous présentons ici le vocabulaire et les principes de base qui seront utiles non seulement tout au long de ce cours, mais tout au long de vos études en mathématiques.

Comme nous l'avons déjà dit, nous n'allons pas définir ce qu'est un *ensemble*. C'est intuitivement une "collection d'objets". On doit imaginer ces objets comme non rangés, sans répétitions possibles. Les objets n'ont pas besoin d'appartenir à une classe d'objets mathématiques précise. Ainsi, $\{1, 2, 3, \text{noir}, \text{blanc}, \text{Paul}, \text{Marie}\}$ est un ensemble. Au sens strict, $\{1, 1, 2\}$ n'en est pas un puisque l'élément 1 est répété. Cependant, par convention, on considèrera que c'est une notation maladroite pour l'ensemble $\{1, 2\}$.

Les ensembles peuvent être finis (c'est à dire qu'ils n'ont qu'un nombre fini d'éléments, comme l'ensemble qui précède), ou bien infinis, comme l'ensemble des entiers naturels, des nombres réels, des nombres complexes, des droites du plan, etc...

Nous allons définir des relations entre objets et ensembles, ainsi que beaucoup de vocabulaire un peu abstrait. On verra bientôt que n'importe quoi ne peut pas être un ensemble.

1.2.1 Définitions

Nous ne définissons pas non plus ce que veut dire le verbe *appartenir*. On dit que a "appartient" à l'ensemble A (et on note $a \in A$) si a est l'un des objets de la collection A. On dit encore que a est un *élément* de A.

On notera $a \notin A$ si a n'appartient pas à A. Nous ne définissons pas ce que veut dire le signe "=". Intuitivement, $a = b$ veut dire que ce sont les mêmes objets (ces objets peuvent eux-même être des ensembles). S'ils sont différents, on note $a \neq b$. Si $a = b$ et $b = c$, alors $a = c$.

Une propriété fondamentale qui permet d'identifier un ensemble est qu'il est entièrement caractérisé par les objets qu'il contient.

Deux ensembles sont égaux s'ils ont les mêmes éléments.

Plus exactement, en langage mathématique,

$$A = B \iff \Big((x \in A) \iff (x \in B)\Big).$$

Cette propriété (intuitivement évidente) est fondamentale dans la pratique. Pour montrer que deux ensembles A et B sont égaux, on montre que tout élément de A est un élément de B, et réciproquement. Les propriétés de l'égalité s'appliquent bien évidement aux ensembles.

L'ensemble vide, noté $\emptyset$, est un ensemble qui n'a aucun élément. D'après la propriété précédente, il est unique. C'est l'ensemble tel que

$$\forall x, x \notin \emptyset.$$

On n'est pas obligé de voir directement qu'un ensemble défini par une certaine propriété est vide. C'est l'impossiblité pour n'importe quel x d'être dans cet ensemble qui fait de cet ensemble l'ensemble vide. Par exemple, l'ensemble des nombres réels tels que $x^2 = -1$ est l'ensemble vide. (Mais il n'est pas vide si on considère des nombres complexes, bien qu'on puisse décrire l'ensemble pas la même formule).

Lorsqu'un ensemble E n'a qu'un nombre fini d'éléments, on appelle ce nombre d'éléments le *cardinal* de l'ensemble, on le notera card E ou parfois $\sharp E$. Ainsi card $\emptyset = 0$, card $\{Blanc, noir\} = 2$. Compter le nombre d'éléments d'un ensemble fini peut être une tâche assez compliquée (surtout quand cet ensemble dépend d'un paramètre $n \in \mathbb{N}$). C'est l'objet du chapitre "dénombrement" ou de "l'analyse combinatoire".

Pour définir un ensemble, on dispose de plusieurs méthodes.

Tout d'abord, si on connait tous les éléments $a_1, \dots, a_n$ de l'ensemble A, on notera $A = \{a_1, \dots, a_n\}$. On dit alors que l'ensemble est défini en *extension*. Rappelons que l'ordre n'a pas d'importance et ainsi que $\{1, 2, 3\}$ et $\{3, 2, 1\}$ désignent le même ensemble.

Il faut faire attention ici aux répétitions. Ainsi, $\{2, 2, 2\}$ désigne en fait l'ensemble dont le seul élément est 2, c'est à dire $\{2\}$. Son cardinal est 1. De même

$$\{1, 2, 3, 2, 4, 2\} = \{1, 2, 3, 4\}.$$

Comme on le voit, il y a diverses façons d'écrire un ensemble en extension.

La possibilité de répéter un élément est parfois utile. Par exemple, dans l'énoncé suivant : Pour tous les réels a, b, c tels que $b^2 \geq 4ac$ et $a \neq 0$, l'ensemble des solutions de l'équation $ax^2 + bx + c = 0$ est

$$\left\{ \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \frac{-b - \sqrt{b^2 - 4ac}}{2a} \right\}.$$

Cet énoncé est vrai, même lorsque $b^2 = 4ac$, auquel cas cet ensemble n'a plus qu'un seul élément, et non pas deux comme il paraît au premier abord.

On peut aussi définir un ensemble par une propriété mathématique. On note $\{x \in E; P(x)\}$ ou $\{x \in E \mid P(x)\}$ l'ensemble des éléments d'un ensemble E qui vérifient la propriété P. Cette notation est souvent plus concise et plus précise que la notation en extension :

$$\{n \in \mathbb{N}; n \leq 55\} = \{0, 1, 2, 3, \dots, 54, 55\}.$$

Lorsque la propriété P est compliquée cette notation permet de définir des ensembles dont on ne connaît pas forcément bien les éléments, comme l'ensemble des nombres premiers par exemple.

Nous n'avons pas vraiment défini ce qu'est une propriété mathématique. Mais il y a des définitions illicites qui ne peuvent pas donner des ensembles. Ainsi, "l'ensemble des nombres entiers qu'on ne peut pas définir en moins de 100 mots dans la langue française" n'est en fait pas un ensemble (l'idée est que cette définition est bien trop vague). Pour toucher du doigt les difficultés qu'on a à définir proprement des ensembles, acceptons cette propriété de l'ensemble $\mathbb{N}$ des entiers naturels : *tout sous ensemble de $\mathbb{N}$ non vide possède un plus petit élément*. Tout le monde conviendra que cette propriété est tout à fait légitime. Alors, si l'ensemble des nombres entiers qu'on ne peut pas définir en moins de 100 mots dans la langue française était vraiment un ensemble (on se convainc sans mal qu'il n'est pas vide, car avec moins de 100 mots, on ne peut définir qu'un nombre fini d'entiers, et il y a un nombre infini d'éléments de $\mathbb{N}$), son plus petit élément serait "le plus petit entier qu'on ne peut pas définir en moins de 100 mots de la langue française". Or cette définition comprend moins de 100 mots, donc cet entier n'appartient pas à l'ensemble dont il est le plus petit élément, ce qui est une contradiction. Comme quoi, on ne peut pas faire n'importe quoi avec les ensembles. Mais nous éviterons bien d'entrer dans ce genre de considérations, et tous les ensembles que nous rencontrerons seront proprement définis.

On peut fabriquer des ensembles dont les éléments sont eux-mêmes des ensembles (on le fera très souvent). Ainsi, si $A = \{1, 2\}$ et $B = \{2, 3\}$, on a l'ensemble $\{A, B\} = \{\{1, 2\}, \{2, 3\}\}$ qui a deux éléments, qu'il ne faut pas confondre avec $\{1, 2, 2, 3\} = \{1, 2, 3\}$ qui en a trois. De même, l'ensemble $\{\{1, 2, 3\}\}$ n'a qu'un seul élément.

Exercice 2. Combien l'ensemble $\{\{\{1, 2\}, \{2, 3\}\}\}$ a-t-il d'éléments ?

Inclusion, sous-ensembles

On dit qu'un ensemble A est *inclus* dans un ensemble B (et on note $A \subset B$) si tous les éléments de A sont des éléments de B, en d'autres termes

$$A \subset B \iff \Big((x \in A) \Longrightarrow (x \in B)\Big).$$

Si A est inclus dans B, on dit que A est un *sous-ensemble* de B, ou que A est une *partie* de B. On note parfois $A \subsetneq B$ pour indiquer que A est inclus dans B mais ne lui est pas égal.

On a toujours $\emptyset \subset A$, et

$$\Big((A \subset B)\, et\, (B \subset C)\Big) \Longrightarrow A \subset C.$$

Si $x \in A$, on peut lui faire correspondre un sous-ensemble $\{x\}$, qui est le sous-ensemble qui contient x et lui seul : c'est un *singleton*. Il ne faut pas le confondre avec l'élément x : x est un élément de A, $\{x\}$ est un sous-ensemble de A.

De même, lorsqu'on a deux éléments x et y de A, on peut lui faire correspondre le sous-ensemble $\{x, y\}$, qui est une paire (et un singleton si $x = y$!). On peut de même considérer des triplets, des quadruplets, etc.

Lorsqu'on a un ensemble A, on peut considérer tous les sous-ensembles (ou parties) de A. Ces parties forment elles-mêmes un ensemble, appellé *l'ensemble des parties de* A et noté $\mathcal{P}(A)$. On retiendra donc que

$$X \in \mathcal{P}(A) \iff X \subset A.$$

Ainsi, si $A = \{1, 2\}$,

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

Ceci s'applique bien sûr aussi à l'ensemble vide (attention, ici, on va commencer à se prendre un peu la tête, mais les considérations qui suivent sont toutes sauf gratuites). On a aussi $\mathcal{P}(\emptyset) = \{\emptyset\}$, qu'il ne faut pas confondre avec $\emptyset$. Ainsi, $\{\emptyset\}$ est un ensemble (qui a un élément, qui est $\emptyset$) et donc il n'est pas vide, et de même $\{\{\emptyset\}\}$, ou même $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$ qui a deux éléments.

Puisque $\mathcal{P}(A)$ est un nouvel ensemble, on peut encore considérer $\mathcal{P}(\mathcal{P}(A))$, et puis encore $\mathcal{P}(\mathcal{P}(\mathcal{P}(A)))$ et ainsi de suite. Comme on le voit, dès qu'on a un ensemble, on en obtient beaucoup d'autres.

Nous verrons dans le chapitre dénombrement que si A est un ensemble fini a n éléments, alors $\mathcal{P}(A)$ a 2^n éléments.

Nous avons bien évidement

$$A \subset B \Longrightarrow \mathcal{P}(A) \subset \mathcal{P}(B).$$

Exercice 3. Démontrez-le à partir des définitions !

Unions et intersections

La *réunion* de A et B est l'ensemble $A \cup B$ dont les éléments sont ceux de A et ceux de B. Autrement dit, x appartient à $A \cup B$ si et seulement si x appartient à A **ou** x appartient à B :

$$\Big(x \in A \cup B\Big) \iff \Big((x \in A)\, ou\, (x \in B)\Big).$$

L'*intersection* de A et B est l'ensemble $A \cap B$ dont les éléments sont formés de ceux qui appartiennent à la fois à A et à B. En d'autres termes, x appartient à $A \cap B$ si et seulement si x appartient à A **et** x appartient à B.

$$\Big(x \in A \cap B\Big) \iff \Big((x \in A)\, et\, (x \in B)\Big).$$

Les éléments de B qui ne sont pas dans A forment la *différence* $B \setminus A := \{x \in B;\, x \notin A\}$.

Lorsque $A \subset B$, l'ensemble formé des éléments de B qui ne sont pas dans A s'appelle le *complémentaire* de A dans B. C'est encore $B \setminus A$. Souvent, l'ensemble B est sous-entendu (par exemple l'ensemble des entiers naturels, ou des nombres réels) et on oublie B dans la notation : on note A^c pour le complémentaire de A.

Attention, la notation A^c suppose qu'on a un ensemble E de référence dans lequel on a pris le complémentaire. Le complémentaire de $\{1, 2\}$ n'est pas le même si l'ensemble de référence est $\{1, 2, 3\}$, auquel cas c'est $\{3\}$, ou bien l'ensemble des entiers naturels $\mathbb{N}$, auquel cas le complémentaire est $\{0\} \cup \{n, n \geq 3\}$. Dans la pratique, l'ensemble de référence E sera toujours bien déterminé et il n'y aura aucune ambiguïté quand on parlera de A^c.

Lorsqu'on a un ensemble de référence E et deux sous-ensembles A et B de E, alors $B \setminus A = B \cap A^c$.

Les règles de calcul sur les ensembles sont calquées sur celles de la logique, ou le passage au complémentaire remplace la négation, l'union remplace *ou* et l'intersection remplace *et*, l'inclusion remplace l'implication. Ainsi,

$$A \cup B = B \cup A,\ A \cap B = B \cap A,$$

$$A \cup (B \cup C) = (A \cup B) \cup C,\ (A \cap B) \cap C = A \cap (B \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),\ A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$
$$A \subset B \iff B^c \subset A^c,$$
$$(A^c)^c = A,\ (A \cup B)^c = A^c \cap B^c,\ (A \cap B)^c = A^c \cup B^c.$$

Exercice 4. Démontrer les trois dernières égalités (en revenant à la définition : deux ensembles sont égaux signifie que tout élément de l'un est un élément de l'autre et réciproquement).

Nous avons vu que $(A \cup B) \cup C = A \cup (B \cup C)$. Pour cette raison, on peut oublier les parenthèses et noter ceci plus simplement $A \cup B \cup C$. C'est aussi la même chose que $C \cup B \cup A$, ou encore $A \cup C \cup B$, etc.

Si nous disposons d'une famille plus grande de sous-ensembles d'un ensemble donné E (fixé une fois pour toutes), disons $A_1, \dots, A_n$ pour fixer les idées, on peut considérer leur union

$$A_1 \cup A_2 \cup \ldots \cup A_n := \bigcup_{i=1}^{n} A_i.$$

C'est l'ensemble des éléments x qui appartiennent à l'un des A_i, pour i de 1 à n. Le résultat ne dépend pas de l'ordre dans lequel on a énuméré ces ensembles. On peut obtenir cette union en réunissant les ensembles deux à deux dans n'importe quel ordre. Par exemple

$$A_1 \cup A_2 \cup A_3 \cup A_4 = (A_1 \cup A_2) \cup (A_3 \cup A_4) = ((A_1 \cup A_2) \cup A_3) \cup A_4 = A_1 \cup ((A_2 \cup A_3) \cup A_4) = \dots.$$

Mais pourquoi s'arrêter à un nombre fini ? Supposons qu'on dispose d'une famille quelconque A_i de sous-ensembles de E, où i est un indice qui appartient lui même à un ensemble I. Alors, nous pouvons considérer l'ensemble $\bigcup_{i \in I} A_i$: c'est l'ensemble des éléments $x \in E$ pour lesquels il existe $i \in I$ tel que x appartient à l'un des A_i. En langage mathématique

$$x \in \bigcup_{i \in I} A_i \iff \exists i \in I,\ x \in A_i.$$

Le résultat ne dépend pas de l'ordre dans lequel sont donnés les A_i (d'ailleurs I étant un ensemble quelconque, il n'y a pas forcément de notion d'ordre naturel pour ses éléments).

De la même manière, on peut considérer l'intersection des A_i, notée $\bigcap_{i \in I} A_i$: c'est l'ensemble des éléments $x \in E$ qui appartiennent à tous les A_i :

$$x \in \bigcap_{i \in I} A_i \iff \forall i \in I, x \in A_i.$$

Une fois de plus, le résultat ne dépend par de l'ordre des A_i.

On a comme pour les unions finies

$$\Big(\bigcup_{i \in I} A_i\Big)^c = \bigcap_{i \in I} A_i^c \ ; \ \Big(\bigcap_{i \in I} A_i\Big)^c = \bigcup_{i \in I} A_i^c.$$

Couples et produits d'ensembles

Une autre façon de construire des ensembles est de considérer des *couples*. Lorsqu'on dispose de deux ensembles A et B, on peut considérer l'ensemble des couples (a, b) avec $a \in A$ et $b \in B$. L'ensemble de ces couples constitue l'ensemble *produit* $A \times B$.

Rien n'interdit de choisir $A = B$, mais **attention** : $A \times A$ n'est pas l'ensemble des couples (a, a) qui est par définition la *diagonale* de $A \times A$.

Ainsi, pour $A = \{1, 2\}$, $A \times A = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$. Contrairement aux points d'un ensemble, il faut faire attention à l'ordre ici : $(1, 2) \neq (2, 1)$.

Si A a n éléments et B a m éléments, alors $A \times B$ possède $n.m$ éléments. On pourra le démontrer par exemple par récurrence sur m.

Relations

Une *relation* sur un ensemble A est une partie $\mathcal{R}$ de $A \times A$: les couples (a, b) qui sont dans $\mathcal{R}$ sont dits en relation, et on note $a\mathcal{R}b$ au lieu de $(a, b) \in \mathcal{R}$.

Un exemple de relation est $x \leq y$ sur $\mathbb{N}$. On fait parfois la distinction entre la relation $\mathcal{R}$ et la partie de $A \times A$ qui la définit, qu'on appelle le *graphe* de la relation, car il n'est souvent pas commode de penser une relation comme un ensemble.

On peut aussi définir des relations entre $a \in A$ et $b \in B$ comme une partie $\mathcal{R}$ de $A \times B$.

Voici quelques propriétés intéressantes que les relations peuvent avoir (ou pas) :

(1) Une relation $\mathcal{R}$ sur A est dite *réflexive* si, $\forall a \in A,\ a\mathcal{R}a$.

(2) Une relation $\mathcal{R}$ sur A est dite *symétrique* si, $a\mathcal{R}b \Longrightarrow b\mathcal{R}a$.

(3) Une relation $\mathcal{R}$ sur A est dite *antisymétrique* si, $\Big((a\mathcal{R}b)\ et\ (b\mathcal{R}a)\Big) \Longrightarrow a = b$.

(4) Une relation $\mathcal{R}$ sur A est dite *transitive* si, $\Big((a\mathcal{R}b)\, et\, (b\mathcal{R}c)\Big) \Longrightarrow a\mathcal{R}c$.

(5) Une relation qui est à la fois, réflexive, symétrique et transitive est une *relation d'équivalence*. (Par exemple, sur l'ensemble des entiers relatifs, "$x - y$ est pair" est une relation d'équivalence).

(6) Une relation qui est à la fois réflexive, antisymétrique et transitive est une *relation d'ordre*. (Par exemple, "$x \leq y$" sur $\mathbb{R}$ est une relation d'ordre.)

Exercice 5. La relation "$x < y$" sur $\mathbb{R}$ est-elle une relation d'ordre ?

Exercice 6. Montrer que la relation de divisibilité est une relation d'ordre sur $\mathbb{N}$. On rappelle qu'un entier naturel a divise un entier naturel b (ce que l'on notera $a|b$) s'il existe $k \in \mathbb{N}$ tel que $b = ka$. Montrer que pour cette relation d'ordre, $\mathbb{N}$ admet un plus petit élément et un plus grand élément.

Exercice 7. On définit sur $\mathbb{Z}$ la relation dite de *congruence modulo 3* en disant que x est en relation avec y si $y - x$ est multiple de 3. Montrer que c'est une relation d'équivalence.

Lorsque l'on a une relation d'équivalence (souvent notée $a \sim b$), on peut considérer un nouvel ensemble, *l'ensemble quotient*. C'est une notion importante, que l'on reverra souvent dans différents contextes.

Tout d'abord, si on considère $a \in A$, on peut regarder l'ensemble des $b \in A$ tels que $a \sim b$. Cet ensemble, noté $\dot{a}$, est la *classe d'équivalence de a*. On remarque que $a \in \dot{a}$ et que si $a \neq b$, ou bien $\dot{a} = \dot{b}$, ou bien $\dot{a} \cap \dot{b} = \emptyset$ (Pourquoi ?). L'ensemble des $\dot{a}$, pour $a \in A$ forme donc une famille $\mathcal{Q}$ de sous-ensembles de A qui a les propriétés suivantes

(i) Aucun des éléments de $\mathcal{Q}$ n'est vide,

(ii) Tout point de A appartient à l'un des éléments de $\mathcal{Q}$,

(iii) L'intersection de deux éléments différents de $\mathcal{Q}$ est toujours vide.

On résume ces trois propriétés en disant que $\mathcal{Q}$ est une *partition* de A. L'ensemble $\mathcal{Q} = \{\dot{a}, a \in A\}$ (qui est un sous ensemble de $\mathcal{P}(A)$) est appelé l'ensemble quotient de A par la relation $\sim$. On le note souvent $A/\sim$. L'opération de passage au quotient revient à ne plus distinguer deux éléments de E qui sont dans la même classe d'équivalence.

Par exemple, si on considère sur l'ensemble $\mathbb{N} \times \mathbb{N}^*$ ($\mathbb{N}^* = \mathbb{N} \setminus \{0\}$) la relation

$$(n, p) \sim (n', p') \iff np' = n'p,$$

on vérifie que c'est une relation d'équivalence (exercice !) et l'ensemble quotient n'est autre que l'ensemble des nombres rationnels positifs ou nuls.

Remarquons que si on se donne une partition de A, on peut fabriquer une relation d'équivalence en décidant que deux points sont en relation si et seulement si ils sont dans le même sous-ensemble de la partition. Il y a donc une analogie étroite entre relations d'équivalence et partitions.

Exercice 8. Démontrer l'assertion qui précède.

1.2.2 Fonctions, applications

Lorsqu'on a deux ensembles A et B (qui peuvent être les mêmes), une *application* entre A et B est une façon de faire correspondre à tout élément $a \in A$ un élément $f(a) \in B$.

Formellement, on peut la définir comme une relation entre A et B, c'est à dire une partie $\mathcal{F}$ de $A \times B$, qui est telle que pour tout point $a \in A$, il existe un unique $b \in B$ tel que $(a, b) \in \mathcal{F}$. Cette partie $\mathcal{F}$ de $A \times B$ s'appelle le graphe de l'application f et peut se décrire comme :

$$\mathcal{F} = \{(x, f(x)); x \in A\}.$$

On dit que A est l'ensemble de *départ* et B l'ensemble d'*arrivée* de l'application f.

On peut aussi dire que f est une *fonction* de A dans B. Nous réserverons le terme de fonction aux applications qui ne sont définies éventuellement que sur un sous-ensemble de A (qu'on appelle *l'ensemble de définition* de la fonction).

On note $f : A \to B$ et aussi $x \mapsto f(x)$.

Lorsque $y = f(x)$, on dit que y est *l'image* de x ou bien que que x est **un** *antécédent* de y. Par définition, **un élément de l'ensemble de départ n'a qu'une image mais un élément de l'ensemble d'arrivée peut avoir plusieurs antécédents, ou aucun**.

L'ensemble de toutes les applications de A dans B (c'est un ensemble !) se note $\mathcal{F}(A, B)$ ou encore B^A. Cette notation sera justifiée au chapitre Dénombrements où l'on montre que si A a n éléments et B a m éléments, alors B^A a m^n éléments.

Parmi les applications les plus simples, il y a l'*identité* $Id_E : E \to E$. C'est l'application qui à $x \in E$ fait correspondre $x \in E$ (elle ne fait rien !).

Les débutants ont souvent tendance à la confondre avec l'application constante qui à $x \in E$ fait correspondre toujours la même valeur $y \in F$.

Par exemple, l'application de $E = \{1, 2, 3\}$ dans lui même qui est telle que $f(1) = f(2) = f(3) = 1$ est constante, alors que pour l'identité, on a $f(1) = 1$, $f(2) = 2$, $f(3) = 3$.

Que répondre à la question "Montrez que f est une application de E dans F" ?

Il s'agit de vérifier les points suivants :

(1) D'abord que $f(x)$ est bien définie pour tout $x \in E$ (par exemple qu'on n'obtient pas 0/0).

(2) Ensuite, que $f(x)$ est définie sans ambiguïté : qu'il n'y a pas plusieurs valeurs possibles pour $f(x)$, comme par exemple "un réel tel que $x^2 = 1$".

(3) Enfin, que la valeur de $f(x)$ est bien dans l'ensemble F. Par exemple, la fonction $x \mapsto x^2$ n'est une application de $\mathbb{R}$ dans $[0, 1]$, alors que c'est bien une application de $\mathbb{R} \to \mathbb{R}$ ou de $[-1, 1] \to [0, 1]$.

Retenir que deux applications f et g sont égales si et seulement si : elles ont même ensemble de départ, même ensemble d'arrivée et pour tout élément x de l'ensemble de départ, $f(x) = g(x)$.

Exercice 9. On considère trois applications f, g, h définies de $E := \{-1, 0, 1\}$ dans $\mathbb{R}$, par
- $f(-1) = 1$, $f(0) = 0$ et $f(1) = 1$,
- $\forall x \in E$, $g(x) = x^2$,
- $\forall x \in E$, $h(x) = x^4 - x^3 + x$.

Montrer que ces trois applications sont égales.

Il ne faut pas croire qu'une application est juste une formule !

Images directes et images réciproques d'ensembles

Une application $f : E \to F$ permet de définir de nouvelles applications $\mathcal{P}(E) \to \mathcal{P}(F)$ et $\mathcal{P}(F) \to \mathcal{P}(E)$ appelées images directes et réciproques :

Image directe. Pour $A \subset E$, l'image directe $f(A)$ est l'ensemble des points de F qui sont images par f d'un point de A

$$f(A) = \{y \in F;\ \exists x \in A,\ y = f(x)\} \text{ aussi notée } \{f(x);\ x \in A\}.$$

Image réciproque. Pour $B \subset F$, l'image réciproque $f^{-1}(B)$ est l'ensemble des points de E dont l'image est dans B (autrement dit, c'est l'ensemble des antécédents des éléments de B)

$$f^{-1}(B) = \{x \in E;\ f(x) \in B\}.$$

L'image réciproque a de bonnes propriétés du point de vue de la théorie des ensembles

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B),\ f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B),\ f^{-1}(A^c) = \Big(f^{-1}(A)\Big)^c.$$

De plus, si $A \subset B$, $f^{-1}(A) \subset f^{-1}(B)$.

Exercice 10. Démontrer les assertions qui précèdent.

En revanche, l'image directe est pleine de pièges et nous invitons les lecteurs à s'en méfier ! Ainsi, on a

$$f(A \cap B) \subset f(A) \cap f(B),$$

mais l'inclusion inverse peut être fausse. De plus, il n'est pas vrai en général que $f(A^c) \subset f(A)^c$, ni que $f(A)^c \subset f(A^c)$.

Composition des applications

Lorsqu'on a deux applications $f : E \to F$ et $g : F \to G$, on peut considérer leur *composition* : c'est l'application $g \circ f : E \to G$ qui à $x \in E$ fait correspondre $g\big(f(x)\big) \in G$. (**Attention : l'écriture de la composition se fait à l'envers de la réprésentation** $E \to F \to G$).

On a $h \circ (g \circ f) = (h \circ g) \circ f$. En d'autres termes, la composition des applications est **"associative"**. Nous avons déjà rencontré ce mot plus haut. Mais on n'a pas en général $f \circ g = g \circ f$ d'autant plus que ceci n'a aucun sens en général (sauf si $E = G$: pourquoi ?). La composition des application **n'est pas commutative**.

L'identité joue un rôle particulier pour la composition. En effet, si on a une application $f : E \to F$, on a $f \circ Id_E = f$ et $Id_F \circ f = f$.

Restriction d'une application

Etant donnés une application $f : E \mapsto F$ et un sous-ensemble $A \subset E$, on peut considérer l'application $f_{|A} : A \to F$ qui à $x \in A$ associe $f(x) \in F$. C'est la *restriction* (au départ) de f à A.

Si un sous-ensemble $B \subset F$ vérifie $f(E) \subset B$, alors on peut construire la restriction à l'arrivée de f à B (attention, dans ce cas, il y a une condition sur B). C'est l'application de $E \to B$ qui a tout $x \in E$ associe $f(x) \in B$. Le plus petit B qu'on puisse choisir est $f(E)$. On peut noter cette nouvelle application $f^{|B}$.

Si $f(A) \subset B$ on peut restreindre f à la fois au départ à A et à l'arrivée à B.

Indicatrice d'un ensemble.

Il est parfois fastidieux de vérifier des identités entre ensembles. Il est alors commode de remplacer le calcul sur les ensembles par du calcul algébrique. Pour cela, introduisons la notion d'*indicatrice* d'un sous-ensemble d'un ensemble E (aussi appelée *fonction caractéristique*). Jusqu'à la fin de ce paragraphe, tous les ensembles considérés seront des sous-ensembles de E. Si $A \subset E$, alors 1_A est la fonction $f : E \to \{0, 1\}$ telle que $f(x) = 1$ si $x \in A$ et $f(x) = 0$ si $x \in A^c$.

En notant $1 = 1_E$ la fonction constante égale à 1 sur E, on a

$$1_{A^c} = 1 - 1_A \text{ et } 1_{A \cap B} = 1_A 1_B$$

(c'est-à-dire le produit terme à terme des fonctions 1_A et 1_B). Vérifiez ces égalités à titre d'exercice. On déduit par exemple de ces relations que :

$$1_{A \cup B} = 1_{(A^c \cap B^c)^c} = 1 - (1 - 1_A)(1 - 1_B) = 1_A + 1_B - 1_{A \cap B}.$$

Ensuite, pour montrer l'identité de deux ensembles, il suffit de montrer l'égalité de leurs indicatrices.

Par exemple, si l'on veut montrer que $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$, (c'est-à-dire la distributivité de $\cup$ par rapport à $\cap$) en prenant les indicatrices, nous trouvons pour le premier terme

$$1_{(A \cap B) \cup C} = 1_{A \cap B} + 1_C - 1_{A \cap B \cap C},$$

tandis que pour le deuxième, nous avons

$$\begin{aligned} 1_{(A \cup C) \cap (B \cup C)} &= \Big(1_A + 1_C - 1_{A \cap C}\Big)\Big(1_B + 1_C - 1_{B \cap C}\Big) \\ &= \quad 1_{A \cap B} + 1_{A \cap C} - 1_{A \cap B \cap C} \\ &\quad + \; 1_{B \cap C} + 1_C - 1_{B \cap C} \\ &\quad - \; (1_{A \cap B \cap C} + 1_{A \cap C} - 1_{A \cap B \cap C}), \end{aligned}$$

et c'est bien la même chose.

Exercices

Exercice 11. Complétez le tableau suivant (à chaque ligne les valeurs des énoncés P et Q sont indiquées par les deux premières cases, à vous de donner la valeur des autres énoncés).

P	Q	$non P$	$P\, et\, Q$	$P\, ou\, Q$	$P \Longrightarrow Q$	$P \iff Q$	$(P\, et\, non Q)\, ou\, (Q\, et\, non P)$
$vrai$	$vrai$						
$vrai$	$faux$						
$faux$	$vrai$						
$faux$	$faux$						

Exercice 12. Dans l'ensemble $\mathbb{N}$ des entiers naturels, lesquels de ces énoncés sont vrais ?

1) $\exists y, \forall x > y, x \leq 1$.
2) $\exists y, \forall x > y, x \geq 1$.
3) $\exists y, \forall x < y, x \leq 1$.
4) $\exists y, \forall x < y, x \geq 1$.
5) $\forall x, \exists y > x, \forall z < y, z \leq x$.
6) $\forall x, \exists y > x, \forall z < y, z < x$.

Exercice 13. (*) Dans l'ensemble $\mathbb{N}$ des entiers naturels, trouvez lorsqu'elles existent les valeurs de a pour lesquelles les propositions suivantes sont vraies

1) $\forall x \leq 3, \exists y \leq a, x \leq y$
2) $\exists y \leq a, \forall x \leq 3, x \leq y$
3) $\forall x \leq a, \exists y \leq 3, x \leq y$
4) $\exists y \leq 3, \forall x \leq a, x \leq y$

Exercice 14. Donnez la négation des énoncés suivants

1) $\forall x \geq 2, \exists y \leq a, y \leq x$
2) $A \Longrightarrow B$
3) $A \iff B$
4) $\exists p \in \mathbb{N}^*$, $\exists n_0 \in \mathbb{N}$, $\forall n \geq n_0, u_{n+p} = u_n$. Quel est le sens de cet énoncé ?
5) (*) J'ai lu tous les livres de math de la BU qui contiennent moins de 300 pages.
6) (*) Dans chaque ville de France, il y a une rue dont les habitants ont moins de 30 ans ou plus de 60 ans.

Exercice 15. (*) Deux frères jumeaux, Jack et John souffrent d'un dédoublement de personnalité. Jack dit toujours la vérité dans son état normal et ment toujours dans son état pathologique. John lui, ment toujours dans son état normal et dit toujours la vérité dans son état pathologique. On ne peut pas les reconnaître par leur apparence. Si vous rencontriez l'un des frères et que vous souhaitiez savoir s'il s'agit de Jack ou de John, quelle question simple (de réponse oui ou non) lui poseriez-vous ?

Indication : on pourra appeler F le frère rencontré, noter A pour "F est Jack" et B pour "F est dans son état normal". On cherche une question Q telle la réponse donnée par F soit en fait la vraie réponse à la question "êtes-vous Jack ?". On pourra trouver Q en complétant la table de vérité suivante

A	B	Réponse de F à Q	Vraie réponse à Q
vrai	vrai	oui	
vrai	faux	oui	
faux	vrai	non	
faux	faux	non	

Exercice 16. (**) Un pays est habité par des gens qui soit ne disent que la vérité, soit mentent systématiquement. Un touriste arrive à une intersection à deux branches dont une seule mène à la capitale. Il n'y a pas de panneau. Un habitant, H, attend à l'intersection. Quelle question, ayant pour réponse oui ou non, peut poser le touriste pour savoir quelle route mène à la capitale ?

Indication : on pourra procéder comme dans l'exercice qui précède.

Exercice 17. Ecrire $\mathcal{P}(A)$ pour $A = \{1, 2, 3\}$, pour $A = \mathcal{P}(\{1\})$ et pour $A = \mathcal{P}(\{\emptyset\})$.

Exercice 18. Soit $A = \{n \in \mathbb{N}^*;\ n \text{ divise } 36\}$ et $B = \{n \in \mathbb{N}^*;\ n \text{ divise } 27\}$. Décrivez en extension les ensembles $A \cap B$ et $A \cup B$.

Exercice 19. Soit A l'ensemble des entiers naturels pairs et B l'ensemble des entiers naturels multiples de 3. Déterminer les ensembles $A \cap B$ et $A \cup B$.

Exercice 20. Soient A, B des ensembles.

1) Montrer que $A = (A \cap B) \cup (A \setminus B)$ et $(A \cap B) \cap (A \setminus B) = \emptyset$. En déduire une partition de A.
2) Montrer que $A \cup B = A \cup (B \setminus A)$ et $A \cap (B \setminus A) = \emptyset$. En déduire une partition de $A \cup B$.

Exercice 21. (*) On définit pour tout $n \in \mathbb{N}$ l'ensemble $I_n = \{0, \ldots, n\}$ et $J_n = \{n, n+1, \ldots\} = \{i \in \mathbb{N};\ i \geq n\}$.

1) Parmi les ensembles $I_n, J_n, I_{n+1}, J_{n+1}$, qui est inclus dans qui ? Lesquels sont disjoints ?
2) Calculez les ensembles
$$\bigcup_{n\in\mathbb{N}} I_n, \quad \bigcap_{n\in\mathbb{N}} I_n, \quad \bigcup_{n\in\mathbb{N}} J_n, \quad \bigcap_{n\in\mathbb{N}} J_n.$$

Exercice 22. (**) On considère un ensemble E et pour tout entier $n \geq 1$, un sous-ensemble A_n de E. On définit pour tout $n \geq 1$ les sous-ensembles $\overline{A}_n = \cup_{p=n}^{\infty} A_p$ et $\underline{A}_n = \cap_{p=n}^{\infty} A_p$

1) Montrez que pour tout $n \geq 1$, $\underline{A}_n \subset \overline{A}_n$.
2) Montrez que pour tout $n \geq 1$, $\overline{A}_{n+1} \subset \overline{A}_n$ et $\underline{A}_n \subset \underline{A}_{n+1}$.
3) On définit $\overline{A} = \cap_{n=1}^{\infty} \overline{A}_n$ et $\underline{A} = \cup_{n=1}^{\infty} \underline{A}_n$. Montrez que $\underline{A} \subset \overline{A}$.
4) Montrez que $\overline{A}$ est l'ensemble de tous les éléments x qui appartiennent à une infinité de A_n.
5) Montrez que $\underline{A}$ est l'ensemble de tous les éléments x qui appartiennent à tous les A_n à partir d'un certain rang.

Exercice 23. On considère les applications f et g de $\mathbb{R}$ dans $\mathbb{R}$ définies pour $x \in \mathbb{R}$ quelconque par $f(x) = x + 1$ et $g(x) = x^2$. Calculer $f \circ g$ et $g \circ f$ et montrer qu'elles ne sont pas égales.

Exercice 24. Trouver un exemple où $A \cap B = \emptyset$ et où $f(A) \cap f(B) \neq \emptyset$.

Exercice 25. (*) Trouver un exemple où les inclusions $f(A^c) \subset f(A)^c$ et $f(A)^c \subset f(A^c)$ sont fausses.

Exercice 26. Soit $f : \mathbb{R} \to \mathbb{R}$ définie pour tout x réel par $f(x) = 3x - x^3$. Décrire les ensembles $f\big(]1, 2[\big)$, $f\big(]0, 2[\big)$, $f\big(]-1, 3/2]\big)$, $f^{-1}(\mathbb{R}^+)$, $f^{-1}\big([0, 18]\big)$.

Exercice 27. Décrire les ensembles $\arctan(\{1\})$, $\tan^{-1}(\{1\})$, $\arccos([0, 1])$, $\cos^{-1}([0, 1])$.

Exercice 28.

1) Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ pour tout x. Déterminer $f^{-1}\big(f(\{1, 2\})\big)$ et $f\big(f^{-1}(\mathbb{R})\big)$.
2) (*) Soit $f : E \to F$ une application.

(a) Soit $A \subset E$, montrer que $A \subset f^{-1}\big(f(A)\big)$.

(b) Soit $A \subset F$. Montrer que $f\big(f^{-1}(A)\big) \subset A$ puis que $f\big(f^{-1}(A)\big) = A \cap f(E)$.

Exercice 29. Lorsque $f : E \to F$ est une application, montrer que $\{f^{-1}(\{y\}), y \in f(E)\}$ forme une partition de E.

(*) Réciproquement, montrer que toute partition de E peut être réalisée ainsi.

Exercice 30. (*) La différence symétrique de deux ensembles A et B est l'ensemble

$$A\Delta B := \big(A \setminus B\big) \cup \big(B \setminus A\big).$$

1) On a vu en cours que l'union d'ensembles correspond au connecteur logique *ou* (ou inclusif). A quelle opération logique correspond Δ ?

2) Montrez que $1_{A\Delta B} = (1_A - 1_B)^2$. En déduire que $(A\Delta B)\Delta C = A\Delta(B\Delta C)$.

Chapitre 2

Dénombrements

2.1 Les bases

Pour commencer, nous présentons des outils indispensables au dénombrement des ensembles.

2.1.1 Principes de récurrence

Rappelons que l'ensemble $\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$ des entiers naturels est muni des opérations d'addition et de multiplication, mais aussi de la relation d'ordre $\leq$ (inférieur ou égal) qui possède la propriété fondamentale suivante :
Toute partie non vide de $\mathbb{N}$ admet un plus petit élément[1].

On peut déduire de cette propriété de $\mathbb{N}$ le *principe de récurrence*. Dans la pratique, on considère ce principe comme une méthode de démonstration. Celle-ci admet au moins trois formes distinctes, que nous allons expliciter et illustrer. On considère un énoncé $P(n)$ qui dépend d'un nombre entier n ; on cherche à montrer qu'il est vrai pour tout $n \geq n_0$ (en pratique n_0 vaudra souvent 0 ou 1).

La preuve par *récurrence simple* est celle-ci :
Soit $P(n)$ un énoncé défini sur $\mathbb{N}$ (pour $n \geq n_0$ suffit) et vérifiant les deux hypothèses suivantes :
- *Initialisation : $P(n_0)$ est vrai.*
- *Hérédité : $\forall n \geq n_0$, $\big(P(n) \Rightarrow P(n+1)\big)$.*

Alors P est vrai pour tout entier n supérieur ou égal à n_0 : $\forall n \geq n_0,\ P(n)$.

Exemple. Montrons que, pour tout $n \in \mathbb{N}$, on a $2^n \geq n + 1$. Nous notons donc, pour tout $n \in \mathbb{N}$:

$$P(n) := (2^n \geq n + 1).$$

Initialisation : la propriété $P(0)$ dit que $2^0 \geq 1$, *i.e.* que $1 \geq 1$, qui est vraie.
Hérédité : supposons $P(n)$ vraie, *i.e.* $2^n \geq n + 1$ (hypothèse de récurrence). Alors :

$$2^{n+1} = 2 \times 2^n \geq 2(n + 1) = 2n + 2 \geq (n + 1) + 1.$$

La première inégalité utilisait l'hypothèse de récurrence. On a bien prouvé $P(n+1)$. Du principe de récurrence (simple), on déduit que $\forall n \in \mathbb{N}$, $P(n)$.

1. On dit que cette relation d'ordre fait de $\mathbb{N}$ un ensemble "bien ordonné".

La preuve par *récurrence double*, ou *récurrence à deux pas*[2] repose sur le principe suivant :
Soit $P(n)$ un énoncé défini sur $\mathbb{N}$ et vérifiant les deux hypothèses suivantes :

— *Initialisation : $P(n_0)$ et $P(n_0+1)$ sont vrais.*

— *Hérédité : $\forall n \geq n_0\ ,\ \Big(\big(P(n) \text{ et } P(n+1)\big) \Rightarrow P(n+2)\Big)$.*

Alors P est vrai sur $\mathbb{N}$ tout entier supérieur ou égal à n_0 : $\forall n \geq n_0\ ,\ P(n)$.

Exemple. Nous allons démontrer que $(1+\sqrt{2})^n + (1-\sqrt{2})^n \in \mathbb{N}$ pour tout entier $n \in \mathbb{N}$. C'est vrai pour $n = 0$ et $n = 1$ (calcul facile). En remarquant que $1+\sqrt{2}$ et $1-\sqrt{2}$ sont solutions de l'équation $X^2 = 2X + 1$, il vient

$$\begin{aligned}
&(1+\sqrt{2})^{n+2} + (1-\sqrt{2})^{n+2}\\
&= (1+\sqrt{2})^n(1+\sqrt{2})^2 + (1-\sqrt{2})^n(1-\sqrt{2})^2\\
&= (1+\sqrt{2})^n\big(2(1+\sqrt{2})+1\big) + (1-\sqrt{2})^n\big(2(1-\sqrt{2})+1\big)\\
&= 2\big((1+\sqrt{2})^{n+1} + (1-\sqrt{2})^{n+1}\big) + \big((1+\sqrt{2})^n + (1-\sqrt{2})^n\big),
\end{aligned}$$

ce que l'on peut écrire $u_{n+2} = 2u_{n+1} + u_n$, en posant $u_n := (1+\sqrt{2})^n + (1-\sqrt{2})^n$. Il est alors clair que $(u_n \in \mathbb{N}$ et $u_{n+1} \in \mathbb{N}) \Rightarrow u_{n+2} \in \mathbb{N}$, et l'on a bien l'hérédité, donc la conclusion par récurrence à deux pas.

L'hérédité dans cette démonstration, repose sur la relation (de récurrence à deux pas !) $u_{n+2} = 2u_{n+1} + u_n$. Notons qu'en posant $v_n := (1+\sqrt{2})^n - (1-\sqrt{2})^n$ on a la relation analogue : $v_{n+2} = 2v_{n+1} + v_n$, donc la même propriété d'hérédité pour l'assertion $v_n \in \mathbb{Z}$. De plus, cette dernière est vraie pour $n = 0$, mais fausse pour $n = 1$: l'initialisation en $n = 0$ est donc insuffisante dans une récurrence à deux pas.

Enfin, nous présentons le principe de preuve par *récurrence complète*, aussi appelée *récurrence forte* :
Soit $P(n)$ un énoncé défini sur $\mathbb{N}$ ($n \geq n_0$ suffit) et vérifiant les hypothèses suivantes :

— *Initialisation : $P(n_0)$ est vrai.*

— *Hérédité :*

$$\forall n \geq n_0\ ,\ \Big(\big(\forall m \in \mathbb{N} \textit{ tel que } n_0 \leq m \leq n\ ,\ P(m)\big) \Rightarrow P(n+1)\Big).$$

Alors P est vrai pour $\mathbb{N}$ tout entier n supérieur ou égal à n_0 : $\forall n \geq n_0\ ,\ P(n)$.

Exemple. Disons qu'un entier $n \in \mathbb{N}\setminus\{0,1\}$ est *irréductible* (ou premier) s'il n'est pas le produit de deux entiers $p, q \in \mathbb{N}\setminus\{0,1\}$. Nous allons montrer que tout entier de $\mathbb{N}\setminus\{0,1\}$ (c'est-à-dire supérieur ou égal à $n_0 := 2$) est produit d'entiers irréductibles.
Nous notons donc, pour tout $n \in \mathbb{N}\setminus\{0,1\}$:

$$P(n) := (n \text{ est produit d'entiers irréductibles}).$$

Initialisation : si n s'écrit comme $n = pq$ avec $p, q \geq 2$ alors $n \geq 4$. Donc 2 est irréductible (3 aussi !)
Hérédité forte : Soit $n \geq 2$. Supposons que tout entier m avec $2 \leq m \leq n$ vérifie $P(m)$, autrement dit, qu'il est produit d'irréductibles. Il s'agit d'en déduire que $n+1$ est lui-même produit d'irréductibles. On distingue deux cas :

(1) Si $n+1$ est irréductible, il est bien entendu produit d'irréductibles !

2. Il existe aussi des récurrence à trois pas et même à k pas, où $k \in \mathbb{N}^*$.

(2) Sinon, il est réductible et l'on peut écrire $n+1 = pq$ avec $p, q \in \mathbb{N} \setminus \{0,1\}$. Comme $p, q > 1$, on a $p, q < n+1$ donc $2 \leq p, q \leq n$. On peut donc leur appliquer l'hypothèse de récurrence forte : $P(p)$ et $P(q)$ sont vraies, *i.e.* p et q sont produits d'irréductibles, donc $n+1 = pq$ aussi.

On a bien démontré $P(n)$, donc l'hérédité forte. Du principe de récurrence forte on tire la conclusion. (Cette démonstration remonte aux Éléments d'Euclide.)

Constructions par récurrence. On peut *définir* des objets par récurrence. Il y a, là encore, les récurrences simple, à deux (ou k) pas et la récurrence forte. Nous allons illustrer chaque cas par un exemple.

Exemple. On peut définir une suite numérique par récurrence simple. C'est le cas de la suite des *factorielles* :

$$0! := 1 \text{ et } \forall n \in \mathbb{N}\ ,\ (n+1)! := (n+1)\, n!.$$

Ainsi, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, etc.

Exemple. La célèbre suite de Fibonacci est définie par récurrence double :

$$F_0 = 0,\ F_1 := 1 \text{ et } \forall n \in \mathbb{N}\ ,\ F_{n+2} := F_{n+1} + F_n.$$

Exemple. On peut aussi définir une suite numérique par récurrence complète, par exemple : $u_0 = 1$ et

$$\forall n \in \mathbb{N}^*\ ,\ u_n := 1 + \sum_{i=0}^{n-1} u_i.$$

On a donc $u_0 = 1$, $u_1 = 2$, $u_2 = 4$, $u_3 = 8$ et ensuite ?

Exercice 31. Démontrer par récurrence sur n que dans l'exemple précédent, $u_n = 2^n$.

Exercice 32. (*) Démontrez le principe de récurrence simple en utilisant la propriété "un sous-ensemble de $\mathbb{N}$ qui n'est pas vide admet un plus petit élément". On pourra considérer l'ensemble $\{n \geq n_0;\ P(n) \text{ est faux}\}$.

2.1.2 Injections, surjections, bijections

Les définitions qui suivent sont **absolument fondamentales**. Il est indispensable de bien les comprendre pour la suite du cours.

Définition. *On dit qu'une fonction $f : E \to F$ est* injective *si tout point $y \in F$ a* **au plus** *un antécédent. On dit aussi que f est* une injection*. Autrement dit*

$$(f(x_1) = f(x_2)) \Longrightarrow x_1 = x_2,$$

ou encore par contraposition

$$(x_1 \neq x_2) \Longrightarrow (f(x_1) \neq f(x_2)).$$

L'exemple le plus simple d'injection s'obtient lorsqu'on a un ensemble E et une partie $A \subset E$. Alors, l'application $f : A \to E$ donnée par $f(x) = x$ est une injection. On l'appelle *l'injection canonique* de A dans E.

Définition. *On dit qu'une fonction* $f : E \to F$ *est* surjective *si tout point* $y \in F$ *a* **au moins** *un antécédent. On dit aussi que* f *est* une surjection. *En d'autres termes,* $f : E \to F$ *et surjective si et seulement si* $f(E) = F$.

Lorsqu'on a $A \subset E$, on peut fabriquer une surjection $f : E \to A$ en choisissant d'abord un élément $a \in A$, puis en considérant la fonction définie de la façon suivante : si $x \in A$, $f(x) = x$, et si $x \notin A$, $f(x) = a$.

Définition. *Lorsqu'une application* $f : E \to F$ *est à la fois injective et surjective, on dit qu'elle est* bijective, *ou que c'est* une bijection. *En d'autres termes, tout* $y \in F$ *admet* **un unique** *antécédent.*

L'exemple le plus simple de bijection est $Id_E : E \to E$.

Définition. *Soit* $f : E \to F$ *une bijection. On définit alors son* application réciproque *comme ceci : c'est l'application de* F *dans* E *qui à tout* $y \in F$ *associe son unique antécédent pour* f. *On la note* f^{-1} *et elle vérifie* $f^{-1} \circ f = \mathrm{Id}_E$ *et* $f \circ f^{-1} = \mathrm{Id}_F$.

Ces deux égalités sont plus une propriété qu'une définition et elles méritent d'être justifiées : pour tout $x \in E$, $f^{-1}(f(x))$ est par définition l'unique antécédent de $f(x)$ par f ; c'est donc x et on a bien $f^{-1}(f(x)) = x$. Pour $y \in F$, comme $f^{-1}(y)$ est l'antécédent de y, $f(f^{-1}(y)) = y$.

Une autre manière de traduire ces relations :

$$y = f(x) \iff x = f^{-1}(y).$$

Ceci montre en particulier que f^{-1} est aussi bijective et que $(f^{-1})^{-1} = f$ (la réciproque de la réciproque est la fonction elle-même).

Nous aurons besoin de la propriété suivante qui relie ces notions avec la composition d'applications :

Proposition 2.1.1. *Soient* $f : E \to F$ *et* $g : F \to G$ *deux applications.*

(1) Si $g \circ f$ *est injective, alors* f *est injective.*

(2) Si $g \circ f$ *est surjective, alors* g *est surjective.*

(3) Si $g \circ f$ *est bijective, alors* f *est injective et* g *est surjective.*

Démonstration. — Pour le premier point : si $x, y \in E$ vérifient que $f(x) = f(y)$, alors en appliquant g on obtient $g(f(x)) = g(f(y))$, c'est-à-dire $(g \circ f)(x) = (g \circ f)(y)$. Comme $g \circ f$ est injective on peut conclure que $x = y$. Nous avons bien montré $(f(x) = f(y)) \Longrightarrow (x = y)$.

Pour le deuxième point : soit $z \in G$. Par hypothèse, $g \circ f$ est surjective, donc il existe $x \in E$ tel que $z = (g \circ f)(x) = g(f(x))$. Ainsi z admet un antécédent pour g (c'est $f(x)$, il y en a peut être d'autres d'ailleurs).

Le dernier point est la conjonction des deux premiers. ∎

Exercice 33. Démontrer que la composition de deux injections est une injection, et que la composition de deux surjections est une surjection. Et pour les bijections ?

Les notions d'injection, surjection et bijection sont essentielles pour le dénombrement. En particulier, si $f : E \to F$ est une bijection, elle définit une correspondance entre les éléments de E et F : à tout élément de E correspond un élément de F et réciproquement. Donc, intuitivement, les deux ensembles ont le même nombre d'éléments. Le résultat suivant, qui permet d'établir facilement que des applications sont des bijections, nous sera très utile en pratique.

Théorème 2.1.2. *Si $f : E \to F$ et $g : F \to E$ sont deux applications telles que $g \circ f = Id_E$ et $f \circ g = Id_F$, alors ces deux applications sont bijectives et $f = g^{-1}$, $g = f^{-1}$.*

Démonstration. — Puisque $g \circ f = Id_E$ est bijective, nous savons par la Proposition 2.1.1 que g est surjective et f est injective. De même, nous savons que $f \circ g$ est bijective et donc que f est surjective et g injective. Donc, f et g sont bijectives et par conséquent, elles admettent des applications réciproques f^{-1} et g^{-1}. Finalement on obtient par composition

$$f^{-1} = Id_E \circ f^{-1} = (g \circ f) \circ f^{-1} = g \circ (f \circ f^{-1}) = g \circ Id_F = g.$$

$$g^{-1} = Id_F \circ g^{-1} = (f \circ g) \circ g^{-1} = f \circ (g \circ g^{-1}) = f \circ Id_E = f.$$

∎

Exercice 34. Supposons que $f : E \to F$ et $g : F \to G$ sont des bijections. Montrer que $g \circ f : E \to G$ est bijective d'application réciproque $f^{-1} \circ g^{-1}$.

Attention : l'application réciproque f^{-1} n'existe que si f est bijective. En revanche, l'image réciproque d'une ensemble B, notée $f^{-1}(B)$, existe toujours (et ce même si f^{-1} n'existe pas !)

Lorsque $f : E \to F$ est bijective et que $B \subset F$ est un sous-ensemble de F, la notation $f^{-1}(B)$ peut désigner l'image réciproque de B par f, mais aussi l'image directe de B par l'application $f^{-1} : F \to E$. Heureusement ces deux notions (apparemment différentes) de $f^{-1}(B)$ décrivent en fait le même sous-ensemble de E.

Exercice 35. (*) Démontrer l'assertion qui précède.

2.2 Cardinaux finis.

La notion de cardinal d'un ensemble fini est claire : c'est le nombre d'éléments qu'il contient. Cette définition intuitive est un peu vague du point de vue mathématique. Nous allons en proposer une plus formelle, pas simplement par souci de rigueur, mais parce qu'elle indique une méthode de dénombrement utilisable en pratique. Dans la suite on note $[\![1, n]\!] := \{1, \ldots, n\}$ l'ensemble des entiers naturels compris, au sens large, entre 1 et n.

Définition. *Un ensemble est dit* fini *s'il est vide ou s'il existe un entier $n \geq 1$ et une bijection $f : [\![1, n]\!] \to E$. Le* cardinal *d'un ensemble fini E, noté* card E *vaut 0 si $E = \emptyset$ et n dans le cas où il existe une bijection $f : [\![1, n]\!] \to E$.*

Remarquons que dans ce dernier cas, on peut écrire $E = \{f(1), \ldots, f(n)\}$. Cette définition correspond donc à ce que l'on fait lorsque l'on compte des objets : on leur attribue (dans un ordre arbitraire) les numéros 1,2,3, etc, en veillant à ne pas compter deux fois le même objet (c'est l'injectivité de f) et à n'oublier aucun objet (surjectivité de f).

2.2.1 Union d'ensembles, addition de cardinaux

Nous prendrons comme point de départ la propriété évidente suivante :
Si A et B sont des ensembles finis disjoints, card $(A \cup B) =$ card $A +$ card B.

Théorème 2.2.1. *Si A et B sont des ensembles finis quelconques, alors :*

$$\text{card } (A \cup B) = \text{card } A + \text{card } B - \text{card } (A \cap B).$$

Démonstration. — On a d'abord, par application de la propriété de départ à la réunion disjointe $A = (A \cap B) \cup (A \setminus B)$, établie à l'exercice 20 :

$$\text{card } A = \text{card } (A \cap B) + \text{card } (A \setminus B) \Longrightarrow \text{card } A - \text{card } (A \cap B) = \text{card } (A \setminus B).$$

On remarque ensuite que $(A \cup B)$ est l'union disjointe de B et de $(A \setminus B)$, auxquels on applique à nouveau la propriété de départ :

$$\text{card } (A \cup B) = \text{card } B + \text{card } (A \setminus B) = \text{card } A + \text{card } B - \text{card } (A \cap B).$$

∎

Essayons le cas de trois ensembles finis A, B, C, en utilisant ce que nous savons pour la réunion de deux ensembles :

$$\text{card } (A \cup B \cup C) = \text{card } \big((A \cup B) \cup C\big) = \text{card } (A \cup B) + \text{card } (C) - \text{card } \big((A \cup B) \cap C\big).$$

On calcule donc $\text{card } (A \cup B) = \text{card } A + \text{card } B - \text{card } (A \cap B)$ et :

$$\begin{aligned} \text{card } \big((A \cup B) \cap C\big) &= \text{card } \big((A \cap C) \cup (B \cap C)\big) \\ &= \text{card } (A \cap C) + \text{card } (B \cap C) - \text{card } \big((A \cap C) \cap (B \cap C)\big) \\ &= \text{card } (A \cap C) + \text{card } (B \cap C) - \text{card } (A \cap B \cap C). \end{aligned}$$

En reportant dans la première égalité, on trouve enfin :

$$\text{card } (A \cup B \cup C) = \text{card } A + \text{card } B + \text{card } C - \text{card } (A \cap B) - \text{card } (A \cap C) - \text{card } (B \cap C) + \text{card } (A \cap B \cap C).$$

Cette formule se généralise à une union de n ensembles, pour n quelconque :

Théorème 2.2.2 (Formule d'inclusion-exclusion ou formule du crible)**.** *Soient $A_1, \ldots, A_n$ des ensembles finis quelconques. Alors :*

$$\begin{aligned} \text{card } (A_1 \cup \cdots \cup A_n) &= \sum_{i=1}^{n} \text{card } A_i - \sum_{1 \leq i < j \leq n} \text{card } (A_i \cap A_j) \\ &+ \sum_{1 \leq i < j < k \leq n} \text{card } (A_i \cap A_j \cap A_k) \\ &- \sum_{1 \leq i < j < k < \ell \leq n} \text{card } (A_i \cap A_j \cap A_k \cap A_\ell) + \cdots \\ &+ (-1)^{n-1} \text{card } (A_1 \cap \cdots \cap A_n) \\ &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \text{card } (A_{i_1} \cap \cdots \cap A_{i_k}). \end{aligned}$$

Démonstration. — Par récurrence sur n ; réservé aux courageux ! ∎

Exemple. Anticipant sur le cours d'arithmétique, nous allons calculer le nombre d'entiers dans $[\![1, 900]\!]$ ayant un facteur commun avec 900. Ces entiers sont nécessairement divisibles par l'un des facteurs premiers de 900, lesquels sont 2, 3 et 5. On pose donc :

$$A := \{n \in [\![1, 900]\!] \mid 2|n\}, \quad B := \{n \in [\![1, 900]\!] \mid 3|n\} \text{ et } C := \{n \in [\![1, 900]\!] \mid 5|n\}.$$

Les éléments de A sont les $2k$ tels que $1 \leq k \leq 900/2$, il y en a donc 450 ; avec le même raisonnement pour B et C, on trouve :

$$\text{card } A = 450, \quad \text{card } B = 300 \text{ et card } C = 180.$$

Les éléments de $A \cap B$ sont les $6k$ tels que $1 \leq k \leq 900/6$, il y en a donc 150 ; avec le même raisonnement pour $A \cap C$ et $B \cap C$, on trouve :

$$\text{card } (A \cap B) = 150, \quad \text{card } (A \cap C) = 90 \text{ et card } (B \cap C) = 60.$$

Enfin, les éléments de $A \cap B \cap C$ sont les $30k$ tels que $1 \leq k \leq 900/30$, il y en a donc 30. Finalement, le nombre cherché est :

$$\begin{aligned}
\text{card } (A \cup B \cup C) &= \text{card } A + \text{card } B + \text{card } C \\
&- \text{card } (A \cap B) - \text{card } (A \cap C) - \text{card } (B \cap C) \\
&+ \text{card } (A \cap B \cap C) \\
&= 450 + 300 + 180 - 150 - 90 - 60 + 30 = 660.
\end{aligned}$$

Il est important de retenir le cas particulier suivant, qui se démontre simplement par récurrence à partir de $(A \cap B = \emptyset) \Longrightarrow (\text{card } (A \cup B) = \text{card } (A) + \text{card } (B))$.

Proposition 2.2.3. *Soient $A_1, \ldots, A_n$ des ensembles finis deux à deux disjoints ($i \neq j \Longrightarrow A_i \cap A_i = \emptyset$), alors*

$$\text{card } \left(\bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} \text{card } (A_i).$$

Remarque importante sur les sommes : La formule ci-dessus fait intervenir une expression de la forme $\sum_{i=1}^{n} u_i = u_1 + \cdots + u_n$, qui pour être précis, est définie par récurrence : $\sum_{i=1}^{1} u_i := u_1$ et pour tout $n \geq 1$, $\sum_{i=1}^{n+1} u_i = \left(\sum_{i=1}^{n} u_i \right) + u_{n+1}$. Il est utile de généraliser la notion de somme finie. Voici le cadre : on considère un ensemble fini A et une application $f : A \to \mathbb{R}$. On définit la somme sur $a \in A$ des $f(a)$ comme ceci : si A est vide la somme vaut 0 par convention, sinon en notant $n = \text{card } (A) \geq 1$, il existe par définition une bijection $\varphi : [\![1, n]\!] \to A$. On pose alors la définition suivante

$$\sum_{a \in A} f(a) := \sum_{i=1}^{n} f(\varphi(i)) = f(\varphi(1)) + \cdots + f(\varphi(n)).$$

La bijection φ nous fournit ainsi un ordre dans lequel on ajoute les nombres $f(a)$, mais par commutativité de l'addition, le résultat ne dépend pas de l'ordre (c'est un bon exercice de le démontrer par récurrence sur n). Donc la quantité définie ne dépend pas du choix de la bijection particulière φ que l'on a pu choisir. Nous allons donner sans détailler les arguments, des conséquences pratiques qui serviront pour traiter les exercices :

— Ce que l'on vient de dire montre que si $\psi : B \to A$ est bijective alors

$$\sum_{a \in A} f(a) = \sum_{b \in B} f(\psi(b)).$$

Les variables a et b sont muettes (nous aurions pu écrire a à la place de b dans la formule de droite). Cette règle revient à effectuer un changement de variable $a := \psi(b)$ dans la somme, donc un changement de l'ordre des termes qui ne modifie pas la valeur de la somme. Un exemple plus concret : en posant $i = n - k + 1$, on a

$$\sum_{i=1}^{n} u_i = \sum_{k=1}^{n} u_{n-k+1}$$

qui signifie $u_1 + u_2 + \cdots + u_n = u_n + u_{n-1} + \cdots + u_1$.

— Présentons un autre exemple de ce principe, tout aussi utile pour les calculs : l'interversion de sommes (finies) :

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{i,j} \right).$$

Si on note $I := [\![1, m]\!]$, $J := [\![1, n]\!]$ et $f : I \times J \to \mathbb{R}$ telle que $f(i, j) = a_{i,j}$, on se rend compte que les deux quantités qui précèdent valent

$$\sum_{x \in I \times J} f(x),$$

mais qu'elles correspondent à deux manières différentes d'énumérer les éléments de $I \times J$. Dans l'exemple concret $I = \{1, 2\}$, $J = \{1, 2, 3\}$ la formule d'interversion dit simplement que

$$(a_{1,1} + a_{1,2} + a_{1,3}) + (a_{2,1} + a_{2,2} + a_{2,3}) = (a_{1,1} + a_{2,1}) + (a_{1,2} + a_{2,2}) + (a_{1,3} + a_{2,3})\,.$$

— Si $A \subset E$ sont des ensembles finis et $1_A : E \to \{0, 1\}$ est la fonction caractéristique de A ($1_A(x)$ vaut 1 si $x \in A$ et 0 si $x \notin A$). Alors on a la relation

$$\text{card } A = \sum_{x \in E} 1_A(x),$$

qui est évidente si on y réfléchit puisque l'on compte 1 pour chaque élément de A.

— Pour finir nous reformulons la proposition 2.2.3 dans le cadre des sommes sur les ensembles : Soit I un ensemble fini et soit une application $i \mapsto A_i$ de I dans $\mathcal{P}(E)$ où E est un ensemble fini. On suppose que $i \neq j \Longrightarrow A_i \cap A_j = \emptyset$. Alors

$$\text{card } \left(\bigcup_{i \in I} A_i \right) = \sum_{i \in I} \text{card } (A_i).$$

Voici une application fondamentale de ce dernier résultat :

Théorème 2.2.4. *Soit $f : E \to F$ une application entre ensembles finis. Alors*

$$\text{card } E = \sum_{y \in F} \text{card } \big(f^{-1}(\{y\})\big).$$

Démonstration. — On note que $E = \cup_{y \in F} f^{-1}(\{y\})$ (l'inclusion $\subset$ vient du fait que tout élément $e \in E$ est antécédent de son image donc $e \in f^{-1}(\{f(e)\})$; l'inclusion dans l'autre sens est évidente). Par ailleurs si $y, z \in F$ et $y \neq z$, $f^{-1}(\{y\}) \cap f^{-1}(\{z\})$ est vide (si x appartient à cet ensemble, $f(x)$ doit valoir à la fois y et z, ce qui est impossible). On peut donc conclure en disant que le cardinal de l'union disjointe vaut la somme des cardinaux. ∎

Ce théorème explique comment compter les éléments d'un ensemble à l'aide d'une application. Voici des exemples de cette méthode :

Théorème 2.2.5. *Soit $f : E \to F$ une application entre ensembles finis.*

(1) Si f est bijective, alors card $E =$ card F.

(2) Si f est injective, alors card $E \leq$ card F.

(3) Si f est surjective, alors card $E \geq$ card F.

(4) On suppose que card $E =$ card F. *Alors f est injective si, et seulement si, elle est surjective. (Naturellement dans ce cas elle est alors bijective.)*

Démonstration. — Pour le premier point : comme f est bijective, tout élément $y \in F$ admet un unique antécédent, ce qui revient à card $\big(f^{-1}(\{y\})\big) = 1$. Le théorème 2.2.4 donne alors

$$\text{card } E = \sum_{y \in F} \text{card } \big(f^{-1}(\{y\})\big) = \sum_{y \in F} 1 = \text{card } F.$$

Pour le second point, comme f est injective, tout élément $y \in F$ admet au plus un antécédent, donc card $\left(f^{-1}(\{y\})\right) \leq 1$ et on termine avec la formule du théorème 2.2.4 :

$$\text{card } E = \sum_{y \in F} \text{card } \left(f^{-1}(\{y\})\right) \leq \sum_{y \in F} 1 = \text{card } F.$$

Lorsque f est surjective, tout élément de l'ensemble d'arrivée a au moins un antécedent, donc card $\left(f^{-1}(\{y\})\right) \geq 1$ et l'on termine comme précédemment.

Pour la preuve du dernier point : par injectivité on sait que card $\left(f^{-1}(\{y\})\right) \leq 1$ est vrai pour tout $y \in F$. Par hypothèse et par le théorème 2.2.4 :

$$\text{card } F = \text{card } E = \sum_{y \in F} \text{card } \left(f^{-1}(\{y\})\right) \leq \sum_{y \in F} 1 = \text{card } F.$$

Il y a donc égalité dans l'inégalité intermédiaire, ce qui oblige à avoir card $\left(f^{-1}(\{y\})\right) = 1$ (et pas < 1) pour tout $y \in F$. Ainsi tout $y \in F$ admet un unique antécédent et f est bijective. ∎

Il faut retenir qu'**un moyen fondamental de démontrer que deux ensembles ont le même nombre d'éléments est d'exhiber une bijection entre les deux**. De même pour montrer que l'ensemble E n'a pas plus d'éléments que F, il suffit d'exhiber une injection de E dans F ; pour montrer que E a au moins autant d'éléments que F, il suffit d'exhiber une surjection de E sur F.

Une autre conséquence remarquable est que si E est fini, toute injection de E dans E est automatiquement bijective. Cette propriété est n'est plus vraie pour les ensembles infinis : par exemple l'application $n \mapsto n+1$ de $\mathbb{N}$ dans $\mathbb{N}$ est injective mais 0 n'est pas dans l'image.

Exercice 36. Soit $f : E \to F$ une application, dont l'ensemble de départ est fini. Montrer que card $(f(E)) \leq$ card (E).

Par contraposition, le second point du théorème précédent donne le résultat suivant, connu sous le nom de "Principe des tiroirs de Dirichlet" et "pigeonhole principle" en anglais ! !

Corollaire 2.2.6 (Principe des tiroirs)**.** *Soit $f : E \to F$ une application entre ensembles finis. Si* card $E >$ card F*, il existe $a \neq b \in E$ tels que $f(a) = f(b)$.*

Exemples. Dans un groupe de 27 personnes, il y en a au moins deux dont le nom commence par la même lettre. Le nombre de cheveux d'une personne est toujours inférieur à 200000 ; il y a donc à Toulouse deux personnes qui ont le même nombre de cheveux.

Voici une autre conséquence immédiate du théorème 2.2.4, utile pour les exercices :

Corollaire 2.2.7 (Principe des bergers)**.** *Soit $f : E \to F$ une application entre ensembles finis. On suppose que toutes les images réciproques $f^{-1}(\{y\})$, $y \in F$, ont le même nombre d'éléments q. Alors*

$$\text{card } E = q \, \text{card } F.$$

Exemple. Pour compter des moutons, il suffit de compter les pattes et de diviser par 4. (Des applications plus significatives suivront !)

Le principe des bergers permet aussi d'établir l'expression du cardinal d'un produit cartésien :

Corollaire 2.2.8. *Si A et B sont des ensembles finis quelconques,*

$$\text{card } (A \times B) = \text{card } A \times \text{card } B.$$

Démonstration. — On considère l'application $f : A \times B \to A$ définie par $f((x, y)) = x$. Pour tout $y \in B$, $f^{-1}(\{y\}) = \{(x, y);\ x \in A\}$ est clairement en bijection avec A (par $x \mapsto (x, y)$) donc a même cardinal que A. On en déduit que le cardinal de l'ensemble de départ vaut card A fois le cardinal de l'ensemble d'arrivée. ∎

2.3 Analyse combinatoire

Nous allons maintenant calculer le cardinal de plusieurs objets mathématiques (ensembles d'applications, de bijections, d'injections, des sous-ensembles d'un ensembles, etc...). Nous nous attacherons à mettre en avant les techniques standard, réutilisables dans les exercices.

2.3.1 Applications

Puissances. Fixons $a \in \mathbb{N}$. On définit a^n par récurrence sur n : $a^0 = 1$ et, pour tout $n \in \mathbb{N}$, $a^{n+1} := a.a^n$. On démontre (par récurrence !) les formules classiques : $a^{m+n} = a^m.a^n$, $(ab)^m = a^m b^m$ et $(a^m)^n = a^{mn}$.

Rappelons que $\mathcal{F}(E, F)$ désigne l'ensemble des applications de E dans F. On le note également F^E, ce qui se peut se justifier par l'analogie avec la formule suivante :

Théorème 2.3.1. *Soient E et F des ensembles finis. On a :* $\text{card } \mathcal{F}(E, F) = (\text{card } F)^{\text{card } E}$.

Démonstration. — Notons $n := \text{card } E$ et $q := \text{card } F$. L'idée de la preuve est simple : si on note $E = \{e_1, \ldots, e_n\}$, une application $f : E \to F$ est déterminée par les valeurs $(f(e_1), f(e_2), \ldots, f(e_n))$. Or il y a $q = \text{card } F$ choix possibles pour $f(e_1)$, q choix possibles pour $f(e_2)$ et ainsi de suite... on arrive à q^n applications différentes de E vers F. Ce qui se cache derrière le "flou des pointillés", c'est une preuve par récurrence. Nous en donnons maintenant une présentation détaillée car elle suit une méthode classique réutilisable.

La preuve se fait par récurrence sur $n = \text{card } E$. On a posé $q = \text{card } F$. Pour $n = 0$, il faut admettre que $\text{card } \mathcal{F}(\emptyset, F) = 1$. Si l'on trouve cette affirmation trop étrange (elle est pourtant rigoureusement exacte), on n'a qu'à l'admettre comme une pure convention et entamer la récurrence avec $n = 1$. Dans ce cas, E est un singleton : $E = \{x\}$ et on vérifie sans peine que l'application $\mathcal{F}(\{x\}, F) \to F$ qui à la fonction f associe la valeur $f(x)$ est une bijection.
Supposons l'affirmation vraie pour $\text{card } E = n$ et prouvons la pour $\text{card } E = n+1$. On écrit $E = E' \cup \{x\}$, où $\text{card } E' = n$ et où $x \notin E'$. L'hypothèse de récurrence nous dit que $\text{card } \mathcal{F}(E', F) = q^n$. Nous allons prouver, que $\text{card } \mathcal{F}(E, F) = q \, \text{card } \mathcal{F}(E', F)$. Considérons en effet l'application de restriction $f \mapsto f_{|E'}$ de $\mathcal{F}(E, F)$ dans $\mathcal{F}(E', F)$. L'image réciproque de $g \in \mathcal{F}(E', F)$ est formée des $f \in \mathcal{F}(E, F)$ qui prennent sur E' les mêmes valeurs que g et qui prennent en x une valeur arbitraire dans F. Il y a donc q telles applications f et le principe des bergers nous donne la conclusion. On a donc : $\text{card } \mathcal{F}(E, F) = q.q^n = q^{n+1}$, vue la définition par récurrence des puissances, ce qui achève la démonstration. ∎

Dans les dénombrements par récurrence, on essaie souvent de trouver une sous-structure de la famille à dénombrer qui soit de même nature. Dans la preuve précédente, on restreint les applications au sous ensemble E' par exemple et l'on remarque que ce sont toutes les applications de E' dans F.

2.3.2 Sous-ensembles

Etant donné un ensemble fini, on veut compter le nombre de ses sous-ensembles. **Un moyen efficace pour calculer le cardinal d'un ensemble est de montrer qu'il est en bijection avec un ensemble dont on connaît déjà le cardinal.** C'est ce que nous allons faire.

À tout sous-ensemble $A \subset E$, on associe sa *fonction caractéristique* 1_A : c'est l'application de E dans $\{0, 1\}$ définie par :

$$\forall x \in E \ , \ 1_A(x) := \begin{cases} 1 \text{ si } x \in A, \\ 0 \text{ si } x \notin A. \end{cases}$$

Théorème 2.3.2. *L'application $A \mapsto 1_A$ est une bijection de $\mathcal{P}(E)$ sur $\mathcal{F}(E, \{0,1\})$.*

Démonstration. — Notons χ l'application $\mathcal{P}(E) \to \mathcal{F}(E, \{0,1\})$ définie par $A \mapsto 1_A$. Pour vérifier que c'est une bijection, nous devons montrer que tout élément de l'ensemble d'arrivée de χ admet un unique antécédent par χ. Soit donc $f \in \mathcal{F}(E, \{0,1\})$ quelconque ; cherchons tous les $A \in \mathcal{P}(E)$ tels que l'image de A par χ vaut f. Nous devons donc résoudre une équation d'inconnue A (qui est un sous-ensemble de $\{0,1\}$) et de paramètre f (qui est une application de E dans $\{0,1\}$). La voici :

$$\begin{aligned} \chi(A) = f &\iff 1_A = f \quad \iff \forall x \in E,\, 1_A(x) = f(x) \\ &\iff \forall x \in A,\, f(x) = 1 \quad \text{et} \quad \forall x \in E \setminus A,\, f(x) = 0 \\ &\iff A \subset f^{-1}(\{1\}) \quad \text{et} \quad E \setminus A \subset f^{-1}(\{0\}) \\ &\iff A \subset f^{-1}(\{1\}) \quad \text{et} \quad E \setminus f^{-1}(\{0\}) \subset A. \end{aligned}$$

Comme $E \setminus f^{-1}(\{0\}) = \{x \in E;\, f(x) \neq 0\} = \{x \in E;\, f(x) = 1\} = f^{-1}(\{1\})$, puisque f est à valeurs dans $\{0,1\}$, nous avons montré

$$\chi(A) = f \iff A = f^{-1}(\{1\}).$$

Ceci montre que f admet un unique antécédent (c'est l'ensemble des points où f vaut 1, ce que l'on aurait pu deviner sans calcul !). Donc χ est une bijection.

∎

Corollaire 2.3.3. *Soit E un ensemble fini. On a :* $\operatorname{card} \mathcal{P}(E) = 2^{\operatorname{card} E}$.

Remarque. *L'intérêt de l'argument qui précède, en plus de montrer que $\mathcal{F}(E, \{0,1\})$ et $\mathcal{P}(E)$ ont le même cardinal, est de l'expliquer par une bijection simple. Remarquons que lorsque* $\operatorname{card} A = \operatorname{card} B = n$, *par définition il existe deux bijections $\phi : [\![1,n]\!] \to A$ et $\psi : [\![1,n]\!] \to B$. Alors $\psi \circ \phi^{-1}$ est une bijection de A dans B. Nous savons qu'elle existe, mais nous ne savons pas forcément la décrire et elle est peut-être compliquée. Un réflexe naturel en combinatoire est, lorsque l'on a montré par des méthodes séparées que deux ensembles ont le même cardinal, de se demander s'il existe une raison simple à cette égalité, c'est-à-dire s'il existe une bijection facile à décrire entre les deux ensembles.*

Exercice 37. (*) On propose une autre méthode pour prouver le thérorème 2.3.2. Soit $\chi : \mathcal{P}(E) \to \mathcal{F}(E, \{0,1\})$ l'application définie par $A \mapsto 1_A$. Considérons l'application $\mathrm{Supp} : \mathcal{F}(E, \{0,1\}) \to \mathcal{P}(E)$, qui à toute application $\phi : E \to \{0,1\}$ associe son *support* $\mathrm{Supp}(\phi) := \phi^{-1}(\{1\}) = \{x \in E,\, \phi(x) = 1\}$. Vérifier les relations suivantes

$$\chi \circ \mathrm{Supp} = \mathrm{Id}_{\mathcal{F}(E,\{0,1\})} \quad \text{et} \quad \mathrm{Supp} \circ \chi = \mathrm{Id}_{\mathcal{P}(E)},$$

puis conclure à l'aide du théorème 2.1.2.

2.3.3 Arrangements, injections

Soient E et F deux ensembles finis ayant respectivement m et n éléments. Nous noterons $I(E,F)$ l'ensemble des applications injectives de E dans F. Si $m > n$, il n'y en a aucune et $I(E,F) = \emptyset$. Nous allons calculer $\operatorname{card} I(E,F)$ lorsque $m \leq n$.

Remarquons d'abord que si $\operatorname{card} E = \operatorname{card} E'$ et $\operatorname{card} F = \operatorname{card} F'$, alors $\operatorname{card} I(E,F) = \operatorname{card} I(E',F')$. Détaillons l'argument (nous l'utiliserons sans répéter la preuve dans d'autres situations) : considérons des bijections $\phi : E \to E'$ et $\psi : F \to F'$. Si $f : E \to F$ est une injection, alors l'application $\psi \circ f \circ \phi^{-1} : E' \to F'$ est encore une injection (voir l'exercice 33). On peut donc définir une application $K : I(E,F) \to I(E',F')$ par $K(f) := \psi \circ f \circ \phi^{-1}$. Le même argument, en inversant les rôles, donne une application $L : I(E', F') \to I(E, F)$ définie par $K(g) := \psi^{-1} \circ g \circ \phi$. Un calcul immédiat montre que $L \circ K = \mathrm{Id}_{I(E,F)}$ et $K \circ L = \mathrm{Id}_{I(E',F')}$. Donc K et L sont des bijections réciproques et on a établi card $I(E, F) = \text{card } I(E', F')$.

Ainsi card $I(E, F)$ ne dépend que de $m = \text{card } E$ et $n = \text{card } F$. On le note A_n^m (donc $A_n^m = 0$ lorsque $m > n$). En particulier, on peut aussi bien supposer que $E = [\![1, m]\!]$. Dans ce cas, se donner une application $f : [\![1, m]\!] \to F$ revient à se donner une suite $(f(1), \ldots, f(m)) \in F^m$. De plus une application f injective correspond à une suite $(f(1), \ldots, f(m))$ dont tous les éléments sont distincts.

On appelle *arrangement de m objets pris parmi les n éléments de F* toute suite $(y_1, \ldots, y_m)$ de m éléments distincts de F. Ainsi compter les injections revient à compter les arrangements.

Exercice 38. Décrire tous les arrangements de 1, 2, 3 objets pris parmi les 4 éléments $\{1, 2, 3, 4\}$.

Lemme 2.3.4. *Si $0 \leq m \leq n-1$, on a : $A_n^{m+1} = (n-m)A_n^m$.*

Démonstration. — À toute suite $(y_1, \ldots, y_{m+1})$ de $(m+1)$ éléments distincts de F, associons la suite $(y_1, \ldots, y_m)$ de m éléments distincts de F. On obtient ainsi une application ϕ de $I([\![1, m+1]\!], F)$ dans $I([\![1, m]\!], F)$. L'image réciproque de $(y_1, \ldots, y_m)$ par ϕ est formée de toutes les suites $(y_1, \ldots, y_m, y)$ telles que $y \in F \setminus \{y_1, \ldots, y_m\}$: cette image réciproque a donc $(n-m)$ éléments. D'après le principe des bergers, card $I([\![1, m+1]\!], F) = (n-m)$ card $I([\![1, m]\!], F)$. ∎

Théorème 2.3.5. *Le nombre d'injections d'un ensemble à m éléments dans un ensemble à n éléments vaut*

$$A_n^m = \begin{cases} 0 & \text{si} \quad m > n \\ \dfrac{n!}{(n-m)!} = \displaystyle\prod_{i=0}^{m-1}(n-i) & \text{si} \quad m \leq n \end{cases}$$

Il est aussi égal au nombre d'arrangements de m objets pris parmi n.

Démonstration. — Il nous reste seulement à démontrer la formule pour A_n^m lorsque $m \leq n$. On le fait par récurrence sur m. Pour $m = 0$, l'unique application de $\emptyset$ dans F est injective et l'on trouve $A_n^0 = 1 = \dfrac{n!}{n!}$, ce qui est correct. Si l'on trouve l'argument trop bizarre, on admet cette valeur comme une convention et l'on initialise la récurrence à $m := 1$. Dans ce cas, E est un singleton et les n applications de E dans F sont injectives : on a bien $A_n^1 = \dfrac{n!}{(n-1)!} = n$. On peut également dire que les arrangements de 1 objet pris parmi n sont ici les n suites (y) de 1 élément de F.
Supposons maintenant que $A_n^m = \dfrac{n!}{(n-m)!}$ pour un certain $m \in [\![0, n-1]\!]$. En combinant le lemme et l'hypothèse de récurrence, on trouve :

$$A_n^{m+1} = (n-m)A_n^m = (n-m)\frac{n!}{(n-m)!} = \frac{n!}{\big(n-(m+1)\big)!},$$

d'où la première égalité. L'égalité $\dfrac{n!}{(n-m)!} = \displaystyle\prod_{i=0}^{m-1}(n-i)$ est immédiate. ∎

Lorsque $m = n$, toute application de E dans F est injective si et seulement si elle est bijective. Donc A_n^n représente le nombre de bijections entre deux ensembles de cardinal n. Le nombre A_n^n représente aussi le nombre d'arrangements $(y_1, \ldots, y_n)$ formés des n éléments de F, chacun étant (bien entendu) présent une fois et une seule. Une telle suite est appelée *permutation* de F. Une définition essentiellement équivalente est celle-ci : une permutation de F est une bijection de F dans lui-même.

Exercice 39. Décrire toutes les permutations de $F := \{1, 2, 3\}$.

Théorème 2.3.6. *Le nombre de bijections entre deux ensembles à n éléments est $n!$.*
Le nombre de permutations d'un ensemble à n éléments est $n!$.

Démonstration. — C'est $A_n^n = \dfrac{n!}{(n-n)!} = \dfrac{n!}{0!} = n!$. ∎

2.3.4 Combinaisons

Soit E un ensemble à n éléments. Lorsque $0 \leq m \leq n$, on appelle *combinaison de m objets pris parmi les n éléments de E* un sous-ensemble $\{y_1, \ldots, y_m\}$ formé de m éléments distincts de E : ce n'est donc rien d'autre qu'un sous-ensemble à m éléments de n. La différence entre une combinaison et un arrangement, c'est que l'ordre importe dans un arrangement mais pas dans une combinaison. Le nombre de ces combinaisons ne dépend évidemment que de m et de n. On le note traditionnellement C_n^m et, de manière plus moderne (influencée par l'univers anglo-saxon !) $\binom{n}{m}$, ce qui se lit : "choix de m parmi n". Les $C_n^m = \binom{n}{m}$ sont appelés *coefficients binomiaux* pour des raisons qui apparaitront à la section 2.4.3. On convient que $\binom{n}{m} = 0$ lorsque $m > n$. (Pourquoi ?) Il est souvent commode de noter $\mathcal{P}_k(E)$ l'ensemble de tous les sous-ensembles de E qui ont pour cardinal k. On retiendra que par définition, pour E fini, on a

$$\operatorname{card} \mathcal{P}_k(E) = \binom{\operatorname{card} E}{k}.$$

Exercice 40. Décrire toutes les combinaisons de 1, 2, 3 objets pris parmi les 4 éléments $\{1, 2, 3, 4\}$.

Théorème 2.3.7. *Lorsque $0 \leq m \leq n$, le nombre de combinaisons de m objets pris parmi n est donné par la formule :*

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!} = \frac{\prod_{i=0}^{m-1}(n-i)}{m!}.$$

Démonstration. — À chaque arrangement $(y_1, \ldots, y_m)$ dans E, associons l'ensemble $\{y_1, \ldots, y_m\}$ sous-jacent, obtenu en oubliant l'ordre. Les arrangements ayant pour image une combinaison $\{y_1, \ldots, y_m\}$ donnée sont les $m!$ permutations de $(y_1, \ldots, y_m)$. D'après le principe des bergers, on a donc $A_n^m = m! \binom{n}{m}$, d'où la conclusion. ∎

Exercice 41. Calculer sans calculatrice le nombre de combinaisons de 3 objets pris parmi 15.

Corollaire 2.3.8. *Le nombre d'applications strictement croissantes de $[\![1, m]\!]$ dans $[\![1, n]\!]$ est $\binom{n}{m}$.*

Démonstration. — L'idée est qu'une application strictement croissante f de $[\![1, m]\!]$ dans $[\![1, n]\!]$ est totalement déterminée par son image $\{y_1, \ldots, y_m\} \subset [\![1, n]\!]$: le plus petit élément est $f(1)$, le suivant est $f(2)$, etc. Nous laissons le lecteur formuler ceci en termes de bijections. ∎

2.4 Propriétés des coefficients binomiaux

2.4.1 Variations

Nous décrivons les variations de $m \mapsto \binom{n}{m}$ pour n fixé. On ne s'intéresse qu'à $0 \leq m \leq n$ car sinon les coefficients sont nuls.

Proposition 2.4.1. *Pour* $0 \leq m \leq n$*, on a :* $\dbinom{n}{m} = \dbinom{n}{n-m}$.

Démonstration. — La formule est immédiate par calcul sur l'expression $\binom{n}{m} = \dfrac{n!}{m!\,(n-m)!}$.

On peut également la justifier en remarquant que, si card $E = n$, l'application $F \mapsto E \setminus F$ (passage au complémentaire dans E) est une bijection de l'ensemble des parties de E à m éléments sur l'ensemble des parties de E à $(n-m)$ éléments. ∎

Proposition 2.4.2. *Si n est pair, l'application $m \mapsto \binom{n}{m}$ croît strictement de $m = 0$ à $m = n/2$ (où elle prend sa valeur maximum) puis décroît strictement de $m = n/2$ à $m = n$*

Si n est impair, cette suite croît strictement de $m = 0$ à $m = (n-1)/2$, prend la même valeur (son maximum) en $m = (n-1)/2$ et $m = (n+1)/2$, puis décroît strictement de $m = (n+1)/2$ à $m = n$.

Démonstration. — Soient m, n tels que $0 \leq m \leq n-1$. On obtient après simplification :

$$\frac{\binom{n}{m+1}}{\binom{n}{m}} = \frac{n-m}{m+1}.$$

On déduit que $\binom{n}{m+1} > \binom{n}{m}$ si et seulement si, $2m \leq n$. On en tire les variations de la suite des $\binom{n}{m}$ à n fixé. ∎

2.4.2 Le triangle de Pascal

Théorème 2.4.3 (Formule de Pascal)**.** *Pour tous $m, n \in \mathbb{N}$ on a :*

$$\binom{n+1}{m+1} = \binom{n}{m+1} + \binom{n}{m}.$$

Démonstration. — Si $m > n$, les deux membres de l'égalité sont nuls. Si $m = n$, $\binom{n+1}{m+1} = \binom{n}{m} = 1$ et $\binom{n}{m+1} = 0$, et les deux membres de l'égalité sont égaux à 1. On peut donc supposer que $0 \leq m < n$. Nous disposons dans ce cas de deux preuves tout à fait différentes. La preuve calculatoire vient immédiatement à l'esprit :

$$\begin{aligned}
\binom{n}{m+1} + \binom{n}{m} &= \frac{n!}{(m+1)!\,\big(n-(m+1)\big)!} + \frac{n!}{m!\,(n-m)!} \\
&= (n-m)\frac{n!}{(m+1)!\,(n-m)!} + (m+1)\frac{n!}{(m+1)!\,(n-m)!} \\
&= (n+1)\frac{n!}{(m+1)!\,(n-m)!} \\
&= \frac{(n+1)!}{(m+1)!\,(n-m)!} = \binom{n+1}{m+1}.
\end{aligned}$$

La preuve combinatoire revient à montrer que les deux nombres comptent la même chose. Considérons un ensemble E à n éléments soit $E' := E \cup \{x\}$ avec $x \notin E$. Donc E' a $(n+1)$ éléments. L'entier $\binom{n+1}{m+1}$ est le nombre de parties de E' qui ont $(m+1)$ éléments. Il y en a de deux types :

(1) Les parties à $(m+1)$ éléments de E : il y en a $\binom{n}{m+1}$.

(2) Les $F \cup \{x\}$, où F est une partie à m éléments de E : il y en a $\binom{n}{m}$.

Au total, on a bien $\binom{n}{m+1}+\binom{n}{m}$ parties à $(m+1)$ éléments de E'. ∎

On peut calculer les coefficients binomiaux à l'aide du *triangle de Pascal* :

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}$$

La règle de formation est la suivante : les côtés du triangle sont formés de 1 ; chaque coefficient du tableau est la somme de ceux qui lui sont supérieurs juste à gauche et juste à droite. Sur la n^{e} ligne on trouve alors de gauche à droite les $\binom{n}{m}$ pour $m=0,1,\ldots,n$.

La formule de Pascal permet également un calcul "récursif" des coefficients binomiaux par l'algorithme suivant :

```
Pasc(n,m)
si m = 0 alors rendre 1 sinon si n = 0 alors rendre 0
sinon rendre Pasc(n-1,m) + Pasc(n-1,m-1) ;;
```

Le lecteur intéressé pourra rechercher la "complexité" de cet algorithme ; par exemple, combien d'additions requiert-il ? Et dans quelle mesure les calculs sont-ils redondants ?

Il est très facile d'obtenir le triangle de Pascal avec un tableur : On remplit la première colonne avec des 1 (par copier-coller bien entendu !). Ensuite on remplit les cases vides de la première ligne avec des 0. Enfin dans toutes les cases restantes, on entre la même formule qui dit que la valeur de la case est la somme de la valeur de la case immédiatement au-dessus et de la case au-dessus à gauche. Le résultat apparaît sous une forme moins symétrique :

1	0	0	0	0	0	0
1	1	0	0	0	0	0
1	2	1	0	0	0	0
1	3	3	1	0	0	0
1	4	6	4	1	0	0
1	5	10	10	5	1	0
1	6	15	20	15	6	1

2.4.3 La formule du binôme de Newton

Théorème 2.4.4 (Formule du binôme de Newton)**.** *Soient a et b deux nombres complexes. On a alors, pour tout $n \in \mathbb{N}$:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Démonstration. — Elle se fait par récurrence sur n. Pour $n=0$, il s'agit de vérifier que $(a+b)^0 = \sum\limits_{k=0}^{0} \binom{0}{k} a^k b^{0-k}$, autrement dit, que $1=\binom{0}{0}a^0b^0$, ce qui est bien vrai. Supposons la formule vraie au rang n. On calcule alors :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n = (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= \sum_{j=1}^{n+1}\binom{n}{j-1}a^j b^{n-j+1} + \sum_{k=0}^{n+1}\binom{n}{k}a^k b^{n-k+1} \\
&= \sum_{j=1}^{n+1}\left(\binom{n}{j-1}+\binom{n}{j}\right)a^j b^{n-j+1} + b^{n+1} \\
&= \sum_{j=1}^{n+1}\binom{n+1}{j}a^j b^{n-j+1} + \binom{n}{0}a^0 b^{n-0+1} \\
&= \sum_{j=0}^{n+1}\binom{n+1}{j}a^j b^{n+1-j}.
\end{aligned}$$

∎

Un peu d'algèbre combinatoire. Si l'on développe $(a+b)^n = (a+b)\cdots(a+b)$ (n facteurs), on voit apparaitre des termes de la forme $a^k b^{n-k}$. Chacun de ces termes apparait autant de fois qu'il y a de choix des k facteurs $(a+b)$ dans lesquels on prend a plutôt que b, donc, au total $\binom{n}{k}$ fois : c'est une autre preuve de la formule de Newton. Notons que les seules propriétés utilisées sont la commutativité et l'associativité de l'addition et de la multiplication, ainsi que la distributivité.

Exemple. Une application directe de la formule du binôme donne

$$2^n = (1+1)^n = \sum_{k=0}^{n}\binom{n}{k}.$$

On peut interpréter (et démontrer) cette formule de manière combinatoire : un ensemble E de cardinal n a au total 2^n parties, dont $\binom{n}{0}$ à 0 éléments, $\binom{n}{1}$ à 1 élément, $\binom{n}{2}$ à 2 éléments,...

Exercice 42. Développer les expressions $(1-x)^3$, $(1+x)^6$ et $(1+x)^7$.

2.5 Compléments

2.5.1 Et bien d'autres formules...

Les coefficients binomiaux satisfont beaucoup de relations, ce qui explique l'attention qui leur a été portée. Nous donnons un exemple parmi tant d'autres, ainsi qu'une preuve combinatoire :

$$\sum_{p=0}^{q}\binom{m}{p}\binom{n}{q-p} = \binom{m+n}{q}.$$

Soient E et F deux ensembles disjoints ayant respectivement m et n éléments. L'ensemble $E \cup F$ a $(m+n)$ éléments, donc $\binom{m+n}{q}$ sous-ensembles à q éléments. Chacun de ces sous-ensembles est de la forme $E' \cup F'$, où $E' \subset E$ a p éléments (pour un p tel que $0 \leq p \leq q$) et où $F' \subset F$ a $(q-p)$ éléments. Pour chaque p, il y a $\binom{m}{p}\binom{n}{q-p}$ tels ensembles $E' \cup F'$, et leur nombre total est bien $\sum_{p=0}^{q}\binom{m}{p}\binom{n}{q-p}$.

Une preuve algébrique est proposée à l'exercice 79. L'exercice 71 présente une autre formule remarquable.

2.5.2 Nombre de surjections

Notons $n := \text{card } E$ et $p := \text{card } F$. Nous allons compter le nombre $S(n,p)$ d'applications surjectives de E dans F. Naturellement, on peut tout aussi bien supposer que $F = [\![1,p]\!]$, ce que nous ferons. Pour tout $i \in F$, soit $\mathcal{F}_i := \{f \in \mathcal{F}(E,F) \mid i \notin \text{Im} f\}$. Par définition, l'ensemble des surjections est égal à $\mathcal{F}(E,F) \setminus \bigcup_{i \in F} \mathcal{F}_i$, de sorte que $S(n,p) = p^n - \text{card } \bigcup_{i \in F} \mathcal{F}_i$. Nous allons calculer le cardinal de $\bigcup_{i \in F} \mathcal{F}_i$ à l'aide de la formule du crible :

$$\text{card } \bigcup_{i \in F} \mathcal{F}_i = \text{card } \bigcup_{i=1}^{p} \mathcal{F}_i = \sum_{k=1}^{p} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq p} \text{card } (\mathcal{F}_{i_1} \cap \cdots \cap \mathcal{F}_{i_k}).$$

L'ensemble $\mathcal{F}_{i_1} \cap \cdots \cap \mathcal{F}_{i_k}$ est formé des applications $f : E \to F$ telles que $i_1, \ldots, i_k \notin \text{Im} f$, autrement dit, des applications de E dans $F \setminus \{i_1, \ldots, i_k\}$. Comme ce dernier ensemble a $p-k$ éléments, on déduit du théorème : $\text{card } (\mathcal{F}_{i_1} \cap \cdots \cap \mathcal{F}_{i_k}) = (p-k)^n$. Puisqu'il y a $\binom{p}{k}$ parties à k éléments $\{i_1, \ldots, i_k\} \subset F$:

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

2.5.3 Combinatoire des ensembles infinis, dénombrabilité

On peut faire de la combinatoire des ensembles infinis en s'inspirant du cas fini. On dira que deux ensembles E et F ont "autant d'éléments" s'il existe une bijection de E dans F. Cette définition correspond bien à ce que nous avons fait pour les ensembles finis, mais elle nous réserve des surprises dans le cas infini. Ainsi, l'application $f : \mathbb{N} \to \mathbb{N}^*$ définie par $n \mapsto n+1$ est une bijection, donc $\mathbb{N}^*$ a autant d'éléments que $\mathbb{N}$ alors même qu'on lui a enlevé l'élement 0 (au fond ce n'est pas si choquant, l'infini moins un reste infini...). On peut aussi voir que $\mathbb{N}$ a autant d'éléments que l'ensemble $2\mathbb{N}$ des entiers pairs (exercice 48) et même que l'ensemble $\mathbb{N} \times \mathbb{N}$ (exercice 49).

Poursuivant l'analogie avec le cas fini (trois premier points du théorème 2.2.5), on a envie de dire que "E n'a pas (strictement) plus d'éléments que F" signifie pour des ensembles généraux (finis ou infinis) qu'il existe une injection de E dans F, et que "E n'a pas (strictement) moins d'éléments que F" signifie qu'il existe une surjection de E dans F. Il n'est pas évident que ces notions soient cohérentes : on voudrait que "E n'a pas plus d'éléments que F" entraine que "F n'a pas moins d'éléments que E". On voudrait aussi que si E n'a ni plus, ni moins d'éléments que F ils aient le même nombre d'éléments ! ! Ce n'est pas évident pour des ensembles infinis, mais c'est vrai comme l'affirme le résultat suivant (que nous admettrons) :

Proposition 2.5.1. *On a les propriétés suivantes*

(1) S'il existe une injection $E \to F$, il existe une surjection $F \to E$.

(2) S'il existe une surjection $E \to F$, il existe une injection $F \to E$.

(3) S'il existe une injection $E \to F$ et une surjection $E \to F$, il existe une bijection $E \to F$.

Cette propriété n'est pas très difficile à montrer en utilisant *l'axiome du choix.* Cet axiome (qu'on ne peut pas démontrer à partir des propriétés déjà énoncées sur les ensembles) nous dit que si on a une application de $f : E \to \mathcal{P}(F)$ telle que, $\forall x \in E,\ f(x) \neq \emptyset$, il existe une application $g : E \to F$ telle que, $\forall x \in E, g(x) \in f(x)$. On le traduit en général par la phrase (beaucoup moins précise) "Si l'on a une famille d'ensembles non vides, on peut choisir un élément de chaque ensemble". Aussi anodin que puisse paraître cet axiome, c'est en fait un outil très puissant (et dangereux à manier sans précautions). On peut par exemple en déduire qu'on peut découper une pomme en un nombre fini de morceaux et qu'en recollant ces morceaux, on obtienne la lune ! Bien sûr, ces morceaux auront une "forme" très difficile à imaginer.

Retenons simplement qu'il existe une bonne notion de cardinal des ensembles infinis et revenons à l'exemple de l'ensemble $\mathbb{N}$. On dira qu'un ensemble E est *dénombrable* s'il a autant d'éléments que $\mathbb{N}$ (s'il est en bijection avec $\mathbb{N}$, ce qui revient à la possibilité de l'écrire $E = \{e_0, e_1, e_2, \ldots\}$). Il est facile de voir que $\mathbb{N}$ est le plus petit des ensembles infinis : tout ensemble infini contient au moins autant d'éléments que $\mathbb{N}$. En effet, si E est infini il n'est pas vide, donc il contient un élément e_1. Il n'est pas non plus de cardinal 1, donc il contient un élément e_2 différent de e_1, il n'est pas de cardinal 3, donc il contient un $e_3 \notin \{e_1, e_2\}$. En répétant le procédé, on construit une injection de $\mathbb{N} \to E$ telle que $n \mapsto e_n$.

Il y a de nombreux ensembles dénombrables qu'on rencontre naturellement. Nous avons évoqué l'ensemble $\mathbb{N} \times \mathbb{N}$ des couples d'entiers. On voit plus facilement que $\mathbb{Z}$ est dénombrable (une manière de l'énumérer est d'écrire $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$). L'ensemble $\mathbb{Q}$ des nombres rationnels est lui aussi dénombrable. Ceci peut surprendre : il semble qu'il y a beaucoup plus de points dans $\mathbb{Q}$ que dans $\mathbb{N}$. Et pourtant, il n'en est rien ! En effet l'application $i : \mathbb{N} \to \mathbb{Q}$, $n \mapsto n$ est injective. Par ailleurs, comme $\mathbb{N}^*$ et $\mathbb{Z}$ sont en bijection avec $\mathbb{N}$,

$\mathbb{N}^* \times \mathbb{Z}$ est en bijection avec $\mathbb{N} \times \mathbb{N}$ donc est aussi dénombrable. Or l'application $\mathbb{N}^* \times \mathbb{Z} \to \mathbb{Q}$, $(p,q) \mapsto q/p$ est clairement surjective. Tout ceci donne une surjection de $\mathbb{N}$ sur $\mathbb{Q}$. La proposition précédente nous assure qu'il existe une bijection entre $\mathbb{N}$ et $\mathbb{Q}$. Enfin mentionons qu'une union finie ou dénombrable d'ensembles dénombrables est dénombrable et qu'un produit fini d'ensembles dénombrables est dénombrable.

Tous les ensembles ne sont pas dénombrables loin de là ! En effet, il n'y a jamais de surjection de E dans $\mathcal{P}(E)$. L'ensemble des parties d'un ensemble a toujours plus d'éléments (beaucoup plus !) que l'ensemble lui-même. Nous le verrons à l'exercice 52. Par conséquent $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable. L'ensemble $\mathbb{R}$ ou $[0,1] = \{x \in \mathbb{R};\ 0 \leq x \leq 1\}$ non plus, voir l'exercice 133. En fait on peut démontrer que $\mathbb{R}$ a autant d'éléments que $\mathcal{P}(\mathbb{N})$.

Exercices

Exercice 43. Démontrer par récurrence que pour tout réel $x \neq 0$ et tout entier $n \geq 0$, on a la formule suivante (où par convention $x^0 = 1$).

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1}.$$

Exercice 44. Montrer par récurrence que les formules suivantes sont valables pour tout entier $n \geq 1$:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Exercice 45. Montrer par récurrence que pour tout $n \geq 1$, $n! \leq n^n$. Montrer ensuite que pour tout $n \geq 1$, $n! \geq \left(\frac{n}{e}\right)^n$. On pourra établir puis utiliser l'inégalité $\ln(x) \leq x - 1$ pour $x > 0$.

Exercice 46. Soit $f : E \to E$ une application. On définit les *itérées* de f par récurrence : $f^{(0)} := \mathrm{Id}_E$ (application identité de E) et pour tout $n \in \mathbb{N}$, $f^{(n+1)} = f \circ f^{(n)}$. Par exemple $f^{(1)} = f \circ \mathrm{Id}_E = f$, $f^{(2)} = f \circ f$.

1) Prouver par récurrence sur n que $\forall n, p \in \mathbb{N}$, $f^{(n+p)} = f^{(n)} \circ f^{(p)}$

2) On prend $E = \mathbb{N}$ et $f(x) = x + 1$. Que vaut $f^{(n)}(x)$?

3) On prend $E = \mathbb{N}$ et $f(x) = 2x$. Que vaut $f^{(n)}(1)$?

Exercice 47. (**) On considère le jeu de la tour de Hanoi de taille $n \geq 1$ (on a trois tiges et n disques de tailles différentes, percés d'un trou. Dans la position initiale tous les disques sont empilés sur une tige, du plus large en bas au plus étroit en haut. Le but du jeu est d'arriver à déplacer la tour sur une autre tige en respectant les règles suivantes : on ne peut déplacer les disques que un par un et d'une tige à une autre ; on ne peut mettre un disque sur une tige que si la tige est vide ou s'il va reposer sur un disque plus large). On note c_n le nombre minimal de coups permettant de déplacer une tour de taille n. Trouver une relation entre c_n et c_{n+1}. En déduire la valeur de c_n.

Exercice 48. Les applications suivantes sont-elles injectives, surjectives, bijectives ?

1) $c : \mathbb{R} \to \mathbb{R}$ telle que $\forall x, c(x) = x^2$.

2) $d : [1, 2] \to [0, 4]$ telle que $\forall x, d(x) = x^2$.

3) $f : \mathbb{N} \to \mathbb{N}$ telle que $\forall n,\ f(n) = n + 1$.

4) $g : \mathbb{N} \to \mathbb{N}^*$ telle que $\forall n,\ g(n) = n + 1$.

5) On note $2\mathbb{N}$ l'ensemble des entiers naturels pairs. On définit $h : \mathbb{N} \to 2\mathbb{N}$ par $n \mapsto 2n$ pour tout $n \in \mathbb{N}$.

6) $k : \mathbb{R} \to \mathbb{R}$ telle que $\forall x, k(x) = x^3 + 3x - 1$.

7) $\ell : \mathbb{R} \to \mathbb{R}$ telle que $\forall x, \ell(x) = x^3 - 3x - 1$.

Exercice 49. (*) Démontrer que l'application $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}^*$ qui au couple (n, p) associe $f(n, p) = 2^n(2p + 1)$ est une bijection.

Exercice 50. (*) Soit E un ensemble non vide, et $A \subset E$. On considère l'application $\phi : \mathcal{P}(E) \to \mathcal{P}(E)$ qui à $B \subset E$ associe $\phi(B) = B \cap A$.

1) Montrez que ϕ est injective si et seulement si $A = E$.

2) Montrez que ϕ est surjective si et seulement si $A = E$.

Exercice 51. (*) Soit E un ensemble non vide, et $A \subset E$. On considère l'application $\phi : \mathcal{P}(E) \to \mathcal{P}(E)$ qui à $B \subset E$ associe $\phi(B) = B \cup A$.

1) Montrez que ϕ est injective si et seulement si $A = \emptyset$.

2) Montrez que ϕ est surjective si et seulement si $A = \emptyset$.

Exercice 52. (**) Dans cet exercice, on montre qu'il ne peut pas avoir d'application surjective de E dans $\mathcal{P}(E)$.

Soit $f : E \to \mathcal{P}(E)$. On considère $A = \{x \in E, x \notin f(x)\}$. Montrez que s'il existe $y \in E$ tel que $A = f(y)$, alors $y \in A \Longleftrightarrow y \notin A$. Conclure.

Exercice 53. Combien de poignées de main échangent n personnes qui se rencontrent ? (on suppose que chacune d'entre elles serre la main de toutes les autres)

Exercice 54. Combien y a-t-il d'applications possibles de $\{1, 2, 3\}$ dans lui même ? Décrire toutes les bijections.

Exercice 55. Dénombrer :

1) les applications de $\mathcal{P}(\{1, 2, 3\})$ à valeurs dans $\{0, 1, 2\}$,

2) les injections de $\{1, 2, 3\}$ dans $\mathcal{P}(\{1, 2, 3, 4, 5\})$.

Exercice 56. Démontrer sans calcul qu'il y a autant d'applications croissantes de $[\![1, m]\!]$ dans $[\![1, n]\!]$ que d'applications décroissantes de $[\![1, m]\!]$ dans $[\![1, n]\!]$.

Exercice 57. En utilisant un alphabet contenant s symboles, combien peut-on écrire de mots de longueur ℓ ? de mots de longueur inférieure ou égale à ℓ ?

Exercice 58. Avec un jeu de 32 cartes, combien y a-t-il de mains (c'est-à-dire d'ensembles) de 5 cartes contenant :

1) un carré d'as ?

2) exactement deux rois ?

3) au moins deux rois ?

4) une suite longue (5 cartes consécutives dans la même famille) ?

5) un full (trois cartes de la même hauteur et deux autres cartes de même hauteur) ?

Exercice 59. On dit qu'un nombre est un *palindrome* si son écriture en base 10 donne le même nombre lorsque qu'on lit de droite à gauche. Par exemple : 131 et 14541 sont des palindromes.

1) Combien y a-t-il de palindromes à deux chiffres ? à trois chiffres ? à n chiffres ?

2) Combien y a-t-il de palindromes entre 1 et 100000 ?

Exercice 60. Combien d'entiers de $[\![100, 999]\!]$ ont au moins un chiffre 7 en écriture décimale ?

Exercice 61. Sur une feuille quadrillée, on considère deux points de la grille : A et B. Par rapport à A, le point B est situé 5 cases plus à droite et 6 cases plus en haut.

1) Quelle est la longueur minimale d'un chemin reliant A et B en restant sur la grille ?

2) Combien y a-t-il de chemins de longueur minimale reliant A à B (en restant sur la grille) ?

Exercice 62. Soit $n \geq 1$. Calculer le nombre d'applications f de $\{1, \ldots, 2n\}$ dans lui-même qui vérifient que pour tout k, les nombres k et $f(k)$ ont la même parité.

Exercice 63. (**) Interpréter les formules $a^{m+n} = a^m.a^n$, $(ab)^m = a^m b^m$ et $(a^m)^n = a^{mn}$ à l'aide de bijections.

Exercice 64. Un carré latin de taille n est un tableau de n lignes et n colonnes, rempli par des nombres de 1 à n et respectant la règle suivante : chaque nombre $k \in \{1, \dots, n\}$ apparaît une fois et une seule dans chaque ligne et dans chaque colonne. Dénombrer les carrés latins de taille 2, puis de taille 3.

Remarques : Un sudoku est un carré latin de taille 9 qui vérifie une règle supplémentaire. Leur dénombrement est une tâche autrement difficile. Il y en a 6709037520210729369600. Par ailleurs, on ne connaît pas à ce jour de formule générale explicite donnant le nombre de carrés latins de taille n.

Exercice 65. (*) Construire explicitement une bijection entre $[\![1, a]\!] \times [\![1, b]\!]$ et $[\![1, ab]\!]$.

Exercice 66. (*) Soit $n \geq 1$ et E un ensemble de cardinal n. On notera $\mathcal{P}_{\text{pair}}(E)$ l'ensemble des parties de E de cardinal pair et $\mathcal{P}_{\text{impair}}(E)$ l'ensemble de celles de cardinal impair. Le but de cet exercice est de démontrer sans calcul que $\text{card}\big(\mathcal{P}_{\text{pair}}(E)\big) = \text{card}\big(\mathcal{P}_{\text{impair}}(E)\big)$.

1) Vérifiez-le lorsque $E = \{1, 2\}$ puis lorsque $E = \{1, 2, 3\}$ (on décrira $\mathcal{P}(E)$).
2) On considère l'application $F : \mathcal{P}(E) \to \mathcal{P}(E)$ définie par $F(A) = A^c$. Montrer que $F \circ F = \text{Id}$. En déduire que F est une bijection de $\mathcal{P}(E)$.
3) Montrer l'application $\widetilde{F} : \mathcal{P}_{\text{pair}}(E) \to F(\mathcal{P}_{\text{pair}}(E))$ définie par $\widetilde{F}(A) = F(A)$ (c'est une restriction de F) est une bijection. En déduire que si n est impair, $\widetilde{F}$ est une bijection entre $\mathcal{P}_{\text{pair}}(E)$ et $\mathcal{P}_{\text{impair}}(E)$ et conclure.
4) Soit $e \in E$ un élément fixé. On considère maintenant l'application $G : \mathcal{P}(E) \to \mathcal{P}(E)$ définie par :

$$G(A) = A \cup \{e\} \quad \text{si } e \notin A,$$
$$G(A) = A \setminus \{e\} \quad \text{si } e \in A.$$

Montrer que $G \circ G = \text{Id}$. Utiliser G pour construire une bijection entre $\mathcal{P}_{\text{pair}}(E)$ et $\mathcal{P}_{\text{impair}}(E)$.

Exercice 67. (*) Soit E un ensemble à n éléments. Calculer le cardinal de

$$\big\{(X, Y) \in \mathcal{P}(E) \times \mathcal{P}(E);\ X \subset Y\big\}.$$

Indication : on pourra commencer par calculer pour un sous-ensemble Y fixé de E, le cardinal de $\big\{X \in \mathcal{P}(E);\ X \subset Y\big\}$.

Exercice 68. (*) Soit $E \subset \mathbb{C}^*$ un ensemble à n éléments et soit $p \in \mathbb{N}^*$. Combien y a-t-il de nombres complexes tels que $z^p \in E$?

Exercice 69. (**) Dans un dé "polyédrique" (nombre quelconque de faces ayant chacune un nombre quelconque de côtés), il y a deux faces qui ont le même nombre de côtés.

Exercice 70. (**) Le but de cet exercice est de montrer que le nombre d'applications croissantes de $[\![1, m]\!]$ dans $[\![1, n]\!]$ est $\binom{m+n-1}{m}$.

1) Soit $f : [\![1, m]\!] \to [\![1, n]\!]$ une application croissante. Montrez que pour tout $k \in [\![1, m]\!]$, $f(k) + k - 1 \leq m + n - 1$. En déduire que l'application $\widetilde{f} : [\![1, m]\!] \to [\![1, m + n - 1]\!]$ telle que pour tout k, $\widetilde{f}(k) = f(k) + k - 1$ est bien définie et est strictement croissante.
2) Utiliser la question précédente pour construire une bijection entre l'ensemble des fonctions croissantes de $[\![1, m]\!]$ dans $[\![1, n]\!]$ d'une part, et l'ensemble des fonctions strictement croissantes de $[\![1, m]\!]$ dans $[\![1, m + n - 1]\!]$.

3) Conclure.

Exercice 71. (*) Soient $q \geq p \geq 0$ des entiers.

1) Démontrer par récurrence sur q la formule : $\sum\limits_{n=p}^{q} \binom{n}{p} = \binom{q+1}{p+1}$.

2) Expliciter la formule ci-dessus pour $p = 0, 1, 2$. En déduire une expression courte de $\sum_{n=1}^{q} n^2$.

Exercice 72. Etant donnés des nombres réels a, b et c, on pose pour $n \in \mathbb{N}$, $u_n := an^3 + bn^2 + cn$.

1) Déterminer des valeurs de a, b et c telles que pour tout $n \in \mathbb{N}^*$, $u_n - u_{n-1} = n^2$.

2) En déduire une expression courte de $\sum_{n=1}^{q} n^2$.

3) (*) Expliquer la méthode utilisée en termes de "primitives discrètes".

Exercice 73. (*) Calculer $\sum\limits_{n=1}^{q} n^3$. On pourra s'inspirer de l'exercice 71 ou de l'exercice 72.

Exercice 74. (*) Soient $n \in \mathbb{N}^*$ et $p \in \mathbb{N}$. Montrer par récurrence sur n que le nombres de n-uplets d'entiers naturels $(x_1, \ldots, x_n) \in \mathbb{N}^n$ tels que $x_1 + \cdots + x_n = p$ est égal à $\binom{p+n-1}{p}$.

Indication : pour l'hérédité on pourra isoler une variable. Par ailleurs on aura besoin de la formule démontrée à la première question de l'exercice 71.

Question subsidiaire : pouvez-vous expliquer pourquoi le résultat est le même que dans l'exercice 70 ?

Exercice 75. Soient $a_{i,j}$ des nombres réels (ici les indices i et j peuvent prendre des valeurs allant de 1 à n). Démontrer que

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{i} a_{i,j} \right) = \sum_{j=1}^{n} \left(\sum_{i=j}^{n} a_{i,j} \right).$$

Exercice 76. (*) Calculer $\sum\limits_{X \in \mathcal{P}([\![1,n]\!])} \operatorname{card}(X)$.

Exercice 77. (*) Calculer $\sum\limits_{(X,Y) \in \mathcal{P}([\![1,n]\!]) \times \mathcal{P}([\![1,n]\!])} \operatorname{card}(X \cap Y)$.

Indication : on pourra utiliser la relation card $(X) = \sum_{i=1}^{n} \mathbf{1}_X(i)$ pour $X \subset [\![1, n]\!]$.

Exercice 78. Soit $n \geq 1$.

1) En calculant de deux manières les quantités $(1 + 1)^n$ et $(1 - 1)^n$ montrer que

$$\sum_{k \in P_n} \binom{n}{k} = \sum_{k \in I_n} \binom{n}{k} = 2^{n-1},$$

où P_n est l'ensemble des entiers pairs inférieurs ou égaux à n et I_n est celui des entiers impairs inférieurs ou égaux à n (par exemple si n est pair $\sum\limits_{k \in P_n} \binom{n}{k} = \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{n}$).

2) En déduire le résultat de l'exercice 66.

Exercice 79. Calculer de deux manières le terme en x^q dans $(1 + x)^m(1 + x)^n = (1 + x)^{m+n}$, et en déduire l'égalité :

$$\sum_{p=0}^{q} \binom{m}{p} \binom{n}{q-p} = \binom{m+n}{q}.$$

Exercice 80. (**) Soit $f : F \to F$ une application bijective (donc une permutation). On dit que c'est un *dérangement* de F si : $\forall y \in F,\ f(y) \neq y$.

1) Décrire tous les dérangements de $\{1, 2, 3, 4\}$.

2) Nous prendrons $F := \{1, \ldots, n\}$. Nous noterons S_n l'ensemble de toutes les permutations de F et D_n l'ensemble de tous les dérangements de F et $d_n :=$ card (D_n). Pour tout $i \in F$, nous noterons $\mathcal{F}_i$ l'ensemble des bijections $f : F \to F$ telles que $f(i) = i$.

   (a) Montrer que $D_n = S_n \setminus \bigcup_{i=1}^n \mathcal{F}_i$.

   (b) Vérifier en appliquant la formule du crible, donnée au Théoreème 2.2.2, que

$$\text{card} \bigcup_{i=1}^{n} \mathcal{F}_i \quad = \quad \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n-k)!$$

   Indication : noter que $\mathcal{F}_{i_1} \cap \cdots \cap \mathcal{F}_{i_k}$ est formé des permutations telles que $f(i_1) = i_1, \ldots, f(i_k) = i_k$. Il est donc facile de les compter.

   (c) Conclure que $d_n = n! - \displaystyle\sum_{k=1}^{n} (-1)^{k-1} \frac{n!}{k!} = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}\cdot$

On verra en cours d'analyse (une autre année !) que $\lim\limits_{n\to+\infty} \sum_{k=0}^{n} \dfrac{(-1)^k}{k!} = \dfrac{1}{e}\cdot$ Donc $d_n \sim \dfrac{n!}{e}\cdot$

Chapitre 3

Probabilités sur un ensemble fini

Ce chapitre est une introduction au langage des probabilités. Nous allons présenter la modélisation d'expériences aléatoires avec un nombre fini de résultats possibles.

3.1 Espace de probabilité

Définition. *Un espace de probabilité fini est la donnée de trois objets :*

1) *Un ensemble fini Ω, appelé* univers, *qui représente l'ensemble des résultats possibles de l'expérience aléatoire.*
2) *L'ensemble $\mathcal{P}(\Omega)$ des parties de Ω. Si $A \subset \Omega$ (c'est-à-dire si $A \in \mathcal{P}(\Omega)$) on dit que A est un* événement. *Un événement peut être vu comme un ensemble de résultats possibles ayant une propriété en commun.*
3) *Une application $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ qui vérifie les propriétés suivantes :*
   — $\mathbb{P}(\Omega) = 1$,
   — $\forall A, B \in \mathbb{P}(\Omega),\ A \cap B = \emptyset \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
   On dit que $\mathbb{P}$ est une probabilité *sur Ω. Elle associe à tout événement $A \subset \Omega$ la probabilité $\mathbb{P}(A)$ que cet événement se réalise, c'est-à-dire que le résultat de l'expérience aléatoire soit dans A.*

Dans le cas très particulier où pour tout A, $\mathbb{P}(A) = \frac{\operatorname{card}(A)}{\operatorname{card}(\Omega)}$, on dit que $\mathbb{P}$ est la probabilité uniforme *sur Ω. En effet, dans ce cas tous les résultats possibles ont la même probabilité d'apparaître :*

$$\forall \omega \in \Omega,\ \mathbb{P}(\{\omega\}) = \frac{1}{\operatorname{card}(\Omega)}.$$

Exemple. Dans tous ce chapitre nous illustrerons les notions sur une même expérience aléatoire. On lance deux fois de suite un dé à 6 faces et l'on note les deux nombres obtenus, dans l'ordre chronologique. L'ensemble des résultats possibles est donc

$$\Omega = \big\{(1,1), (1,2), \dots, (1,6), (2,1), (2,2), \dots, (6,6)\big\} = [\![1,6]\!] \times [\![1,6]\!] = [\![1,6]\!]^2.$$

L'ensemble des événements est donc $\mathcal{P}\big([\![1,6]\!]^2\big)$. Considérons trois événements particuliers :

— L'événement correspondant à la propriété "le premier lancer vaut 1" :

$$A_1 := \big\{(1,1), (1,2), \dots, (1,6)\big\} = \{1\} \times [\![1,6]\!],$$

— l'événement "la sommes des lancers vaut 3" :

$$A_2 := \{(1,2), (2,1)\},$$

— l'événement "les deux lancers ont donné le même chiffre" :

$$A_3 := \{(1,1),(2,2),\dots,(6,6)\}.$$

On suppose que les dés ne sont pas pipés, ce qui conduit à choisir la probabilité uniforme sur Ω, telle que pour tout événement A, $\mathbb{P}(A) = \mathrm{card}(A)/\mathrm{card}(\Omega) = \mathrm{card}(A)/36$. On ontient alors que

$$\mathbb{P}(A_1) = \frac{6}{36} = \frac{1}{6}, \quad \mathbb{P}(A_2) = \frac{2}{36} = \frac{1}{18}, \quad \mathbb{P}(A_3) = \frac{6}{36} = \frac{1}{6}.$$

Nous présentons maintenant des propriétés utiles des probabilités. Elles découlent facilement de la définition ci-dessus.

Proposition 3.1.1. *Soit $\mathbb{P}$ une probabilité sur un ensemble fini Ω. Soient $A, B \subset \Omega$ des événements quelconques, alors*
- $\mathbb{P}(\emptyset) = 0$,
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$,
- $\mathbb{P}(A^c) = 1 - P(A)$,
- $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})$.

La dernière propriété assure que la connaissance des probabilités de tous les événements *élémentaires* de la forme $\{\omega\}$ suffit à calculer la probabilité de tous les événements. Cette remarque est à la base du résultat suivant :

Proposition 3.1.2. *Soit Ω un ensemble fini et $p : \Omega \to [0,1]$ une application qui vérifie que*

$$\sum_{\omega \in \Omega} p(\omega) = 1.$$

Alors l'application $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ définie pour tout $A \subset \Omega$ par la formule

$$\mathbb{P}(A) := \sum_{\omega \in A} p(\omega)$$

est une probabilité sur Ω.

Remarque. *Le fait qu'une probabilité soit déterminée par les probabilités des événements élémentaires n'est pas vrai pour les expériences aléatoires en général. Si l'on tente de modéliser le choix aléatoire et uniforme d'un nombre réel compris entre 0 et 1, on est amené à poser $\Omega = [0,1]$ qui est bien entendu infini. L'hypothèse d'uniformité suggère que l'on a autant de chance de tirer un nombre inférieur à $\frac{1}{2}$ qu'un nombre plus grand. Ceci revient à $\mathbb{P}([0,\frac{1}{2}]) = \mathbb{P}([\frac{1}{2},1]) = \frac{1}{2}$. Plus généralement une probabilité uniforme devrait satisfaire que pour $0 \le a \le b \le 1$, $\mathbb{P}([a,b]) = b - a$. Mais ceci donne par exemple que $\mathbb{P}(\{\omega\}) = 0$ pour tout $\omega \in [0,1]$. Ainsi la probabilité de tirer au sort un nombre donné est toujours nulle. Il est alors clair que la donnée des valeurs des $\mathbb{P}(\{\omega\})$ (qui valent toutes 0), ne permet pas de retrouver les valeurs des probabilités de tous les événements par sommation.*

3.2 Variables aléatoires

Définition. *Une* variable aléatoire *sur un espace de probabilités fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est une application $X : \Omega \to E$ où E est un sous-ensemble fini de $\mathbb{R}$ (ou de $\mathbb{R}^d$ et on peut alors parler de vecteur aléatoire).*

Une variable aléatoire n'est donc rien d'autre qu'une fonction du résultat de l'expérience aléatoire. L'idée est que si le résultat de l'expérience est $\omega \in \Omega$, considérer $X(\omega)$ revient à ne s'occuper que d'une partie de l'information contenue dans ω.
Exemple. Dans notre exemple, tout résultat $\omega \in \Omega = [\![1,6]\!]^2$ s'écrit $\omega = (i,j)$ avec $1 \le i \le 6$ et $1 \le j \le 6$. On peut considérer la variable aléatoire $X_1 : [\![1,6]\!]^2 \to [\![1,6]\!]$ définie simplement par

$X_1((i,j)) = i$. Cette variable donne la valeur obtenue lors du premier lancer. De même on peut définir une variable aléatoire X_2 par $X_2((i,j)) = j$, qui à un résultat possible associe la valeur du second lancer. On peut considérer des variables plus compliquées comme $X_1 + X_2$:

$$(X_1 + X_2)((i,j)) = i + j,$$

qui prend ses valeurs dans $[\![2, 12]\!]$ et donne la somme des deux chiffres obtenus lors des lancers.

Définition. *L'*espérance *d'une variable aléatoire X sur un espace de probabilité fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est définie comme suit :*

$$\mathbb{E}X := \sum_{\omega \in \Omega} \mathbb{P}\big(\{\omega\}\big)\, X(\omega).$$

L'espérance est donc la moyenne de la valeur de X, pondérée par la probabilité des résultats possibles de l'expérience aléatoire.

Exemples. Dans une situation générale, si $A \subset \Omega$ est un événement on peut considérer la variable aléatoire $\mathbf{1}_A : \Omega \to \{0, 1\}$ qui vaut 1 si et seulement si l'événement se réalise. On vérifie alors que

$$\mathbb{E}\mathbf{1}_A = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\})\mathbf{1}_A(\omega) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \mathbb{P}(A).$$

Dans l'exemple des deux lancers de dé, l'espérance du premier chiffre obtenu est

$$\mathbb{E}X_1 = \sum_{(i,j) \in [\![1,6]\!]^2} \frac{1}{36}\, i = \frac{1}{36} \sum_{i=1}^{6} \left(\sum_{j=1}^{6} i \right) = \frac{1}{6} \sum_{i=1}^{6} i = \frac{1}{6} \frac{7 \times 6}{2} = 3,5.$$

Un calcul similaire donne que $\mathbb{E}X_2 = 3,5$. On peut obtenir par le calcul que $\mathbb{E}(X_1 + X_2) = 7$, mais il est plus simple d'utiliser la linéarité de l'espérance, énoncée maintenant :

Proposition 3.2.1. *Soit $\lambda \in \mathbb{R}$ et X, Y des variables aléatoires, toutes les deux à valeurs dans $\mathbb{R}$ (ou $\mathbb{R}^d$). Alors*

$$\mathbb{E}(X + \lambda Y) = \mathbb{E}X + \lambda \mathbb{E}Y.$$

Définition. *Soit $X : \Omega \to E$ une variable aléatoire. La* loi *de X est probabilité $\mathbb{P}_X$ sur l'espace E définie comme ceci : pour tout $B \subset E$,*

$$\mathbb{P}_X(B) := \mathbb{P}\big(X^{-1}(B)\big) = \mathbb{P}\big(\{\omega \in \Omega;\ X(\omega) \in B\}\big).$$

On utilisera souvent une notation plus concise de cette dernière expression, qui ne mentionne pas l'ensemble Ω ni la variable ω, en écrivant simplement $\mathbb{P}(X \in B)$. Cette notation peut porter à confusion au début, car elle consiste à noter un ensemble par la propriété qui le définit.

Démonstration. — La définition affirme que $\mathbb{P}_X$ est une probabilité sur E. On le montre facilement en vérifiant les deux propriétés requises. D'abord $\mathbb{P}_X(E) = \mathbb{P}(X^{-1}(E)) = \mathbb{P}(\Omega) = 1$. Ensuite si A, B sont des parties de E avec $A \cap B = \emptyset$, il vient $X^{-1}(A) \cap X^{-1}(B) = \emptyset$. Ainsi en utilisant que $\mathbb{P}$ est une probabilité, on obtient que

$$\begin{aligned} \mathbb{P}_X(A \cup B) &= \mathbb{P}\big(X^{-1}(A \cup B)\big) = \mathbb{P}\big(X^{-1}(A) \cup X^{-1}(B)\big) = \mathbb{P}\big(X^{-1}(A)\big) + \mathbb{P}\big(X^{-1}(B)\big) \\ &= \mathbb{P}_X(A) + \mathbb{P}_X(B). \end{aligned}$$

∎

Remarque. *La loi de X décrit les probabilités que X prenne certaines valeurs. Elle est caractérisée par la donnée des événements élémentaires de E, c'est à dire par la donnée pour tout $e \in E$ de*

$$\mathbb{P}_X(\{e\}) = \mathbb{P}\big(\{\omega \in \Omega;\ X(\omega) \in \{e\}\}\big) = \mathbb{P}\big(\{\omega \in \Omega;\ X(\omega) = e\}\big) = \mathbb{P}(X = e).$$

Exemple. La loi de X_1 est caractérisée par les six nombres $\mathbb{P}(X_1 = 1), \dots, \mathbb{P}(X_1 = 6)$. Pour $i_0 \in [\![1, 6]\!]$, on a

$$\mathbb{P}_{X_1}(\{i_0\}) = \mathbb{P}(X_1 = i_0) = \mathbb{P}\big(\{(i, j) \in [\![1, 6]\!]^2;\ i = i_0\}\big) = \mathbb{P}\big(\{(i_0, 1), (i_0, 2), \dots, (i_0, 6)\}\big) = \frac{6}{36} = \frac{1}{6}.$$

On obtient donc que la loi de X_1 est la probabilité uniforme sur $[\![1, 6]\!]$. On dit que X_1 suit la loi uniforme sur $[\![1, 6]\!]$. Un calcul similaire montre que X_2 suit aussi la loi uniforme sur $[\![1, 6]\!]$. On note au passage que des variables aléatoires différentes peuvent avoir la même loi.

La loi de la variable $X_1 + X_2$, qui prend ses valeurs dans $[\![2, 12]\!]$, est plus compliquée. Pour la caractériser nous devons calculer pour pour nombre $k \in [\![2, 12]\!]$ la quantité

$$\mathbb{P}_{X_1+X_2}(\{k\}) = \mathbb{P}\big(\{\omega \in [\![1, 6]\!]^2;\ X_1(\omega) + X_2(\omega) = k\}\big) = \frac{\operatorname{card}\big(\{(i, j) \in [\![1, 6]\!]^2;\ i + j = k\}\big)}{36}.$$

Si $k = 2$ il n'y a qu'une seule manière d'écrire $2 = i + j$ avec i, j entiers entre 1 et 6 (2=1+1). Donc $\mathbb{P}_{X_1+X_2}(\{2\}) = 1/36$. Il y a deux manières d'écrire 3 (3=1+2=2+1), donc $\mathbb{P}_{X_1+X_2}(\{3\}) = 2/36 = 1/18$... on obtient finalement que

$$\mathbb{P}_{X_1+X_2}(\{k\}) = \begin{cases} \frac{k-1}{36} & \text{si} \quad k \in [\![2, 7]\!], \\ \frac{13-k}{36} & \text{si} \quad k \in [\![7, 12]\!]. \end{cases}$$

On note en particulier que la loi de $X_1 + X_2$ n'est pas uniforme (la valeur la plus probable est 7, alors que les valeurs 1 et 12 sont les moins probables).

La loi de X permet de calculer l'espérance de X et de toutes les variables aléatoires qui dépendent seulement de X :

Proposition 3.2.2. *Soit $X : \Omega \to E$ une variable aléatoire. Soit $f : E \to \mathbb{R}^d$ une application. Alors*

$$\mathbb{E}f(X) = \sum_{e \in E} f(e)\mathbb{P}_X(\{e\}) = \sum_{e \in E} f(e)\mathbb{P}(X = e).$$

En particulier $\mathbb{E}X = \sum_{e \in E} e\, \mathbb{P}(X = e)$.

Démonstration. — On applique la définition de l'espérance de la variable aléatoire $f(X)$ et l'on regroupe les termes correspondant à des ω de même image par X :

$$\begin{aligned} \mathbb{E}f(X) &= \sum_{\omega \in \Omega} f(X(\omega))\mathbb{P}(\{\omega\}) = \sum_{e \in E} \Big(\sum_{\omega \in X^{-1}(\{e\})} f(X(\omega))\mathbb{P}(\{\omega\}) \Big) \\ &= \sum_{e \in E} f(e) \Big(\sum_{\omega \in X^{-1}(\{e\})} \mathbb{P}(\{\omega\}) \Big) = \sum_{e \in E} f(e)\mathbb{P}\big(X^{-1}(\{e\})\big). \end{aligned}$$

∎

Exemple. Si l'on utilise que la loi de X_1 est uniforme sur $[\![1, 6]\!]$, on peut calculer son espérance de manière plus rapide que précédemment :

$$\mathbb{E}X_1 = \sum_{i=1}^{6} i\, \mathbb{P}(X_1 = i) = \frac{1}{6} \sum_{i=1}^{6} i = 3,5.$$

3.3 Indépendance

Il d'agit d'une notion fondamentale en théorie des probabilités, la première de ce cours qui soit vraiment nouvelle. Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace de probabilité fini.

Définition. *On dit que deux événements $A, B \subset \Omega$ sont* indépendants *si* $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Si $\mathbb{P}(A) \neq 0$ on peut récrire la condition comme $\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)} = \mathbb{P}(B)$. Le premier terme est appelé *probabilité conditionnelle* de B sachant A, c'est la probabilité "réactualisée" pour que l'événement B se produise si l'on sait que l'événement A s'est produit. En ce sens l'indépendance de A et B signifie que la connaissance de A n'a pas changé la probabilité de B.

Définition. *On dit que deux variables aléatoires $X : \Omega \to E$ et $Y : \Omega \to F$ sont* indépendantes *si une des assertions suivantes (qui sont équivalentes) est vérifiée :*

— *Pour tout $x \in E$ et tout $y \in F$, les événements $X^{-1}(\{x\})$ et $Y^{-1}(\{y\})$ sont indépendants,*

— *Pour tout $x \in E$ et tout $y \in F$,*

$$\mathbb{P}(X = x \text{ et } Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y),$$

— *Pour tous les ensembles $A \subset E$, $\forall B \subset F$,*

$$\mathbb{P}(X \in A \text{ et } Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B),$$

— *Pour toutes les applications $\varphi : E \to \mathbb{R}$ et $\psi : F \to \mathbb{R}$,*

$$\mathbb{E}\left(\varphi(X)\psi(Y)\right) = \mathbb{E}\varphi(X)\,\mathbb{E}\psi(Y).$$

Démonstration. — Démontrons l'équivalence de ces quatre assertions. Les deux premières ont le même sens. La deuxième découle de la troisième en choisissant des singletons. La troisième est un cas particulier de la quatrième pour des fonctions caractéristiques d'ensembles. Donc pour démontrer l'équivalence, il suffit par exemple de supposer la deuxième propriété et d'en déduire la quatrième. Soient $\varphi : E \to \mathbb{R}$ et $\psi : F \to \mathbb{R}$, alors en appliquant la proposition précédente à la variable aléatoire (X, Y) à valeurs dans $E \times F$, on obtient

$$\begin{aligned}
\mathbb{E}\left(\varphi(X)\psi(Y)\right) &= \sum_{(x,y)\in E\times F} \varphi(x)\psi(y)\,\mathbb{P}\big((X,Y) = (x,y)\big) \\
&= \sum_{(x,y)\in E\times F} \varphi(x)\psi(y)\,\mathbb{P}(X = x \text{ et } Y = y) \\
&= \sum_{(x,y)\in E\times F} \varphi(x)\psi(y)\,\mathbb{P}(X = x)\,\mathbb{P}(Y = y) \\
&= \Big(\sum_{x\in E} \varphi(x)\,\mathbb{P}(X = x)\Big)\Big(\sum_{y\in F} \psi(y)\,\mathbb{P}(Y = y)\Big) \\
&= \mathbb{E}\varphi(X)\,\mathbb{E}\psi(Y).
\end{aligned}$$

∎

Remarque. *Il est clair que, bien que toutes ses conditions soient équivalentes, les premières correspondent à des cas particuliers des suivantes. On utilisera donc plutôt les premières pour établir l'indépendance de variables et les dernières pour tirer des conséquences de l'indépendance quand elle est connue.*

Exemple. Montrons que pour les lancers de dé, X_1 et X_2 sont indépendantes : soient i, j quelconques dans $[\![1, 6]\!]$. On a vu que $\mathbb{P}(X_1 = i) = \frac{1}{6}$ et $\mathbb{P}(X_2 = j) = \frac{1}{6}$. Par ailleurs

$$\mathbb{P}(X_1 = i \text{ et } X_2 = j) = \mathbb{P}(\{(i, j)\}) = \frac{1}{36}.$$

On a donc bien que

$$\mathbb{P}(X_1 = i \text{ et } X_2 = j) = \mathbb{P}(X_1 = i)\,\mathbb{P}(X_2 = j) = \frac{1}{36}.$$

Ceci montre l'indépendance de X_1 et X_2.

On peut en tirer parti pour calculer l'espérance du produit des deux chiffres obtenus lors du lancer des deux dés :

$$\mathbb{E}(X_1 X_2) = \mathbb{E}X_1\,\mathbb{E}X_2 = 3,5 \times 3,5 = 12,25.$$

Nous terminons en mentionnant que l'on peut définir l'indépendance pour plusieurs variables aléatoires. On dit que des variables $X_1, \ldots, X_n$ à valeurs dans des ensembles $E_1, \ldots, E_n$ sont mutuellement indépendantes si pour tous les choix de $x_1 \in E_1, \ldots, x_n \in E_n$ on a

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

Cette définition est plus contraignante que de demander que pour tous $i \neq j$, X_i et X_j soient indépendantes (dans ce cas ont dit que les X_i sont indépendantes deux à deux).

3.4 Le jeu de pile ou face

On lance une pièce n fois de suite. L'étude probabiliste de ce simple jeu remplit des livres entiers. Nous en présentons uniquement le tout début. On considère des variables aléatoires $X_1, \ldots, X_n$ (définies sur un espace de probabilité $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$), qui sont mutuellement indépendantes et chacune de loi uniforme sur $\{0, 1\}$. Ainsi $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$. La valeur de $X_i(\omega)$ représente le résultat du i-ème lancer (disons 0 pour pile, 1 pour face).

Pour $k \leq n$, $X_1(\omega) + \cdots + X_k(\omega)$ compte le nombre de 1 (ou de face) obtenus après k lancers.

Proposition 3.4.1. *Soit $n \in \mathbb{N}^*$. Alors pour chaque k entier tel que $0 \leq k \leq n$,*

$$\mathbb{P}\big(X_1 + \cdots + X_n = k\big) = \binom{n}{k} 2^{-n}.$$

On dit que $S_n := X_1 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}\big(n, \frac{1}{2}\big)$.

Démonstration. — Il faut bien comprendre l'événement "$S_n = k$" : exactement k lancers ont donné 1 et les autres 0. Les indices correspondant à ces lancers forment un sous-ensemble de $[\![1, n]\!]$ de cardinal k. On décompose l'événement suivant ce que vaut cet ensemble :

$$\begin{aligned}
\mathbb{P}(S_n = k) &= \mathbb{P}\big(\exists S \in \mathcal{P}_k([\![1, n]\!]); \forall i \in S, X_i = 1 \text{ et } \forall i \in S^c, X_i = 0\big) \\
&= \sum_{S \in \mathcal{P}_k([\![1,n]\!])} \mathbb{P}\big(\forall i \in S, X_i = 1 \text{ et } \forall i \in S^c, X_i = 0\big) \\
&= \sum_{S \in \mathcal{P}_k([\![1,n]\!])} 2^{-n} = \text{card } \mathcal{P}_k([\![1, n]\!]) \times 2^{-n} = \binom{n}{k} 2^{-n}.
\end{aligned}$$

On peut aussi comprendre cette preuve graphiquement en dessinant un arbre (comme en terminale) : de la racine partent deux branches (suivant que $X_1 = 0$ ou $X_1 = 1$), puis chaque branche se divise en deux suivant la valeur de X_2,.... Cependant plusieurs branches donnent le même nombre de "face". ∎

Exercices

Exercice 81. On lance successivement deux dés équilibrés. On note X_1 le résultat du premier lancer et X_2 celui du second. Calculer $\mathbb{P}(X_1 = X_2)$, $\mathbb{P}(X_1 \geq X_2)$, $\mathbb{P}(X_1 < X_2)$ et $\mathbb{P}(X_1 \leq X_2)$.

Exercice 82. Une urne contient a boules blanches et b boules noires. On en extrait deux boules au hasard. Quelle est la probabilité pour que ces deux boules soient de couleurs différentes ?

Exercice 83. (*) Soit $q > 0$. Pour quelle valeur de C l'application $p : \mathcal{P}([\![1, n]\!]) \to \mathbb{R}$ définie par

$$p(A) = Cq^{\text{card}(A)}$$

peut-elle être étendue pour définir une mesure de probabilités sur $\mathcal{P}([\![1, n]\!])$?

Exercice 84. On choisit au hasard, de manière équiprobable, un sous-ensemble de $[\![1, n]\!]$. On notera S l'ensemble aléatoire ainsi choisi.

1) Calculer $\mathbb{P}(1 \in S)$ et plus généralement $\mathbb{P}(i \in S)$ pour un entier fixé $i \in [\![1, n]\!]$.
2) Calculer $\mathbb{E}\,\text{card}(S)$ en utilisant $\text{card}(S) = \sum_{i=1}^{n} 1_{i \in S}$.
3) (*) Déterminer la loi de card(S). Calculer $\mathbb{E}\,\text{card}(S)$ en utilisant la loi de S. Calculer ensuite $\mathbb{E}\,(\text{card}(S)(\text{card}(S) - 1))$

Exercice 85. Soient $1 \leq m \leq n$ des entiers. On choisit au hasard et de manière équiprobable une application allant de $[\![1, m]\!]$ dans $[\![1, n]\!]$.

1) Calculer la probabilité pour que l'image de 1 vaille 1.
2) Calculer la probabilité pour que 1 ne soit pas dans l'image de l'application.
3) Calculer la probabilité pour que l'application soit injective.
4) Quelle est la probabilité pour que dans une groupe de 23 personnes, deux aient la même date d'anniversaire ?

Exercice 86. Un questionnaire présente dix questions, dont la réponse peut être soit oui soit non. Un étudiant insuffisamment préparé décide de choisir les réponses au hasard, uniformément.

1) Quelle est l'espérance de sa note si une bonne réponse donne un point et une mauvaise réponse ne donne aucun point ?
2) Quelle est l'espérance si une bonne réponse donne un point et une mauvaise réponse fait perdre un point ?

Exercice 87. Le loto (simplifié) consiste à tirer au sort un ensemble de 5 nombres parmi $[\![1, 49]\!]$, toutes les combinaisons ayant la même probabilité. Un joueur a parié sur la combinaison $\{1, 2, 3, 4, 5\}$.

1) Quelle est la probabilité pour qu'il ait exactement k bons numéros ? (pour k allant de 0 à 5). On pourra effectuer les calculs avec un tableur comme celui d'OpenOffice.
2) Le joueur mise 2€ et remporte 5€ pour 2 bons numéros, 50€ pour 3 bons numéros, 5000€ pour 4 bons numéros et 1000000€ pour 5 bons numéros. Quelle est l'espérance de son gain (i.e. la différence entre de qu'il a remporté et ce qu'il a misé) ?
3) Ces résultats dépendent-ils de la combinaison choisie par le joueur ? En est-il de même pour le vrai loto où les gains sont partagés entre les gagnants ?

Exercice 88. On lance n fois de suite une pièce et l'on note le résultat comme une suite de 0 et 1 (correspondant à pile ou face). L'ensemble des résultats possibles est $\Omega = \{0, 1\}^n$, que l'on munit de la probabilité uniforme. Pour $\omega = (\omega_1, \ldots, \omega_n) \in \Omega$, on note $X_i(\omega) = \omega_i$, $1 \leq i \leq n$.

1) Montrer que $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = \frac{1}{2}$.

2) Calculer $\mathbb{E}X_1$ puis $\mathbb{E}(X_1 + \cdots + X_n)$.

3) Déterminer la loi de $X_1 + \cdots + X_n$.

4) Montrer que $X_1, \ldots, X_n$ sont mutuellement indépendantes.

Exercice 89. Soient $2 \leq k \leq n$ des entiers. Les événements $\{1 \in S\}$ et $\{2 \in S\}$ sont-ils indépendants, lorsque

1) S est choisi uniformément dans $\mathcal{P}(\llbracket 1, n \rrbracket)$?

2) S est choisi uniformément dans $\mathcal{P}_k(\llbracket 1, n \rrbracket)$?

Exercice 90. (*) Soit $p \in]0, 1[$. Soient $X_1, \ldots X_n$ des variables aléatoires indépendantes et telles que pour tout i, $\mathbb{P}(X_i = 1) = p$ et $\mathbb{P}(X_i = 0) = 1 - p$ (on dit que X_i suit une loi de Bernoulli de paramètre p).

1) En utilisant l'exercice précédent, calculer la fonction génératrice de X_i puis de $X_1 + \cdots + X_n$.

2) En déduire la loi de $X_1 + \cdots + X_n$.

3) Retrouver ce résultat par une méthode directe.

Exercice 91. Pour une variable aléatoire X à valeurs dans un sous-ensemble fini de $\mathbb{R}$, on définit la variance de X par

$$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}X)^2\right).$$

1) Montrer que $\mathrm{Var}(X) = \mathbb{E}\big(X^2\big) - \big(\mathbb{E}X\big)^2$.

2) Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes (à valeurs dans une partie finie de $\mathbb{R}$) ayant la même loi. Montrer que

$$\mathbb{E}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \mathbb{E}X_1 \quad \text{et} \quad \mathrm{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\mathrm{Var}(X_1)}{n}.$$

Exercice 92. On considère un jeu de pile ou face avec n lancers. Comme dans le cours, on note $X_1, \ldots, X_n$ les résultats des lancers (0 ou 1). Pour $1 \leq k \leq n$, calculer la probabilité pour que le premier 1 apparaisse au k-ième lancer, puis la probabilité pour que le deuxième 1 apparaisse lors du k-ième lancer.

Exercice 93. (*) Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace de probabilité fini et $X : \Omega \to \llbracket 0, n \rrbracket$ une variable aléatoire. On définit la *fonction génératrice* de X par la formule suivante : pour $r \in \mathbb{R}$:

$$G_X(r) := \mathbb{E}\big(r^X\big).$$

1) Montrer que $G_X(r)$ est une fonction polynomiale en r, dont les coefficients caractérisent la loi de X.

2) Montrer que si X et Y sont des variables aléatoires indépendantes (sur Ω et à valeurs dans $\mathbb{N}$) alors pour tout $r \geq 0$,

$$G_{X+Y}(r) = G_X(r) G_Y(r).$$

Chapitre 4

Nombres réels

Pour les mathématiciens grecs de l'école de Pythagore (VI^e siècle av. J.C.), les seuls nombres étaient les nombres rationnels. Bien qu'ils aient su que $\sqrt{2}$ n'est pas rationnel et que cette grandeur est la longueur de la diagonale d'un carré de côté 1, l'existence de telles grandeurs "incommensurables" était pour eux un secret bien gardé. Il a fallu attendre la fin du XIX^e siècle pour que les mathématiciens comme Peano, Dedekind et Cantor notamment aboutissent par une démarche rigoureuse à la première construction du corps des nombres réels. Dans la chaine d'inclusions suivante :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

c'est l'inclusion $\mathbb{Q} \subset \mathbb{R}$ qui est la plus mystérieuse et la plus délicate. Nous allons tenter de l'explorer dans ce chapitre.

La construction des nombres réels n'est pas au programme de ce cours, mais compte tenu de leur importance en Mathématique, il nous a paru utile de donner, en fin de ce chapitre, une idée de cette construction par la méthode des coupures de Dedekind. Historiquement, c'est la première construction rigoureuse, elle date de la fin du XIX^e siècle. Elle repose sur des considérations d'ordre et notamment sur la notion de **borne supérieure** que nous allons introduire et étudier en détail dans ce chapitre. En effet, il s'agit d'un outil essentiel à l'analyse mathématique.

Afin de rendre notre étude plus concrète, commençons par évoquer une autre approche possible des nombres réels, par leur écriture décimale. Ce point de vue permet une autre construction des nombres réels, qui bien qu'intuitive et naturelle, est techniquement délicate (voir par exemple l'ouvrage de référence Cours de Mathématiques L1 tout en un). Le paragraphe qui suit n'est pas rigoureux et se veut purement illustratif. Nous donnerons des preuves précises au chapitre suivant, en utilisant la notion de limite d'une suite.

L'écriture décimale des nombres entiers relatifs est connue depuis l'école primaire. Les nombres décimaux (qui sont le quotient d'un entier relatif par une puissance de 10) se notent aussi facilement en base 10 : ils ont un nombre fini de chiffres après la virgule. Il n'en va pas de même pour tous les nombres rationnels : si l'on pose la division de 1 par 3, l'opération ne s'arrête pas et l'on note le résultat avec des pointillés $1/3 = 0,33333\ldots$ (mais cette notation n'a pas été vraiment justifiée, nous le ferons au chapitre suivant). Plus généralement pour tout rationnel $r = p/q$ (avec disons $p, q \in \mathbb{N}^*$) si l'on pose la division de p par q soit l'opération s'arrête et r est un nombre décimal, soit elle ne s'arrête pas mais dans ce cas les décimales qui apparaîssent sont périodiques (il n'y a qu'un nombre fini de restes possibles à chaque étape $\{1, 2, ..., q-1\}$; comme il y a une infinité d'étapes un reste finit par se répéter et la suite du calcul aussi). Ainsi un nombre au développement décimal infini mais pas périodique comme

$$0,1010010001000010000010000001\ldots$$

ne peut pas correspondre à un rationnel. En fait, nous verrons que tout nombre réel admet un développement décimal illimité, mais pas forcément unique ! C'est ainsi que $1 = 0,999999999\cdots$. D'autres exemple célèbres $\sqrt{2} = 1,414213562\cdots$, $e = 2,718281828\cdots$ et $\pi = 3,141592654\cdots$

4.1 Le corps des nombres rationnels

L'ensemble des nombres rationnels, c'est à dire des fractions $\frac{p}{q}$, avec $p, q \in \mathbb{Z}$, est muni de deux opérations, l'addition $+$ et la multiplication, notée $\times$ ou $\cdot$, ainsi que d'un ordre $\leq$. Il possède les propriétés suivantes.

(1) $\mathbb{Z} \subset \mathbb{Q}$. Les opérations $+$ et $\times$, ainsi que la relation d'ordre $\leq$ prolongent celles de $\mathbb{Z}$.

(2) L'addition $+$ sur $\mathbb{Q}$ vérifie les propriétés suivantes :

  (a) $\forall x \in \mathbb{Q}, \forall y \in \mathbb{Q}, \forall z \in \mathbb{Q}, \quad x + (y + z) = (x + y) + z$,
  (*associativité* de l'addition).

  (b) $\forall x \in \mathbb{Q}, \quad x + 0 = 0 + x = x$,
  (0 est l'*élément neutre* pour l'addition).

  (c) Pour tout $x \in \mathbb{Q}$, il existe un nombre réel unique, noté $-x \in \mathbb{Q}$ et appelé l'*opposé* de x, tel que $x + (-x) = (-x) + x = 0$,

  (d) $\forall x \in \mathbb{Q}, \forall y \in \mathbb{Q}, \quad x + y = y + x$,
  (*commutativité* de l'addition).

  Ces quatre propriétés se résument en disant que $(\mathbb{Q}, +)$ est un "groupe abelien" (ou commutatif).

(3) La multiplication (ou produit) notée $\cdot$ vérifie les propriétés suivantes :

  (a) $\forall x \in \mathbb{Q}, \forall y \in \mathbb{Q}, \forall z \in \mathbb{Q}, \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z$,
  ("associativité" de la multiplication),

  (b) $\forall x \in \mathbb{Q}, \quad x \cdot 1 = 1 \cdot x$,
  (1 est l'"élément unité"),

  (c) tout $x \in \mathbb{Q} \setminus \{0\}$ admet un *inverse* unique noté $x^{-1} \in \mathbb{Q}$ tel que $x \cdot x^{-1} = x^{-1} \cdot x = 1$,

  (d) $\forall x \in \mathbb{Q}, \forall y \in \mathbb{Q}, \forall z \in \mathbb{Q}, \quad x \cdot (y + z) = x \cdot y + x \cdot z$,
  ("distributivité" de la multiplication par rapport à l'addition),

  (e) $\forall x \in \mathbb{Q}, \forall y \in \mathbb{Q}, \quad x \cdot y = y \cdot x$,
  ("commutativité" de la multiplication)

  Les propriétés (2) à (3e) se résument en disant que $(\mathbb{Q}, +, \cdot)$ est un "corps commutatif".

  Nous allons maintenant décrire les propriétés de la relation d'ordre $\leq$ sur le corps $\mathbb{Q}$ des nombres rationnels et sa compatibilité avec les opérations algébriques.

(4) La relation d'ordre $\leq$ sur $\mathbb{Q}$ vérifie les propriétés suivantes :

  (a) Pour tout $x \in \mathbb{Q}$ et $y \in \mathbb{Q}$, on a ou bien $x \leq y$ ou bien $y \leq x$,
  ($\leq$ est une *relation d'ordre total*).

  (b) Pour tout $x \in \mathbb{Q}, y \in \mathbb{Q}$ et $z \in \mathbb{Q}$ on a $x \leq y \Longrightarrow x + z \leq y + z$,
  (*compatibilité de l'addition* avec la relation d'ordre).

  (c) Pour tout $x \in \mathbb{Q}, y \in \mathbb{Q}$ et $a \in \mathbb{Q}^+$ on a $x \leq y \Longrightarrow a \cdot x \leq a \cdot y$,
  (*compatibilité de la multiplication* avec la relation d'ordre).

(d) Pour tout $x \in \mathbb{Q}$ et tout $y \in \mathbb{Q}_+^*$ il existe un entier $N > 1$ tel que $N \cdot y > x$ ("Propriété d'Archimède").

Ces propriétés à elles seules ne suffisent pas à caractériser entièrement l'ensemble $\mathbb{Q}$. Nous verrons que le corps des nombres réels satisfait les mêmes propriétés, et aussi une propriété supplémentaire que $\mathbb{Q}$ n'a pas. Mais c'est le plus petit corps qui possède ces propriétés.

Comment construire $\mathbb{Q}$ à partir de $\mathbb{Z}$?

Nous avons déjà signalé que l'on peut construire $\mathbb{Q}$ à partir de $\mathbb{Z}$ en considérant sur $\mathbb{Z} \times \mathbb{Z}^*$ la relation d'équivalence

$$(p, q) \equiv (p', q') \iff pq' = p'q$$

et les classes d'équivalence pour cette relation. Pour construire l'addition, on définit $(p, q)+(p', q') := (pq'+qp', qq')$ et on montre que la classe du résultat ne change pas si l'on remplace (p, q) par n'importe quel autre élément équivalent à lui, et de même pour (p', q'). Ceci permet de définir l'addition des classes. On définit de même la multiplication par $(p, q) \times (p', q') = (pp', qq')$ et on voit qu'elle satisfait les mêmes propriétés d'indépendance de la classe du résultat par rapport aux classes de (p, q) et (p', q'). Ensuite, pour l'ordre, on remarque que tout couple (p, q) est équivalent à un couple (p', q'), avec $q' > 0$ et on dit que, pour deux tels couples avec second membres positifs, $(p, q) \leq (p', q')$ lorsque $pq' \leq qp'$, et cette relation ne dépend pas non plus des représentants choisis. Il reste ensuite à faire la vérification fastidieuse de toutes les propriétés que nous avons énoncées.

4.2 Insuffisance des nombres rationnels

4.2.1 Nombres irrationnels

Il existe des grandeurs "naturelles" qui ne sont pas rationnelles. Ainsi, d'après le théorème de Pythagore, la diagonale d'un carré de côté 1 a une longueur x telle que $x^2 = 1 + 1 = 2$. Mais il n'existe pas de nombre rationnels dont le carré est 2 (comme nous l'avons démontré au chapitre 1) :

$$\sqrt{2} \notin \mathbb{Q}.$$

Ainsi, on ne peut pas "tout mesurer" avec des nombres rationnels. C'est pourquoi nous sommes amenés à considérer un ensemble de nombres plus riches, qui est l'ensemble $\mathbb{R}$ des nombres réels. Les nombres qui ne sont pas rationnels sont dits *irrationnels.*

Nous avons observé que $\sqrt{2} \notin \mathbb{Q}$. Cependant nous allons voir que $\mathbb{Q}$ contient des nombres qui en sont très proches. Dans la pratique, lorsqu'on a affaire à un nombre comme $\sqrt{2}$, nous avons besoin de connaitre des valeurs rationnelles approchées de ces nombres avec une certaine précision. Nous le faisons d'habitude en écriture décimale, c'est à dire qu'on donne une valeur approchée avec un certain nombre de chiffres "après la virgule". Ces approximations sont faites par des nombres *décimaux*, c'est à dire des nombres rationnels particuliers.

Définition. *Un rationnel $r \in \mathbb{Q}$ est dit décimal s'il existe un entier $n \in \mathbb{N}$ tel que $10^n r \in \mathbb{Z}$.*

Nous n'avons pas encore décrit mathématiquement ce qu'est l'écriture décimale d'un nombre réel ou même rationnel, mais le lecteur (la lectrice) comprend bien ce qu'est un nombre décimal : c'est un nombre dont l'écriture décimale s'arrête au bout d'un certain rang. Tous les nombres décimaux sont rationnels, mais il y a des nombres rationnels qui ne sont pas décimaux, comme par exemple 1/3.

Nous allons présenter un procédé assez simple permettant par exemple de construire des nombres rationnels (décimaux) dont le carré est arbitrairement voisin de 2 par défaut et des nombres rationnels (décimaux) dont le carré est arbitrairement voisin de 2 par excès. Ces nombres décimaux seront appelés respectivement des *approximants décimaux par défaut* et des *approximants décimaux par excès* de la grandeur réelle ℓ. C'est en ce sens que l'on peut dire que le nombre irrationnel $\sqrt{2}$ existe ! Nous allons appliquer le "procédé de dichotomie" pour démontrer cette propriété.

Observons d'abord que si $a, b \in \mathbb{Q}$ sont des nombres décimaux, leur moyenne arithmétique $m := (a + b)/2 \in \mathbb{Q}$ est un nombre décimal (pourquoi ?) vérifiant $a < m < b$. Si l'on représente les nombres rationnels par des points situés sur une droite orientée, le nombre rationnel m correspond au milieu du segment qui joint les points représentant les nombres a et b respectivement. Ce qui coupe en deux ce segment, d'où le nom de "procédé de dichotomie" utilisé pour qualifier cette méthode.

Choisissons deux nombres décimaux $a_0, b_0 \in \mathbb{Q}^+$ tels $a_0^2 < 2 < b_0^2$: par exemple $a_0 = 1$ et $b_0 = 2$. Le nombre a_0 (resp. b_0) peut être considéré comme un premier approximant décimal par défaut (resp. par excès) de la grandeur géométrique ℓ solution de l'équation $x^2 = 2$.

Considérons ensuite le nombre décimal $m_0 := \frac{a_0+b_0}{2}$. Alors, puisque l'équation $x^2 = 2$ n'a pas de solution dans $\mathbb{Q}$, il n'y a que deux cas possibles.

$$\begin{cases} \text{Ou bien } m_0^2 < 2, \text{ dans ce cas on pose } a_1 := m_0 \text{ et } b_1 := b_0. \\ \text{Ou bien } m_0^2 > 2, \text{ auquel cas on pose } a_1 := a_0 \text{ et } b_1 := m_0. \end{cases}$$

Dans tous les cas on obtient un nouveau couple (a_1, b_1) de nombres décimaux positifs vérifiant $a_0 \leq a_1 < b_1 \leq b_0$, $a_1^2 < 2 < b_1^2$ et tels que $b_1 - a_1 = \frac{b_0 - a_0}{2}$.

On obtient ainsi un nouvel approximant décimal par défaut a_1 de la grandeur ℓ tel que a_1^2 soit une valeur approchée par défaut de 2 et un nouvel approximant rationnel par excès b_1 de la grandeur ℓ tel que b_1^2 soit une valeur approchée par excès de 2 vérifiant $b_1 - a_1 = \frac{b_0 - a_0}{2}$, ce qui implique que l'une au moins des deux valeurs décimales approchées ainsi obtenues est plus précise que chacune des deux valeurs approchées précédentes.

En itérant ce procédé de dichotomie n fois, on construit successivement des approximants décimaux par défaut $a_0 \leq a_1 \leq \ldots \leq a_n$ et des approximants décimaux par excès $b_0 \geq b_1 \geq \ldots \geq b_n$ de la même grandeur ℓ tels qu'au rang n on ait $a_n^2 < 2 < b_n^2$ et l'écart entre les deux approximants est donné par $e_n := b_n - a_n = \dfrac{b_0 - a_0}{2^n}$.

Mettons en pratique cette méthode en considérant comme premières valeurs approchées décimales par défaut et par excès de $\sqrt{2}$, les deux nombres réels suivants : $a_0 := 1$ et $b_0 := 2$. On a vu qu'au bout de n itérations, l'erreur par défaut ou par excès est dominée par $(b_0 - a_0) \cdot 2^{-n} = 2^{-n}$. Si l'on souhaite déterminer une valeur approchée décimale de $\sqrt{2}$ à 10^{-3} près par exemple, il nous suffit de faire cette opération 10 fois : en effet, $2^{10} = 1024 > 10^3$, et $a_{10} < \sqrt{2} < b_{10}$, avec une différence entre les deux qui sera au plus de 10^{-3}. On obtient successivement pour $(a_n; b_n)$:

$(1; 2)$, $(1; 1,5)$, $(1,25; 1,5)$, $(1,375; 1,5)$, $(1,375; 1,4375)$, $(1,40625; 1,4375)$, $(1,40625; 1,421875)$, $(1,4140625; 1,421875)$, $(1,4140625; 1,417968750)$, $(1,4140625; 1,416015625)$, $(1,4140625; 1,415039062)$.

Il nous faudrait encore une itération pour avoir la garantie que $\sqrt{2}$ est inférieur à $1,415$.

Nous verrons au chapitre suivant les outils qui permettent de formaliser la notion de convergence d'une suite vers une limite. Nous y présenterons aussi un procédé plus rapide pour calculer les décimales de $\sqrt{2}$.

4.2.2 Borne supérieure

Certaines des définitions de cette partie seront données pour $(X, \leq)$ un ensemble ordonné. En fait nous les utiliserons uniquement pour $X = \mathbb{Z}$, $X = \mathbb{Q}$ ou $X = \mathbb{R}$. Cette notation nous évitera des répétitions.

Commençons par rappeler quelques définitions :

Définition (Majorant, minorant)**.** *Une partie non vide $A \subset X$ est majorée s'il existe $M \in X$ tel que $\forall x \in A,\ x \leq M$. On dit alors que M est un majorant de A.*

Une partie non vide $A \subset X$ est minorée s'il existe $m \in X$ tel que $\forall x \in A,\ m \leq x$. On dit alors que m est un minorant de A.

Définition (Maximum, minimum)**.** *On dit qu'une partie $A \subset X$ admet un plus grand élément s'il existe $a \in A$ tel que $\forall x \in A, x \leq a$. Cet élément est forcément unique et on le note $a = \max(A)$.*

Une partie $A \subset X$ admet un plus petit élément s'il existe $a \in A$ tel que $\forall x \in A, a \leq x$. Cet élément est forcément unique et on le note $a = \min(A)$.

Donc un plus grand élément est un majorant de A qui appartient à A. **Toute partie non vide et majorée $A \subset \mathbb{Z}$ admet un plus grand élément. Ce n'est plus vrai dans $\mathbb{Q}$!**

Exemple. $\mathbb{Q}_-^* := \{x \in \mathbb{Q};\ x < 0\}$ est clairement une partie non-vide de $\mathbb{Q}$ majorée par 0. Mais aucun de ces éléments n'est le plus grand. En effet si $x < 0$, on a aussi $x < x/2 < 0$ (multiplier par $x < 0$ l'inégalité $1/2 < 1$). Donc étant donné $x \in \mathbb{Q}_-^*$ on peut toujours trouver un élément plus grand dans $\mathbb{Q}_-^*$.

Dans cet exemple, on voit tout de même que 0 joue un rôle particulier : il n'est pas dans l'ensemble mais il en est en quelque sorte la frontière. La définition suivante en rend compte :

Définition (Borne supérieure, borne inférieure)**.** *Soit $A \subset X$ une partie non vide et majorée. Soit $\mathcal{M}$ l'ensemble des majorants de A. Si l'ensemble $\mathcal{M}$ admet un plus petit élément alors on dit que A admet une borne supérieure, notée* $\sup(A)$ *qui vaut par définition* $\min(\mathcal{M})$*. Entre d'autres termes* $\sup(A)$ *est* **le plus petit majorant de** A*.*

Si $A \subset X$ est une partie non-vide et minorée, et s'il existe un plus grand minorant de A, on l'appelle borne inférieure de A et on le note $\inf(A)$*.*

Proposition 4.2.1. *Si $A \subset X$ a un plus grand élément alors il admet aussi une borne supérieure et* $\sup(A) = \max(A)$*.*

Démonstration. — Il suffit de noter que si M est un majorant de A, il est plus grand que $\max(A) \in A$. Par ailleurs $\max(A)$ est par définition un majorant de A. C'est donc le plus petit majorant. ∎

Cependant, des parties qui n'ont pas de plus grand élément peuvent admettre une borne supérieure. Dans ce cas, la borne supérieure de l'ensemble n'est pas dans l'ensemble !

Exemple. Nous avons vu que $\mathbb{Q}_-^*$ n'a pas de plus grand élément. Nous allons maintenant établir que cet ensemble admet une borne supérieure : $\sup(\mathbb{Q}_-^*) = 0$. Premièrement il est évident que 0 est un majorant (car $x \in \mathbb{Q}_-^*$ signifie $x < 0$, donc implique $x \leq 0$). Ensuite il faut montrer que tout majorant de $\mathbb{Q}_-^*$ est positif ou nul. En d'autres termes, nous devons prouver que

$$(\forall x \in \mathbb{Q}_-^*, x \leq M) \Longrightarrow 0 \leq M.$$

Il est plus commode d'établir l'assertion contraposée :

$$M < 0 \Longrightarrow (\exists x \in \mathbb{Q}_-^*, M < x).$$

Cette assertion est facile à établir : si $M < 0$ alors nous avons déjà vu que $M < M/2 < 0$. Donc il existe bien un élément $x \in \mathbb{Q}_-^*$ tel que $M < x$: on peut prendre $x = M/2$.

Une autre insuffisance de l'ensemble des rationnels $\mathbb{Q}$ est que certains de ses sous-ensembles qui intuitivement devraient avoir une borne supérieure, n'en ont pas :

Exemple. L'ensemble $A := \{x \in \mathbb{Q};\ \ x > 1 \text{ et } x^2 < 2\}$ n'a pas de borne supérieure dans $\mathbb{Q}$.

Intuitivement, la raison en est claire : on peut aussi écrire que $A = \{x \in \mathbb{Q};\ 1 < x < \sqrt{2}\}$. Ainsi c'est $\sqrt{2}$ qui semble délimiter A à droite. Mais ce nombre n'est pas dans $\mathbb{Q}$! Très informellement, on peut dire que l'ensemble des rationnels possède des "trous" qu'il faut "boucher" pour obtenir les réels.

Nous donnons maintenant une preuve détaillée du fait que $A \subset \mathbb{Q}$ n'a pas de borne supérieure : *Démonstration.* — Pour cela, commençons par montrer que pour tout $r \in A$, il existe $r' \in A$ tel que $r' > r$. En effet, si $r \in A$, alors $r > 1$ et $r^2 < 2$. Choisissons $n \in \mathbb{N}$, $n \geq 1$, tel que $n(2 - r^2) > 2r + 1$, ce qui est possible puisque $\mathbb{Q}$ est archimédien. Alors $r' = r + \frac{1}{n}$ est bien un élément de $\mathbb{Q}$ et vérifie $r' > r$ et $r' \in A$. La dernière propriété vient de ce que

$$r'^2 = r^2 + \frac{2r}{n} + \frac{1}{n^2} \leq r^2 + \frac{2r}{n} + \frac{1}{n} = r^2 + \frac{2r+1}{n} < 2.$$

Donc, aucun élément de A ne peut être un majorant de A.

Par ailleurs, A est majoré (par 2 par exemple), il est non vide (il contient 1). Soit alors $M \in \mathbb{Q}$ un majorant de A. Alors $M^2 > 2$ En effet, d'après ce qu'on vient de voir, si $M^2 \leq 2$, alors $M^2 < 2$, puisqu'il n'existe pas de rationnel tel que $M^2 = 2$, et alors $M \in A$ et M ne peut pas être un majorant de A. En fait, tout rationnel $M > 0$ tel que $M^2 > 2$ est un majorant de A.

Or, si $M^2 > 2$, alors $N = \frac{1}{2}(M + \frac{2}{M})$ est un rationnel qui satisfait à la fois $N^2 > 2$ et $N < M$. La première inégalité vient de ce que

$$N^2 - 2 = \frac{1}{4}(M^2 + \frac{4}{M} - 4) = \frac{1}{4}(M - \frac{2}{M})^2,$$

tandis que la seconde vient de ce que

$$N < M \iff M + \frac{2}{M} < 2M \iff \frac{2}{M} < M \iff 2 < M^2.$$

Donc, pour tout majorant M de A dans $\mathbb{Q}$, il existe un majorant N de A dans $\mathbb{Q}$ tel que $N < M$. L'ensemble A n'admet donc pas de borne supérieure (dans $\mathbb{Q}$). ∎

4.3 Le corps des nombres réels

Comme d'habitude, nous allons procéder de manière axiomatique, c'est à dire que nous allons décrire l'ensemble des nombres réels par ses propriétés. Qu'un tel ensemble existe bien peut-être établi rigoureusement par une "construction". Il y a plusieurs façons de procéder, nous en décrivons une à la fin de ce chapitre. L'important, c'est que les propriétés que nous décrivons permettent de définir uniquement l'ensemble des nombres réels (si deux ensembles ont ces propriétés, ils sont en bijection et les opérations d'addition, de multiplication et la relation d'ordre se correspondent dans cette bijection, si bien que les deux ensembles sont "indiscernables").

Nous admettrons dans ce cours qu'il existe un ensemble $\mathbb{R}$ contenant $\mathbb{Q}$, muni de deux opérations internes (l'addition notée $+$ et la multiplication notée $\cdot$ ou bien $\times$) et d'une relation d'ordre total notée $\leq$ qui étendent les opérations internes et la relation d'ordre correspondantes sur $\mathbb{Q}$, de telle sorte que $(\mathbb{R}, +, \cdot, \leq)$ soit un "corps commutatif archimédien complet". Ces propriétés sont les mêmes que celles de $\mathbb{Q}$, à l'exception d'une seule, **la propriété de la borne supérieure**. Voici la liste détaillée de ces propriétés :

(1) $\mathbb{Q} \subset \mathbb{R}$. Les opérations $+$ et $\times$, ainsi que la relation d'ordre sur $\mathbb{R}$, prolongent celles de $\mathbb{Q}$.

(2) L'addition $+$ sur $\mathbb{R}$ vérifie les propriétés suivantes :

   (a) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \forall z \in \mathbb{R}, \quad x + (y + z) = (x + y) + z$,
   (*associativité* de l'addition).

   (b) $\forall x \in \mathbb{R}, \quad x + 0 = 0 + x = x$,
   (0 est l'*élément neutre* pour l'addition).

   (c) Pour tout $x \in \mathbb{R}$, il existe un nombre réel unique, noté $-x \in \mathbb{R}$ et appelé l'*opposé* de x, tel que $x + (-x) = (-x) + x = 0$,

   (d) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad x + y = y + x$,
   (*commutativité* de l'addition).

   Ces quatre propriétés se résument en disant que $(\mathbb{R}, +)$ est un "groupe abelien" (ou commutatif).

(3) La multiplication (ou produit) notée $\cdot$ vérifie les propriétés suivantes :

   (a) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \forall z \in \mathbb{R}, \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z$,
   ("associativité" de la multiplication),

   (b) $\forall x \in \mathbb{R}, \quad x \cdot 1 = 1 \cdot x$,
   (1 est l'"élément unité"),

   (c) tout $x \in \mathbb{R} \setminus \{0\}$ admet un *inverse* unique noté $x^{-1} \in \mathbb{R}$ tel que $x \cdot x^{-1} = x^{-1} \cdot x = 1$,

   (d) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \forall z \in \mathbb{R}, \quad x \cdot (y + z) = x \cdot y + x \cdot z$,
   ("distributivité" de la multiplication par rapport à l'addition),

   (e) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad x \cdot y = y \cdot x$,
   ("commutativité" de la multiplication)

Les propriétés (2) à (3e) se résument en disant que $(\mathbb{R}, +, \cdot)$ est un "corps commutatif".

Nous allons maintenant décrire les propriétés de la relation d'ordre $\leq$ sur le corps $\mathbb{R}$ des nombres réels et sa compatibilité avec les opérations algébriques.

(4) La relation d'ordre $\leq$ sur $\mathbb{R}$ vérifie les propriétés suivantes :
   (a) Pour tout $x \in \mathbb{R}$ et $y \in \mathbb{R}$, on a ou bien $x \leq y$ ou bien $y \leq x$, ($\leq$ est une *relation d'ordre total*).
   (b) Pour tout $x \in \mathbb{R}, y \in \mathbb{R}$ et $z \in \mathbb{R}$ on a $x \leq y \Longrightarrow x + z \leq y + z$, (*compatibilité de l'addition* avec la relation d'ordre).
   (c) Pour tout $x \in \mathbb{R}, y \in \mathbb{R}$ et $a \in \mathbb{R}^+$ on a $x \leq y \Longrightarrow a \cdot x \leq a \cdot y$, (*compatibilité de la multiplication* avec la relation d'ordre).
   (d) Pour tout $x \in \mathbb{R}$ et tout $y \in \mathbb{R}_+^*$ il existe un entier $N > 1$ tel que $N \cdot y > x$ ("Propriété d'Archimède").

(5) (*Propriété de la borne supérieure*).
   Toute partie $A \subset \mathbb{R}$ non vide et majorée de $\mathbb{R}$ possède une borne supérieure.
   Les propriétés de (2) à (5) se résument en disant que $(\mathbb{R}, +, \cdot, \leq)$ est un "corps commutatif archimédien complet".

Remarque. *Attention la propriété* (4c) *n'est valable que si $a \geq 0$. Dans le cas où $a < 0$, on obtient l'inégalité renversée i.e.*

$$x \leq y \text{ et } a < 0 \Longrightarrow a \cdot x \geq a \cdot y.$$

En effet si $x \leq y$ on obtient d'après (4b), *$x+(-y) \leq y+(-y)$ et donc $x-y \leq 0$. En appliquant de nouveau la propriété* (4b), *on obtient $(-x)+x-y \leq -x$ et donc par associativité, on en déduit que $-y \leq -x$. En multipliant chaque membre de cette inégalité par $-a > 0$, on obtient l'inégalité $(-a)\cdot(-y) \leq (-a)\cdot(-x)$ i.e. $a \cdot x \geq a \cdot y$.*

Exercice 94. (*) Montrer que si $A \subset \mathbb{R}$ est une partie non-vide et minorée, alors l'ensemble $-A := \{x \in \mathbb{R};\ -x \in A\}$ est majoré. En déduire que A admet une borne inférieure donnée par $\inf(A) = -\sup(-A)$.

On retiendra de cet exercice que **toute partie non-vide et minorée de $\mathbb{R}$ admet une borne inférieure**.

4.3.1 La propriété de la borne supérieure en pratique

Nous revenons un peu sur cette propriété, qui est la différence majeure entre $\mathbb{R}$ et $\mathbb{Q}$. La propriété qui suit est souvent très utile pour vérifier qu'un réel est bien la borne supérieure d'un ensemble A donné.

Proposition 4.3.1 (Caractérisation de la borne supérieure)**.** *Soit $A \subset \mathbb{R}$ une partie non vide et majorée de $\mathbb{R}$, alors le nombre réel* $\sup A$ *est caractérisé par les conditions suivantes : $S = \sup A$ si et seulement si*

(i) Pour tout $x \in A, x \leq S$,

(ii) pour tout $\varepsilon > 0$, il existe $a \in A$ tel que $S - \varepsilon < a$.

Démonstration. — La propriété (i) traduit le fait que S est un majorant de A et la propriété (ii) traduit le fait que tout nombre réel $S'(= S - \varepsilon) < S$ n'est pas un majorant de A autrement dit : S est un majorant de A et tous les majorants de A sont $\geq S$, donc S est le plus petit des majorants de A.
∎

Corollaire 4.3.2. *Soit $b \in \mathbb{R}$. Alors* $\sup\big(]-\infty, b[\big) = b$.

Démonstration. — On vérifie les deux points de la caractérisation. Premièrement, il est évident que b est un majorant de $B :=]-\infty, b[= \{x \in \mathbb{R};\ x < b\}$. Deuxièmement : pour tout $\varepsilon > 0$, $b - \varepsilon < b - \varepsilon/2 < b$. Ainsi $a := b - \varepsilon/2$ vérifie $a \in B$ et $b - \varepsilon < a$. ∎

Il faut bien retenir que $\sup A$ **n'appartient pas forcément à** A comme le montre l'exemple précédent. Lorsque $\sup A \in A$, nous avons vu que c'est le aussi plus grand élément de A.

Exercice 95. Que vaut $\inf\big(]0, 1]\big)$?.

Nous donnons ci-dessous des résultats utiles pour les estimations (majoration, minoration). Nous commençons par un résultats simple :

Proposition 4.3.3. *Soit $A \subset \mathbb{R}$ un ensemble qui a un plus grand élément et $M \in \mathbb{R}$. Alors*

$$\max(A) \leq M \iff \forall x \in A,\ x \leq M,$$

$$\max(A) < M \iff \forall x \in A,\ x < M.$$

On a des résultats analogues si A admet un plus petit élément pour $\min(A)$.

Démonstration. — Pour les implication $\Longrightarrow$ il suffit d'utiliser que pour tout $x \in A$, $x \leq \max(A)$. Pour le sens $\Longleftarrow$, on applique la propriété pour le choix particulier $x = \max(A) \in A$. ∎

Les bornes supérieures se manipulent presque de la même manière :

Proposition 4.3.4. *Soit $A \subset \mathbb{R}$ un ensemble non-vide et majoré et $M \in \mathbb{R}$. Alors*

$$\sup(A) \leq M \iff \forall x \in A,\ x \leq M,$$

$$\sup(A) < M \Longrightarrow \forall x \in A,\ x < M,$$

$$\sup(A) \leq M \Longleftarrow \forall x \in A,\ x < M.$$

On a des résultats analogues pour $\inf(A)$ *si A est non-vide et minoré.*

Attention : la différence entre la proposition 4.3.4 et la proposition 4.3.3 est que les deux derniers points ne sont pas des équivalences

Démonstration. — Pour les implications $\Longrightarrow$ il suffit d'utiliser que pour tout $x \in A$, $x \leq \sup(A)$, qui vient du fait que $\sup(A)$ est un majorant de A.

Réciproquement, si $\forall x \in A$, $x \leq M$ alors M est un majorant de A et donc $M \geq \sup(A)$ qui est le plus petit majorant de A. La dernière implication s'en déduit immédiatement puisque $x < M \Longrightarrow x \leq M$. ∎

Notons que l'on ne peut pas espérer de meilleur résultat si $\forall x \leq A,\ x < M$. En effet pour $M = 1$ et $A =]0, 1[$ cette hypothèse est vérifiée et $\sup(]0, 1[) = 1$. On retiendra que pour majorer un sup (au sens large), il suffit de majorer tous les éléments.

En fait la bonne condition pour avoir $\sup(A) < M$ est la suivante :

$$\sup(A) < M \iff \exists \varepsilon > 0, \forall x \in A, x \leq M - \varepsilon.$$

Elle se déduit de la première équivalence de la propriété et de l'observation simple valable pour tous $s, M \in \mathbb{R}$: $s < M \iff \exists \varepsilon > 0, s \leq M - \varepsilon$. En prenant la contraposée des deux implications de cette équivalence et de la première équivalence de la proposition 4.3.4, on obtient le résultat suivant qui explique comment minorer une borne supérieure :

Proposition 4.3.5. *Soit $A \subset \mathbb{R}$ un ensemble non-vide et majoré et $M \in \mathbb{R}$. Alors*

$$\sup(A) > M \iff \exists x \in A,\ x > M,$$

$$\sup(A) \geq M \iff \forall \varepsilon > 0, \exists x \in A,\ x \geq M - \varepsilon.$$

On a des résultats analogues pour $\inf(A)$ *si A est non-vide et minoré.*

4.3.2 Les intervalles

Il y a dans $\mathbb{R}$ plusieurs types d'intervalles :

— Les intervalles fermés bornés de la forme $[a, b] := \{x \in \mathbb{R};\ a \leq x \leq b\}$. Noter que $[a, a] = \{a\}$ et que $[a, b] = \emptyset$ si $a > b$.

— Les intervalles ouverts bornés $]a, b[:= \{x \in \mathbb{R};\ a < x < b\}$. On peut aussi utiliser la notation anglo-saxonne (a, b). Cet ensemble est vide si $a \geq b$.

— Les intervalles semi-ouverts bornés $]a, b] := \{x \in \mathbb{R};\ a < x \leq b\}$ et $[a, b[:= \{x \in \mathbb{R};\ a \leq x < b\}$ (non-vides si et seulement si $a < b$).

— Les demi-droites fermées $[a, +\infty[:= \{x \in \mathbb{R};\ x \geq a\}$ et $] - \infty, a] := \{x \in \mathbb{R};\ x \leq a\}$.

— Les demi-droites ouvertes $]a, +\infty[:= \{x \in \mathbb{R};\ x > a\}$ et $] - \infty, a[:= \{x \in \mathbb{R};\ x < a\}$.

— Il ne faut pas oublier $\mathbb{R} =] - \infty, +\infty[$.

Mais nous pouvons tous les décrire par une propriété commune, qui peut être fort utile en pratique

Proposition 4.3.6. *Une partie $I \subset \mathbb{R}$ est un intervalle si et seulement si, $\forall x, y \in I,\ [x, y] \subset I$.*

Démonstration. — Bien sûr, tous les intervalles que nous avons décrits vérifient cette propriété. Il n'est pas évident que si I vérifie cette propriété, alors c'est l'un des intervalles. Pour le voir, on commence par se restreindre aux cas où I est non vide, puis on sépare les cas où I est majoré ou non, minoré ou non. Si I est à la fois majoré et minoré, on regarde si sa borne supérieure est ou non un max, et de même pour sa borne inférieure.

Montrons par exemple que si I est borné et ne contient pas sa borne inférieure a ni sa borne supérieure b, alors $I =]a, b[$. Montrons donc que $]a, b[\subset I$ (l'inclusion inverse est automatique par le fait que a est un minorant et que b est un majorant, et que ni a ni b ne sont dans I). Soit donc x tel que $a < x < b$, et montrons que $x \in I$. Puisque a est la borne inférieure de I, et que $a \notin I$, alors il existe $x_1 \in I$ tel que $a < x_1 < x$. Maintenant, pour la même raison, il existe $x_2 \in I$ tel que $x < x_2 < b$. Puisque $[x_1, x_2] \subset I$, et que $x \in [x_1, x_2]$, alors $x \in I$.

On procède de même dans les 8 autres cas (nous en laissons la rédaction en exercice). ∎

Notons que cette propriété permet de voir immédiatement que l'intersection de deux intervalles (ou d'une famille quelconque d'intervalles) est un intervalle (éventuellement vide). L'union de deux intervalles n'est en général pas un intervalle, mais elle l'est si l'intersection des deux est non vide (pourquoi ?).

4.3.3 La droite réelle et la fonction valeur absolue

L'ensemble ordonné $\mathbb{R}$ peut être représenté géométriquement par une droite orientée munie d'une origine O symbolisant le nombre réel 0. Chaque nombre réel x est alors représenté par un point unique M de la droite de telle sorte que si $x > 0$ (resp. $x < 0$) le segment OM soit orienté positivement (resp. négativement) et sa longueur soit égale à la valeur absolue de x notée $|x|$.

Rappelons que pour tout $x \in \mathbb{R}$, on définit la valeur absolue de x comme étant le plus grand des deux nombres réels x et $-x$:

$$|x| := \max\{x, -x\} = \begin{cases} x & \text{si} \quad x \geq 0, \\ -x & \text{si} \quad x \leq 0. \end{cases}$$

Ainsi $|x|$ est toujours un nombre réel positif. Voici les propriétés essentielles de la valeur absolue :

(i) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad |x \cdot y| = |x| \cdot |y|$,

(ii) $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad |x + y| \leq |x| + |y|$, (*inégalité triangulaire*)

(iii) $|x| = 0 \Longleftrightarrow x = 0$,

(iv) $|x| \leq t \Longleftrightarrow -t \leq x \leq t$.

Remarque. *La propriété* (iii) *de la valeur absolue admet une version plus raffinée, dont nous servirons souvent.*

$$\textit{Si, } \forall \varepsilon > 0, \ |x| \leq \varepsilon, \textit{ alors } x = 0.$$

En effet, si tel n'était pas le cas, alors il existerait un entier $n \in \mathbb{N}$ tel que $n|x| > 1$ (car $\mathbb{R}$ est archimédien) et alors le nombre $\varepsilon = 1/n$ contredirait l'hypothèse.

Ces propriétés permettent d'exprimer la notion de voisinage et de proximité dans $\mathbb{R}$. En effet soit $a \in \mathbb{R}$ un nombre réel fixé et $\varepsilon > 0$ un nombre réel positif donné. Alors, un nombre réel variable x vérifie $|x - a| \leq \varepsilon$ si et seulement si x vérifie la double inégalité $a - \varepsilon \leq x \leq a + \varepsilon$. Cela veut dire que x est à une distance de a au plus égale à ε ou encore que x est voisin de a à ε près. L'ensemble ainsi obtenu :

$$\bar{I}(a, \varepsilon) := \{x \in \mathbb{R}; a - \varepsilon \leq x \leq a + \varepsilon\} = [a - \varepsilon, a + \varepsilon],$$

est appelé l'intervalle fermé de centre a et de rayon ε. On peut aussi considérer l'intervalle ouvert de centre de centre a et de rayon ε défini comme suit :

$$I(a, \varepsilon) := \{x \in \mathbb{R}; a - \varepsilon < x < a + \varepsilon\} =]a - \varepsilon, a + \varepsilon[.$$

D'une manière générale, si $a \in \mathbb{R}$ et $b \in \mathbb{R}$ sont tels que $a < b$, l'intervalle ouvert $]a, b[$ coïncide avec l'intervalle ouvert $I(c, \varepsilon)$, de centre $c := (a+b)/2$ et de rayon $\varepsilon := (b-a)/2$ (à vérifier !).

4.3.4 L'inclusion $\mathbb{Z} \subset \mathbb{R}$ et la fonction partie entière

Tout nombre réel est compris entre deux nombres entiers relatifs. C'est une conséquence immédiate de de la propriété d'Archimède de $\mathbb{R}$. En effet,

Proposition 4.3.7. *Soit $x \in \mathbb{R}$. Alors il existe un entier $n \in \mathbb{Z}$ unique vérifiant la propriété suivante :*

$$(E) \qquad n \leq x < n + 1.$$

Le nombre entier n vérifiant la propriété (E) sera noté $n = \lfloor x \rfloor$ et appelé la *partie entière* de x. Par exemple $\lfloor 2 \rfloor = 2 = \lfloor 2,5 \rfloor$, $\lfloor -2 \rfloor = -2 = \lfloor -1,5 \rfloor$ et $\lfloor -2,5 \rfloor = -3$.

Démonstration. — Observons que si $x \in \mathbb{Z}$ est un nombre entier, alors $n := x$ vérifie la propriété requise et donc $\lfloor n \rfloor = n$.

Supposons maintenant que $x \notin \mathbb{Z}$. D'après la propriété d'Archimède, il existe un entier $N \in \mathbb{N}$ tel que $N > x$.

— Si $x \geq 0$, alors l'ensemble des entiers naturels $p \in \mathbb{N}$ tels que $p \leq x$ est non vide (il contient 0) et majoré par N, il admet donc un plus grand élément $n \in \mathbb{N}$. Cet entier vérifie clairement la propriété requise.

— Si $x < 0$, alors $-x > 0$ est un nombre réel qui n'est pas un entier et donc l'entier $N := \lfloor -x \rfloor \in \mathbb{N}$ vérifie la propriété $N < -x < N+1$. Il en résulte que l'entier $n := -N-1$ vérifie les inégalités $n < x < n + 1$, ce qui prouve que $\lfloor -x \rfloor = n = -\lfloor -x \rfloor - 1$.

∎

4.3.5 L'inclusion $\mathbb{Q} \subset \mathbb{R}$ et la densité des rationnels

Bien que les nombres rationnels ne remplissent pas tout $\mathbb{R}$, il y en a partout :

Proposition 4.3.8. *Soient a et b deux réels avec $a < b$. Dans l'intervalle $]a, b[$, il existe au moins un rationnel r.*

On exprime cette propriété en disant que l'ensemble $\mathbb{Q}$ est dense dans $\mathbb{R}$.

Démonstration. — On cherche deux nombres entiers $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$ tels que $a < p/q < b$ i.e. $qa < p < qb$. Pour ce faire, grâce à la propriété d'Archimède, on commence par choisir un entier $q > 1$ tel que $q(b-a) > 1$. Alors l'entier $p := \lfloor qa \rfloor + 1$ vérifie $qa < \lfloor qa \rfloor + 1 \leq qa + 1 < qb$ et donc $qa < p < qb$ de sorte que le nombre rationnel p/q convient. ∎

Remarque. *S'il y a au moins un rationnel r dans l'intervalle $]a, b[$ avec $a < b$, alors il y en a une infinité, puisqu'il y en a dans $]a, r[$ et un dans $]r, b[$, et en itérant le processus, on en trouve autant que l'on veut.*

L'ensemble des nombres irrationnels $\mathbb{R} \setminus \mathbb{Q}$ possède la même propriété.

Corollaire 4.3.9. *Pour tous nombres réels a, b tels que $a < b$, il existe un nombre irrationnel $c \in \mathbb{R} \setminus \mathbb{Q}$ (et donc une infinité) tel que $a < c < b$.*

Démonstration. — En effet, comme $a < b$, on $a/\sqrt{2} < b/\sqrt{2}$ et d'après le théorème précédent, il existe un nombre rationnel $r \in \mathbb{Q}$ tel que $a/\sqrt{2} < r < b/\sqrt{2}$, autrement dit $a < r\sqrt{2} < b$. Comme $\sqrt{2} \notin \mathbb{Q}$, on en déduit que $r\sqrt{2} \notin \mathbb{Q}$, ce qui prouve la propriété voulue. ∎

Remarque. *Nous avons déjà signalé que l'ensemble $\mathbb{Q}$ est dénombrable, c'est à dire en bijection avec $\mathbb{N}$. On peut par contre démontrer que $\mathbb{R}$ ne l'est pas, non plus qu'aucun intervalle $]a, b[$, avec $a < b$. Il en résulte que $\mathbb{R} \setminus \mathbb{Q}$ n'est pas dénombrable, ce qui en quelque sorte signifie qu'il y a beaucoup plus de nombres irrationnels que de nombres rationnels. Ainsi, il y a dans chaque intervalle $]a, b[$ non vide une infinité de nombres irrationnels (il y en a même "beaucoup plus" que de nombres rationnels).*

4.4 Compléments

4.4.1 Exemples d'autres corps de nombres

Il y a bien d'autres corps archimédiens K plus riches que l'ensemble $\mathbb{Q}$ et moins riches que $\mathbb{R}$. Nous donnons ici deux exemples :

L'ensemble $K = \{r + \sqrt{2}r' \; ; \; r, r' \in \mathbb{Q}\}$. Pour vérifier qu'un tel ensemble est bien un corps totalement ordonné archimédien, il n'y a pas besoin de redémontrer toutes les propriétés de l'addition, de la multiplication, et de l'ordre. En effet, ces propriétés étant vraies sur $\mathbb{R}$, elles le restent sur K. Les seules choses à démontrer sont que les opérations d'addition, de multiplication, d'opposé et d'inverse laissent stable K. Plus précisément, si $x, y \in K$, alors $x + y \in K$ et $xy \in K$, et de plus, si $x \in K$, alors $-x \in K$ et si $x \in K, x \neq 0$ alors $1/x \in K$. On pourra à titre d'exercice vérifier ces propriétés pour l'exemple considéré. Un tel corps n'a jamais la propriété de la borne supérieure. Ici, par exemple, on pourra considérer l'ensemble des $x \in K$ tels que $1 < x < \sqrt{3}$ et montrer que sa borne supérieure (dans $\mathbb{R}$), qui est $\sqrt{3}$, n'est pas dans K. Il n'a donc pas de borne supérieure dans K.

L'ensemble de nombres qui sont solution d'une équation algébrique (c'est-à-dire polynomiale) à coefficient entiers $a_n x^n + \cdots + a_1 x + a_0 = 0$. Cet ensemble s'appelle l'ensemble des nombres *algébriques*. Le nombre $\sqrt{2}$ en fait partie (c'est une solution de $x^2 - 2 = 0$), ainsi que toutes les diagonales des rectangles à côtés entiers. L'ensemble des nombres algébriques forme lui aussi un corps (ce n'est pas du tout évident !). On pourrait donc se contenter de considérer ces nombres. Mais cela ne suffirait pas pour calculer toutes les grandeurs "naturelles". Ainsi, le nombre π, qui mesure la surface d'un cercle de rayon 1, n'est pas algébrique. On dit que π est *transcendant*. C'est un résultat très difficile qui n'a été démontré qu'en 1882 par Ferdinand von Lindeman, mettant ainsi un point final au problème de *la quadrature du cercle*. Auparavant, Charles Hermite avait démontré en 1873 que le nombre e, base des logarithmes népériens, est lui aussi transcendant.

4.4.2 Une construction des nombres réels

Nous allons présenter ici la construction de Dedekind basée sur la notion de coupures. Elle fournit une définition simple et assez intuitive de l'ensemble $\mathbb{R}$ des nombres réels tout en donnnant un accès assez direct à certaines de ses propriétés fondamentales (densité de $\mathbb{Q}$ dans $\mathbb{R}$ et théorème de la borne supérieure).

L'idée de Dedekind part de l'observation simple selon laquelle tout nombre rationnel $s \in \mathbb{Q}$ découpe l'ensemble des rationnels en deux parties : la partie $C(s) := \{r \in \mathbb{Q};\ r < s\}$ et son complémentaire $C(s)^c = \{r \in \mathbb{Q};\ r \geq s\}$. L'ensemble $C(s)$ n'a pas de plus grand élément et a la propriété que si $x \in C(s)$ alors tout nombre inférieur à x est encore dans $C(s)$. Ceci conduit à poser la définition suivante :

Définition. *On appelle* coupure *toute une partie $A \subset \mathbb{Q}$ vérifiant les propriétés suivantes :*

1. $A \neq \emptyset,\ \ A \neq \mathbb{Q}$,
2. $\forall x \in A,\ \forall y \in \mathbb{Q},\ y \leq x \Longrightarrow y \in A$,
3. *A n'a pas de plus grand élément.*

Ainsi à chaque nombre rationnel s on peut associer la coupure $C(s) = \{r \in \mathbb{Q};\ r < s\}$. Mais il existe d'autres coupures comme $D := \mathbb{Q}_- \cup \{r \in \mathbb{Q}_+; r^2 < 2\}$. Si l'on disposait déjà des nombres réels et de $\sqrt{2}$ on pourrait écrire $D = \{r \in \mathbb{Q};\ r < \sqrt{2}\}$. L'idée est justement d'utiliser D comme définition de $\sqrt{2}$, et de construire le corps des nombres réels à partir de l'ensemble $\mathcal{C}$ de toutes les coupures !

Il est clair que l'ensemble $\mathcal{C}$ ainsi construit "contient" les nombres rationnels (puisque pour tout $r \in \mathbb{Q}$ on dispose de la coupure $C(r)$). Il est aussi facile de définir une relation d'ordre sur l'ensemble des coupures (on pose $A \leq B$ si $A \subset B$, ce qui est cohérent avec le fait que si deux rationnels vérifient $r_1 \leq r_2$ alors $C(r_1) \subset C(r_2)$). Il faut travailler plus pour construire l'addition, la multiplication et pour vérifier que l'ensemble obtenu a toutes les propriétés requises et notamment la propriété de la borne supérieure. Nous admettrons le résultat suivant :

Théorème 4.4.1 (Théorème de Dedekind)**.** *L'ensemble $\mathcal{C}$ des coupures de $\mathbb{Q}$ au sens de Dedekind peut être muni d'une addition $+$ et d'une multiplication $\cdot$ compatible avec sa relation d'ordre $\leq$ de telle sorte que $(\mathcal{C}, +, \cdot, \leq)$ soit un corps commutatif archimédien complet.*

Ce nouvel ensemble $\mathcal{C}$ muni de sa structure de corps commutatif, archimédien et complet sera noté $\mathbb{R}$ et appelé le *corps des nombres réels.*

Exercices

Exercice 96. Démontrer que la somme d'un nombre rationnel et d'un nombre irrationnel est un nombre irrationnel.

Exercice 97. (*) Soient x et y deux nombres rationnels strictement positifs. On suppose de plus que $\sqrt{x}$ est irrationnel. Démontrer que $\sqrt{x}+\sqrt{y}$ est irrationnel.

Exercice 98. Les sous-ensembles de $\mathbb{R}$ décrits ci-dessous sont-ils majorés, minorés ? Ont-il un plus grand élément, un plus petit élément, une borne supérieure, une borne inférieure (en cas de réponse positive, donner leur valeur) ? Les ensembles suivants sont-ils des intervalles ?

1) $[0,3]\cup]2,4[$,
2) $]-\infty,3]\cap]2,\infty[$,
3) $[0,2[\cup]3,4]$,
4) $\mathbb{R}\setminus[0,3[$,
5) $\mathbb{Q}\cap[0,1]$,
6) $\mathbb{Q}\cap[\sqrt{2},2]$,
7) $\left\{2-\frac{1}{n}\,;n\in\mathbb{N}^*\right\}$,
8) $\{x\in\mathbb{R};\ |x-1|<2\}$,
9) $\{x\in\mathbb{R};\ |x-1|\leq|x+1|\}$,
10) $\{x\in\mathbb{R};\ \cos(x)\leq\frac{1}{2}\}$.

Exercice 99. (*) Donner une expression simple des ensembles suivants :

1) $\bigcup_{n\in\mathbb{N}}[0,n]$,
2) $\bigcup_{n\in\mathbb{N}}]n,n+1]$,
3) $\bigcap_{n\in\mathbb{N}}[n,+\infty[$,
4) $\bigcup_{n\in\mathbb{N}^*}\left[0,1-\frac{1}{n}\right]$,
5) $\bigcup_{n\in\mathbb{N}^*}\left[0,1-\frac{1}{n}\right[$,
6) $\bigcap_{n\in\mathbb{N}^*}\left[0,\frac{1}{n}\right[$,
7) $\bigcap_{n\in\mathbb{N}^*}\left]0,\frac{1}{n}\right[$,
8) $\bigcap_{n\in\mathbb{N}^*}\left]-\frac{1}{n},1\right[$,
9) $\bigcap_{n\in\mathbb{N}^*}\left[-\frac{1}{n},1\right]$.

Exercice 100.

1) Démontrer que $A:=\left\{x\in\mathbb{R}^*\,|\,x+\dfrac{1}{x}<2\right\}$ est un intervalle majoré et en déduire $\sup A$.
2) Montrer que l'ensemble $B:=\left\{x\in\mathbb{R}^*\,|\,x+\dfrac{1}{x}\leq 2\right\}$ est majoré mais n'est pas un intervalle. Calculer $\sup B$.

Exercice 101. (**) On étudie $a:=\inf_{n\in\mathbb{N}}\cos(\ln n)$.

1) Montrer que a est bien défini et que $a\in[-1,1]$.
2) A l'aide d'une calculatrice, tentez de deviner la valeur de a.
3) Montrer que $a=\inf_{n,k\in\mathbb{N}}\cos(\ln(n)-2k\pi)$.
4) Montrer que l'ensemble $A:=\{\ln(n)-2k\pi\mid(n,k)\in\mathbb{N}^2\}$ est dense dans $\mathbb{R}$.
   Indication : on se donnera $a<b$ quelconques et on cherchera un élément de A dans $]a,b[$ de la manière suivante : trouver un entier N tel que pour $n\geq N$ on a $\log(n+1)-\log(n)<b-a$. Ensuite trouver $K\in N$ tel que $\log(N)-2K\pi\leq a$. Ensuite faire augmenter n et constater que nécessairement $\log(n)-2K\pi\in]a,b[$ pour un $n\geq N$ (on pourra faire un dessin).
5) En déduire la valeur de a.

Exercice 102. Soient $A\subset B\subset\mathbb{R}$ des ensembles non vides. On suppose que B est majoré.

1) Montrer que A et majoré et qu'il admet une borne supérieure.
2) En déduire que $\sup(A)\leq\sup(B)$.

Exercice 103. Soient A et B des sous-ensembles de $\mathbb{R}$, non vides et majorés. Montrer que

$$\sup(A \cup B) = \max\big(\sup(A), \sup(B)\big).$$

Que peut-on dire pour $\sup(A \cap B)$?

Exercice 104. (*) Soit $x \in \mathbb{Q}$ tel que $x > 0$ et $x^2 > 2$. On pose : $x' := \dfrac{1}{2}\left(x + \dfrac{2}{x}\right)$.

1) Démontrer que $x' \in \mathbb{Q}$ et vérifie $0 < x' < x$ et $x'^2 > 2$.
2) On pose $A := \{x \in \mathbb{Q}^+ ; x^2 > 2\}$.
   a) Démontrer que A est une partie non vide et minorée de $\mathbb{Q}$ qui n'a pas de plus petit élément dans $\mathbb{Q}$ (on pourra utiliser le résultat de la question précédente).
   b) Montrer que A n'a pas de borne inférieure dans $\mathbb{Q}$. Quelle est sa borne inférieure dans $\mathbb{R}$?

Exercice 105. (*) Soit A un sous-ensemble de $\mathbb{Q}$ qui est non-vide et majoré. Démontrer que s'il a une borne supérieure au sens de $\mathbb{Q}$ alors elle coincide avec la borne supérieure de A vu comme un sous-ensemble de $\mathbb{R}$. Indication : pour établir que $\sup_{\mathbb{Q}}(A) \leq \sup_{\mathbb{R}}(A)$ on pourra raisonner par l'absurde et utiliser la densité de $\mathbb{Q}$ dans $\mathbb{R}$.

En déduire une autre preuve du fait que l'ensemble $\{x \in \mathbb{Q}^+;\ x^2 < 2\}$ n'a pas de borne supérieure dans $\mathbb{Q}$.

Exercice 106. Soient $x, y \in \mathbb{R}$.

1) Démontrer que : $\big||x| - |y|\big| \leq |x - y|$.
2) On suppose que $x \leq y$. Démontrer l'équivalence suivante :

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

Exercice 107. Soient x et a deux nombres réels tels que $a \neq 0$ et $|x - a| < |a|$. Démontrer que $a - |a| < x < a + |a|$ et en déduire que x est du signe de a.

Exercice 108.

1) Soit $x \in \mathbb{R}$, démontrer qu'il existe un entier unique N tel que $x < N \leq x + 1$; Exprimer N en fonction de la partie entière de x, notée $\lfloor x \rfloor$.
2) Calculer $\lfloor x + n \rfloor$ lorsque $x \in \mathbb{R}$ et $n \in \mathbb{Z}$.

Exercice 109.

1) Démontrer que pour tout $n \in \mathbb{N}^*$,

$$\sqrt{n+1} - \sqrt{n} < \frac{1}{2\sqrt{n}} < \sqrt{n} - \sqrt{n-1}.$$

   Indication : on pourra utiliser les quantités "conjuguées".
2) Calculer la partie entière du nombre réel

$$a := \frac{1}{2}\sum_{n=1}^{10000} \frac{1}{\sqrt{n}} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{10000}}\right).$$

Exercice 110. Soient x, y deux nombres réels tels que $x < y$ et soient $z \in \mathbb{R}$ tel que $0 < z < y - x$. Démontrer qu'il existe un entier $m \in \mathbb{Z}$ tel que $x < mz < y$ (utiliser la propriété d'Archimède).

Exercice 111. (**) On définit $\alpha := \sqrt{2} - 1$ et l'ensemble suivant :

$$A := \left\{ x \in \mathbb{R}; \exists\, p \in \mathbb{Z}, \exists\, q \in \mathbb{Z}, x = p + q\sqrt{2} \right\}.$$

1) Démontrer que pour tout $n \in \mathbb{Z}$ et $x \in A$, on a $nx \in A$.
2) Démontrer par récurrence que pour tout $n \in \mathbb{N}$, on a $\alpha^n \in A$.
3) Démontrer que $0 < \alpha < 1/2$. En déduire que pour tout $n \in \mathbb{N}^*$ on a $0 < \alpha^n < 1/n$.
4) Soient $x, y \in \mathbb{R}$ tels que $x < y$. Démontrer qu'il existe un entier $n \in \mathbb{N}$ tel que $\alpha^n < y - x$ (utiliser la propriété d'Archimède).
5) En utilisant l'exercice précédent, démontrer qu'il existe $a \in A$ tel que $x < a < y$. On dit alors que l'ensemble A est dense dans $\mathbb{R}$. (Comparer avec la proposition 4.3.8).

Chapitre 5

Convergence des suites

Ce chapitre propose une étude rigoureuse des suites de nombres réels. Son objectif principal est d'acquérir une bonne maîtrise des techniques de base de l'analyse réelle : calculer, majorer, minorer, approcher. On introduira la notion de "convergence" qui est un concept fondamental en analyse et on insistera sur son utilisation pour approcher par exemple certains nombres irrationnels par des nombres rationnels ou décimaux.

La notion de convergence n'est pas facile à appréhender. Prenons un exemple pour commencer : chacun sait que le nombre $1/3$, en écriture décimale, est $0,3333\ldots$. Qu'est-ce que cela signifie précisément ?

Si on écrit n décimales, le nombre $0,3333\ldots3$ est par définition $33\ldots3/10^n$ (nous y reviendrons). Maintenant, lorsqu'on augmente le nombre de décimales dans l'écriture, par exemple si l'on passe de 100 à 1000 décimales, le nombre qu'on écrit se "rapproche" de plus en plus de $1/3$. En d'autre termes, lorsque le nombre de décimales augmente infiniment (on dira "tend vers l'infini"), le nombre écrit se rapproche "infiniment" de $1/3$. C'est assez facile à comprendre sur cet exemple : ce nombre x_n des écritures avec n décimales vaut en fait

$$x_n = \frac{1}{10^n}\big(3 + 10\times 3 + \ldots + 10^{n-1}\times 3\big) = \frac{3}{10^n}\big(1 + 10 + \ldots + 10^{n-1}\big).$$

Maintenant, nous appliquons la formule fondamentale de sommation des termes d'une suite géométrique de raison $q \neq 1$:

$$1 + q + q^2 + \cdots + q^N = \frac{q^{N+1}-1}{q-1}. \tag{5.1}$$

Nous obtenons donc

$$x_n = \frac{3}{10^n}\frac{10^n-1}{9} = \frac{1}{3}(1-\frac{1}{10^n}) = \frac{1}{3} - \frac{1}{3\times 10^n}.$$

Le nombre $\frac{1}{3\times 10^n}$ est de plus en plus petit lorsque n est de plus en plus grand, et on voit bien que le résultat se rapproche de plus en plus de $1/3$. On dira que $1/3$ est la limite de la suite x_n. Notre but dans ce chapitre est de définir correctement ce qu'on entend par là .

Exercice 112. En utilisant le même raisonnement que pour $1/3$, justifiez l'écriture de $1/7$

$$\frac{1}{7} = 0,142857142857142857\ldots.$$

L'écriture décimale est ici le nombre 142857 répété une infinité de fois.

5.1 Notion de convergence et limite d'une suite

On rappelle qu'une *suite* de nombres réels est une application $u : \mathbb{N} \longrightarrow \mathbb{R}$ qui à chaque entier $n \in \mathbb{N}$ associe un nombre réel $u(n)$ encore noté u_n. On parle alors de la suite $(u_n)_{n\geq 0}$ de terme général u_n, appelé aussi le *terme de rang n* de la suite $(u_n)_{n\geq 0}$. Il arrive que l'application u soit définie sur une partie infinie $I \subset \mathbb{N}$, on parle dans ce cas de la suite $(u_n)_{n\in I}$ indexée par I.

5.1.1 Suites et sous-suites

Une suite n'est pas nécessairement une suite de nombres. Elle peut prendre ses valeurs dans n'importe quoi. Mais c'est pour les suites à valeurs dans $\mathbb{R}$, ou dans $\mathbb{C}$, que nous parlerons de convergence.

Définition (Suite)**.** *Soit E un ensemble. Une suite à valeurs dans E est une application $x = \mathbb{N} \mapsto E$. On note en général x_n pour $x(n)$. La suite elle même sera notée $(x_n)_{n\geq 0}$, ou plus simplement (x_n).*

Lorsque $E = \mathbb{R}$, on parle de suite réelle, et lorsque $E = \mathbb{C}$, on parle de suite complexe. Mais on peut bien sûr envisager des suites à valeurs dans bien d'autres ensembles, comme des suites de fonctions.

Dans ce qui suit, nous nous intéresserons essentiellement aux suites réelles. Nous en avons déjà vu plusieurs exemples, comme $x_n = n!$ ou bien $x_n = 1/n$. Dans ce dernier cas, la suite n'est définie que pour $n \geq 1$.

De façon générale, il se peut que la suite ne soit définie que pour $n \geq n_0$. On peut alors considérer la suite $y_n = x_{n+n_0}$, qui est elle définie pour tout $n \geq 0$. Pour les problèmes de convergence qui vont nous occuper dans la suite, cela ne fait aucune différence. Dans ce cas, (y_n) est ce que l'on appelle *une suite extraite* de (x_n). Comme cette notion sera très utilisée dans ce cours, nous en donnons une définition précise.

Définition (Suite extraite, ou sous-suite)**.** *Soit (x_n) une suite, et soit $p : \mathbb{N} \mapsto \mathbb{N}$ une suite strictement croissante (c'est à dire, pour tout n, $p(n+1) > p(n)$). Alors la suite $y_n = x_{p(n)}$ est appelée suite extraite de x_n. On la note souvent x_{p_n}. On dit aussi que x_{p_n} est une sous-suite de la suite x_n.*

Les valeurs prises par une suite extraite sont donc choisies parmi celles de la suite x_n, dans l'ordre croissant. Par exemple, la suite 2^{-n} est extraite de la suite $1/n$, mais la suite $n2^{-n}$ ne l'est pas.

Nous exigeons que la suite p soit strictement croissante pour éviter par exemple que la suite $(x_1, x_1, x_1, \ldots, x_1, \ldots)$ soit considérée comme une suite extraite de x_n.

Remarque. *Dans toute la suite, nous nous intéresserons essentiellement aux suites réelles. C'est en ce sens que nous utiliserons le mot suite dans ce texte.*

5.1.2 Convergence d'une suite réelle

Définition (Convergence)**.** *On dit que la suite réelle $(u_n)_{n\geq 0}$ converge vers $\ell \in \mathbb{R}$ si pour tout $\varepsilon > 0$, il existe un rang $N \in \mathbb{N}$ tel que pour tout $n \geq N$, on ait $|u_n - \ell| \leq \varepsilon$. On notera dans ce cas $x_n \to_{n\to\infty} \ell$, ou encore $x_n \to \ell$. En d'autres termes :*

$$x_n \to_{n\to\infty} \ell \quad \Longleftrightarrow \quad \forall\varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |u_n - \ell| \leq \varepsilon.$$

Nous dirons que la suite (x_n) converge lorsqu'il existe un réel ℓ pour lequel $x_n \to \ell$. Si la suite est définie seulement pour $n \geq n_0$, on adapte la définition en remplaçant $N \in \mathbb{N}$ par $N \geq n_0$ (et $N \in \mathbb{N}$ que l'on sous-entend souvent pour ne pas alourdir les notations).

Ce qu'il faut comprendre dans cette définition, c'est que quel que soit ε *aussi petit que l'on veut*, à partir d'un certain rang N (qui peut être très grand), on peut affirmer que la distance de x_n à ℓ est inférieure à ε.
Exemple. La suite $x_n = 1/n$, définie pour $n \geq 1$ converge vers 0. En effet, étant donné $\varepsilon > 0$, il existe N tel que $N\varepsilon \geq 1$ (ceci n'est rien d'autre que le fait que $\mathbb{R}$ soit archimédien). Alors, pour tout $n \geq N$, on a $n\varepsilon \geq 1$, soit $x_n \leq \varepsilon$. Mais $x_n = |x_n| = |x_n - 0|$ car $x_n \geq 0$, et donc $\lim_n x_n = 0$.

La définition de la notion de limite est assez compliquée. Dans la pratique, nous ne l'utiliserons que dans des cas particuliers. Ensuite, on se donnera des règles qui permettent de ramener l'étude de la convergence de suites compliquées à des suites plus simples, pour lesquelles la convergence ne posera pas de problème. Il y a plusieurs remarques importantes à faire à propos de cette définition. Commençons par vérifier l'unicité de la limite.

Proposition 5.1.1. *Soit* $(u_n)_{n\geq 0}$ *une suite de nombres réels qui converge vers* ℓ. *Alors le nombre réel* ℓ *est unique. Ceci justifie la notation* $\lim_{n\to\infty} u_n$ *pour* **la** *limite de la suite* (u_n).

Démonstration. — En effet, supposons que la suite vérifie à la fois $u_n \to \ell_1$ et $u_n \to \ell_2$. Nous voulons montrer que $\ell_1 = \ell_2$.

Soit $\varepsilon > 0$ un nombre réel arbitraire. En appliquant la définition de la convergence à chacun avec $\varepsilon/2$ à chacune des limites, on aboutit à l'existence d'un entier $N_1 \geq 1$ et d'un entier $N_2 \geq 1$ tels que pour tout $n \geq N_1$ on aie on ait $|u_n - \ell_1| \leq \varepsilon/2$ et pour tout $n \geq N_2$, on aie $|u_n - \ell_2| \leq \varepsilon/2$.

Posons $N := \max\{N_1, N_2\}$ et écrivons $\ell_1 - \ell_2 = (\ell_1 - u_N) + (u_N - \ell_2)$. En appliquant l'inégalité triangulaire, on obtient $|\ell_1 - \ell_2| \leq |\ell_1 - u_N| + |u_N - \ell_2|$. Il en résulte grâce au choix de N que $|\ell_1 - \ell_2| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$.

Nous pouvons alors appliquer la remarque 4.3.3, page 60, pour en conclure que $|\ell_1 - \ell_2| = 0$, d'où $\ell_1 = \ell_2$. ∎

L'énoncé suivant se déduit immédiatement de la définition.

Proposition 5.1.2. *Soit* (x_n) *une suite réelle et* $\ell \in \mathbb{R}$. *Alors*

$$x_n \to \ell \iff x_n - \ell \to 0 \iff |x_n - \ell| \to 0.$$

En pratique on utilisera souvent cet autre énoncé

Proposition 5.1.3. *Soient* (x_n), (y_n) *deux suites réelles* $k \in \mathbb{N}$ *et* $M \in \mathbb{R}^+$. *Si* $\forall n \geq k$, $|x_n| \leq M|y_n|$ *et* $\lim_n y_n = 0$ *alors* $\lim_n x_n = 0$.

Démonstration. — Remarquons pour $\varepsilon > 0$ donné, il existe un entier N (que l'on peut choisir plus grand que k) tel que pour $n \geq N$, $|y_n| \leq \varepsilon/M$, et donc, toujours pour $n \geq N$, $|x_n| \leq \varepsilon$. ∎

Remarque. *Pour appliquer la définition de la convergence d'une suite, encore faut-il connaître par avance le nombre* ℓ. *Nous verrons bientôt des critères permettant d'affirmer que la suite* (x_n) *converge sans connaître* ℓ.

Dans les applications, lorsqu'une suite $(u_n)_{n\geq 0}$ *converge, il arrive souvent qu'elle donne naissance à un nouveau nombre réel* ℓ *que l'on ne connait pas à priori. Lorsque* $\varepsilon > 0$ *est donné, l'entier* N *à partir duquel on a l'inégalité* $|u_n - \ell| \leq \varepsilon$, *est alors important d'un point de vue "qualitatif", puisque* $u_N - \varepsilon \leq \ell \leq u_N + \varepsilon$. *On dira que* u_N *est une valeur approchée (ou un approximant) de* ℓ *à* $\varepsilon-$*près. Le nombre réel* ε *doit être assez petit et représente l'erreur maximale commise dans l'approximation de* ℓ *par* u_N. *Il est dans ce cas important de trouver le plus petit entier* N *(qui dépend de* ε*) vérifiant cette propriété. Il représente le nombre minimum d'opérations permettant de calculer* ℓ *avec une erreur au plus égale à* ε.

Cependant d'un point de vue "qualitatif", pour démontrer qu'une suite converge, il n'est pas nécessaire de trouver le plus petit entier N *satisfaisant aux exigences de la définition, ce qui peut être assez compliqué. Il suffit d'en trouver un.*

Remarque. *Lorsqu'une suite converge, tous ses termes ont tendance à "s'accumuler" arbitrairement près (i.e. à ε près, ε étant arbitraire) autour d'un même nombre réel à savoir sa limite, à l'exception d'au plus un nombre fini d'entre eux N.*
Il en résulte que la nature d'une suite (convergence ou non) ainsi que sa limite, lorsqu'elle existe, ne change pas si l'on modifie ou supprime un nombre fini de termes de cette suite. Par exemple, si $p \geq 1$ est un entier fixé, la suite $(u_{n+p})_{n\geq 0}$, dite tronquée au rang p, est de même nature que la suite $(u_n)_{n\geq 0}$ et a la même limite lorsque celle-ci existe (à vérifier !).

Suites complexes

Il peut arriver que nous ayons affaire à une suite à valeurs dans le corps $\mathbb{C}$ des nombres complexes. La définition de la convergence est alors la même, à ceci près que nous devons remplacer la fonction valeur absolue $|x|$ par le module $|z|$ du nombre complexe, c'est à dire, pour $z = x + iy$, $|z| = \sqrt{x^2 + y^2}$.

Définition. *Une suite (z_n) à valeurs dans $\mathbb{C}$ converge vers une limite $\ell \in \mathbb{C}$ si*

$$\forall \varepsilon > 0,\ \exists N,\ \forall n \geq N,\ |z_n - \ell| \leq \varepsilon.$$

Si $z_n = x_n + iy_n$ et $\ell = \ell_1 + i\ell_2$, (où x_n, y_n, ℓ_1, ℓ_2 sont réels), alors z_n converge vers ℓ si et seulement si x_n converge vers ℓ_1 et y_n converge vers ℓ_2. Pour le voir, il suffit de remarquer dans un sens que $|x_n - \ell_1| \leq |z_n - \ell|$, $|y_n - \ell_2| \leq |z_n - \ell|$, et dans l'autre que $|z_n - \ell| \leq 2(|x_n - \ell_1| + |y_n - \ell_2|)$. On applique ensuite la remarque 5.1.3 de la page 68, puis le fait que la somme de deux suites qui convergent vers 0 converge vers 0 (proposition 5.2.3, page 73).

Exemples.

(i) Une suite $(u_n)_{n\geq 0}$ est dite *constante* s'il existe $c \in \mathbb{R}$ tel que $u_n = c$ pour tout $n \geq 0$. La suite $(u_n)_{n\geq 0}$ est dite *stationnaire* s'il existe un rang $p \geq 0$ et un nombre réel $c \in \mathbb{R}$ tels que $u_n = c$ pour tout $n \geq p$. Dans ce cas la suite converge vers c (à démontrer en utilisant la définition).

(ii) Si $p \geq 1$ est un entier fixé, la suite $(1/n^p)_{n\geq 0}$ converge vers 0.
En effet, en posant $N = \lfloor 1/\varepsilon^{1/p} \rfloor + 1$, on obtient pour tout $n \geq N$, $1/n^p \leq 1/N^p \leq \varepsilon$.

(iii) Posons $u_n := \frac{2n}{n+\sqrt{n}}$ pour $n \geq 1$ et montrons que $\lim_{n\to+\infty} u_n = 2$. Pour $n \geq 1$, on a :

$$|u_n - 2| = \left| \frac{2n - 2(n + \sqrt{n})}{n + \sqrt{n}} \right| = \frac{2\sqrt{n}}{n + \sqrt{n}} \leq \frac{2}{\sqrt{n}}.$$

Soit $\varepsilon > 0$, pour avoir l'inégalité $|u_n - 2| \leq \varepsilon$, il suffit d'avoir l'inégalité $\frac{2}{\sqrt{n}} \leq \varepsilon$, i.e. $n > 4/\varepsilon^2$. Pour cela il suffit de poser $N := \lfloor 4/\varepsilon^2 \rfloor + 1$. Alors pour $n \geq N$, on a $\frac{2}{\sqrt{n}} \leq \varepsilon$ et donc pour $n \geq N$, on a $|u_n - 2| \leq \varepsilon$. Ce qui prouve notre assertion. Observer que l'entier N trouvé ici n'est pas le plus petit possible, mais cela suffit à prouver la convergence !

Remarque. *Dans la définition de la limite, il est important que la propriété soit vraie pour tout $\varepsilon > 0$. Si on avait demandé qu'elle soit vraie pour $\varepsilon \geq 0$, c'est à dire*

$$\forall \varepsilon \geq 0,\ \exists N, \forall\ n \geq N,\ |x_n - \ell| \leq \varepsilon,$$

alors on pourrait l'appliquer avec $\varepsilon = 0$, et cela signifierait que pour n assez grand, $x_n = \ell$, ce qui n'est pas la même chose. En revanche, d'autres définitions peuvent être équivalentes, comme les lecteurs pourront le vérifier à titre d'exercice. Par exemple

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

ou
$$\forall \varepsilon > 0,\ \exists N,\ \forall n > N,\ |x_n - \ell| < \varepsilon,$$
ou bien encore
$$\forall \varepsilon \in \mathbb{Q}_+^*,\ \exists N,\ \forall n \geq N,\ |x_n - \ell| \leq \varepsilon,$$
et même
$$\forall p \in \mathbb{N}^*,\ \exists N,\ \forall n \geq N,\ |x_n - \ell| < \frac{1}{p}\cdot$$
Remarquons que la définition peut aussi sécrire sous la forme équivalente :
$$\forall \varepsilon > 0,\ \exists N,\ n \geq N \Longrightarrow |x_n - \ell| \leq \varepsilon.$$

5.1.3 Sous-suites et limites

La première remarque à faire est qu'une sous-suite d'une suite convergente est convergente.

Proposition 5.1.4. *Soit* (y_n) *une suite extraite de la suite* (x_n). *Si* $\lim_n x_n = \ell$ *alors* $\lim_n y_n = \ell$.

Démonstration. — Dire que y_n est extraite de x_n revient à dire qu'il existe une fonction $p : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que que $y_n = x_{p_n}$. La fonction p étant strictement croissante et à valeurs dans $\mathbb{N}$, on vérifie facilement par récurrence que pour tout n, on a $p(n) \geq n$.

Soit donc $\varepsilon > 0$: il existe N tel que, si $n \geq N$, $|x_n - \ell| \leq \varepsilon$. Mais alors, si $n \geq N$, $p(n) \geq n \geq N$, et donc
$$|y_n - \ell| = |x_{p(n)} - \ell| \leq \varepsilon,$$
et par conséquent, y_n converge vers ℓ.

∎

Ainsi, dans l'exemple (ii) de la page 69, la suite $x_n = 1/n^p$ est extraite de la suite $1/n$, et donc converge vers 0.

Remarque. *En pratique, ce résultat permet de montrer qu'une suite ne converge pas. Il suffit de trouver deux sous-suites qui convergent vers des limites différentes. Ainsi, la suite* $u_n = (-1)^n$ *ne converge pas, car* $u_{2n} = 1$ *et donc* (u_{2n}) *converge vers* 1, *tandis que* $u_{2n+1} = -1$ *et donc* (u_{2n+1}) *converge vers* -1.

A titre d'exemple d'application de la proposition 5.1.4, nous avons

Corollaire 5.1.5. *Soit* (x_n) *une suite qui converge vers* ℓ *et soit* $p \in \mathbb{N}$ *un entier. Alors :*
-la suite $(x_{n+p})_{n\in\mathbb{N}}$ *converge aussi vers* ℓ,
-les suites (x_{2n}) *et* (x_{2n+1}) *convergent vers* ℓ.

(Voir l'exercice 115 pour une réciproque de ce dernier point).

5.1.4 Limites infinies

Lorsqu'une suite $(u_n)_{n\geq 0}$ ne converge pas, on dira par définition qu'elle *diverge*. En fait lorsqu'une suite diverge, elle peut avoir des comportements très variés. Nous allons décrire un type de particulièrement simple.

Définition.

(1) *On dit qu'une suite* $(u_n)_{n\geq 0}$ *de nombres réels* a pour limite $+\infty$ *si pour tout nombre réel* $A > 0$ *il existe un rang* $N \geq 0$ *tel que pour tout* $n \geq N$, *on ait* $u_n \geq A$. *On dira aussi que la suite* $(u_n)_{n\in\mathbb{N}}$ *tend vers* $+\infty$ *et on écrira dans ce cas* $\lim_{n\to+\infty} u_n = +\infty$ *ou encore* $\lim_n x_n = +\infty$.

(2) On dit que la suite $(u_n)_{n\geq 0}$ a *pour limite $-\infty$ si pour tout nombre réel $A > 0$ il existe un rang $N \geq 0$ tel que pour tout $n \geq N$, on ait $u_n \leq -A$. On dira aussi que la suite $(u_n)_{n\in\mathbb{N}}$ tend vers $-\infty$ et on écrira dans ce cas $\lim_{n\to+\infty} u_n = -\infty$, ou encore $\lim_n x_n = -\infty$.*

On peut faire les mêmes remarques à propos de cette définition que celles faites à propos de la convergence. Elle suivent les mêmes règles en ce qui concerne les sous-suites.

Observons que la suite $(u_n)_{n\geq 0}$ tend vers $-\infty$ si et seulement si la suite $(-u_n)_{n\geq 0}$ tend vers $+\infty$.

Une suite tend vers $+\infty$ lorsque ses termes deviennent arbitrairement grands à partir d'un certain rang. Cette propriété ne change pas si on modifie ou supprime un nombre fini de termes de la suite. Le lien entre les deux notions de limites (finies et infinies) sera établi plus loin.

Donnons quelques exemples pour illustrer cette définition.

Exemples.

(i) Pour $x_n = n$: $\lim_{n\to\infty} n = +\infty$ (heureusement !)

(ii) Soit $p \geq 1$ un entier fixé. Alors on a : $\lim\limits_{n\to+\infty} n^p = +\infty$. C'est en effet une suite extraite de la précédente.

(iii) Soit $a > 1$ un nombre réel. Alors $\lim_n a^n = +\infty$.
En effet, soit $A > 0$ quelconque. Comme $a > 1$, on sait que $\ln a > 0$ donc

$$a^n \geq A \Longleftrightarrow n \ln a \geq \ln A \Longleftrightarrow n \geq \frac{\ln A}{\ln a}\cdot$$

Si l'on choisit $N := 1 + \lfloor | \frac{\ln A}{\ln a} | \rfloor$, on a bien que tout $n \geq N$ vérifie $n \geq \frac{\ln A}{\ln a}$ et donc $a^n \geq A$.

(iv) Soit $0 < q < 1$. Il résulte de l'exemple précédent (à démontrer !) que

$$\lim_{n\to+\infty} q^n = 0.$$

On pose pour $n \geq 1$, $S_n := \sum_{k=0}^{n} q^k = 1 + q + \ldots + q^n$. La formule fondamentale (5.1) de la page 66 nous donne

$$\forall n \geq 1, \quad (1-q)S_n = 1 - q^{n+1}.$$

On en déduit alors que

$$\lim_{n\to+\infty} S_n = \frac{1}{1-q}\cdot$$

De plus on a l'inégalité fondamentale suivante

$$\forall n \geq 1,\ 1 + q + \ldots + q^n < \frac{1}{1-q}\cdot$$

5.2 Propriétés élémentaires des suites et règles de calcul

Les propriétés énoncées dans ce paragraphe sont élémentaires mais sont fondamentales pour simplifier l'étude de la convergence des suites.

5.2.1 Suites bornées

Commençons par rappeler quelques définitions déja vues pour les parties de $\mathbb{R}$.

Définition.

(i) On dit que la suite $(u_n)_{n\geq 0}$ est majorée dans $\mathbb{R}$ s'il existe un nombre réel B tel que $\forall n \in \mathbb{N},\ u_n \leq B$. On dit alors que B est un majorant de la suite $(u_n)_{n\geq 0}$ dans $\mathbb{R}$ ou que celle-ci est majorée par B.

(ii) On dit que la suite $(u_n)_{n\geq 0}$ est minorée dans $\mathbb{R}$ s'il existe un nombre A tel que $\forall n \in \mathbb{N},\ u_n \geq A$.

(iii) On dit que la suite $(u_n)_{n\geq 0}$ est bornée si la suite est à la fois majorée et minorée.

Remarque. *Observons qu'une suite $(u_n)_n$ est bornée dans $\mathbb{R}$ si et seulement si il existe un réel $M > 0$ tel que $|u_n| \leq M$ pour tout $n \in \mathbb{N}$: pour le voir, observons que si $\forall n \in \mathbb{N}, A \leq u_n \leq B$, la propriété est vraie avec $M := \max\{|A|, |B|\}$.*

On a alors les propriétés élémentaires suivantes.

Proposition 5.2.1.

(i) Toute suite de nombres réels qui converge dans $\mathbb{R}$ est une suite bornée dans $\mathbb{R}$.

(ii) Si (u_n) converge vers $b > 0$, alors il existe un entier N tel que, pour tout $n \geq N$, $u_n \geq \frac{b}{2} > 0$. Donc, à partir d'un certain rang, les termes de la suite sont strictement positifs.

(iii) Toute suite de nombres réels qui a pour limite $+\infty$ est minorée et non majorée dans $\mathbb{R}$.

On pourrait énoncer des résultats analogues pour une suite convergente de limite strictement négative, ou pour une suite tendant vers $-\infty$.

Démonstration. — En effet, supposons d'abord que $\ell := \lim_{n\to+\infty} u_n \in \mathbb{R}$ et fixons $\varepsilon = 1$. Il existe alors par définition un entier $N \geq 1$ tel que $\forall n \geq N, |u_n - \ell| \leq 1$. Par l'inégalité triangulaire, on a donc $|u_n| \leq |\ell| + 1$ pour $n \geq N$. En posant

$$M := \max\{|\ell| + 1, |u_0|, \ldots, |u_{N-1}|\},$$

on a $\forall n \in \mathbb{N}, |u_n| \leq M$, et donc la suite (u_n) est bornée.

Supposons maintenant que $\lim_n u_n = b > 0$. Appliquons la définition de la limite avec $\varepsilon = b/2 > 0$: il existe un entier N tel que, pour $n \geq N$ $|u_n - b| \leq b/2$, d'où $u_n \geq b - b/2 = b/2$.

Supposons maintenant que $\lim_{n\to+\infty} u_n = +\infty$. Alors par définition pour tout $A > 0$, il existe $N \in \mathbb{N}$ tel que $\forall n \geq N, u_n > A$. Il en résulte en particulier qu'aucun nombre réel $A > 0$ n'est un majorant de la suite (u_n) et donc celle-ci n'est pas majorée dans $\mathbb{R}$.

Par ailleurs en appliquant la définition de la limite $+\infty$ avec $A = 1$ on obtient un rang n_0 tel que pour tout $n \geq n_0$, $u_n \geq 1$. Posons alors $m := \min\big(1, u_0, u_1, \ldots, u_{n_0-1}\big)$. Alors pour tout $n \in \mathbb{N}$, $u_n \geq m$, donc la suite est minorée. ∎

Remarque. *La réciproque de cette proposition est fausse : il ne suffit pas que la suite u_n soit bornée pour qu'elle converge, et il ne suffit pas non plus qu'elle soit non majorée pour converger vers $+\infty$. On le voit sur les exemples suivants*

(1) La suite définie par $u_n := (-1)^n$ pour $n \geq 0$ est bornée mais elle n'a pas de limite (voir exercice 1).

(2) La suite définie par $u_n := (-1)^n \cdot n$ pour $n \geq 0$ n'est ni majorée ni minorée et n'a pas de limite.

(3) La suite $n(1 + (-1)^n)$ est minorée, mais ne converge pas vers $+\infty$ (La sous-suite u_{2n} converge vers $+\infty$ tandis que la sous-suite u_{2n+1} vaut identiquement 0, et donc converge vers 0.

Le résulat suivant est utile dans la pratique.

Proposition 5.2.2. *Soient $(a_n)_{n\in\mathbb{N}}$ une suite qui converge vers 0 et $(u_n)_{n\in\mathbb{N}}$ une suite bornée. Alors la suite $(a_n u_n)_{n\in\mathbb{N}}$ converge vers 0.*

Démonstration. — Il suffit d'appliquer la proposition 5.1.3 de la page 68 ∎

Exemple. Ainsi, la suite $u_n = \dfrac{(-1)^n}{n}$ converge vers 0. En effet, $a_n = 1/n$ converge vers 0, et $u_n = (-1)^n$ est bornée.

5.2.2 Opérations algébriques sur les limites

Dans la pratique, on n'a pas besoin de revenir à la définition pour démontrer qu'une suite donnée converge vers une limite ℓ. On part de suites dont on connaît bien la limite, et on construit avec elles de nouvelles suites plus compliquées dont on peut identifier la limite.

Proposition 5.2.3. *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites de nombres réels qui convergent vers les nombres réels a et b respectivement. Alors on a les propriétés suivantes :*

(i) (Addition des limites) La suite $(u_n + v_n)_{n\in\mathbb{N}}$ converge vers $a + b$ i.e. *:*

$$\lim_{n\to+\infty}(u_n + v_n) = \lim_{n\to+\infty} u_n + \lim_{n\to+\infty} v_n.$$

(ii) (Multiplication des limites) La suite $(u_n \cdot v_n)_{n\in\mathbb{N}}$ converge vers $a \cdot b$ i.e. *:*

$$\lim_{n\to+\infty}(u_n \cdot v_n) = (\lim_{n\to+\infty} u_n) \cdot (\lim_{n\to+\infty} v_n).$$

(iii) (Quotient de limites) Si $b \neq 0$, il existe un rang $p \geq 0$ tel que $v_n \neq 0$ pour tout $n \geq p$ et la la suite $(u_n/v_n)_{n\geq p}$ converge vers a/b i.e. *:*

$$\lim_{n\to+\infty}(u_n/v_n) = (\lim_{n\to+\infty} u_n)/(\lim_{n\to+\infty} v_n).$$

Démonstration. —

Commençons par le point (i) : $\varepsilon > 0$ étant donné, il existe N_1 tel que, pour $n \geq N_1$, $|x_n - \ell_1| \leq \varepsilon/2$. De même, il existe N_2 tel que, pour $n \geq N_2$, $|y_n - \ell_2| \leq \varepsilon/2$. Alors, pour $n \geq \max(N_1, N_2)$, on a

$$|x_n + y_n - \ell_1 - \ell_2| \leq |x_n - \ell_1| + |y_n - \ell_2| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Pour la point (ii), c'est un peu plus compliqué. Tout d'abord, les suites (u_n) et (v_n) sont convergentes, et donc bornées. Ensuite, on écrit

$$u_n v_n - ab = u_n(v_n - b) + b(u_n - a).$$

Les suites $(v_n - b)$ et $(u_n - a)$ convergent vers 0. Par la proposition 5.2.2 de la page 72, on en déduit que $u_n(v_n - b)$ et $v_n(u_n - a)$ convergent vers 0, et donc par le point (i), la somme converge : $u_n v_n - ab$ converge donc vers 0 et par suite $u_n v_n$ converge vers ab.

Pour le point (iii). On commence par se ramener au cas $b = 1$ en changeant v_n en v_n/b. Grâce à la propriété (ii), on se ramène ensuite au cas $u_n = 1$ pour tout n.

Puisque $\lim_n v_n = 1$, la proposition 5.2.1 fournit un rang N_0 tel que, pour tout $n \geq N_0$, $v_n \geq 1/2$. Pour $n \geq N_0$, on a ainsi

$$\left|\frac{1}{v_n} - 1\right| = \frac{|1 - v_n|}{|v_n|} \leq 2|1 - v_n|.$$

Comme $2|v_n - 1|$ converge vers 0, $|\frac{1}{v_n} - 1|$ converge vers 0, grâce à la proposition 5.1.3. ∎

Les formules d'opérations sur les limites s'étendent partiellement aux limites infinies, grâce au résultat suivant :

Proposition 5.2.4 (Mélange de limites finies et infinies)**.**

(i) Soit $(u_n)_{n\in\mathbb{N}}$ une suite **minorée** *de nombres réels et $(v_n)_{n\in\mathbb{N}}$ une suite de nombres réels qui tend vers $+\infty$. Alors la suite $(u_n + v_n)_{n\in\mathbb{N}}$ tend vers $+\infty$.*

Ceci s'applique en particulier lorsque (u_n) converge, ou tend vers $+\infty$.

(ii) *Supposons qu'il existe un rang n_0 et un réel $a > 0$ tel que $\forall n \geq n_0,\ u_n \geq a$. Si de plus $\lim_n v_n = +\infty$, alors $\lim_n(u_n \cdot v_n) = +\infty$.*

Ceci s'applique en particulier lorsque (u_n) converge vers un nombre réel strictement positif, ou lorqu'elle tend vers $+\infty$

(iii) *Si (u_n) tend vers $+\infty$ alors $(1/u_n)$ tend vers 0.*

Démonstration. — Nous montrons le premier point et laissons les autres en exercice. Soit m un minorant de la suite (u_n). Soit $A > 0$ quelconque. Soit $A' := A + |m| > 0$. Comme (v_n) tend vers $+\infty$, il existe n_0 tel que pour $n \geq n_0$ on a toujours $v_n \geq A'$. Ainsi pour $n \geq n_0$, $u_n + v_n \geq m + A' = A + m + |m| \geq A$. ∎

Les règles de calcul pour les mélanges de limites finies et infinies se déduisent du résultat précédent, en effectuant les changement de signes appropriés. Nous les résumons dans un tableau, avec les conventions suivantes : $a \in \mathbb{R}$ (donc fini) et on note ? quand la réponse est *indéterminée*, c'est à dire qu'elle dépend des cas.

u_n	v_n	$u_n + v_n$	$u_n \cdot v_n$	u_n/v_n	v_n/u_n
$+\infty$	$+\infty$	$+\infty$	$+\infty$	?	?
$+\infty$	$-\infty$	?	$-\infty$	?	?
$a > 0$	$+\infty$	$+\infty$	$+\infty$	0	$+\infty$
$a > 0$	$-\infty$	$-\infty$	$-\infty$	0	$-\infty$
0	$+\infty$	$+\infty$	?	0	?
0	$-\infty$	$-\infty$	?	0	?
$a < 0$	$+\infty$	$+\infty$	$-\infty$	0	$-\infty$
$a < 0$	$-\infty$	$-\infty$	$+\infty$	0	$+\infty$
$-\infty$	$+\infty$	?	$-\infty$	?	?
$-\infty$	$-\infty$	$-\infty$	$+\infty$	?	?

L'indétermination des deux cases centrales de la colonne de droite du tableau peut être levée avec des hypothèses de signe. En effet, si $u_n > 0$ avec $\lim_n u_n = 0$ et si $\lim_n v_n = +\infty$, alors $\lim_n(v_n/u_n) = +\infty$ (avec les règles évidentes si les signes sont changés). Mais il est possible que (u_n) converge vers 0, que v_n converge vers $+\infty$ sans que v_n/u_n ne tende vers $+\infty$ ni ne tende vers $-\infty$: considérer par exemple $u_n = (-1)^n/n$ et $v_n = n$.

5.2.3 Passage à la limite dans les inégalités

Voici maintenant des propriétés d'encadrement très utiles dans les calculs de limites. Ces propositions résultent facilement des définitions.

On notera $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, où $-\infty$ et $+\infty$ sont deux nouveaux éléments représentant les limites infinies avec la relation d'ordre suivante : pour tout $x \in \mathbb{R}$, $-\infty < x < +\infty$.

Proposition 5.2.5 (Principe de conservation des inégalités larges)**.** *Soient $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites convergentes avec $\lim_n u_n = a$ et $\lim_n v_n = b$. Supposons qu'il existe un rang $p \in \mathbb{N}$ tel que $u_n \leq v_n$ pour tout $n \geq p$. Alors on a $a \leq b$.*

Démonstration. — Commençons par le cas où $u_n = 0$ pour tout n. Dans ce cas $a = 0$ et il nous faut démontrer que $b \geq 0$.

Supposons $b < 0$, et choisissons $\varepsilon = -b/2$ dans la définition de la convergence de b_n. Il existe donc N tel que si $n \geq N$, on aie $|v_n - b| \leq -b/2$, et pour un tel n, on a $v_n \leq b - b/2 = b/2 < 0$. Ceci est impossible par hypothèse si $n \geq p$.

Le cas général se ramène à celui-ci en considérant la suite $v_n - u_n$. ∎

Attention : une erreur fréquente consiste à affirmer que, si $u_n < v_n$ et $\lim u_n = a$ et $\lim v_n = b$, alors $a < b$. C'est évidemment faux : prendre $u_n = 0$ et $v_n = 1/n$. Dans ce cas, on peut seulement conclure que $a \leq b$. **Le principe de passage à la limite dans les inégalités n'est pas valable pour les inégalités strictes.**

Proposition 5.2.6 (Principe des gendarmes)**.** *Soient (u_n) une suite de nombres réels et $\ell \in \mathbb{R}$. On suppose qu'il existe deux suites de nombres réels (a_n) et (b_n) et un entier $p > 1$ tels que pour tout $n \geq p$ on ait $a_n \leq u_n \leq b_n$. Alors*

(i) Si $\lim_n a_n = \ell = \lim_n b_n$ alors $\lim_n u_n = \ell$

(ii) Si $\lim_n a_n = +\infty$, alors $\lim_n u_n = +\infty$. (Dans ce cas, la suite (b_n) ne sert à rien : on peut choisir $b_n = u_n$).

(iii) Si $\lim_n b_n = -\infty$, alors $\lim_n u_n = -\infty$. (Même remarque : (a_n) ne sert pas).

Démonstration. — Nous ne traitons que le cas où ℓ est finie. Les autres cas sont laissés aux lecteurs à titre d'exercice.

La suite $b_n - a_n$ converge vers 0. Mais

$$|u_n - a_n| = u_n - a_n \leq b_n - a_n = |b_n - a_n|.$$

Donc $u_n - a_n$ converge vers 0 (par la proposition 5.1.3). Or $u_n = (u_n - a_n) + a_n$ et c'est donc la somme d'une suite qui tend vers 0 et d'un suite qui converge vers ℓ. Par conséquent u_n converge vers ℓ. ∎

Attention, en général, on ne peut passer à la limite dans une inégalité que si on sait que chaque membre de cette inégalité a une limite. Dans le cas du théorème des suites encadrées, c'est le fait que les deux "suites extrêmes" tendent vers la même limite qui implique l'existence de la limite de la suite encadrée.

Exemple. Posons pour $n \geq 1$:

$$u_n := \sum_{k=1}^{n} \frac{n}{k + n^2} = \frac{n}{1 + n^2} + \ldots + \frac{n}{n + n^2}.$$

Ainsi u_n est la somme de n termes dont le plus petit est $\frac{n}{n+n^2}$ et le plus grand est $\frac{n}{1+n^2}$. Il en résulte que :

$$\forall n \geq 2,\ n \cdot \frac{n}{n + n^2} < u_n < n \cdot \frac{n}{1 + n^2}.$$

Posons $a_n = \frac{n^2}{n+n^2}$ et $b_n := \frac{n^2}{1+n^2}$. On montre facilement (exercice!) que $\lim_n a_n = 1 = \lim_n b_n$. Comme $a_n \leq u_n \leq b_n$ pour tout $n \geq 1$, on en déduit grâce au principe des gendarmes que $\lim_n u_n = 1$.

5.3 Suites monotones

L'utilisation de la définition pour démontrer qu'une suite converge suppose que l'on en connaisse la limite à l'avance, ce qui n'est pas toujours le cas. Il est donc souhaitable de trouver des conditions suffisantes (appelés "critères de convergence") permettant de décider qu'une suite converge sans en connaître a priori la limite. Le critère le plus simple concerne les suites monotones. Pour énoncer ce critère, nous aurons besoin de quelques définitions.

Définition. *Soit (x_n) une suite de nombre réels.*

(i) On dit que la suite $(x_n)_{n\in\mathbb{N}}$ *est* croissante *(resp.* strictement croissante*) si pour tout entier* $n \in \mathbb{N}$*, on a* $x_n \leq x_{n+1}$ *(resp.* $x_n < x_{n+1}$*).*

(ii) On dit que la suite $(x_n)_{n\in\mathbb{N}}$ *est* décroissante *(resp.* strictement décroissante*) si pour tout entier* $n \in \mathbb{N}$*, on a* $x_n \geq x_{n+1}$ *(resp.* $x_n > x_{n+1}$*).*

(iii) On dit que la suite $(x_n)_{n\in\mathbb{N}}$ *est* monotone *(resp.* strictement monotone*) si elle est soit croissante (resp. strictement croissante), soit décroissante (resp. strictement décroissante).*

Remarque. *Si une suite est croissante, alors pour tous* $n \leq p$, $x_n \leq x_p$*. Cela se voit immédiatement par récurrence sur p. Il y a bien sûr une propriété identique pour les suites décroissantes.*

Rappelons que nous avons déjà défini ce que veut dire suite majorée ou minorée (définition 5.2.1 page 71). Dire qu'une suite $x = (x_n)$ est majorée revient à dire que l'ensemble $x(\mathbb{N}) = \{x_n, n \in \mathbb{N}\}$ (c'est à dire l'ensemble des valeurs prises par la suite, ou bien l'ensemble image $x(\mathbb{N})$ par la suite si l'on se rappelle qu'une suite x est une application de $\mathbb{N}$ dans $\mathbb{R}$) est un ensemble majoré.

On peut maintenant énoncer le résultat fondamental suivant qui est une conséquence simple de la propriété de la borne supérieure de $\mathbb{R}$.

Théorème 5.3.1 (Critère de convergence des suites monotones)**.**

(i) Si (x_n) *est une suite croissante majorée, elle converge (avec une limite finie).*

(ii) Si (x_n) *est une suite croissante non majorée, alors* $\lim_n x_n = +\infty$*.*

(iii) De même, une suite décroissante et minorée converge dans $\mathbb{R}$*.*

(iv) Une suite décroissante et non minorée converge vers $-\infty$*.*

Démonstration. —

Nous ne traitons que les deux premiers points, les deux autres s'en déduisent en utilisant que si (x_n) est décroissante alors et $(-x_n)$ est croissante.

Commençons par le cas où la suite $x = (x_n)$ est croissante majorée. Nous savons que l'image $x(\mathbb{N})$ de la suite est un ensemble non vide et majoré. Il admet une borne supérieure M. Nous allons montrer que $\lim_n x_n = M$.

Tout d'abord, par définition, nous savons que pour tout $n \in \mathbb{N}$, $x_n \leq M$. Ensuite, par la propriété de la borne supérieure de la proposition 4.3.1 de la page 57, pour tout $\varepsilon > 0$, il existe un entier N tel que $x_N \geq M - \varepsilon$. Alors, pour $n \geq N$, nous avons, puisque la suite (x_n) est croissante,

$$M - \varepsilon \leq x_N \leq x_n \leq M,$$

et donc $|x_n - M| = M - x_n \leq \varepsilon$. La suite converge donc bien vers M.

Passons au cas où x est non majorée. Cela veut dire que pour tout $A \in \mathbb{R}$, il existe un élément x_N de la suite tel que $x_N \geq A$. Alors, pour $n \geq N$, nous avons

$$x_n \geq x_N \geq A,$$

et nous avons bien montré que la suite (x_n) converge vers $+\infty$ ∎

Remarque. *La démonstration du théorème nous en apprend un peu plus : si une suite est croissante et majorée, sa limite est sa borne supérieure.*

Donnons des exemples qui illustrent ce théorème.

Exemples.

(I) (Un cas de convergence).
Soit $(x_n)_{n\geq 1}$ la suite définie pour $n \geq 1$ par :

$$x_n := 1 + \frac{1}{1!} + \cdots + \frac{1}{n!}.$$

(i) Il est clair que la suite $(x_n)_{n\geq 1}$ est strictement croissante puisque $x_{n+1} - x_n = \frac{1}{(n+1)!} > 0$, pour tout $n \geq 1$.

(ii) Montrons qu'elle est majorée. En effet on vérifie facilement par récurrence que :

$$\forall n \geq 1, \quad n! \geq 2^{n-1}.$$

Il en résulte que pour tout $n \geq 1$, on a $x_n \leq 1 + \frac{1}{1} + \cdots \frac{1}{2^{n-1}} < 1 + 2 = 3$. Donc la suite $(x_n)_{n\geq 1}$ converge et sa limite, notée e, vérifie les inégalités $2,5 < e \leq 3$.

(II) (Un cas de divergence).
Soit (u_n) la suite définie pour $n \geq 1$ par la formule suivante :

$$u_n := \sum_{k=1}^{n} \frac{1}{k} = 1 + \frac{1}{2} + \cdots + \frac{1}{n}.$$

Nous allons montrer que la suite (u_n) tend vers $+\infty$.

(i) La suite $(u_n)_n$ est croissante puisque $u_{n+1} - u_n = 1/(n+1) \geq 0$ pour tout $n \geq 1$.

(ii) Montrons qu'elle n'est pas majorée. En effet, observons d'abord que si $n \geq 1$ et $p \geq 1$, alors

$$u_{n+p} - u_n = \frac{1}{n+1} + \cdots + \frac{1}{n+p} \geq \frac{p}{n+p},$$

puisque cette expression est la somme de p termes dont le plus petit est $1/(n+p)$. Par suite $u_{2n} - u_n \geq 1/2$ pour tout $n \geq 1$ et en particulier on a :

$$\forall k \geq 1, \quad u_{2^k} - u_{2^{k-1}} \geq \frac{1}{2}.$$

En fixant un entier $p \geq 2$ et en additionnant membre à membre les p inégalités obtenues pour $k = 1, 2, \cdots, p$ on obtient :

$$\forall p \geq 2, \quad (u_2 - u_1) + \cdots + (u_{2^p} - u_{2^{p-1}}) \geq \frac{p}{2}.$$

Les termes se simplifient deux à deux par "téléscopage" et l'on obtient $u_{2^p} - u_1 \geq p/2$ pour tout $p \geq 2$ et donc $u_{2^p} \geq 1 + p/2$ pour tout $p \geq 2$, ce qui prouve que la suite (u_n) n'est pas majorée. Par conséquent d'après le théorème précédent, elle tend vers $+\infty$ i.e.

$$\lim_{n\to+\infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) = +\infty.$$

5.4 Suites adjacentes

Le théorème suivant repose sur le critère de convergence des suites monotones.

Théorème 5.4.1 (Théorème des suites adjacentes)**.** *Soient* $(a_n)_{n\in\mathbb{N}}$ *et* (b_n) *deux suites de nombres réels telles que*

$$\forall n \in \mathbb{N},\ a_n \leq a_{n+1} \leq b_{n+1} \leq b_n. \tag{5.2}$$

Alors les deux suites $(a_n)_{n\in\mathbb{N}}$ *et* $(b_n)_{n\in\mathbb{N}}$ *convergent vers des nombres réels* α *et* β *respectivement qui vérifient les inégalités suivantes :*

$$\forall n \in \mathbb{N}, a_n \leq \alpha \leq \beta \leq b_n.$$

Si de plus $\lim_{n\to+\infty}(b_n - a_n) = 0$, *alors* $\alpha = \beta$.

Démonstration. — C'est une conséquence immédiate du critère des suites monotones : la suite (a_n) est croissante par définition, et majorée par b_1. Donc elle converge vers $\alpha := \sup\{a_k; k \in \mathbb{N}\}$. De même, (b_n) est décroissante et minorée par a_1, donc elle converge vers $\beta := \inf\{b_k; k \in \mathbb{N}\}$. On peut donc passer à la limite en n dans l'inégalité $a_n \leq b_n$, ce qui donne $\alpha \leq \beta$. Pour le dernier point on utilise l'encadrement $0 \leq \beta - \alpha \leq b_n - a_n$. ∎

Si les deux suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ vérifient les inégalités (5.2) et si $\lim_{n\to+\infty}(b_n - a_n) = 0$, on dira que ce sont des *suites adjacentes.* Le théorème affirme en particulier que deux suites adjacentes convergent vers une même limite, donnant ainsi naissance à un nombre réel ℓ dont $(a_n)_{n\geq 0}$ est une suite d'approximants par défaut et (b_n) est une suite d'approximants par excès.

Remarque. *Les hypothèses faites dans ce théorème traduisent le fait géométrique que les segments de la droite réelle à savoir* $I_n := [a_n, b_n]$ *pour* $n \in \mathbb{N}$, *forment une suite de segments emboîtés les uns dans les autres i.e.* $I_{n+1} \subset I_n$ *pour tout* $n \in N$ *(faire un dessin). La conclusion affirme que ces segments ont une intersection qui est encore un segment et que celui-ci est réduit à un point dans le cas où les longueurs des segments* (I_n) *deviennent infiniment petites lorque* n *devient grand. En raison de cette propriété géométrique, on appelle également ce théorème le théorème des segments emboîtés.*

Nous avons déja vu que le procédé de dichotomie appliqué à l'approximation de $\sqrt{2}$ donnait naissance à de telles suites. Donnons maintenant un autre exemple intéressant illustrant cette situation.

Exemple *(Le nombre réel e est irrationnel).* Posons pour $n \geq 1$:

$$a_n := 1 + \frac{1}{1!} + \frac{1}{2!} \ldots + \frac{1}{n!}, \quad b_n := a_n + \frac{1}{n \cdot n!}.$$

Nous allons démontrer que ces deux suites sont adjacentes et convergent vers un nombre réel *irrationnel* noté e en l'honneur du fameux mathématicien L. Euler : c'est la base du logarithme neperien.

(i) La suite (a_n) est strictement croissante car $a_{n+1} - a_n = 1/(n+1)! > 0$ pour tout $n \geq 1$.

(ii) La suite $(b_n)_{n\geq 1}$ est strictement décroissante. En effet pour $n \geq 1$, on a

$$\begin{aligned} b_{n+1} - b_n &= a_{n+1} - a_n + \frac{1}{(n+1)(n+1)!} - \frac{1}{nn!} \\ &= \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)^2} - \frac{1}{n}\right) \\ &= -\frac{1}{n(n+1)(n+1)!} < 0 \end{aligned}$$

(iii) Il est clair que $\lim_{n\to+\infty}(b_n - a_n) = 0$. D'après le théorème des suites adjacentes, on en déduit qu'il existe un nombre réel, noté e tel que :

$$e := \lim_{n\to+\infty} a_n = \lim_{n\to+\infty} b_n.$$

Il satisfait l'encadrement suivant : pour tout $n \in \mathbb{N}$, $a_n < a_{n+1} \leq e \leq b_{n+1} < b_n$, donc

$$1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} < e < 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{n\,n!}. \tag{5.3}$$

(iv) Montrons que le nombre réel e est irrationnel i.e. $e \in \mathbb{R} \setminus \mathbb{Q}$. On raisonne par l'absurde en supposant le contraire à savoir que e est rationnel. Dans ce cas, il existe deux entiers naturels $p \geq 1$ et $q \geq 1$ premiers entre eux tels que pour tout entier $n \in \mathbb{N}$,

$$a_n < \frac{p}{q} < a_n + \frac{1}{n\,n!}.$$

Prenons $n = q$ et multiplions les deux membres par $q \cdot q!$. On observe que par sa définition, $q \cdot q! a_q \in \mathbb{N}$, et en posant $N := q\,q! a_q$ on obtient $N < p\,q! < N + 1$, ce qui est absurde.

Les inégalités (5.3) permettent de donner une valeur approchée de e avec une majoration précise de l'erreur d'approximation. En effet on a :

$$0 < e - \left(1 + \frac{1}{1!} + \ldots + \frac{1}{n!}\right) < \frac{1}{n\ n!}, \quad \forall n \geq 1.$$

Ainsi pour obtenir une valeur approchée de e à 10^{-8} près par exemple, il suffit de choisir un entier n (le plus petit possible) tel que $n\ n! > 10^8$. Il est facile de vérifier que $n = 10$ suffit et qu'alors en calculant $a_{10} = 1 + \frac{1}{1!} + \ldots + \frac{1}{10!}$, on en déduit que $e \simeq 2,71828182\cdots$ à 10^{-8} près par défaut, ce qui signifie que $2,71828182 < e < 2,71828183$, autrement dit $e = 2,71828182\ldots$, les sept premières décimales obtenues étant exactes.

Note historique : L'irrationnalité du nombre réel e a été démontrée au 18^{eme} siècle par Leonhard Euler (1707-1783) et Johann Heinrich Lambert $(1728 - 1777)$. Cela signifie que e n'est pas solution d'une équation linéaire (i.e. équation algébrique de degré 1) $ax + b = 0$ à coefficients entiers $a, b \in \mathbb{Z}, a \neq 0$. Plus tard vers 1844, Joseph Liouville démontre que e n'est pas solution d'une équation quadratique (i.e. équation algébrique de degré 2) $ax^2 + bx + c = 0$ à coefficients entiers $a, b, c \in \mathbb{Z}, a \neq 0$ et conjecture que e est un nombre *transcendant* dans le sens où il n'est solution d'aucune équation algébrique $a_n x^n + \ldots a_1 x + a_0 = 0$ à coefficients entiers $a_n, a_{n-1}, \ldots, a_0 \in \mathbb{Z}, a_n \neq 0$. Cette conjecture a été démontrée vers 1873 par Charles Hermite $(1822 - 1901)$.

5.5 Développement décimal d'un nombre réel

Nous avons l'habitude de représenter les nombres réels par leur développement en base 10 (leur écriture décimale). Lorsque cette écriture comporte une infinité de décimales (par exemple celle de $1/3$ que nous avons vue page 66), il s'agit en fait d'une limite. Nous allons la définir rigoureusement.

Rappelons que si nous écrivons un nombre x en base 10 sous la forme

$$x = \pm c_k c_{k-1} \cdots c_0, d_1 d_1 \ldots d_p,$$

où les c_i et les d_i sont des chiffres compris entre 0 et 9, cela signifie que

$$x = (\pm 1)(c_k 10^k + c_{k-1} 10^{k-1} + \ldots + c_0 + d_1 10^{-1} + \ldots d_p 10^{-p}).$$

Il s'agit de l'écriture décimale de x. Le nombre $N := 10^p x$ est un entier, obtenu en décalant la virgule de p places vers la droite. Comme $x = N/10^p$ c'est un nombre décimal.

Les nombres décimaux, sont des nombres rationnels particuliers, dont l'écriture décimale contient un nombre fini de chiffres après la virgule. Nous allons voir que toute écriture décimale "illimitée" définit un nombre réel unique, et que tout nombre réel admet une écriture décimale *presque unique.*

Proposition 5.5.1.

(i) Soit $(d_n)_{n\geq 1}$ *une suite d'entiers prenant les valeurs* $0, 1, \ldots, 9$. *Alors la suite*

$$x_n = d_1 10^{-1} + \ldots + d_n 10^{-n} = \sum_{k=1}^{n} d_k 10^{-k}$$

converge en croissant vers un nombre réel x, *qu'on écrit par convention*

$$0, d_1 d_2 \ldots d_k \ldots \qquad \textit{ou} \qquad \sum_{k=1}^{\infty} d_k 10^{-k}.$$

(ii) Soit x *un nombre réel positif. Il existe un entier* P *et une suite d'entiers* $(d_k)_{k\geq 1}$ *avec* $d_k \in \{0, 1, \ldots, 9\}$ *tels que*

$$x = P + \sum_{k=1}^{\infty} d_k 10^{-k}.$$

(iii) Si $x \geq 0$ *n'est pas un nombre décimal alors il admet une unique écriture de cette forme. En revanche si* x *est décimal il en admet deux : une avec un nombre fini de décimales (son écriture décimale* propre*) et une qui se termine par une infinité de* 9.

Démonstration. — Commençons par le premier point. Considérons tout d'abord le cas particulier où tous les chiffres d_i valent 9 :

$$9 \times 10^{-1} + \ldots + 9 \times 10^{-n} = 9 \times 10^{-1}\big(1 + 10^{-1} + \ldots + 10^{-(n-1)}\big) = \frac{9}{10}\frac{1 - 10^{-n}}{1 - 10^{-1}} = 1 - 10^{-n}.$$

Cette suite est majorée par 1 et tend vers 1 quand n tend vers l'infini. On retiendra pour la suite que

$$1 = 0,999999\ldots$$

Considérons maintenant le cas général. La suite (x_n) est croissante. Il nous suffit de montrer qu'elle est majorée pour établir sa convergence. En majorant chaque d_i par 9 et en utilisant le calcul précédent, on obtient bien

$$x_n \leq 9 \times 10^{-1} + \ldots + 9 \times 10^{-n} = 1 - 10^{-n} \leq 1.$$

Pour le second point, on pose $P := \lfloor x \rfloor$ (la partie entière de x), c'est un entier positif car $x \geq 0$. On définit $y := x - \lfloor x \rfloor \in [0, 1[$. Notre but est d'écrire y sous la forme $\sum_{k=1}^{\infty} d_k 10^{-k}$ pour des chiffres d_k bien choisis. Définissons d'abord pour $n \in \mathbb{N}$,

$$y_n := \frac{\lfloor 10^n y \rfloor}{10^n}.$$

Comme $y \in [0, 1[$ on note que $y_0 = 0$. Par définition de la partie entière, on sait que

$$\lfloor 10^n y \rfloor \leq 10^n y < \lfloor 10^n y \rfloor + 1 \tag{5.4}$$

En divisant par 10^n on obtient que $y_n \leq y < y_n + 10^{-n}$. En particulier $|y_n - y| \leq 10^{-n}$ ce qui donne que

$$\lim_{n\to\infty} y_n = y.$$

Par ailleurs, en multipliant par 10 les termes de (5.4) on obtient que

$$10\lfloor 10^n y\rfloor \leq 10^{n+1}y < 10\lfloor 10^n y\rfloor + 10.$$

Comme les termes de gauche et de droite dont des entiers, on en déduit que

$$10\lfloor 10^n y\rfloor \leq \lfloor 10^{n+1}y\rfloor \leq 10\lfloor 10^n y\rfloor + 9.$$

On peut alors définir $d_{n+1} := \lfloor 10^{n+1}y\rfloor - 10\lfloor 10^n y\rfloor$ qui est un entier, et l'encadrement précédent permet de conclure que $d_{n+1} \in \{0, 1, \ldots, 9\}$. De plus, par construction

$$y_{n+1} = \frac{\lfloor 10^{n+1}y\rfloor}{10^{n+1}} = \frac{\lfloor 10^n y\rfloor}{10^n} + \frac{d_{n+1}}{10^{n+1}} = y_n + \frac{d_{n+1}}{10^{n+1}}.$$

Par récurrence, on en déduit que pour tout $n \geq 1$, $y_n = d_1 10^{-1} + \cdots + d_n 10^{-n}$. Ainsi on obtient l'écriture souhaitée du nombre réel x :

$$x = P + y = P + \lim_n y_n = P + \sum_{k=1}^{\infty} d_k 10^{-k}.$$

Nous débutons la preuve du troisième point en montrant que les nombres décimaux ont deux écritures possibles. Le point de départ est la relation $1 = 0,99999\ldots$. On peut en déduire d'autres en divisant par des puissances de 10 et par addition : par exemple $0,01 = 0,0099999\ldots$ et $1,24 = 1,2399999\ldots$, ou encore $2400 = 2399,99999\ldots$. De manière plus générale, si $x = c_m \ldots c_1 c_0, d_1, \ldots d_{k-1} d_k$ avec $d_k \neq 0$ on peut aussi l'écrire sous la forme

$$x = c_m \ldots c_1 c_0, d_1, \ldots d_{k-1}(d_k - 1)99999\ldots$$

Supposons maintenant qu'un nombre réel positif x admet deux écritures. Quitte à diviser x par une puissance de 10, nous pouvons supposer que ces écritures n'ont pas de chiffre à gauche de la virgule. Notre point de départ est le suivant :

$$x = 0, c_1 c_2 \ldots c_k \ldots \quad \text{et} \quad x = 0, d_1 d_2 \ldots d_k \ldots.$$

Notons ℓ la première décimale qui n'est pas la même dans les deux écritures et supposons (quitte à échanger les rôles) que $c_\ell < d_\ell$. On a donc $c_i = d_i$ si $i < \ell$. De plus $c_\ell \leq d_\ell - 1$. Nous utilisons aussi, pour $i > \ell$ les inégalités $c_i \leq 9$ et $0 \leq d_i$, qui nous permettent d'écrire

$$\begin{array}{rcl} x & = & 0, c_1 c_2 \ldots c_k \ldots = 0, d_1 \ldots d_{\ell-1} c_\ell c_{\ell+1} \ldots c_k \ldots \\ & \leq & 0, d_1 \ldots d_{\ell-1}(d_\ell - 1)99999\ldots = 0, d_1 \ldots d_{\ell-1} d_\ell \\ & \leq & 0, d_1 \ldots d_{\ell-1} d_\ell d_{\ell+1} \ldots d_k \ldots = x. \end{array}$$

Tous ces termes sont forcément égaux à x, ce qui signifie qu'aucune des inégalités que nous venons d'appliquer ne peut être stricte. Ainsi nous avons forcément : $c_\ell = d_\ell - 1$ et pour $i > \ell$, $c_i = 9$ et $d_i = 0$. Les deux écritures sont donc de la forme

$$x = 0, d_1 d_2 \ldots d_{\ell-1}(d_\ell - 1)99999\ldots \quad \text{et} \quad x = 0, d_1 d_2 \ldots d_{\ell-1} d_\ell.$$

Ainsi x est un nombre décimal.

∎

Pour finir mentionnons que les nombres rationnels sont ceux qui ont une écriture décimale périodique à partir d'un certain rang. D'une part il est facile de voir, en généralisant l'approche de l'exercice 112 qu'une écriture décimale infinie périodique à partir d'un certain rang est égale à un rationnel. On se convainc facilement de la réciproque : si $x = p/q$, on "pose" la division de p par q. Soit la division s'arrête et x est décimal. Soit elle ne s'arrête pas ; dans ce cas il y a une infinité de restes successifs qui ne peuvent prendre que des valeurs entières entre 1 et $q - 1$. Donc une valeur du reste se répète et à partir de là l'opération devient périodique.

Remarque. *Ce qu'on a fait avec l'écriture décimale (en base* 10*), peut bien évidement se faire avec une base b quelconque.*

5.6 Complément : Deux théorèmes fondamentaux

Nous ne nous servirons pas beaucoup des résultats qui suivent dans la suite de ce cours, mais ils jouent un rôle central dans de nombreux pans des mathématiques. Le premier concerne les suites bornées :

Théorème 5.6.1 (Bolzano-Weierstrass)**.** *De toute suite bornée, on peut extraire une sous-suite convergente.*

Démonstration. — Soit x_n une suite bornée, à valeurs par exemple dans l'intervalle $[a, b]$, avec $a < b$. Quitte à remplacer u_n par $\frac{u_n - a}{b-a}$ qui est à valeurs dans $[0, 1]$, on peut supposer que $a = 0$ et $b = 1$.

Nous allons procéder par dichotomie, comme dans l'exemple de la page 53. Coupons l'intervalle $I = [0, 1]$ en 2 parts égales : $[0, \frac{1}{2}]$ et $[\frac{1}{2}, 1]$. L'un au moins de ces intervalles contient une infinité de points de la suite : appelons le I_1. Il est de longueur 1/2. Coupons à nouveau I_1 en deux parts égales. L'un au moins de ces deux nouveaux intervalles contient une infinité de points de la suite. Appelons le I_2 et recommençons. On construit ainsi par récurrence une suite d'intervalles $I \supset I_1 \supset I_2 \supset \ldots \supset I_n$ qui contiennent chacun une infinité de points de la suite. La longueur de I_n vaut $1/2^n$ (on a coupé n fois en 2).

Choisissons alors n_1 tel que $x_{n_1} \in I_1$. Il existe $n_2 > n_1$ tel que $x_{n_2} \in I_2$. On recommence l'opération, et on construit ainsi par récurrence $n_{k+1} > n_k$ tel que $x_{n_k} \in I_k$ (c'est possible justement parce que I_k contient une infinité de points de la suite). Appelons $I_k = [a_k, b_k]$. On a $b_n - a_n = 2^{-n}$ et $a_k \leq x_{n_k} \leq b_k$. D'après le théorème des intervalles emboités, a_k et b_k convergent vers la même limite, et par le théorème des gendarmes, il en va de même de x_{n_k} : on a bien construit une sous-suite convergente.

∎

Le deuxième résultat fondamental est le critère de Cauchy. A part le résultat sur les suites monotones, c'est l'un des rares résultats dont nous disposions qui permet d'affirmer l'existence d'une limite sans la connaître. Mais il est beaucoup plus puissant que le résultat sur les suites monotones, car il s'agit d'une condition nécessaire et suffisante. Nous commençons par une définition

Définition. *Soit (x_n) une suite de nombres réels. Nous dirons que (x_n) est une suite de Cauchy si*

$$\forall \varepsilon > 0,\ \exists N,\ \forall n, p \geq N,\ |x_n - x_p| \leq \varepsilon.$$

Remarque. *La notion de suite de Cauchy n'est pas si facile à comprendre. Elle signifie que les distances mutuelles entre deux points de la suite deviennent petites quand N grandit, mais ceci* de façon uniforme*, c'est à dire d'une manière qui ne dépend que de N, mais pas de n et p plus grands que N.*

Notons que toute suite convergente est de Cauchy : en effet, si $\lim_n x_n = \ell$, pour $\varepsilon > 0$ donné, il existe un N tel que si $n \geq N$, alors $|x_n - \ell] \leq \varepsilon/2$. Dans ce cas, en écrivant

$$|x_n - x_p| \leq |x_n - \ell| + |x_p - \ell|,$$

on voit que si n et p sont supérieurs à N, alors $|x_n - x_p| \leq \varepsilon$. La suite est donc bien de Cauchy.

Le théorème fondamental suivant assure que la réciproque est vraie :

Théorème 5.6.2 (Critère de Cauchy)**.** *Toute suite de Cauchy est convergente.*

Démonstration. — Soit (x_n) une suite de Cauchy. Montrons d'abord qu'elle est bornée. En choisissant $\varepsilon = 1$, nous trouvons un N tel que, si $n, p \geq N$, on a $|x_n - x_p| \leq 1$. En particulier, si $n \geq N$, $|x_n - x_N| \leq 1$. Alors, $|x_n| \leq |x_N| + 1$, et la suite (x_n) est bornée par

$$\max\{|x_0], |x_1|, \ldots, |x_N|, |x_N| + 1\}.$$

Comme (x_n) est bornée, par le théorème de Bolzano-Weierstrass, elle admet une sous-suite (x_{n_k}) qui converge vers un nombre ℓ. Montrons alors que x_n converge vers ℓ.

Pour $\varepsilon > 0$ donné, il existe un k_0 tel que, si $k \geq k_0$, $|x_{n_k} - \ell| \leq \varepsilon/2$. Maintenant, puisque (x_n) est de Cauchy, il existe un N, tel que si $n, p \geq N$, $|x_n - x_p| \leq \varepsilon/2$. Choisissons alors $k \geq k_0$ tel que $n_k \geq N$. Pour $n \geq n_k$, on a

$$|x_n - \ell| \leq |x_n - x_{n_k}| + |x_{n_k} - \ell| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

C'est bien ce qu'on voulait démontrer.

$\blacksquare$

Puisqu'une suite de Cauchy (x_n) est bornée, pour tout n, l'ensemble $A_n = \{|x_p - x_q|,\ p, q \geq n\}$ est un ensemble majoré. Si l'on appelle $S_n = \sup(A_n) \geq 0$, dire que la suite est de Cauchy revient à dire que la suite (S_n) converge vers 0 lorsque $n \to \infty$. Remarquons que (S_n) est décroissante (car $A_{n+1} \subset A_n$), et pour qu'elle converge vers 0, il suffit qu'elle admette une sous-suite qui converge vers 0. Si (x_n) est une suite bornée, on peut aussi considérer l'ensemble majoré

$$B_n = \{[x_n - x_p|,\ p \geq n\} \subset A_n$$

et sa borne supérieure : $s_n = \sup(B_n) \geq 0$. Puisque $B_n \subset A_n$, alors $s_n \leq S_n$. Mais par ailleurs, on a pour $p, q \geq n$

$$|x_p - x_q| \leq |x_p - x_n| + |x_q - x_n| \leq 2s_n.$$

Si bien que $S_n \leq 2s_n$. Ces deux encadrements nous montrent que $\lim_n S_n = 0$ si et seulement si $\lim_n s_n = 0$. Donc une suite est de Cauchy si et seulement si $\lim_n s_n = 0$.

Exercices

Exercice 113. Démontrer que si (u_n) tend vers $+\infty$ alors il existe un rang $N \geq 1$ tel que pour tout $n \geq N$, $u_n > 0$ et que la suite $(1/u_n)_{n\geq N}$ converge vers 0.

Exercice 114. Soit $(u_n)_{n\geq 0}$ une suite de nombres réels qui converge vers 0. On suppose qu'il existe un rang $N \geq 1$ tel que $u_n > 0$ pour tout $n \geq N$. Démontrer que la suite $(1/u_n)_{n\geq N}$ converge vers $+\infty$. Que peut-on dire si $(u_n)_{n\geq 0}$ converge vers 0 en changeant de signe ?

Exercice 115. Soit (x_n) une suite de nombres réels. Montrez que si les sous-suites (x_{2n}) et (x_{2n+1}) convergent vers la même limite ℓ, alors (x_n) converge vers ℓ

Exercice 116. (*) Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs positives et convergeant vers $\ell \in \mathbb{R}$. Démontrer (sans invoquer la continuité de la fonction racine carrée) que $\lim \sqrt{u_n} = \sqrt{\ell}$. On pourra traiter à part le cas $\ell = 0$.

Exercice 117. Etudier la convergence des suites définies pour $n \geq 1$ par

$$a_n = 2 - \frac{5}{n}, \quad b_n = 1 - n^2, \quad c_n = (-1)^n, \quad d_n = \frac{\sin(2\pi n/3)}{n}, \quad e_n = \frac{(-1)^{n^2}}{n}, \quad f_n = \frac{3^n}{5^{n-1}}.$$

Exercice 118. Etudier la convergence des suites défines pour n assez grand par :

$$u_n = \frac{3n^3 - 5n^2 - 7}{2n^3 - 8n - 11}, \quad v_n = \frac{70n^2 + 10n^3 + 50n + 170}{2n^3 - 90n - 110}, \quad w_n = \frac{n^5 - 30n^2 - 50\cos(n) - 750}{200n^3 + 18n^2 + 150},$$

$$x_n = \frac{n\sin(n!)}{n^2 + n - 1}, \quad y_n = \frac{2n^2 + (-1)^n n}{7\sqrt{5n} + 3n^2}, \quad z_n = (-1)^n \frac{3n-2}{n+5}.$$

Exercice 119. Posons $u_n := \sqrt{n+1} - \sqrt{n}$ pour $n \geq 0$.
1) Démontrer que

$$u_n = \frac{1}{\sqrt{n+1} + \sqrt{n}}, \quad \forall n \in \mathbb{N}^*.$$

En déduire $\lim_{n\to+\infty} u_n = 0$.
2) Calculer la limite suivante :

$$\lim_{n\to+\infty} \sqrt{n}(\sqrt{n+1} - \sqrt{n}).$$

Exercice 120. (*) En utilisant la méthode de l'exercice 119, étudiez la convergence des suites définies pour $n \geq 0$ par

$$v_n = \sqrt{n^2 + 2n - 1} - \sqrt{n^2 + 5n - 6}, \qquad w_n = n^2 \left(1 + \frac{1}{2n} - \sqrt{1 + \frac{1}{n}}\right).$$

Exercice 121. Calculer la limite suivante :

$$\lim_{n\to+\infty} (n^\alpha - n),$$

où $\alpha \in \mathbb{R}$ est un paramètre.

Exercice 122. (*) Soient P, Q des polynômes non nuls à coefficients réels. Montrez que pour $n \in \mathbb{N}$ assez grand, $u_n := P(n)/Q(n)$ est bien définie et étudiez-en la convergence lorsque n tend vers l'infini.

Exercice 123. On se propose de revenir sur l'exercice 112 en utilisant les résultats sur les limites que nous avons vus depuis. Soit $\alpha = 0,142857142857142857\ldots$, le nombre dont l'écriture décimale est constitué de la séquence 142857 répétée une infinité de fois.

1) Montrer qu'il existe un nombre entier β tel que $10^6\alpha = \beta + \alpha$.

2) En déduire l'expression de α comme une fraction que l'on simplifiera.

Exercice 124. Démontrer que pour tout entier $n \geq 1$:

$$\left(1 + \frac{1}{\sqrt{n}}\right)^n \geq 1 + \sqrt{n}.$$

En déduire $\lim_{n\to+\infty} \left(1 + \frac{1}{\sqrt{n}}\right)^n$.

Exercice 125. Soit a un nombre réel tel que $0 \leq a < 1$.
1) Démontrer que pour tout entier $n \geq 1$, on a $1 - a^n \leq n(1-a)$ et en déduire que :

$$0 \leq 1 - \left(1 - \frac{1}{n^2}\right)^n \leq \frac{1}{n}, \quad \forall n \in \mathbb{N}^*.$$

2) Calculer la limite suivante

$$\lim_{n\to+\infty} \left(1 - \frac{1}{n^2}\right)^n.$$

Exercice 126. On définit la suite $(x_n)_{n\geq1}$ par la formule suivante :

$$x_n := 2\cos(1/n^2) + 3\sum_{p=1}^{n} \frac{1}{\sqrt{p+1}}, \ n \in \mathbb{N}^*.$$

1) Calculer x_1, x_2, x_3.
2) Démontrer que pour si n et p sont des entiers tels que $1 \leq p \leq n$, on a

$$\frac{1}{\sqrt{n+1}} \leq \frac{1}{\sqrt{p+1}} \leq \frac{1}{\sqrt{2}}.$$

3) Démontrer que pour tout $n \geq 1$, $x_n \geq 3\sqrt{n+1} - 5$ et calculer la limite de la suite $(x_n)_{n\geq1}$. Trouver un entier N tel que $x_N \geq 235$.

Exercice 127. (*) Pour $n \geq 1$ on définit $u_n = \displaystyle\sum_{k=1}^{n} \frac{1}{n+k}$.

1) Explicitez les trois premiers termes. Etudiez le signe de $u_{n+1} - u_n$.

2) Montrer que (u_n) converge.

Exercice 128. (*)

1) Calculer

$$\lim_{n\to+\infty} \frac{n!}{n^n}.$$

2) Soit $b > 1$. Démontrer que

$$\lim_{n\to+\infty} \frac{b^n}{n!} = 0.$$

Indication : Considérer l'unique entier $p \geq 1$ tel que $p \leq b < p+1$ et vérifier que pour $n \geq p+1$ on a $n! \geq (p+1)^{n-p}p!$.

3) Soit $p \in \mathbb{N}^*$ fixé. En posant $b = 1 + a$ et en utilisant la formule du binôme, calculer

$$\lim_{n\to+\infty} \frac{b^n}{n^p}\cdot$$

4) Soit $b \in \mathbb{R}$ tel que $b > 1$ et $p \in \mathbb{N}^*$. Dans l'échelle de croissance à l'infini comparer les suites $(b^n)_{n\geq 1}$, $(n^p)_{n\geq 1}$, $(n!)_{n\geq 1}$ et $(n^n)_{n\geq 1}$.

Exercice 129. (*) Soit $(u_n)_{n\geq 0}$ une suite de nombres réels strictement positifs.
1) On suppose qu'il existe un rang $p \geq 1$ et un nombre réel $\alpha \in]0,1[$ tels que pour tout $n \geq p$ on ait $u_{n+1} \leq \alpha u_n$. Démontrer par récurrence que pour tout $n \geq p$, $0 \leq u_n \leq \alpha^{n-p} u_p$ et en déduire que $\lim_{n\to+\infty} u_n = 0$.
2) Démontrer que si $\lim_{n\to+\infty} u_{n+1}/u_n < 1$ alors $\lim_{n\to+\infty} u_n = 0$. En déduire que si l'on a $\lim_{n\to+\infty} u_{n+1}/u_n > 1$, alors $\lim_{n\to+\infty} u_n = +\infty$.
3) *Application :* Soit $c > 1$ un nombre réel et $p \geq$ un entier. Calculer les limites suivantes :

$$\alpha := \lim_{n\to+\infty} \frac{c^n}{n!} \text{ et } \beta := \lim_{n\to+\infty} \frac{c^n}{n^p}.$$

Exercice 130. Etudier la convergence de la suite $(y_n)_{n\in\mathbb{N}^*}$ définie par $y_n := \frac{x_1+\ldots+x_n}{n}$, $n \geq 1$, dans les cas suivants :

1) $x_n = 2^{-n}$,
2) $x_n = (-1)^n$,
3) $x_n = (-1)^n n$.

Exercice 131. (*) Soit $(x_n)_{n\in\mathbb{N}^*}$ une suite de nombres réels. On définit la suite de ses moyennes arithmétiques en posant :

$$y_n := \frac{x_1 + \ldots + x_n}{n}, \quad n \geq 1.$$

1) On suppose que $(x_n)_{n\geq 1}$ converge vers 0. Montrer que (y_n) converge aussi vers 0, en revenant à la définition des limites.
   Indication : Si pour $k \geq N$ on a $|x_k| \leq \delta$ alors on peut en déduire que pour $n \geq N$,
   $$|y_n| \leq \frac{|x_1 + \cdots + x_N|}{n} + \delta.$$
2) Démonter que si la suite $(x_n)_{n\geq 1}$ converge vers ℓ, alors la suite $(y_n)_{n\geq 1}$ converge vers ℓ. Etudier la réciproque.
3) Démontrer que le résultat est encore valable si $\ell = \pm\infty$.
4) On suppose que la suite $(x_n)_{n\geq 1}$ est monotone. Démontrer que si la suite des moyennes arithmétiques $(y_n)_{n\geq 1}$ converge vers ℓ, alors la suite $(x_n)_{n\geq 1}$ converge vers ℓ.

Exercice 132. (**) On rappelle que le nombre e a été défini comme $e := \lim_n x_n$ avec $x_n = \sum_{k=0}^{n} \frac{1}{k!}$. Le but de cet exercice est de démontrer que :

$$\lim_{n\to+\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

On considère la suite définie pour $n \geq 1$ par la formule $y_n := \left(1 + \frac{1}{n}\right)^n$.

1) Montrer que $y_n = 1 + \sum_{k=1}^{n} a_n(k)$ avec $a_n(k) := \frac{1}{k!}\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)$.
2) Montrer que $0 \leq a_n(k) < a_{n+1}(k) < \frac{1}{k!}$.

3) En déduire que la suite $(y_n)_{n\in\mathbb{N}}$ est croissante et majorée par e.

4) Montrer que pour $1 \leq p < n$ on a $1 + \sum_{k=1}^{p} a_n(k) < y_n < x_n$. Qu'en déduit-on lorsque n tend vers l'infini.

5) Conclure que $\lim_n y_n = \lim_n x_n$.

Exercice 133. (**) ($\mathbb{R}$ **n'est pas dénombrable**). Soit $I \subset \mathbb{R}$ un intervalle non vide et non réduit à un point, $(x_n)_{n\geq 0}$ une suite de nombres réels de l'intervalle I.
1) En s'inspirant du procédé de dichotomie, démontrer par récurrence sur $n \geq 0$ qu'il existe une suite décroissante de segments $[a_n, b_n] \subset I$ telle que pour tout $n \geq 0$ on ait $x_0, \cdots, x_{n+1} \notin [a_n, b_n]$ et $b_n - a_n \leq (b_0 - a_0)/3^n$.
En déduire qu'il existe un nombre réel $c \in I$ tel que $x_n \neq c$ pour tout $n \geq 0$.
2) En raisonnant par l'absurde, démontrer qu'il n'existe pas d'application surjective de $\mathbb{N}$ sur I et en déduire que I n'est pas dénombrable. En particulier $\mathbb{R}$ n'est pas dénombrable.

Chapitre 6

Suites récurrentes

Il est rare dans les applications qu'une suite soit donnée par une "formule explicite" permettant d'en calculer la limite. Nous allons étudier ici les suites données par un "procédé itératif" dans lequel on connait le premier terme x_0 de la suite et une relation de récurrence :

$$(R) \qquad x_{n+1} = f(x_n),\ n \in \mathbb{N}$$

faisant intervenir une fonction f définie sur un intervalle $I \subset \mathbb{R}$. Une telle relation, dite relation de récurrence d'ordre 1, permet de calculer tous les termes de la suite $(x_n)_{n\geq 0}$ de proche en proche à partir du premier terme x_0. En effet, connaissant le premier terme $x_0 \in I$, on peut calculer le terme suivant x_1 en appliquant la formule (R) pour $n = 0$, qui donne $x_1 = f(x_0)$. Connaissant le terme x_1, on peut alors calculer le terme suivant x_2 grâce à la relation de récurrence (R) appliquée avec $n = 1$ à condition que $x_1 \in I$, d'où $x_2 = f(x_1)$ et ainsi de suite ... connaissant le terme x_n on peut calculer le terme x_{n+1} par la formule $x_{n+1} = f(x_n)$ à condition que $x_n \in I$.

On dira que l'ensemble I est **stable** par f si $f(I) \subset I$, ce qui signifie que $x \in I \Longrightarrow f(x) \in I$. Par conséquent si $x_0 \in I$ et si I est stable par f, on peut définir par récurrence une suite $(x_n)_{n\in\mathbb{N}}$ unique vérifiant la relation de récurrence (R).

Il est possible en théorie de déduire de (R) une formule donnant le terme x_n en fonction de n et de x_0. En effet on définit les fonctions itérées de f en posant $f^{(1)} := f$, $f^{(2)} := f \circ f, \ldots,$ $f^{(n+1)} = f^{(n)} \circ f = f \circ f^{(n)}$ pour $n \in \mathbb{N}^*$. On a alors $x_n = f^{(n)}(x_0)$ pour $n \geq 1$. Cette formule est en général difficilement exploitable car calculer les itérées $f^{(n)}$ de f peut être très compliqué lorsque n devient grand, même si la fonction de départ est simple (voir les exemples ci-dessous).

L'étude générale du comportement de telles suites en fonction de la valeur initiale x_0 s'appelle la "théorie des systèmes dynamiques". Elle fait l'objet de recherches actuelles très actives. Il n'y a pas de méthode générale pour étudier les suites récurrentes. Nous nous contenterons ici de donner quelques techniques permettant d'étudier des exemples simples : lorsque la fonction f est continue, ou monotone sur l'intervalle où la suite est définie, ou bien lorsqu'elle est "contractante". Nous montrerons aussi comment les suites peuvent être utilisées pour approcher un nombre réel solution d'une équation de la forme $x = f(x)$, où f est une fonction convenable définie et *continue* sur un intervalle $I \subset \mathbb{R}$ tel que $f(I) \subset I$.

6.1 Compléments sur les fonctions : limites, continuité

Définition. *Soit $J \subset \mathbb{R}$ et $f : J \to \mathbb{R}$ une fonction. Soit $c \in \mathbb{R}$ un point tel que pour tout $\delta > 0$, $]c-\delta, c+\delta[\cap J \neq \emptyset$ (on dit que c est adhérent à J). On dit que la fonction f admet $\ell \in \mathbb{R}$ comme limite au point c si :*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in J \cap]c-\delta, c+\delta[,\ |f(x) - \ell| \leq \varepsilon.$$

Dans ce cas, le nombre ℓ est unique, c'est la limite de $f(x)$ lorsque x tend vers ℓ (en restant dans J), notée

$$\ell = \lim_{x \to c; x \in J} f(x)$$

que l'on note simplement $\lim_{x\to c} f(x)$ ou $\lim_c f$ lorsqu'il n'y a pas ambiguité sur l'ensemble J.

En d'autres termes, pour toute précision $\varepsilon > 0$, il existe un (petit) intervalle ouvert autour de c dans lequel f ne prend que des valeurs qui ne diffèrent pas de ℓ de plus de ε. Notons que le nombre c peut ne pas appartenir à J (par exemple $J = [0, 1[$ et $c = 1$) ou peut y appartenir (et dans ce cas nécessairement $\ell = f(c)$).

Nous utiliserons la proposition suivante, dite de caractérisation de la limite par les suites.

Proposition 6.1.1. *Les propriétés suivantes sont équivalentes :*

i) $\lim_{x\to c,\, x\in J} f(x) = \ell$,

ii) pour toute suite (x_n) à valeurs dans J et de limite c, on a $\lim_{n\to\infty} f(x_n) = \ell$.

Ce résultat permet d'adapter les résultats sur les limites de suites aux limites de fonctions.

Exercice 134. En utilisant la proposition précédente et les résultats connus sur les limites de suites, démontrer que si $\lim_c f = a$ et $\lim_c g = b$ alors $\lim_c(f + g) = a + b$. On pourra aussi démontrer ce résultat directement à partir de la définition de la limite d'une fonction

Définition. *Soit $J \subset \mathbb{R}$ et $f : J \mapsto \mathbb{R}$ une fonction.*

i) Soit $c \in J$. On dit que f est continue au point c si $\lim_{x\to c} f(x) = f(c)$, ce qui revient à écrire

$$\forall \varepsilon > 0,\ \exists \eta > 0, \forall x \in J \cap]x-\eta, x+\eta[,\ |f(x) - f(\ell)| \leq \varepsilon.$$

ii) On dit que f est continue sur J si f est continue en tous les points de J.

Les résultats d'opération sur les limites permettent de voir que la somme et le produit de fonctions continues sont encore continus, que l'inverse d'une fonction continue qui ne s'annule pas est continue.

La plupart des fonctions simples que nous connaissons sont continues. Par exemple

(i) Tous les polynômes sont continus sur $\mathbb{R}$.

(ii) Une fraction rationnelle, de la forme $\frac{P(x)}{Q(x)}$ où P et Q sont des polynômes, est continue sur son ensemble de définition. Ainsi par exemple $x \mapsto \frac{2x+5}{x-2}$ est continue sur $I =]-\infty, 2[\cup]2, +\infty[$.

(iii) Les fonctions $\sin x$, $\cos x$, $\exp(x)$ sont continues sur $\mathbb{R}$.

(iv) La fonction $\tan(x)$ est continue sur $]-\pi/2, \pi/2[$.

(v) La fonction $\log(x)$ est continue sur $]0, +\infty[$.

(vi) Plus généralement, si $f : I \mapsto J$ et $g : J \mapsto \mathbb{R}$ sont continues sur les intervalles I et J, la fonction $g \circ f : I \mapsto \mathbb{R}$ est continue sur I.

Exercice 135. Montrer que la fonction partie entière est continue sur $\mathbb{R} \setminus \mathbb{Z}$.

Une conséquence immédiate des définitions est la suivante

Proposition 6.1.2. *Soit $f : I \mapsto \mathbb{R}$ une fonction définie sur I et (x_n) une suite à valeurs dans I. Si (x_n) converge vers ℓ, avec $\ell \in I$, et si f est continue au point ℓ, alors la suite $(f(x_n))$ converge vers $f(\ell)$.*

Démonstration. — La preuve de cette proposition est très simple et est laissée à titre d'exercice. ∎

Nous terminons cette partie en présentant une preuve du théorème des valeurs intermédiaires.

Théorème 6.1.3. *Soit $I \subset \mathbb{R}$ un intervalle et $f : I \to \mathbb{R}$ une fonction continue. Alors $f(I)$ est un intervalle.*

Ainsi toutes les valeurs comprises entre deux valeurs de f sont encore des valeurs prises par f. Le théorème se déduit facilement de la proposition suivante, plus concrète :

Proposition 6.1.4. *Soit $f : [a, b] \to \mathbb{R}$ une fonction continue telle que $f(a) < 0 < f(b)$. Alors il existe $c \in [a, b]$ tel que $f(c) = 0$.*

Démonstration. — On suit le principe de dichotomie. Construisons deux suites (a_n) et (b_n) comme suit : $a_0 := a$ et $b_0 := b$. Pour $n \geq 0$,
— si $f\left(\frac{a_n+b_n}{2}\right) \leq 0$, on pose $a_{n+1} := \frac{a_n+b_n}{2}$ et $b_{n+1} = b_n$,
— si $f\left(\frac{a_n+b_n}{2}\right) > 0$, on pose $a_{n+1} := a_n$ et $b_{n+1} := \frac{a_n+b_n}{2}$.

On vérifie facilement que pour tout $n \in \mathbb{N}$, $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$. De plus $b_{n+1} - a_{n+1} = (b_n - a_n)/2$ donc $b_n - a_n = 2^{-n}(b - a)$ tend vers 0 lorsque n tends vers l'infini. Les suites (a_n) et (b_n) sont adjacentes, donc par le théorème 5.4.1 elles convergent vers une même limite notée $c \in [a, b]$.

Comme (a_n) converge vers c et que f est continue en c, on en déduit que la suite $(f(a_n))$ converge vers $f(c)$. Le même argument montre que $(f(b_n))$ converge vers $f(c)$. Or par construction, pour tout n, $f(a_n) \leq 0 \leq f(b_n)$. Donc en passant à la limite dans les inégalités, nous obtenons que $f(c) \leq 0 \leq f(c)$. Ainsi $f(c) = 0$. ∎

6.2 Représentation graphique

Pour étudier une suite définie par la donnée de x_0 et par la relation $x_{n+1} = f(x_n)$, il est utile de construire le graphe de la fonction f. On peut alors construire sur le dessin les premiers termes de la suite; ils apparaissent naturellement sur la droite déquation $(y = x)$. Le procédé de construction est illustré par la figure suivante, page 91. On place d'abord le point de coordonnées (x_0, x_0) ; la droite verticale passant par ce point rencontre le graphe de f au point de coordonnées $(x_0, f(x_0)) = (x_0, x_1)$. La droite horizontale passant par ce dernier point coupe la bissectrice d'équation $(y = x)$ au point de coordonnées (x_1, x_1). Ces opérations nous ont permis de passer de (x_0, x_0) à (x_1, x_1). Il ne reste plus qu'à répéter l'opération...

6.3 Fonctions f continues, points fixes

Définition. *Un point $\ell \in I$ tel que $f(\ell) = \ell$ s'appelle un point fixe de f. Sur notre construction graphique, un point fixe correspond à un point d'intersection du graphe de f et de la droite d'équation $(y = x)$.*

Le résultat suivant est fondamental. Il permet d'obtenir des informations sur l'éventuelle limite d'une suite récurrente, en termes des points fixes de la fonction f associée.

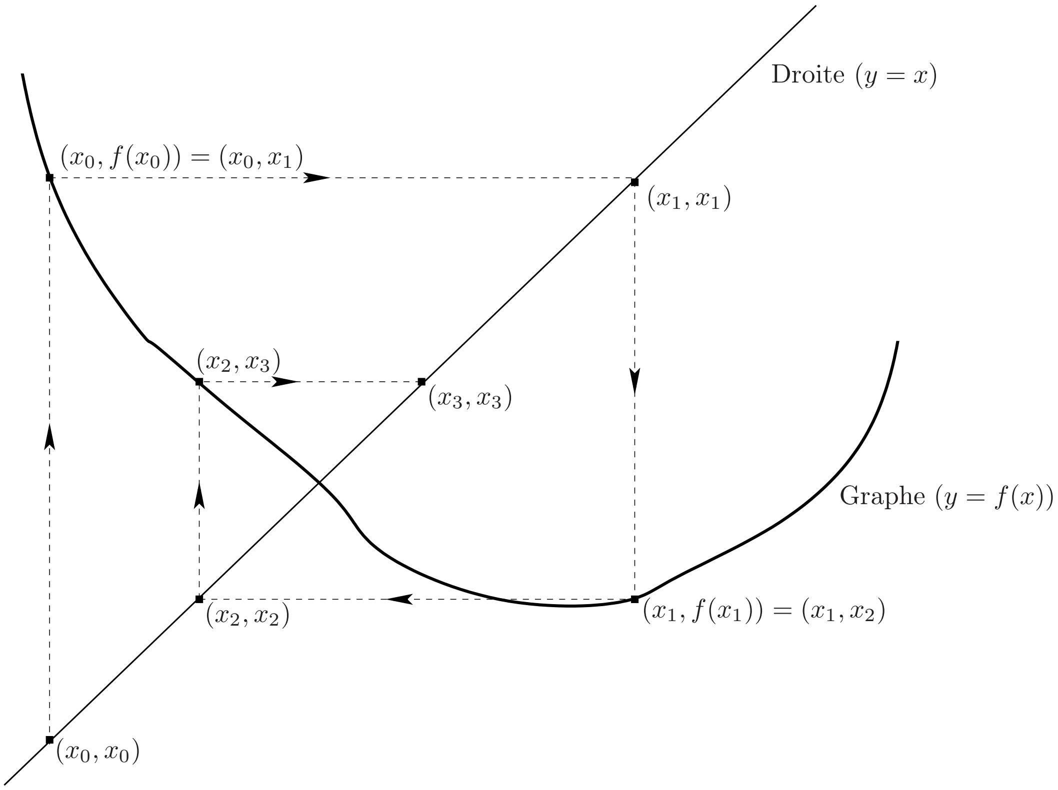

FIGURE 6.1 – Construction graphique d'une suite récurrente $x_{n+1} = f(x_n)$

Proposition 6.3.1. *Nous considérons un intervalle I et une fonction $f : I \mapsto I$. On choisit $x_0 \in I$ et on définit par récurrence la suite (x_n) par $x_{n+1} = f(x_n)$. Alors, si $\lim_n x_n = \ell \in I$ et si f est continue sur I, ℓ est solution de l'équation $f(\ell) = \ell$.*

Démonstration. — On passe à la limite dans $x_{n+1} = f(x_n)$, en appliquant la proposition 6.1.2.∎
Exemple. Soient a et b deux paramètres réels. On considère la suite (x_n) de premier terme x_0 et définie l'équation de récurrence linéaire

$$x_{n+1} = ax_n + b, \ n \in \mathbb{N}.$$

La proposition précédente nous dit que **si** la suite converge vers un nombre ℓ alors forcément $\ell = a\ell + b$. Quelle information peut-on déduire de cette équation portant sur ℓ ?

(i) Si $a = 1$ et $b = 0$, l'équation devient $0 = 0$ et ne nous donne aucun renseignement sur ℓ.
(ii) Si $a = 1$ et $b \neq 0$, l'équation devient $0 = b \neq 0$. Cette contradiction montre, par l'absurde, que la suite est divergente.
(iii) Enfin, si $a \neq 1$ l'équation obtenue a une seule solution $\ell = \frac{b}{1-a}$. Nous avons obtenu que **si** la suite converge, c'est vers $\frac{b}{1-a}$.

Il se trouve que le terme général de la suite (x_n) peut être calculé explicitement. Nous allons faire ce calcul, afin de pouvoir aller plus loin dans l'étude de la convergence :

(i) Si $a = 1$ et $b = 0$, l'équation de récurrence devient $x_{n+1} = x_n$. La suite est donc constante égale à x_0, qui est donc sa limite.
(ii) Si $a = 1$ et $b \neq 0$, l'équation de récurrence devient $x_{n+1} = x_n + b$. On reconnaît une **suite arithmétique** de raison b. On sait alors que pour tout n, $x_n = x_0 + nb$. La suite tend vers $+\infty$ si $b > 0$ et vers $-\infty$ si $b < 0$.

(iii) Enfin, si $a \neq 1$ notons $\ell := \frac{b}{1-a}$ la seule solution de l'équation de point fixe. Observons d'abord que si $b = 0$, la relation de récurrence devient $x_{n+1} = ax_n$ et l'on reconnaît une **suite géométrique** de raison a dont le terme général vaut alors $a^n x_0$. On peut en fait trouver une formule courte pour x_n, même si $b \neq 0$: en soustrayant terme à terme les relations $x_{n+1} = ax_n + b$ et $\ell = a\ell + b$, on obtient que $x_{n+1} - \ell = a(x_n - \ell)$. Donc la suite $(x_n - \ell)$ est géométrique de raison a et son terme général vaut $a^n(x_0 - \ell)$. On obtient ainsi que pour tout n,

$$x_n = a^n \left(x_0 - \frac{b}{1-a} \right) + \frac{b}{1-a}.$$

Si $|a| < 1$, on sait que $\lim_{n\to+\infty} a^n = 0$ et donc $\lim_{n\to+\infty} x_n = \frac{b}{1-a}$. Mais la suite n'est pas convergente pour toutes les valeurs des paramètres. Par exemple, si $a > 1$ et $x_0 > \frac{b}{1-a}$ alors la suite tend vers $+\infty$. Les autres cas sont laissés en exercice.

Faisons plusieurs remarques importantes sur la proposition 6.3.1.

Remarque. *Cette proposition nous montre que si l'on connaît f, alors les seules limites possibles de la suite (x_n) sont déterminées par la l'équation $\ell = f(\ell)$. Même si cette équation n'a qu'une solution dans I, cela ne veut pas dire que la suite x_n converge effectivement vers ℓ. Il se peut que la suite ne converge pas, ou bien qu'elle converge vers un point ℓ qui n'est pas dans I, par exemple lorsque l'un des bords de I est à l'infini, ou bien que $I =]a, b[$ et que $\ell = a$.*

Remarque. *Le choix de l'intervalle I sur lequel on travaille est important en pratique : il faut que $x_0 \in I$, mais surtout que $f(I) \subset I$. On a intérêt à chercher I le plus petit possible, car alors on peut localiser l'endroit où la suite se trouve et donc où sont les solutions possibles. C'est important par exemple lorsque l'équation $f(\ell) = \ell$ a plusieurs racines : dans ce cas, on peut avoir à chercher un intervalle I qui ne contienne qu'une seule de ces racines (et x_0).*

Par ailleurs, on aura toujours intérêt à chercher pour I un intervalle fermé borné $[a, b]$. En effet, si une suite x_n est à valeurs dans $[a, b]$ et si elle converge vers ℓ, alors $\ell \in [a, b]$ par la propriété de conservation des inégalités larges à la limite. Si l'intervalle était ouvert, alors la limite pourrait se trouver au bord de l'intervalle qui n'est pas dans I.

Remarque. *Il n'est pas nécessaire d'avoir $x_0 \in I$. Il suffit que x_1 soit dans I, ou bien qu'au bout d'un certain temps, $x_n \in I$: une fois que la suite est arrivée dans I, elle ne peut plus en sortir.*

6.4 Fonctions f monotones

Lorsque la fonction f est croissante sur I, la suite récurrente associée est monotone :

Proposition 6.4.1. *Soit I un intervalle de $\mathbb{R}$, $f : I \mapsto I$ une fonction croissante. On suppose que $x_0 \in I$. Alors on a les propriétés suivantes :*

i) Si $f(x_0) > x_0$, la suite définie par récurrence par $x_{n+1} = f(x_n)$ est croissante.

ii) Si $f(x_0) < x_0$, la suite définie par récurrence par $x_{n+1} = f(x_n)$ est décroissante.

iii) Si $f(x_0) = x_0$ la suite définie par récurrence par $x_{n+1} = f(x_n)$ est constante.

En particulier, si $I = [a, b]$ est un intervalle fermé et borné, et que f est croissante et continue sur I, la suite (x_n) converge vers une solution $\ell \in I$ de $f(\ell) = \ell$. En particulier, cette équation a nécessairement une solution dans I.

Démonstration. — Puisque f est croissante, pour tout $n \geq 1$, le nombre réel $x_{n+1} - x_n = f(x_n) - f(x_{n-1})$ est du signe de $x_n - x_{n-1}$. Donc ce signe est constant. Il est du signe de de $x_1 - x_0 = f(x_0) - x_0$. Par conséquent si $f(x_0) \leq x_0$, la suite $(x_n)_{n\geq 0}$ est décroissante et si $f(x_0) \geq x_0$, la suite est croissante.

Enfin, si $I = [a, b]$, la suite est soit croissante, majorée par b, soit décroissante, minorée par a. Elle converge donc par le résultat fondamental sur les suites monotones, vers une valeur $\ell \in I$. Puisque f est continue sur I, on a $f(\ell) = \ell$. ∎

Exemple. On considère la suite définie par son premier terme $x_0 \geq 0$ et la relation

$$x_{n+1} = x_n^2 + \frac{1}{9}, \; n \in \mathbb{N}.$$

Ici on a $x_{n+1} = f(x_n)$, où $f(x) := x^2 + 1/9$. Comme $x_0 \geq 0$ et que, si $x \geq 0$, alors $f(x) \geq 0$, on peut choisir $I = [0, \infty[$. Alors I est fermé et $f(I) \subset I$. Sur cet intervalle, la fonction est croissante et continue. Cherchons ses points fixes. L'équation $f(x) = x$ s'écrit $x^2 - x + 1/9 = 0$. Elle a deux racines positives

$$\ell_1 = \frac{3 - \sqrt{5}}{6}, \; \ell_2 = \frac{3 + \sqrt{5}}{6} \text{ avec } 0 < \ell_1 < \ell_2.$$

D'après la proposition 6.4.1, la suite (x_n) est une suite monotone dans l'intervalle I dont le sens de monotonie dépend du signe de $f(x_0) - x_0$ et en cas de convergence, sa limite vérifie l'équation $f(x) = x$ dans I, c'est à dire que c'est ℓ_1 ou ℓ_2.

Comme $f(x) - x = (x - \ell_1)(x - \ell_2)$, le signe de $f(x_0) - x_0$ dépend de la position relative de x_0 par rapport aux deux points fixes ℓ_1 et ℓ_2.

(i) Si $0 \leq x_0 < \ell_1$, on a $f(x_0) - x_0 > 0$. On sait donc que la suite est croissante.
Par ailleurs, on voit que si $x \leq \ell_1$, alors $f(x) \leq f(\ell_1) = \ell_1$. On en déduit par récurrence que pour tout n, $x_n \leq \ell_1$.
La suite est alors croissante majorée. Elle converge vers ℓ_1 ou bien ℓ_2. Comme $x_n \leq \ell_1$, par passage à la limite, on a aussi $\lim_n x_n \leq \ell_1$ et la limite est donc ℓ_1.

(ii) Si $\ell_1 < x_0 < \ell_2$, on a $f(x_0) - x_0 < 0$. On en déduit que x_n est décroissante. Elle est minorée par 0 car f est à valeurs positives. Donc (x_n) converge, vers ℓ_1 ou ℓ_2. Comme pour tout n, $x_n \leq x_0$ la limite est inférieure ou égale à $x_0 < \ell_2$. Donc (x_n) converge vers ℓ_1.

(iii) Si $x_0 > \ell_2$, on a $f(x_0) - x_0 > 0$ et donc la suite est croissante. On a toujours $x_n \geq x_0 > \ell_2$. Si la suite convergeait on aurait $\lim x_n \geq x_0 > \ell_2 > \ell_1$ mais aussi $\lim x_n \in \{\ell_1, \ell_2\}$. Cette contradiction nous assure que (x_n) diverge. On peut donc conclure que $\lim_n x_n = +\infty$ (puisqu'une suite croissante convergente ou tend vers $+\infty$).

(iv) Si $x_0 = \ell_1$ ou $x_0 = \ell_2$, la suite x_n est constante.

Remarque. *Dans cet exemple, selon la position de x_0 par rapport à ℓ_1 et ℓ_2, nous aurions pu choisir $I = [0, \ell_1]$, ou bien $[\ell_1, \ell_2]$ ou bien $[\ell_2, +\infty[$, qui sont aussi des intervalles tels que $f(I) \subset I$.*

A titre de comparaison, lorsque la fonction est décroissante, la situation est un peu différente :

Proposition 6.4.2. *Soient I un intervalle de $\mathbb{R}$ et $f : I \mapsto I$ une fonction décroissante, et on suppose $x_0 \in I$ Alors pour tout $x_0 \in I$, les deux suites $(x_{2n})_{n \geq 0}$ et $(x_{2n+1})_{n \geq 0}$ sont monotones.*

Si de plus f est continue sur I et si ces deux suites convergent dans I, leurs limites respectives sont des solutions de l'équation $f^{(2)}(x) = x$ dans I, où $f^{(2)} = f \circ f$.

Attention : ne pas confondre $f^{(2)}(x) = f(f(x))$ et $f^2(x) = f(x)^2$.

Démonstration. — En effet puisque f est décroissante sur I et que $f(I) \subset I$, la fonction $g := f^{(2)} = f \circ f$ est une fonction croissante sur I telle que $g(I) \subset I$. En effet, on a, pour $x \leq y$, $f(x) \geq f(y)$, donc $f(f(x)) \leq f(f(y))$.

Si on considère la suite $(y_n) := (x_{2n})$ on voit que $y_{n+1} = x_{2n+2} = f(x_{2n+1}) = f(f(x_{2n})) = g(y_n)$. Donc, la suite y_n est une suite récurrente associée à la fonction $g = f^{(2)}$. On peut alors appliquer la proposition 6.4.1 à la suite y_n. De la même façon la suite $(z_n) := (x_{2n+1})$ satisfait elle aussi l'équation de récurrence $z_{n+1} = g(z_n)$ et elle est elle aussi monotone. ∎

Ce résultat ne dit rien a priori sur la convergence de la suite (x_n). Cependant, si les deux suites extraites (x_{2n}) et (x_{2n+1}) convergent vers la même limite ℓ, alors la suite (x_n) elle même converge (voir l'exercice 115).

En particulier, si l'intervalle $I=[a,b]$ est fermé borné, et que $g=f^{(2)}$ est continue et n'a qu'un point fixe ℓ sur I, alors la suite (x_n) converge vers ℓ.
Exemple. Soit $(x_n)_{n\geq 0}$ la suite récurrente définie par son premier terme $x_0>0$ et la relation de récurrence $x_{n+1}:=1+1/x_n$ pour $n\geq 0$. Ici, la fonction f est $f(x)=1+1/x$. On peut choisir $I=]0,+\infty[$, intervalle (ouvert) sur lequel la fonction f est continue et décroissante.

Les points fixes de f sont donnés par l'équation $x=1+1/x$. Il n'y en a qu'un seul dans l'intervalle I, qui vaut $\ell=\frac{1+\sqrt{5}}{2}$. On l'appelle le nombre d'or.

La fonction $g=f^{(2)}=f\circ f$ est une fonction continue croissante sur $]0,+\infty[$. On peut faire le calcul exact, et on obtient

$$g(x)=f(f(x))=\frac{2x+1}{x+1},\quad x>0.$$

Dans ce cas, il nous faut regarder les sous-suites $(y_n)=(x_{2n})$ et $(z_n)=(x_{2n+1})$. On remarque que la fonction g prend ses valeurs dans l'intervalle $[0,2]$. De plus, g est croissante (nous le savions par la proposition 6.4.2) et est continue sur $[0,\infty[$ (intervalle fermé). On voit que y_n et z_n sont bornées, puisque $y_{n+1}=g(y_n)$ et $z_{n+1}=g(z_n)$. Ce sont donc des suites monotones bornées. Par conséquent, elles possèdent des limites finies ℓ_1 et ℓ_2 respectivement, qui sont dans $[0,+\infty[$, et qui doivent être solution de l'équation $g(x)=x$.

Mais le seul point fixe de g dans $[0,\infty[$ est ℓ. Les deux sous-suites (y_n) et (z_n) convergent donc vers la même limite ℓ, et donc la suite (x_n) converge vers $\ell=\frac{1+\sqrt{5}}{2}$.

Dans cet exemple, on peut déterminer sans calcul quand les suites (x_{2n}) et (x_{2n+1}) sont croissantes ou décroissantes, selon les valeurs de x_0. Si $x_0<\ell$, alors puisque f est décroissante, on a $x_1>\ell$. Mais puisque x_{2n} converge vers ℓ, et que sa valeur initiale x_0 est inférieure à ℓ, alors nécessairement elle est croissante. De même, x_{2n+1} est décroissante. Le résultat est inversé si $x_0>\ell$. Dans ce cas, la limite ℓ est prise en sandwich entre les termes pairs et les termes impairs de la suite.

6.5 Fonctions f contractantes

Nous présentons une méthode qui peut être appliquée dans des situations très générales. Supposons que l'on a une suite (x_n) avec $x_0\in I$ et pour tout n, $x_{n+1}=f(x_n)$ où $f(I)\subset I$. Supposons de plus que $\ell\in I$ est un point fixe de f et qu'il existe une constante $k\in[0,1[$ telle que pour tout $x\in I$ on ait

$$|f(x)-\ell|\leq k|x-\ell|. \tag{6.1}$$

Alors pour tout n, $|x_{n+1}-\ell|=|f(x_n)-\ell|\leq k|x_n-\ell|$. Ceci signifie que chaque terme de la suite est plus proche du point fixe que le précédent. On en déduit par récurrence que pour tout n,

$$|x_n-\ell|\leq k^n|x_0-\ell|.$$

Comme $k\in[0,1[$, on sait que $\lim_n k^n=0$ et donc on en déduit que (x_n) converge vers ℓ. Observons qu'il est fondammental de supposer $k<1$ pour conclure et que $k=1$ dans (6.1) ne permettrait pas de montrer la convergence. Cette approche est mise en œuvre à l'exercice 142.

Il n'est pas toujours facile de montrer directement (6.1), notamment lorsque l'on ne connaît pas la valeur de ℓ. On peut dans ce cas chercher à vérifier une hypothèse plus forte mais plus accessible :

Définition. *Une fonction f définie sur un ensemble J pour laquelle il existe une constante $k<1$ telle que $|f(x)-f(y)|\leq k|x-y|$ pour tout couple de points (x,y) dans J^2 est appelée une application* **contractante**. *On dit alors que f est contractante de constante k.*

Dans ce cas, si ℓ est un point fixe de f, on a pour tout $x\in J$,

$$|f(x)-\ell|=|f(x)-f(\ell)|\leq k|x-\ell|$$

et l'on peut appliquer la méthode précécente (sur un sous ensemble stable par f).

La propriété de contraction peut se vérifier en étudiant la dérivé de f. En effet si f est définie et dérivable sur un intervalle I et si $|f'|$ est majorée par $k < 1$ sur I alors f est contractante de constante k. C'est une conséquence du théorème des accroissements finis qui assure que pour $x \neq y \in I$, il existe θ compris entre x et y tel que $f(x) - f(y) = f'(\theta)(x-y)$. Sachant $|f'(\theta)| \leq k$ on obtient $|f(x) - f(y)| \leq k|x-y|$.

Remarque. *Supposons que f définie sur un intervalle I est dérivable et que sa dérivée est continue. Si, de plus, ℓ est une solution de $f(\ell) = \ell$ et que $|f'(\ell)| < 1$, alors il existe un intervalle autour de ℓ, de la forme $I = [\ell-\varepsilon, \ell+\varepsilon]$, sur lequel cette fonction dérivée $|f'|$ est majorée par une constante $k < 1$. Alors, dès que la suite (x_n) arrive dans l'intervalle I, la suite va converger vers ℓ. En effet f est contractante de constante $k < 1$ sur I et de plus I est stable par f (car $|x-\ell| \leq \varepsilon \Longrightarrow |f(x) - \ell| = |f(x) - f(\ell)| \leq k|x-\ell| \leq \varepsilon$).*

Par contre, ce n'est plus le cas si $|f'(\ell)| > 1$. Dans ce cas, on ne pourra jamais avoir la convergence de (x_n) vers ℓ, à moins qu'il n'existe un entier n_0 pour lequel $x_{n_0} = \ell$ (et la suite reste constante et égale à ℓ à partir de $n = n_0$). En effet $x_{n+1} - \ell = f(x_n) - f(\ell) = f'(\ell)(x_n - \ell) + o(x_n - \ell)$.

On voit donc qu'en pratique, ce qui importe est la position de $|f'(\ell)|$ par rapport à 1. Si $|f'(\ell)| < 1$ on dit que ℓ est un **point fixe stable** *ou attractif et si $|f'(\ell)| > 1$ on dit qu'il est un* **point fixe instable** *ou répulsif (et on se garde bien de dire quoi que ce soit lorsque $|f'(\ell)| = 1$).*

6.6 Exemple d'application : l'approximation de $\sqrt{2}$

Nous avons déjà décrit la méthode de dichotomie pour construire deux suites de nombres rationels approchant par défaut et par excès le nombre irrationnel $\sqrt{2}$. Mais l'algorithme que nous avions donné n'est pas très rapide. Nous allons présenter une méthode plus performante : on considère la suite (x_n) définie par $x_0 = 2$ et $x_{n+1} = f(x_n)$ où

$$f(x) := \frac{1}{2}\left(x + \frac{2}{x}\right).$$

On vérifie facilement que $\sqrt{2}$ et $-\sqrt{2}$ sont les seules solutions de l'équation $f(x) = x$ et que f est croissante et continue sur $I := [\sqrt{2}, +\infty[$. Donc si $x \geq \sqrt{2}$, alors $f(x) \geq f(\sqrt{2}) = \sqrt{2}$ et l'on a bien que I est stable par f. Sur I, f est croissante et pour $x > \sqrt{2}$, on a $f(x) - x = \frac{1}{2}(\frac{2}{x} - x) < 0$. Donc, si $x_0 > \sqrt{2}$ la suite (x_n) est décroissante, minorée (par $\sqrt{2}$) et converge donc vers $\sqrt{2}$ (unique point fixe de f sur I).

Etudions maintenant la vitesse de convergence de (x_n) vers $\sqrt{2}$. Elle va être rapide car $f'(\sqrt{2}) = 0$ (donc la dérivé de f au voisinage de $\sqrt{2}$ est très petite, voir le paragraphe précédent). Cependant nous allons estimer la vitesse directement sur la fonction f qui vérifie pour $x \geq \sqrt{2}$,

$$0 \leq f(x) - \sqrt{2} = \frac{1}{2x}\big(x - \sqrt{2}\big)^2 \leq \frac{1}{2}\big(x - \sqrt{2}\big)^2.$$

On en déduit, que $x_{n+1} - \sqrt{2} \leq \frac{1}{2}\big(x_n - \sqrt{2}\big)^2$. Ceci permet de montrer par récurrence que

$$x_n - \sqrt{2} \leq 2^{1-2^n}.$$

En effet, c'est vrai pour $n = 0$ car $x_0 - \sqrt{2} = 2 - \sqrt{2} < 1$. De plus, si cette propriété est vraie à l'ordre n, on a à l'ordre $n+1$

$$x_{n+1} - \sqrt{2} \leq \frac{1}{2}\big(x_n - \sqrt{2}\big)^2 \leq \frac{1}{2}2^{2(1-2^n)} = 2^{1-2^{n+1}}.$$

Le procédé par dichotomie nous donnait au bout de n étapes une erreur de l'ordre de 2^{-n}, tandis que celui-ci nous donne au bout de n étapes une erreur de l'ordre de 2^{-2^n}, ce qui est considérablement plus petit. Si on calcule explicitement les valeurs de x_n en partant de $x_0 = 2$, on obtient $x_1 = 1,5$, $x_2 = 1,466666$, $x_3 = 1,414215686$, $x_4 = 1,414213562$ alors que $\sqrt{2} \simeq 1,414213562$. En 4 étapes, nous avons une écriture décimale approchée à 9 décimales exactes ! Avec le procédé de dichotomie, il nous avait fallu 10 étapes pour obtenir 3 décimales.

La méthode utilisée ici porte le nom de "méthode de Héron". C'est un cas particulier de la "méthode de Newton" qui permet de déterminer une approximation de la solution de certaines equations du type $F(x) = 0$. La méthode de Newton consiste à écrire une suite récurrente qui remplace x_n par l'intersection de la tangente à la courbe $y = F(x)$ au point x_n avec l'axe $\{y = 0\}$. Dans ce cas précis, avec $F(x) = x^2 - 2$, l'équation de la tangente au point x_n est $Y - (x_n^2 - 2) = 2x_n(X - x_n)$ et l'intersection avec l'axe des X est donné en faisant $Y = 0$, c'est à dire en remplaçant x_n par la solution X de l'équation obtenue

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

Exercice 136. Donnez une méthode similaire pour approcher $\sqrt{a}$.

6.7 Complément sur les contractions

Par rapport à ce que nous avons vu, le résultat suivant montre en plus que, sur un intervalle stable fermé, un application contractante admet un point fixe. Il s'agît d'un premier exemple d'un principe trés général en mathématiques, qui est utilisé pour résoudre de nombreuses équations et notamment des équations différentielles.

Théorème 6.7.1. *Soit $f : I \mapsto I$ une fonction définie sur un intervalle fermé $I \subset \mathbb{R}$ (de la forme $[a, b]$, $[a, \infty[$, $]-\infty, b]$ ou bien $]-\infty, +\infty[$). Supposons qu'il existe un nombre réel $0 \leq k < 1$ tel que pour tout $x \in I$ et $y \in I$:*

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

Alors pour tout $x_0 \in I$ la suite de premier terme x_0 définie par la relation de récurrence $x_{n+1} = f(x_n)$ pour $n \geq 0$ converge vers un nombre réel $\ell \in I$ vérifiant l'équation $f(\ell) = \ell$. De plus ℓ est l'unique point fixe de f dans I.

Démonstration. — Remarquons tout d'abord que l'hypothèse faite sur f montre qu'elle est continue sur I (nous laissons cette vérification aux lecteurs). Nous étudions d'abord le cas simple où nous savons qu'il existe dans I un point fixe ℓ pour f, c'est à dire une solution de $f(\ell) = \ell$. Dans ce cas, nous avons vu

$$|x_{n+1} - \ell| = |f(x_n) - f(\ell)| \leq k|x_n - \ell|.$$

En itérant cete inégalité, on en déduit que pour tout $n \geq 0$, $|x_n - \ell| \leq k^n|x_0 - \ell|$. Puisque $k < 1$, nous voyons donc que $\lim_n x_n = \ell$. Ceci montre en plus que le point fixe est unique, puisqu'il est la limite de la suite (x_n).

En fait, la seule hypothèse que f est contractante suffit à assurer l'existence d'un point fixe ℓ. Pour le voir, nous allons utiliser le critère de Cauchy (théorème 5.6.2, page 82). En effet, pour $n \in \mathbb{N}^*$, écrivons $|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})|$. D'après l'hypothèse sur f, on a $|x_{n+1} - x_n| \leq k|x_n - x_{n-1}|$. En itérant cette inégalité, on en déduit par récurrence que pour tout $n \in \mathbb{N}^*$,

$$|x_{n+1} - x_n| \leq k^n|x_1 - x_0| \tag{6.2}$$

On va montrer qu'alors (x_n) est une suite de Cauchy. Ce qu'il nous faut montrer, c'est que $s_n := \sup_{p \geq n} |x_p - x_n|$ converge vers 0. Il suffit donc de majorer (s_n) par une suite qui converge vers 0.

Soit donc n et $p \geq n$ que nous écrivons $p = n + q$. On a

$$|x_{n+q} - x_n| \leq |x_{n+1} - x_n| + |x_{n+2} - x_{n+1}| + \ldots + |x_{n+q} - x_{n+q-1}|.$$

Par la majoration (6.2), on obtient

$$|x_{n+q} - x_n| \leq |x_1 - x_0|\big(k^n + k^{n+1} + \ldots + k^{n+q-1}\big) = |x_1 - x_0|k^n\big(1 + k + \ldots + k^{q-1}\big).$$

Or,

$$1 + k + \ldots + k^{q-1} = \frac{1 - k^q}{1 - k} \leq \frac{1}{1 - k},$$

et on obtient

$$|x_{n+q} - x_n| \leq k^n \frac{|x_1 - x_0|}{1 - k}.$$

On en déduit que $s_n = \sup_{q \geq 1} |x_{n+q} - x_n|$ est majorée par $k^n \frac{|x_1 - x_0|}{1 - k}$. C'est une suite qui converge vers 0 et la suite est donc de Cauchy. Grâce au théorème 5.6.2 page 82, nous voyons qu'elle converge et puisque f est continue, elle converge vers une limite ℓ telle que $f(\ell) = \ell$.

∎

Exercices

Exercice 137. Calculer les limites des suites suivantes, définies pour $n \in \mathbb{N}^*$:

$$u_n := n \sin\left(\frac{\pi}{2n}\right), \qquad v_n := n \ln\left(1 + \frac{1}{\sqrt{n}}\right).$$

On utilisera que $\lim_{x \to 0,\, x \neq 0} \frac{\sin x}{x} = 1$ et $\lim_{x \to 0,\, x \neq 0} \frac{\ln(1+x)}{x} = 1$.

Exercice 138. Soient $a < b$ des réels. Soit $f : [a, b] \to [a, b]$ une fonction continue. Montrer que f admet un point fixe.

Exercice 139. Soient $a < b$ des réels. Soit $g : [a, b] \to \mathbb{R}$ une fonction continue. On suppose que $[a, b] \subset g\big([a, b]\big)$. Montrer que g admet un point fixe.

Exercice 140. On considère la suite définie par son premier terme x_0 et par la relation de récurrence :

$$x_{n+1} = \frac{x_n}{1 + n x_n^2},\ n \geq 0.$$

1) Montrez que la suite est de signe constant.
2) On suppose que $x_0 > 0$.
Montrez que la suite $(x_n)_{n \geq 0}$ est monotone et qu'elle converge vers une limite que l'on calculera.

Exercice 141. (*) Soit $(u_n)_{n \geq 0}$ la suite de nombres réels définie par $u_0 = 1$ et $u_{n+1} = \sqrt{u_n^2 + 2^{-n}}$ pour tout $n \in \mathbb{N}$.
1) Démontrer que pour tout $n \geq 1$,

$$1 \leq u_n < u_{n+1} < u_n + \frac{1}{2^{n+1}}.$$

En déduire que pour entier tout $n \geq 1$, $u_n < 2$.
2) En déduire que la suite $(u_n)_{n \geq 0}$ converge et encadrer sa limite.
3) Soit $v_n := u_n^2$. Calculer explicitement v_n et en déduire la limite de (u_n).

Exercice 142. On fixe un nombre réel $a \geq 0$.
1) Démontrer que l'on définit une suite de nombres réels positifs $(u_n)_{n \geq 0}$ en posant :

$$u_0 = a, \quad u_{n+1} = \frac{u_n}{1 + u_n} + 1, \quad \text{si } n \geq 0.$$

2) Montrer que la suite $(u_n)_{n \geq 0}$ est monotone.
3) Démontrer que la suite $(u_n)_{n \geq 0}$ est convergente et déterminer sa limite ℓ.
4) Démontrer que $|u_{n+1} - \ell| \leq \frac{1}{1+\ell}\ |u_n - \ell|$.
En déduire une majoration de $|u_{n+1} - \ell|$ en fonction de n et de a.

Exercice 143. Soit l'application $f : \mathbb{R}^+ \to \mathbb{R}^+$ définie par $f(x) = 2 + \sqrt{x}$. Soit $a \geq 0$ et la suite $(u_n)_{n \in \mathbb{N}}$ définie par $u_0 = a$ et pour tout $n \geq 0$, $u_{n+1} = f(u_n)$.

1) Résoudre dans $\mathbb{R}^+$ l'équation $f(x) = x$ et l'inéquation $f(x) \geq x$.

2) Tracer le graphe de f. Représenter sur le graphe les valeurs de la suite pour $a = 1$ et $a = 9$.

3) Montrer que $(u_n)_{n \in \mathbb{N}}$ est monotone. Préciser son sens de variation en fonction de a.

4) Montrer que la suite converge vers une limite ℓ à préciser.

5) Montrer que $|u_{n+1} - \ell| \leq |u_n - \ell|/2$ et en déduire une majoration explicite de $|u_n - \ell|$.

Exercice 144. (*) Soit $\lambda \in [0,4]$ un paramètre. On considère les suites $(u_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence

$$u_{n+1} = \lambda u_n(1-u_n), \quad n \geq 0.$$

Ces suites sont un modèle simple de dynamique des populations. Le nombre λ représente un taux de croissance de la population. Le terme $1-u_n$ fait que si la taille de la population approche de 1 (vu comme la capacité maximale de l'environnement, dans une unité bien choisie), alors le taux de croissance diminue, faute de ressources.

1) Montrer que si $u_0 \in [0,1]$ alors pour tout n, $u_n \in [0,1]$.
2) Montrer que si $\lambda \in [0,1]$ et $u_0 \in [0,1]$, alors la suite converge vers 0 en décroissant.
3) Etudier les points fixes de $x \mapsto \lambda x(1-x)$. Pour quelles valeurs de λ sont-ils attractifs ?
4) Montrer que si $\lambda \in]1,2[$ et $u_0 \in]0,1[$ alors la suite converge vers $1-1/\lambda$.

Remarque : le comportement de cette suite a fait l'objet de nombreuses recherches récentes. Lorsque $\lambda > 3$, à part pour des valeurs initiales très particulières, la suite ne converge pas car les points fixes sont répulsifs. Si λ est assez proche de 3 la suite peut être "asymptotiquement périodique", mais pour λ proche de 4 elle devient chaotique.

Exercice 145. (*) Soit $a \in \mathbb{R}$ un paramètre réel tel que $a > 4$.

1) Démontrer qu'on peut définir une suite $(x_n)_{n\geq 0}$ par récurrence en posant $x_0 = a > 4$ et

$$x_{n+1} = \frac{x_n - 6}{x_n - 4}, \quad n \geq 0.$$

   Que se passe-t-il si $a \leq 4$?
2) Démontrer que l'application $x \longmapsto \frac{x-6}{x-4}$ possède deux points fixes α, β avec $\alpha < \beta$ que l'on déterminera.
3) On suppose que $a \notin \{\alpha, \beta\}$. On pose :

$$y_n := \frac{x_n - \alpha}{x_n - \beta}, \quad n \geq 0.$$

   Calculer y_{n+1} en fonction de y_n et en déduire la limite de $(x_n)_{n\geq 0}$.

Exercice 146. (*) Soit $(x_n)_{n\geq 0}$ la suite définie par son premier terme $x_0 \neq -1$ et la relation de récurrence :

$$x_{n+1} = \frac{3x_n - 1}{x_n + 1}, \quad n \geq 0.$$

1) Démontrer que l'application $x \longmapsto \frac{3x-1}{x+1}$ possède un seul point fixe $q \in \mathbb{R}$.
2) On suppose que $x_0 \neq q$ et on pose $y_n := \frac{1}{x_n - q}$ pour $n \geq 0$. Calculer y_{n+1} en fonction de y_n et en déduire la limite de la suite $(y_n)_{n\geq 0}$.
3) Calculer la limite de la suite $(x_n)_{n\geq 0}$.

Exercice 147. (*) Vous empruntez une somme D_0 à la banque, au taux d'intérêt mensuel τ (par exemple $\tau = 0,01$ pour un taux mensuel de un pour cent). A la fin de chaque mois, les intérêts sont calculés (et s'ajoutent au montant de la dette) et vous remboursez une somme fixe R. On note D_n le montant de la dette au début du n-ième mois après la réalisation du prêt.

1) Montrez que la suite (D_n) vérifie la relation

$$D_{n+1} = (1+\tau)D_n - R, \quad n \geq 0.$$

2) Donnez une expression explicite de D_n en fonction de D_0, R, τ et n. On pourra s'inspirer de l'exemple suivant la Proposition 6.3.1.

3) Vous souhaitez rembourser le prêt en N mensualités, ce qui ce traduit par $D_N = 0$. Calculez le montant R des remboursements mensuels en fonction de D_0, N et τ.

Exercice 148. (*) La suite de Fibonacci est définie par ses deux premiers termes $u_0 = 1, u_1 = 1$ et la relation de récurrence linéaire d'ordre 2 suivante :

$$(F) \qquad u_{n+2} = u_{n+1} + u_n, \quad n \geq 0.$$

Cette formule permet de calculer les termes de la suite, de proche en proche. L'objet de cet exercice est d'établir une formule explicite donnant u_n directement en fonction de n.
1) Soit $q \in \mathbb{R}^*$. Démontrer que la suite géométrique $(q^n)_{n\in\mathbb{N}}$ vérifie la condition (F) si et seulement si le nombre q satisfait l'équation $q^2 - q - 1 = 0$.
2) Trouver les deux valeurs q_1 et q_2 du paramètre $q \in \mathbb{R}$ tel que la suite géométriques $(q^n)_{n\in\mathbb{N}}$ vérifie la relation (F).
3) Trouver deux constantes numériques α, β tels que la suite de Fibonacci s'écrive : $u_n = \alpha q_1^n + \beta q_2^n$ pour tout $n \in \mathbb{N}$.
4) En déduire la formule suivante donnant les nombres de Fibonacci :

$$u_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right), \forall n \geq 0.$$

Observons que par définition les nombres de Fibonacci sont des entiers !
5) Etudier la convergence de $\left(\frac{u_{n+1}}{u_n}\right)$.
La suite de Fibonacci est une suite strictement croissante d'entiers naturels, qui tend donc vers $+\infty$. Elle a des propriétés importantes et intervient dans plusieurs domaines des sciences.

Exercice 149. (*) Soit $(u_n)_{n\in\mathbb{N}}$ une suite de nombres réels telle que pour tout $n \in \mathbb{N}$:

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

1) Démontrer que si la suite $(u_n)_{n\in\mathbb{N}}$ converge, sa limite est nécessairement égale à 1.
2) Démontrer que si $u_0 = 1$, la suite $(u_n)_{n\in\mathbb{N}}$ est constante.
3) Démontrer que pour tout $n \in \mathbb{N}$,

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

4) Supposons que $0 \leq u_0 < 1$. Démontrer que pour tout $n \in \mathbb{N}$, $u_n < 1$ et en déduire que la suite $(u_n)_{n\in\mathbb{N}}$ converge.
5) Supposons qu'il existe un entier $p \geq 1$ tel que $1 < u_p \leq p^2$. Démontrer que pour tout $n \geq p$ on a $1 < u_{n+1} \leq u_n$ et en déduire que la suite $(u_n)_{n\in\mathbb{N}}$ converge.
6) Supposons que $1 < u_0 < \sqrt[4]{7}$. Démontrer que $1 < u_2 \leq 4$ puis que la suite $(u_n)_{n\in\mathbb{N}}$ converge.
7) On suppose que $u_0 \geq 2\sqrt{2}$. Démontrer par récurrence que pour tout $n \in \mathbb{N}^*$, $u_n \geq 8n^2$.

Exercice 150. (*) Soient $a, b > 0$ deux nombres réels positifs. On définit par récurrence deux suites $(u_n)_{n\geq 0}$ et $(v_n)_{n\geq 0}$ en posant $u_0 = a, v_0 = b$ et pour tout $n \geq 0$:

$$u_{n+1} := \sqrt{u_n \cdot v_n}, \quad v_{n+1} := \frac{u_n + v_n}{2}.$$

(Le nombre réel u_{n+1} est appelé la *moyenne géométrique* des nombres réels u_n et v_n et le nombre reél v_{n+1} est leur *moyenne arithmétique*).
1) Calculer u_1 et v_1 et les comparer.
2) Démontrer que pour tout $n \geq 1$:

$$u_n \leq u_{n+1} \leq v_{n+1} \leq v_n.$$

3) Démontrer que (u_n) et (v_n) convergent, puis que leurs limites sont égales.
4) En déduire que $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 0}$ sont des suites adjacentes et que leur limite commune ℓ vérifie :

$$\sqrt{a \cdot b} \ \leq \ \ell \ \leq \ \frac{a+b}{2}.$$

Le nombre réel ℓ est appelé la *moyenne arithmético-géométrique* de a et b.

Exercice 151. (*) Soient $a, b > 0$ deux nombres réels positifs. On définit par récurrence deux suites $(x_n)_{n\geq 0}$ et $(y_n)_{n\geq 0}$ en posant $x_0 = a, y_0 = b$ et pour tout $n \geq 0$:

$$\frac{1}{x_{n+1}} := \frac{1}{2}\left(\frac{1}{x_n} + \frac{1}{y_n}\right), \quad y_{n+1} := \frac{x_n + y_n}{2}.$$

(Le nombre réel x_{n+1} est appelé la *moyenne harmonique* de x_n et y_n). On suppose que $a < b$.
1) Démontrer que pour tout $n \in \mathbb{N}$, $y_{n+1} - x_{n+1} = \dfrac{(x_n - y_n)^2}{2(x_n + y_n)}$.
2) Démontrer que pour tout $n \in \mathbb{N}$, $0 < y_{n+1} - x_{n+1} \leq \dfrac{1}{2}(x_n - y_n)$.
3) Démontrer que la suite $(x_n)_{n\in\mathbb{N}}$ est croissante et que la suite $(y_n)_{n\in\mathbb{N}}$ est décroissante.
4) Démontrer que les deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ sont convergentes et ont la même limite.
5) Démontrer que la suite $(x_n \cdot y_n)_{n\in\mathbb{N}}$ est constante. En déduire la limite des suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$.