Mathématiques

2021-2022

Lycée Clemenceau Nantes MP*

Maryna VIAZOVSKA – Empilements de sphères (conjecture de Kepler) en dimensions 8 et 24

François SAUVAGEOT

CHAPITRE 1

Structures mères

Véritable génie des mathématiques, né en 1903 et mort en 1987, Andrei Kolmogorov a traversé le vingtième siècle en y laissant des contributions mathématiques considérables. On retrouve son nom aux fondements de la théorie des probabilités, des systèmes dynamiques, de la théorie de l'information. En particulier il a montré en 1941 que les propriétés statistiques des mouvements de fluides « turbulents » (ronds de fumée, torrents de montagne etc.) obéissent à des lois universelles, de type autosimilaire : le « spectre de Kolmogorov », observé expérimentalement dans beaucoup d'écoulements turbulents, indique ainsi la loi selon laquelle l'énergie d'un fluide est transmise vers des échelles de plus en plus petites. Cette loi, dite du K41 (pour « Kolmogorov 1941 »), est aux fondements de l'étude de la turbulence hydrodynamique qui est très étudiée aujourd'hui, car encore largement incomprise.

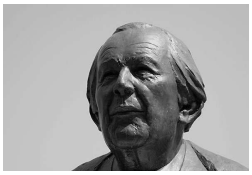
Andrei Kolmogorov fut un mathématicien exceptionnel, également passionné d'histoire et de littérature, féru d'exploits sportifs, et engagé jusqu'à la fin de sa vie dans la pédagogie des sciences. Il fut également très proche du pouvoir soviétique, tout comme son ami Pavel Aleksandrov, dénonçant notamment son patron de thèse Nikolai Luzin dans un procès Stalinien. Ce dernier ne fut réhabilité qu'en 2012.

L'objectif de ce chapitre est de revoir quelques éléments du programme de MPSI et de préciser le vocabulaire mathématique. Il ne s'agit nullement de refaire le cours de première année, mais simplement de s'assurer que nous allons parler la même langue ! On introduira au passage quelques éléments nouveaux, utiles notamment en théorie des probabilités.

Programme

- Rudiments de logique : quantificateurs, implication, contraposition, équivalence. Modes de raisonnement : par récurrence (faible et forte), par contraposition, par l'absurde, par analyse-synthèse.
- Ensembles, appartenance, inclusion. Ensemble $\mathscr{P}(E)$ des parties de E. Opérations sur les parties : intersection, réunion, différence, complémentaire. Produit (cartésien) de deux ensembles.
- Application, graphe d'une application, familles d'éléments.
- Fonction indicatrice. Restriction et prolongement, image directe et réciproque. Composition.
- Applications injectives, surjectives, bijectives. Application réciproque d'une bijection. Composée de deux injections, de deux surjections, de deux bijections. Réciproque de la composée.
- Loi de composition interne, associativité, commutativité, élément neutre, inversibilité, distributivité. Partie stable.
- Groupes, anneaux, corps. Calcul dans un anneau.
- Relation binaire, relation d'équivalence, classe d'équivalence, relation d'ordre, ordre total, ordre partiel.
- Cardinal d'un ensemble fini, listes et combinaisons.
- Ensemble dénombrable, ensemble fini ou dénombrable, produit cartésien fini d'ensembles dénombrables, réunion finie ou dénombrable d'ensembles finis ou dénombrables. Exemples de $\mathbf{N}^2$, $\mathbf{Z}$, $\mathbf{Q}$ et $\mathbf{R}$.
- Tribu, événements. Espace probabilisable, probabilité, espace probabilisé. Continuité croissante, continuité décroissante, sous-additivité. Événements négligeables, événements presque sûrs. Réunion finie ou dénombrable d'événements négligeables
- Probabilité conditionnelle, formule des probabilités composées, formule des probabilités totales, formules de Bayes. Couple d'événements indépendants. Famille quelconque d'événements mutuellement indépendants.
- Variables aléatoires discrètes. Loi P_X de la variable aléatoire X. Notations $(X \geqslant x)$, $(X \leqslant x)$, $(X < x)$, $(X > x)$ pour une variable aléatoire réelle X.
- Loi géométrique $\mathcal{G}(p)$. Interprétation comme rang du premier succès. Caractérisation comme loi sans mémoire : $\mathbf{P}(X > n+k \mid X > n) = \mathbf{P}(X > k)$.
- Loi de Poisson $\mathcal{P}(\lambda)$. Approximation de la loi binomiale par la loi de Poisson. Événements rares.

Introduction



Jean PIAGET

> La méthode axiomatique permet, lorsqu'on a affaire à des êtres mathématiques complexes, d'en dissocier les propriétés et de les regrouper autour d'un petit nombre de notions, c'est-à-dire [...] de les classer suivant les structures auxquelles elles appartiennent[...] ; pour définir une structure, on se donne une ou plusieurs relations, où interviennent ces éléments [...] ; on postule ensuite que la ou les relations données satisfont à certaines conditions (qu'on énumère) et qui sont les axiomes de la structure envisagée. Faire la théorie axiomatique d'une structure donnée, c'est déduire les conséquences logiques des axiomes de la structure, en s'interdisant toute autre hypothèse sur les éléments considérés (en particulier, toute hypothèse sur leur « nature » propre).
>
> – N. BOURBAKI

C'est le groupe de mathématiciens publiant sous le pseudonyme de N. BOURBAKI qui a développé pour la première fois la théorie des structures de manière explicite et rigoureuse dans ses *Éléments de mathématique* à partir des années 1930.

La notion de structure dérive de la méthode axiomatique adoptée par BOURBAKI. Cette axiomatique permet de mettre au jour une unité profonde entre diverses branches des mathématiques, considérées comme distinctes dans la classification traditionnelle des disciplines mathématiques (arithmétique, algèbre, analyse, géométrie) :

Nous croyons que l'évolution interne de la science mathématique a, malgré les apparences, resserré plus que jamais l'unité de ses diverses parties, et y a créé une sorte de noyau central plus cohérent qu'il n'a jamais été. L'essentiel de cette évolution a consisté en une systématisation des relations existant entre les diverses théories mathématiques, et se résume en une tendance qui est généralement connue sous le nom de « méthode axiomatique ».

BOURBAKI observe : « *Dans cette nouvelle conception, les structures mathématiques deviennent, à proprement parler, les seuls « objets » de la mathématique.* » et distingue principalement trois types de « structures-mères » : la structure algébrique, dont les relations sont des lois de composition, la structure d'ordre, et la structure topologique.

Ce terme est à l'origine de ce que l'on a appelé le structuralisme mathématique. Cette façon de penser a intéressé notamment les psychanalystes (Jacques LACAN), les anthropologues (Claude LEVI-STRAUSS), les psychologues (Jean PIAGET). Le structuralisme est né avec le cercle (linguistique) de Prague, en se fondant sur des travaux de linguistique (notamment de Ferdinand DE SAUSSURE).

Attention ! Ce principe d'exposition est en fait artificiel. BOURBAKI avoue trois inconvénients de cette théorie des structures : « elle est à la fois schématique, idéalisée et figée. » Schématique, car dans le détail il existe « d'inattendus retours en arrière », comme l'intervention des nombres réels pour fonder la topologie. Idéalisée, car « dans certaines théories (par exemple en théorie des nombres), il subsiste de très nombreux résultats isolés qu'on ne sait jusqu'ici classer ni relier de façon satisfaisante à des structures connues » et figée, car les structures ne sont pas « immuables », et peuvent se prêter à des inventions ou reformulations futures.

1 Logique

En sus de la théorie des ensembles, on a besoin de structurer le raisonnement. Ainsi on a besoin de la notion d'assertion (proposition, prédicat) et de connecteur logique.

Notation

On a les **connecteurs logiques** :

1. $\neg A$ est vrai si (et seulement si) A est faux.
2. $A \wedge B$ est vrai si A et B sont vrais.
3. $A \vee B$ est faux si A et B sont faux.
4. $A \Rightarrow B$ est vrai si A est faux ou si B est vrai.
5. $A \Leftrightarrow B$ si A et B sont simultanément vrais ou faux.

Une tautologie est un énoncé qui est vrai quelque soient les valeurs de vérité des assertions considérées. Autrement dit une tautologie est une simple reformulation.

♥

Les **tautologies** les plus importantes, pour le raisonnement, sont :

1. $A \equiv \neg\neg A$ (principe du **tiers exclus** ou **double négation**).
2. $\neg(A \Rightarrow B) \equiv (A \wedge \neg B)$.
3. $(A \Rightarrow B) \equiv (\neg B \Rightarrow \neg A)$ (**contraposition**).

Enfin on a les quantificateurs : $\forall$ et $\exists$. Quand un prédicat A dépend d'une variable x, l'assertion $\forall x\ A(x)$ signifie que A est vraie pour toutes les valeurs de x. Par contre l'assertion $\exists x\ A(x)$ signifie qu'il existe un x pour lequel $A(x)$ est vrai. La difficulté est parfois de rendre effective cette affirmation d'existence : peut-on réellement trouver le x dont on parle ? On y reviendra lors de l'axiome du choix. On prendra garde que, dans une assertion quantifiée, la variable quantifiée est muette. Par exemple, en fait, dans l'assertion $\forall x\ A(x)$, **la variable x n'existe pas** !

♥

Les tautologies mettant en jeu des **quantificateurs** qui sont les plus utiles sont :

1. $\neg(\exists x\ A(x)) \equiv \forall x\ \neg A(x)$.
2. $\neg(\forall x\ A(x)) \equiv \exists x\ \neg A(x)$.
3. $\neg(\forall x\ (P(x) \Rightarrow Q(x))) \equiv \exists x\ (P(x) \wedge \neg Q(x))$.

2 Théorie des ensembles

En mathématiques, une structure désigne une théorie « plus forte » que la théorie des ensembles, c'est-à-dire une théorie qui en contient tous les axiomes, signes et règles. C'est donc une théorie « fondée » sur la théorie des ensembles, mais contenant également des contraintes supplémentaires, qui lui sont propres, et qui permettent également de définir de nouvelles structures qu'elle inclut.

La théorie des ensembles utilise notamment les symboles $=$ et $\in$. Le premier signifie que deux objets sont identiques ($a = b$) et le second qu'un objet appartient à un

ensemble ($x \in A$). On prendra garde que toutes les propriétés ne définissent pas un ensemble, comme le montre le célèbre paradoxe de Bertrand Russel (1872–1970).

Pour aller plus loin

Paradoxe de la théorie des ensembles. – Russel

La notion toute intuitive d'ensemble est en fait une notion particulièrement difficile à axiomatiser. En effet si l'idée de « collection » d'objets est a priori satisfaisante, elle conduit à un paradoxe dans la théorie : la collection de tous les ensembles n'est pas un ensemble. Supposons le contraire, et notons $\mathcal{E}$ cet ensemble de tous les ensembles. Un tel ensemble a alors une particularité majeure : ses éléments sont aussi des parties de cet ensemble. On peut alors définir $\mathcal{A} = \{x \in \mathcal{E} \,|\, x \notin x\}$. Supposons un instant $\mathcal{A} \in \mathcal{A}$. Alors puisque $\mathcal{A}$ ne satisfait pas à la définition de $\mathcal{A}$, on a $\mathcal{A} \notin \mathcal{A}$, ce qui est absurde. Mais alors $\mathcal{A} \notin \mathcal{A}$ ce qui conduit au même type d'absurdité !

Remarque 1 - 1

Pour autant il existe des objets qui se contiennent eux-mêmes dans un sens strict. On peut penser aux objets fractals par exemple. Un autre exemple est donné par les entiers naturels. Dans le monde anglo-saxons les entiers se comptent à partir de 1, dans le monde francophone c'est à partir de 0. Pour autant les deux ensembles sont en bijection et l'un contient strictement l'autre.

Les tautologies permettant de travailler avec les ensembles, et donc de démontrer des assertions, sont les suivantes :

♥

1. $A \subset B \equiv \forall x \; (x \in A \Rightarrow x \in B)$.
2. $A = B \equiv (A \subset B \wedge B \subset A)$ ou encore $A = B \equiv \forall x \; (x \in A \Leftrightarrow x \in B)$.

Définition 1 - 1

Quelques ensembles :

1. Une **paire** est un ensemble $E = \{a, b\}$. On a $x \in A \Leftrightarrow (x = a \vee x = b)$. On notera qu'il peut arriver que a soit égal à b et que E n'ait donc qu'un seul élément.
2. L'écriture $\{x \in E \,|\, A(x)\}$ signifie $\{x \,|\, x \in E \wedge A(x)\}$ (définition en **compréhension**).
3. L'**ensemble vide** peut être défini par $\emptyset = \{x \in E \,|\, x \neq x\}$. Il est en fait indépendant de E.
4. L'**ensemble des parties** d'un ensemble est noté $\mathscr{P}(E)$. On a donc $A \in \mathscr{P}(E) \equiv A \subset E$.

Définition 1 - 2

Quelques opérations sur les ensembles :

1. La **différence** de deux ensembles est définie par $A \setminus B = \{x \in A \,|\, x \notin B\}$.
2. Le **complémentaire** dans E d'un sous-ensemble A de E est $E \setminus A$. On le note aussi $\complement_E^A$ ou encore $\overline{A}$ quand E est implicite.
3. La **réunion** est définie par $A \cup B = \{x \,|\, x \in A \vee x \in B\}$.
4. L'**intersection** est définie par $A \cap B = \{x \,|\, x \in A \wedge x \in B\}$.

Définition 1 - 3

Plus généralement

1. Une **réunion d'une famille** d'ensembles X appartenant eux-mêmes à un ensemble E (qui est donc alors un ensemble d'ensembles) se définit par
$$\bigcup_{X\in E} X = \{x \,|\, \exists X \in E\,, x \in X\}\ .$$
2. Une **intersection d'une famille** d'ensembles X appartenant eux-mêmes à un ensemble E (qui est donc encore un ensemble d'ensembles) se définit par
$$\bigcap_{X\in E} X = \{x \,|\, \forall X \in E\,, x \in X\}\ .$$

On a les propriétés élémentaires :

Propriétés 1 - 1

1. Si $A' \subset A$ et $B' \subset B$, alors $A' \cap B' \subset A \cap B$ et $A' \cup B' \subset A \cup B$.
2. Si $A \subset E$ et $B \subset E$, alors $E \setminus (A \cup B) = (E \setminus A) \cap (E \setminus B)$ et $E \setminus (A \cap B) = (E \setminus A) \cup (E \setminus B)$.

Propriétés 1 - 2

Lois de de Morgan

1. $A \cap (B \cap C) = (A \cap B) \cap (A \cap C)$
2. $A \cup (B \cup C) = (A \cup B) \cup (A \cup C)$
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Augustus de Morgan 1806-1871.

3 Fonctions

Comme souvent en mathématiques, ce sont plus les transformations qui sont intéressantes et pertinentes que les objets eux-mêmes. Au niveau des ensembles, une transformation est une fonction. Elle ne diffère pas grandement d'un ensemble, mais c'est surtout la façon d'y penser qui diffère. La plupart des notions de cette section sont à la limite du programme et seules celles qui sont mises en exergues sont à maîtriser.

Pour commencer, il faut définir le **couple**. Un couple est une notion compliquée, même si elle est d'apparence simple. Techniquement l'écriture (a, b) est un raccourci pour $\{\{a\}, \{a, b\}\}$. On a donc $(a, b) = (a', b') \Leftrightarrow (a = a' \wedge b = b')$. On note également, si $c = (a, b)$, $a = pr_1(c)$ et $b = pr_2(c)$. Ce sont les première et seconde projections du couple c.

On définit alors le **produit cartésien** de deux ensembles :

$$A \times B = \{x \,|\, \exists a \in A\,, \exists b \in B\,, x = (a, b)\}\ .$$

On peut ainsi définir $A \times B \times C = (A \times B) \times C$ etc. On remarque au passage $A' \times B' \subset A \times B \Leftrightarrow (A' \subset A \wedge B' \subset B)$ et $A \times B = \emptyset \Leftrightarrow (A = \emptyset \vee B = \emptyset)$.

Un **graphe** est un ensemble de couples : $\Gamma \subset A \times B$. On définit la première projection de Γ comme l'ensemble $pr_1(\Gamma) = \{x \mid \exists y\,(x,y) \in \Gamma\}$. Et de même pour la seconde projection. On a donc $\Gamma \subset pr_1(\Gamma) \times pr_2(\Gamma)$.

Une **correspondance** est la donnée de deux ensembles et d'un graphe inclus dans leur produit cartésien : (Γ, E, F) avec $\Gamma \subset E \times F$. On peut l'interpréter comme une relation entre des éléments de E et de F n'ayant aucune contrainte ni d'existence ni d'unicité. Par exemple $\Gamma = \left\{(x,y) \in \mathbf{R}^2 \;\middle|\; x^2+y^2=1\right\}$ est un graphe et $(\Gamma, \mathbf{R}, \mathbf{R})$ est une correspondance.

Un **graphe fonctionnel** est un graphe dans lequel tout élément a au plus une image. Il n'est pas nécessaire néanmoins que tous les éléments aient effectivement une image. Autrement dit Γ est un graphe vérifiant $\forall x\ ((x,y) \in \Gamma \wedge (x,y') \in \Gamma) \Rightarrow y = y'$. On peut interpréter cette condition comme l'injectivité de pr_1.

Au sens strict, une fonction est une correspondance associée à un graphe fonctionnel. C'est donc un triplet (Γ, E, F) avec $\Gamma \subset E \times F$ et Γ graphe fonctionnel. L'ensemble E est appelé ensemble **source** ou ensemble de départ, tandis que l'ensemble F est appelé ensemble **image** ou ensemble d'arrivée.

Définition 1 - 4

— Une **application** est une fonction, au sens strict précédent, telle que tous les éléments de E ont effectivement une image. Autrement dit $pr_1(\Gamma) = E$, i.e. pr_1 est surjective et elle est donc bijective.

— On note $f : E \to F$ une application. Pour x dans E, on note alors $f(x)$ l'unique élément de F tel que $(x, f(x)) \in \Gamma$.

— On dit que $f(x)$ est l'**image** de x par l'application f et x est l'**antécédent** de $f(x)$ par f.

Remarques 1 - 2

On peut aussi écrire une application sous la forme $x \mapsto f(x)$, et alors la variable x est muette et les ensembles E et F sont implicites.

Dans la suite **on confondra les notions d'application et de fonction**, conformément au programme.

L'égalité des applications est l'égalité des triplets. L'ensemble des applications de E dans F est noté $\mathscr{F}(E,F)$ est peut-être vu comme un sous-ensemble de $\mathscr{P}(E \times F) \times \mathscr{P}(E) \times \mathscr{P}(F)$. L'ensemble des graphes d'applications de E dans F est noté F^E. Par abus de notation F^E et $\mathscr{F}(E,F)$ peuvent être utilisés l'un pour l'autre. Enfin $\mathscr{F}(E,E)$ est noté $\mathscr{F}(E)$.

On a les notions habituelles sur les applications :

Définition 1 - 5

Application identique (Id_E, E, E) est définie par $\mathrm{Id}_E(x) = x$ pour $x \in E$.

Composition si $f : E \to F$ et $g : F \to G$, on définit $g \circ f : E \to G$ par $g \circ f(x) = g(f(x))$. La composition est associative. Mais attention ! dans le cas des fonctions, elle n'est pas toujours définie.

Injection f est injective si pr_2 l'est, i.e. $\forall (x,x') \in E \times E,\ f(x) = f(x') \Rightarrow x = x'$.

Surjection f est surjective si pr_2 l'est, i.e. $\forall y \in F,\ \exists x \in E,\ y = f(x)$.

Bijection f est bijective si elle est injective et surjective, tout comme pr_2.

Coïncidence $f : E \to F$ et $g : G \to H$ coïncident sur A si $A \subset E \cap G$ et $\forall x \in A$, $f(x) = g(x)$.

Équipotence Deux ensembles E et F sont équipotents s'il existe une bijection de l'un dans l'autre.

et les résultats importants :

Propriétés 1 - 3

1. La composée de deux injections (surjections, bijections) en est une.
2. Si $g \circ f$ est injective, alors f aussi.
3. Si $g \circ f$ est surjective, alors g aussi.
4. L'application $f : E \to F$ est bijective si et seulement s'il existe $g : F \to E$ tel que $g \circ f = \mathrm{Id}_E$ et $f \circ g = \mathrm{Id}_F$. L'application g est alors unique et est notée f^{-1}.
5. Une **involution** ($f \circ f = Id_E$, avec $f : E \to E$) est bijective.
6. La composée de deux bijections en est une et on a $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. On dit que le passage à l'inverse est contravariant.

On peut reformuler ces notions en étudiant, pour $y \in F$, l'équation $(E_y) : y = f(x)$ en l'inconnue x.

Propriétés 1 - 4

1. f est injective si pour tout y de F, l'équation (E_y) admet **au plus une solution**.
2. f est surjective si pour tout y de F, l'équation (E_y) admet **au moins une solution**.
3. f est bijective si pour tout y de F, l'équation (E_y) admet **une et une seule solution**, ce que l'on écrit parfois

$$\forall y \in F \quad \exists! x \in E \qquad y = f(x)\ .$$

Les applications manipulent des éléments d'ensembles, mais on peut aussi les étendre à des transformations des parties d'un ensemble. On introduit ainsi les applications « image directe » et « image réciproque ».

Définition 1 - 6

Si $f : E \to F$, on définit

1. $f_* : \mathcal{P}(E) \to \mathcal{P}(F)$ par

$$f_*(A) = \{y \in F \mid \exists x \in A\,, y = f(x)\}\ .$$

   On dit que $f_*(A)$ est l'**image directe** de A par f.
2. $f^* : \mathcal{P}(F) \to \mathcal{P}(E)$ par

$$f^*(B) = \{x \in E \mid f(x) \in B\}\ .$$

   On dit que $f^*(B)$ est l'**image réciproque** de B par f.

Les propriétés immédiates de ces applications sont :

Propriétés 1 - 5

1. $f_*(\emptyset) = \emptyset$ et $f^*(\emptyset) = \emptyset$.
2. $f^*(F) = E$.
3. $f_*(E) = F \Leftrightarrow f$ est surjective.
4. Si f est bijective, $f^* = (f^{-1})_*$.
5. Par composition, on a $(g \circ f)_* = g_* \circ f_*$ (**covariance**) et $(g \circ f)^* = f^* \circ g^*$ (**contravariance**).

Danger

Par abus de notation, on note f_* et f de la même façon. Pire ! On note f^{-1} au lieu de f^*. Il ne faut pourtant pas en déduire que f_* prend ses valeurs dans F, ni que f est toujours bijective ! Ainsi, si A est un singleton, $A = \{x\}$, alors $f(A)$ est l'ensemble $\{f(x)\}$ et non pas l'élément $f(x)$. Avec ces abus de notations la propriété $f^* = (f^{-1})_*$ s'écrit simplement $f^{-1} = f^{-1}$!

Les résultats importants sur les images directe et réciproque sont :

Propriétés 1 - 6

1. $f(A \cup B) = f(A) \cup f(B)$.
2. $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.
3. $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.
4. $f^{-1}(F \setminus B) = E \setminus f^{-1}(B)$.
5. $(g \circ f)(A) = g\left(f(A)\right)$.
6. $(g \circ f)^{-1}(A') = f^{-1}\left(g^{-1}(A')\right)$.

mais attention

Danger

1. $f(A \cap B) \subset f(A) \cap f(B)$.
2. $f\left(f^{-1}(A')\right) \subset A'$.
3. $f^{-1}\left(f(A)\right) \supset A$.
4. Il n'y a aucun lien **a priori** entre $F \setminus f(A)$ et $f\left(E \setminus A\right)$.

Exercice

Soit $f : E \to F$ une application. Montrer que les assertions suivantes sont équivalentes :

1. f est surjective ;
2. $\forall y \in F \quad f_*(f^*(\{y\})) = \{y\}$;
3. $\forall Y \subset F \quad f_*(f^*(Y)) = Y$;
4. $\forall Y \subset F \quad (f^*(Y) = \emptyset \Rightarrow Y = \emptyset)$.

Trouver un énoncé analogue pour les applications injectives.

Exercice

Soit $f : E \to F$ une application. Montrer : f est injective $\Leftrightarrow \forall X \times Y \subset E^2 \quad f_*(X \cap Y) = f_*(X) \cap f_*(Y)$.

Exercice

1. Soit $f : X \to Y$ une application entre ensembles non vides. Montrer : f injective $\Leftrightarrow \exists g \ : \ Y \to X$, $g \circ f = \mathrm{Id}_X$. Autrement dit f est injective si et seulement si elle est **inversible à gauche**.
2. Soit $f : X \to Y$ une application entre ensembles non vides. Montrer : f surjective $\Leftrightarrow \exists h \ : \ Y \to X$, $f \circ h = \mathrm{Id}_Y$. Autrement dit f est surjective si et seulement si elle est **inversible à droite**.
3. Soit $f : X \to Y$ une application entre ensembles non vides. Montrer : f bijective $\Leftrightarrow \exists g \ : \ Y \to X$, $g \circ f = \mathrm{Id}_X \wedge f \circ g = \mathrm{Id}_Y$.

Exercice

1. Soit $f : X \to Y$ et $g : Y \to Z$. Montrer : ($g \circ f$ injective $\wedge$ f surjective) $\Rightarrow$ g injective.
2. Soit $f : X \to Y$ et $g : Y \to Z$. Montrer : ($g \circ f$ surjective $\wedge$ g injective) $\Rightarrow$ f surjective.

On a les notions de **restriction** pour une fonction (de la source, de l'image ou des deux ensembles) :

Définition 1 - 7

Soit $f : E \to F$, A une partie de E et B une partie de F.

Restriction la restriction de f à A est l'application notée $f_{|A}$ définie de A dans F et qui coïncide avec f sur A.

Co-restriction si B contient $f(E)$, la (co)restriction de f à B est l'application notée $f^{|B}$ définie de E dans B et qui coïncide avec f sur E.

Bi-restriction si B contient $f(A)$, la (bi)restriction de f à A et B est l'application notée $f_{|A}^{|B}$ définie de A dans B et qui coïncide avec f sur A.

Lorsque h est restriction de f, on dit que h est un **prolongement** de f, ou encore une **extension** de f.

Une fonction étant avant tout une transformation, certaines notions liées aux fonctions sont des notions faisant référence aux ensembles, ou parties.

Définition 1 - 8

On parle de partie :

Stable si $f(A) \subset A$,

Invariante (globalement) si $f(A) = A$,

Fixe (point par point) si $\forall x \in A$, $f(x) = x$.

Si A est stable, la restriction de f à A est à valeurs dans A. On définit alors l'**application induite** par f sur A comme l'application de A dans A coïncidant avec f sur A.

Remarque 1 - 3

Une application f de E dans $A \times B$ définit deux applications données par $f_1 = pr_1 \circ f$ et $f_2 = pr_2 \circ f$. De la sorte on construit une bijection $(A \times B)^E \cong A^E \times B^E$, ou encore $\mathscr{F}(E, A \times B) \cong \mathscr{F}(E, A) \times \mathscr{F}(E, B)$.

Pour aller plus loin

Soit Γ une partie de $E \times F$, i.e. un graphe quelconque.

Si pour tout x de E on peut trouver au moins un y de F vérifiant $(x, y) \in \Gamma$, on **admettra** que l'on peut définir une application $f : E \to F$ telle que $\forall x \in E$, $(x, f(x)) \in \Gamma$.

Cela signifie que parmi tous les y possibles de F tels que $(x, y) \in \Gamma$, et on sait qu'il en existe au moins un, on a pu en choisir un arbitrairement.

Cela n'est en rien évident. En fait on a démontré que cette propriété est indécidable, c'est-à-dire qu'on peut la rajouter ou non à la théorie. Il s'agit donc d'un « axiome » supplémentaire à la théorie générale, connu sous le nom d'**axiome du choix**.

4 Familles

Une famille n'est rien d'autre qu'un graphe d'application. On appelle l'ensemble de départ ensemble des indices et on écrit $(A_i)_{i\in I}$ plutôt que $i \mapsto A(i)$.

On définit ainsi la réunion et l'intersection de familles et on étend les lois de DE MORGAN aux familles d'ensemble. On montre également l'associativité de l'intersection et de la réunion, ainsi que la possibilité de réindexation :

Propriétés 1 - 7

1. Si $I = \bigcup_{k\in K} I_k$, alors $\bigcup_{i\in I} A_i = \bigcup_{k\in K}\left(\bigcup_{i\in I_k} A_i\right)$.
2. Si $I = \bigcup_{k\in K} I_k$, alors $\bigcap_{i\in I} A_i = \bigcap_{k\in K}\left(\bigcap_{i\in I_k} A_i\right)$.
3. Si $\varphi : J \to I$ est une bijection, alors $\bigcup_{i\in I} A_i = \bigcup_{j\in J} A_{\varphi(j)}$ et $\bigcap_{i\in I} A_i = \bigcap_{j\in J} A_{\varphi(j)}$.

Définition 1 - 9

Un **recouvrement** de E est une famille $(A_i)_{i\in I}$ dont la réunion contient E

$$\bigcup_{i\in I} A_i \supset E\ .$$

Une **partition** de E est un recouvrement ayant les propriétés suivantes :

1. $\forall i \in I,\ A_i \neq \emptyset$.
2. $\forall (i,j) \in I \times I,\ (i \neq j \Rightarrow A_i \cap A_j = \emptyset)$.
3. $\bigcup_{i\in I} A_i = E$.

Exercice

Soit f une application de E dans F. À quelles conditions nécessaires et suffisantes

a. l'image réciproque de toute partition de F est-elle une partition de E ?

b. l'image directe de toute partition de E est-elle une partition de F ?

♠

On étend également aux familles les résultats sur l'inclusion, les complémentaires, le passage à l'image directe et à l'image réciproque. On étend enfin la notion de produit cartésien et les résultats sur les applications de E à valeurs dans un produit $\prod_{i\in I} A_i$. Si, pour tout i dans I, on a $A_i = B$, on note $\prod_{i\in I} A_i = B^I$, ce qui est consistent avec les notations précédentes.

Notation

Si $I = [\![1; n]\!]$, on note $B^I = B^n$.

5 Relation binaire

Une relation binaire est une autre façon d'interpréter un graphe inclus dans $E \times E$. Au lieu de dire qu'un couple (x, y) d'éléments de E appartient au graphe, on dit que x et y sont reliés, et on note $x\mathcal{R}y$. Tout comme le couple est une paire ordonnée, i.e. (x, y) et (y, x) désignent a priori des choses distinctes, les relations $x\mathcal{R}y$ et $y\mathcal{R}x$ sont a priori des expressions distinctes.

Danger

Attention ! le graphe d'une relation binaire n'a aucune raison d'être un graphe fonctionnel. C'est même plutôt l'exception !

On distingue les propriétés suivantes pour une relation $\mathcal{R}$ sur l'ensemble E :

Définition 1 - 10

Réflexivité : $\mathcal{R}$ est réflexive si $\forall x \in E$, $x\mathcal{R}x$.
Symétrie : $\mathcal{R}$ est symétrique si $\forall (x, y) \in E^2$, $x\mathcal{R}y \Leftrightarrow y\mathcal{R}x$.
Antisymétrie : $\mathcal{R}$ est antisymétrique si $\forall (x, y) \in E^2$, $(x\mathcal{R}y \wedge y\mathcal{R}x) \Rightarrow x = y$.
Transitivité : $\mathcal{R}$ est transitive si $\forall (x, y, z) \in E^3$, $(x\mathcal{R}y \wedge y\mathcal{R}z) \Rightarrow x\mathcal{R}z$.

On a deux types principaux de relations binaires :

Définition 1 - 11

Relation d'équivalence c'est une relation réflexive, symétrique et transitive.
Relation d'ordre c'est une relation réflexive, antisymétrique et transitive.

Relation d'équivalence

Définition 1 - 12

Soit $\mathcal{R}$ une relation d'équivalence définie sur E. Pour x élément de E, on appelle **classe d'équivalence** de x modulo $\mathcal{R}$ et on note $cl_{\mathcal{R}}(x)$ (ou $\overline{x}$ en l'absence d'ambiguïté) l'ensemble défini par

$$cl_{\mathcal{R}}(x) = \{y \in E \,|\, x\mathcal{R}y\} \ .$$

♥

Bien qu'il soit parfois commode de ne retenir dans une classe d'équivalence que l'un de ses éléments, il est en fait plus profond de penser que les **classes d'équivalence sont des ensembles.**

Théorème 1 - 1

Les classes d'équivalence de E modulo $\mathcal{R}$ forment une partition de E.

Réciproquement, pour toute partition $(A_i)_{i \in I}$ de E, la relation binaire $\mathcal{R}$ définie par

$$x\mathcal{R}y \equiv \exists i \in I \ (x, y) \in A_i^2$$

est une relation d'équivalence dont les classes d'équivalence sont les A_i.

Exemple 1 - 1

Soit $f : E \to F$ une application de E sur F. Alors la famille $\left(f^{-1}(y)\right)_{y \in f(E)}$ est une partition de E. La relation d'équivalence qui lui est associée est définie par

$$x_1\mathcal{R}x_2 \Leftrightarrow f(x_1) = f(x_2) \ .$$

♠

Bien que signalé hors programme, il est bon de savoir que l'ensemble des classes d'équivalence de E modulo $\mathcal{R}$ est appelé ensemble quotient et est noté $E/\mathcal{R}$.

Exemples 1 - 2

Si G est un groupe additif et H un sous-groupe de G, la relation $x\mathcal{R}y \equiv x - y \in H$ est une relation d'équivalence sur G et, par commodité, on note G/H l'ensemble quotient. Il peut être muni d'une structure de groupe comme en attestent les deux principaux exemples :

— Un exemple important pour l'arithmétique est $\mathbf{Z}/n\mathbf{Z}$. Un représentant de chaque classe d'équivalence est donné par le reste dans la division euclidienne par n et est donc formé par les nombres de 0 à $n-1$. La classe d'un entier k se note $\overline{k}$ et on a $\overline{k} = k+n\mathbf{Z}$ avec la notation $k+n\mathbf{Z} = \{k + np \,|\, p \in \mathbf{Z}\}$.

— Un exemple important en trigonométrie est $\mathbf{R}/\mathbf{Z}$ ou encore $\mathbf{R}/2\pi\mathbf{Z}$. Un élément de $\mathbf{R}/\mathbf{Z}$ est formé de l'ensemble des réels qui ont la même partie fractionnaire, i.e. qui ne diffèrent que d'un nombre entier relatif. De même, pour x réel, sa classe $\overline{x}$ dans $\mathbf{R}/2\pi\mathbf{Z}$ est donnée par $\overline{x} = x + 2\pi\mathbf{Z} = \{x + 2k\pi \,|\, k \in \mathbf{Z}\}$.

♠

Un exemple de groupe abélien muni d'une structure supplémentaire est celui d'espace vectoriel. Lorsque G est un espace vectoriel E et H est un sous-espace vectoriel F de E, l'ensemble quotient E/F peut être muni d'une structure d'espace vectoriel et c'est en fait la bonne notion pour interpréter celle de supplémentaire de F. Par exemple le théorème du rang s'écrit simplement $E/\operatorname{Ker}(u) \cong \operatorname{Im}(u)$. Voir exercice 1 - 66.

Quand le groupe n'est plus abélien, la relation d'équivalence associée à un sous-groupe H ne permet pas en général de définir une structure de groupe sur G/H. On a alors besoin de la notion de sous-groupe distingué (ou normal). Voir exercice 15 - 39.

Relation d'ordre

Pour une relation d'ordre deux éléments ne sont pas nécessairement en relation.

Définition 1 - 13

On dit que x et y sont **comparables** si $x\mathcal{R}y \vee y\mathcal{R}x$. Quand tous les éléments sont comparables deux à deux, on dit que l'ordre est **total**. Sinon on dit qu'il est **partiel**. Une relation d'ordre (total ou partiel) est souvent notée $\leq$. On note alors $x < y$ pour $(x \leq y \wedge x \neq y)$ et on dit que $<$ est un **ordre strict**.

Exemple 1 - 3

La divisibilité dans $\mathbf{N}^*$ est une relation d'ordre partiel.

On a les notions importantes suivantes pour une relation d'ordre :

Définition 1 - 14

1. Un **majorant** d'une partie A est un élément x de E tel que $\forall a \in A$, $a \leq x$.
2. Un **plus grand élément** de A est un majorant de A qui appartient à A.
3. Une **borne supérieure** de A est le plus petit des majorants de A, i.e.

$$\alpha = \sup A \equiv \begin{cases} \forall a \in A \quad a \leq \alpha \\ \forall M \in E \quad (\forall a \in A \quad a \leq M) \Rightarrow \alpha \leq M \end{cases}$$

4. On a de même les notions de **minorant**, de **plus petit élément**, d'élément **minimal** et de **borne inférieure**.
5. Si $f : E \to F$ est une application entre ensembles ordonnés, on dit qu'elle est (strictement) (dé)croissante si elle préserve (renverse) l'ordre (strict). On parle de fonction (strictement) **monotone** pour une fonction qui est (strictement) croissante ou décroissante.

Remarques 1 - 4

Une partie A de E peut ne pas admettre de majorant bien sûr ! Si elle en admet on dit qu'elle est majorée.

La deuxième propriété satisfaite par la borne supérieure (quand elle existe) est très importante dans la pratique, car pour établir une inégalité du type $\sup A \leq M$ il suffit de vérifier $a \leq M$ pour tout a de A.

Proposition 1 - 1

S'il existe, le plus grand (resp. petit) élément est unique. On le note $\max A$ (resp. $\min A$).

Danger

On prendra garde qu'une fonction qui n'est pas croissante n'a aucune raison d'être décroissante, même un tout petit peu. Par exemple la fonction caractéristique des rationnels prend les valeurs 0 ou 1 et n'est monotone sur aucun intervalle de $\mathbf{R}$.

Proposition 1 - 2

Une fonction qui est monotone et injective est strictement monotone.

Danger

La réciproque n'est pas vraie en général. Elle l'est si l'ordre sur l'ensemble de départ est total. Un contre-exemple est donné par la fonction cardinal sur l'ensemble des parties d'un ensemble fini.

♠

Un élément **maximal** de A est un élément a de A tel que $\forall b \in A$, $a \leq b \Rightarrow a = b$. On définit de même un élément **minimal**.

La notion d'élément maximal ne présente d'intérêt que dans un ensemble muni d'une relation d'ordre partiel. En effet si l'ordre est total a est un élément maximal de A si et seulement si $a = \max A$. L'exemple qui suit montre l'intérêt de cette notion.

Exemple 1 - 4

Dans $\mathbf{N}^*$ ordonné par la divisibilité, l'ensemble $A = \mathbf{N} \setminus \{0, 1\}$ ne possède pas de plus petit élément mais une infinité d'éléments minimaux : les nombres premiers.

On peut construire des relations d'ordre à partir d'autres relations d'ordre :

Pour aller plus loin

Ordre opposé : c'est l'ordre défini par $x\mathcal{R}y \equiv y \leq x$. On le note $x \geq y$.

Ordre induit : l'ordre induit sur une partie A est l'ordre sur A obtenu par intersection du graphe $\mathcal{R}$ avec $A \times A$.

Ordre fonctionnel si X est un ensemble quelconque et si E est un ensemble ordonné, E^X l'est aussi. On ordonne les fonctions par leurs valeurs, i.e. $f \leq g \equiv \forall x \in X,\ f(x) \leq g(x)$. On prendra garde à ne pas utiliser l'ordre strict dans ce cas, de peur de confondre $f < g$ avec $\forall x \in X,\ f(x) < g(x)$, ces deux notions étant différentes.

Ordre produit : si pour tout i dans I, E_i est un ensemble ordonné, l'ordre produit sur $\prod_{i\in I} E_i$ est donné par $(x_i)_{i\in I} \leq (y_i)_{i\in I} \equiv \forall i \in I,\ x_i \leq y_i$. C'est en fait un ordre fonctionnel.

Ordre lexicographique : si I est un ensemble totalement ordonné, on peut ordonner $\prod_{i\in I} E_i$ différemment en posant : $(x_i)_{i\in I} \leq (y_i)_{i\in I} \equiv (\forall i \in I,\ x_i = y_i) \vee (\exists i_0 \in I,\ \forall i \in I,\ (i < i_0 \Rightarrow x_i = y_i) \wedge x_{i_0} < y_{i_0})$.

Exemple 1 - 5

N, **Z**, **Q**, **R** avec l'ordre habituel ; **C** avec l'ordre lexicographique ; **N** avec l'ordre (partiel) donné par la relation de divisibilité ; $\mathscr{P}(E)$ avec l'inclusion ; $\mathscr{F}(\mathbf{R})$ avec l'ordre fonctionnel habituel.

Exemple 1 - 6

Dans $\mathbf{K}[X]$ la divisibilité n'est pas une relation d'ordre. On peut toutefois lui associer la relation d'équivalence $\sim$ définie par

$$P \sim Q \equiv (P \mid Q \wedge Q \mid P)\ .$$

Les classes d'équivalence sont les classes de polynômes associés : ils ne diffèrent qu'à un multiple scalaire non nul près. On peut par exemple choisir dans chaque classe de polynômes associés l'unique polynôme normalisé.

Cet ensemble quotient, ou cet ensemble de représentants des classes d'équivalence, est alors ordonné par la divisibilité et si on lui retire la classe de 1 (i.e. des polynômes constants non nuls) il possède une infinité d'éléments minimaux : les polynômes irréductibles (unitaires).

Plus généralement on peut construire une relation d'ordre par passage au quotient.

Pour aller plus loin

Soit $\prec$ une relation de préordre définie sur E, c'est-à-dire une relation binaire réflexive et transitive. La relation binaire définie sur E par

$$x \asymp y \Leftrightarrow (x \prec y \wedge y \prec x)$$

est une relation d'équivalence. Pour $x \in E$, on note $\overline{x}$ sa classe d'équivalence. On note $\widehat{E}$ l'ensemble des classes d'équivalence. La relation définie sur $\widehat{E}$ par

$$\overline{x} \leq \overline{y} \Leftrightarrow x \prec y$$

est une relation d'ordre.

6 Lois

Une loi est une autre façon de penser aux graphes ou aux applications. C'est une application d'un produit d'ensembles (éventuellement différents) dans un ensemble. Les notions de magma et d'opération sur un ensemble ne sont pas au programme, mais interviennent naturellement dans les définitions des objets du programme.

Une loi de composition interne définie sur un ensemble E est une application de $E \times E$ vers E. Si $\intercal$ (lire truc) est une telle loi de composition interne, pour $(x, y) \in E^2$, on note $x \intercal y$ au lieu de $\intercal(x, y)$. On dit alors que E est un **magma**. Si E est muni de plusieurs lois internes, on parle de **multi-magma**.

Une loi de composition externe définie sur un ensemble E et à opérateurs dans un ensemble X est une application de $X \times E$ vers E. Si $\star$ (lire étoile) est une telle loi de composition externe, pour $(\lambda, x) \in X \times E$, on note $\lambda \star x$ au lieu de $\star(\lambda, x)$.

Dans la suite, on note $\intercal$ et $\perp$ (lire anti-truc) des lois internes sur un ensemble E et $\star$ une loi externe de l'ensemble X sur l'ensemble E.

Exemples 1 - 7

Par exemple $\mathscr{P}(E)$ est muni de deux lois internes : $\cap$ et $\cup$; $\mathscr{F}(E)$ est muni de la loi interne $\circ$; un espace vectoriel est muni de deux lois, une interne et une externe : $+$ et $\cdot$.

Les notions importantes pour les lois internes sont :

♥

Commutativité : la loi $\intercal$ est commutative si $\forall (x, y) \in E^2$, $x \intercal y = y \intercal x$. Un magma commutatif est aussi appelé **abélien** en hommage à Niels Henrik Abel (1802–1829).

Associativité : la loi $\intercal$ est associative si $\forall (x, y, z) \in E^3$, $(x \intercal y) \intercal z = x \intercal (y \intercal z)$. Il en résulte, mais ce n'est pas évident, que l'ordre des calculs pour $\intercal$ est indifférent, voir exercice 1 - 8

Élément neutre : c'est un élément $e_\intercal$ de E tel que $\forall x \in E$, $x \intercal e_\intercal = e_\intercal \intercal x = x$. L'élément neutre, s'il existe, est unique.

Distributivité : on dit que $\perp$ est distributive par rapport à $\intercal$ si $\forall (x, y, z) \in E^3$, $(x \intercal y) \perp z = (x \perp z) \intercal (y \perp z)$ et $z \perp (x \intercal y) = (z \perp x) \intercal (z \perp y)$.

Symétrique : un élément x de E est dit symétrisable ou inversible par rapport à $\intercal$ s'il existe y dans E tel que $x \intercal y = y \intercal x = e_\intercal$. L'inverse, s'il existe, est unique et on le note $\overset{-1}{\intercal} x$.

Régularité : un élément x de E est dit régulier ou simplifiable si $\forall (y, z) \in E \times E$, $x \intercal y = x \intercal z \Rightarrow y = z$ et $\forall (y, z) \in E \times E$, $y \intercal x = z \intercal x \Rightarrow y = z$. Un élément inversible est simplifiable.

Remarque 1 - 5

Une loi externe peut être distributive par rapport à une loi interne : dans ce cas on a $\forall x \in X$, $\forall (a, b) \in E \times E$, $x \star (a \intercal b) = (x \star a) \intercal (x \star b)$.

On appelle **monoïde** tout couple $(E, \intercal)$ constitué d'un ensemble et d'une loi de composition interne définie sur cet ensemble qui est associative et est pourvue d'un élément neutre.

Notation

Quand, dans un monoïde, la loi de composition interne est notée + (loi additive) elle est nécessairement commutative et son élément neutre est noté 0. L'inverse de a (quand il existe) est appelé opposé et noté $-a$.

Quand, dans un monoïde, la loi de composition interne est notée $\times$ ou $\cdot$ (loi multiplicative), son élément neutre est noté 1. On écrit alors ab au lieu de $a \times b$. L'inverse de a (quand il existe) est noté a^{-1}. Si, de plus, la loi est commutative, l'inverse de a peut se noter $\frac{1}{a}$.

Exemple 1 - 8

$(\mathbf{N}, +)$ est un monoïde d'élément neutre 0.

La notion de groupe s'est dégagée progressivement des travaux d'Évariste Galois (1811–1832) sur le groupe symétrique. Son but était de résoudre les équations polynomiales par des formules explicites ne mettant en jeu que des radicaux (extractions de racines n^{e}). C'est Arthur Cayley (1821–1895) qui a donné la bonne définition de groupe fini.

Les groupes apparaissent également en géométrie, notamment dans le programme d'Erlangen de Felix Klein (1849–1925) et les travaux de Sophus Lie (1842–1899), et en théorie des nombres, dans les travaux de Leopold Kronecker (1823–1891). La synthèse a été opérée par Camille Jordan (1838–1922) et la définition définitive de groupe abstrait a probablement été donnée par Walther von Dyck (1856–1934).

Définition 1 - 15

Groupe – Galois, Cayley, C. Jordan, von Dyck

Un **groupe** est un monoïde (i.e. un ensemble muni d'une loi interne associative et admettant un élément neutre) dans lequel tout élément est inversible.

Exemple 1 - 9

- les ensembles classiques $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ ou tout espace vectoriel muni de l'addition.
- Les bijections d'un ensemble E (noté S_E) muni de la composition.
- Les matrices à coefficients dans $\mathbf{K}$ pour l'addition.
- Le groupe des rotations (vectorielles) ou des similitudes (directes vectorielles).
- Le groupe des translations affines.
- Le groupe des isométries qui préservent un polygone, un polyèdre ou un polytope.

Les anneaux ont été formalisés par Richard Dedekind (1831–1916), bien que le terme ait été introduit par David Hilbert (1862–1943), et étudiés systématiquement par Emmy Noether (1882–1935, mathématicienne allemande chassée par le régime nazi en 1933 et morte peu après aux États-unis).

Définition 1 - 16

Anneau – Dedekind, Noether

Soit A un ensemble muni de deux lois internes, notées $\intercal$ et $\perp$, ayant au moins deux éléments. Alors A est un **anneau** si $(A, \intercal)$ est un groupe abélien et si la loi $\perp$ vérifie les propriétés d'associativité, de distributivité (par rapport à $\intercal$) et admet un élément neutre. On note 0 l'élément neutre de $(A, \intercal)$ et 1 celui de $(A, \perp)$. On a alors $1 \neq 0$. Si $\perp$ est commutative, on dit que A est commutatif.

Exemple 1 - 10

Un anneau est un ensemble dans lequel on peut effectuer des opérations semblables à l'addition, la soustraction et la multiplication. En dehors des corps habituel (le corps $\mathbf{Q}$ des nombres rationnels, le corps $\mathbf{R}$ des nombres réels et le corps $\mathbf{C}$ des nombres complexes), en voici quelques autres :

- les entiers relatifs $\mathbf{Z}$ et les nombres décimaux (parfois noté $\mathbf{D}$),
- les matrices carrées d'ordre n à coefficients dans un corps $\mathscr{M}_n(\mathbf{K})$,
- les endomorphismes d'un espace vectoriel $\mathrm{End}(E)$,
- les (homo)morphismes d'un groupe G abélien (commutatif) $\mathrm{Hom}(G)$.
- les entiers de Gauss $\mathbf{Z}[i]$, i.e. les nombres complexes de parties réelle et imaginaire entières.
- l'algèbre associative $\mathbf{K}[X]$ des polynômes à coefficients dans le corps $\mathbf{K}$.

Les corps ont également été formalisés par Dedekind et c'est Ernst Steinitz (1871–1928) qui en donne la définition actuelle.

Définition 1 - 17

Corps – Dedekind, Steinitz

Un **corps** est un anneau commutatif $\mathbf{K}$ dont les éléments non nuls sont inversibles. Tout comme un anneau, il a au moins deux éléments et on a $1 \neq 0$.

Exemple 1 - 11

On désigne par $\mathbf{F}_2$ l'ensemble à deux éléments 0 et 1 muni des deux lois de composition $+$ et $\times$ définies par leurs tables :

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

Muni de ces deux lois, $\mathbf{F}_2$ possède une structure de corps. On le note aussi $\mathbf{Z}/2\mathbf{Z}$.

Plus généralement, lorsque p est un nombre premier, $\mathbf{Z}/p\mathbf{Z}$ admet une structure de corps. Comme on le fait du plan $\mathbf{R}^2$, on peut munir $(\mathbf{Z}/p\mathbf{Z})^2$ d'une structure de corps et, même $(\mathbf{Z}/p\mathbf{Z})^r$ pour tout entier naturel non nul r. Il y a unicité, à isomorphisme près, d'un corps fini à cardinal fixé. Si $q = p^r$, on note $\mathbf{F}_q$ le corps fini à q éléments.

La structure des corps infinis est plus complexe. En dehors des classiques $\mathbf{Q}$, $\mathbf{R}$ et $\mathbf{C}$, on peut citer le corps des fractions d'un anneau intègre, obtenu comme $\mathbf{Q}$ à partir de $\mathbf{Z}$. Par exemple $\mathbf{K}(X)$ ou $\mathbf{Q}[i]$, obtenus à partir de $\mathbf{K}[X]$ et $\mathbf{Z}[i]$ respectivement.

La notion d'espace vectoriel a germé dans les travaux de Hermann Günther Grassmann (1809–1877) et a été formalisée par Giuseppe Peano (1858–1932).

Définition 1 - 18

Espace vectoriel – Grassmann, Peano

Un **espace vectoriel** E sur un corps $\mathbf{K}$ est un groupe abélien pour une loi notée $+$, muni d'une loi externe de $\mathbf{K}$ sur E, notée $\star$, qui est distributive par rapport à $+$, associative au sens suivant :

$$\forall(\lambda,\mu) \in \mathbf{K}^2 \quad \forall x \in E \quad \lambda \star (\mu \star x) = (\lambda \cdot \mu)x$$

et compatible avec l'élément neutre de $\mathbf{K}$ (noté 1) : $\forall x \in \mathbf{K}$, $1 \star x = x$.

Exemple 1 - 12

Si $\mathbf{K}$ est un corps et n un entier naturel, $\mathbf{K}^n$ et $\mathscr{M}_n(\mathbf{K})$ admettent une structure canonique de $\mathbf{K}$-espace vectoriel. Il en va de même de $\mathbf{K}[X]$.

L'ensemble des solutions d'un problème linéaire (équation différentielle linéaire, suite récurrente linéaire etc.) forme un espace vectoriel.

La notion d'algèbre a une histoire très complexe. Elle puise ses sources à la fois dans la logique, notamment avec les travaux de George Boole (1815–1864) complétés par ceux de Benjamin Peirce (1809–1880) et Charles Sanders Peirce (1839–1914), et l'algèbre linéaire, avec la recherche de multiplication sur des n-uplets. C'est notamment William Rowan Hamilton (1805–1865) et Grassmann qui explorent les algèbres extérieures et amènent aux quaternions, bientôt suivis des octaves, découverts par John Thomas Graves (1806–1870) puis par Arthur Cayley (1821–1895). Leurs généralisations sont étudiées par William Kingdon Clifford (1845-1879). L'ensemble de ces structures porte initialement le nom de systèmes hypercomplexes, puisqu'ils généralisent en quelque sorte la notion de nombre complexe. Le terme d'algèbre est donné par Bartel Leendert van der Waerden (1903—1996) dans son ouvrage fondateur *Moderne Algebra* en 1930.

Définition 1 - 19

Algèbre – G. Boole, Grassmann, ..., van der Waerden

Une **algèbre** sur un corps $\mathbf{K}$, ou $\mathbf{K}$-algèbre, est un $\mathbf{K}$-espace vectoriel $(A, +, \star)$ muni d'une multiplication interne qui est bilinéaire. Dans le cadre du programme, on demande en sus que $(A, \times)$ soit unifère. Autrement dit $(A, +, \times)$ est un anneau où l'axiome d'associativité de la multiplication est remplacé par

$\forall \lambda \in \mathbf{K}, \forall (x, y) \in A^2 \ \lambda \star (x \times y) = (\lambda \star x) \times y = x \times (\lambda \star y)$.

Exemple 1 - 13

Très souvent les algèbres sont aussi associatives, donc $(A, +, \times)$ est vraiment un anneau, et contiennent $\mathbf{K}$, ce qui fait que l'associativité externe résulte de l'associativité. C'est le cas pour les $\mathbf{K}$-algèbres $\mathbf{K}[X]$, $\mathscr{M}_n(\mathbf{K})$ ou $\mathscr{F}(X, \mathbf{K})$ pour un ensemble X quelconque.

Le corps non commutatif des quaternions ou l'agèbre non associative des octonions forment les seules $\mathbf{R}$-algèbres de dimension finie dites à division, i.e. sans diviseurs de 0. Elles ne sont que le début d'une suite infinie d'algèbres obtenue par la construction de Cayley-Dickson (Leonard Eugene Dickson, 1874-1954), avec par exemple les sedenions qui forment une algèbre de dimension 16 sur $\mathbf{R}$.

Il existe bien d'autres types d'algèbres : algèbre extérieure (avec comme exemple celle fournie par le produit vectoriel de vecteurs en dimension 3), algèbre de Lie, algèbre de Jordan, algèbre de Banach. Les plus importantes dans ce cours seront toutefois $\mathbf{K}[X]$ et $\mathscr{M}_n(\mathbf{K})$.

Définition 1 - 20

Soit $A \in \mathscr{P}(E)$, on définit $\mathbb{1}_A : E \to \{0, 1\}$ la fonction **indicatrice** de A par $\mathbb{1}_A(x) = 1$ si $x \in A$ et $\mathbb{1}_A(x) = 0$ sinon. On la note aussi parfois χ_A.

C'est pour cette raison que George Boole et, à sa suite, Alfréd Rényi (1921-1970), notent xy l'intersection de deux ensembles x et y, 1 l'ensemble total, 0 l'ensemble vide, $1 - x$ le complémentaire de l'ensemble x etc. Le plus subtil est $x + y$ pour la réunion disjointe, puisque la réunion est plutôt $x + y - xy$ ou encore $1 - (1 - x)(1 - y)$. Avec ces notations booléennes, il existe des diviseurs de 0 : $xy = 0$ signifie que x et y sont disjoints et ainsi $x(1 - x) = 0$ ou encore $x = x^2$.

7 L'ensemble N et les théorèmes de récurrence

L'existence de l'ensemble **N** des entiers naturels est supposée acquise. Il est muni des lois + et ·, ainsi que d'une relation d'ordre ≤, dont les propriétés essentielles sont supposées connues. En particulier on retiendra :

Propriétés 1 - 8

1. Tout entier naturel n admet un « successeur » : $n+1$. Tout entier naturel non nul n admet un « prédécesseur » : $n-1$.
2. **N** est un ensemble ordonné dont toute partie non vide admet un plus petit élément. On dit aussi que **N** est *bien ordonné*.
3. Toute partie non vide majorée de **N** admet un plus grand élément.

Les traces sur **N** des intervalles de **R** seront notés avec une double barre. Par exemple $[\![p;q]\!] = [p;q] \cap \mathbf{N}$ et $[\![p;+\infty[\![= [p;+\infty[\cap \mathbf{N}$.

Théorème 1 - 2

Récurrence – Grassmann 1861

Soit n_0 un entier naturel et $\mathcal{A}$ une partie de **N** qui satisfait aux axiomes de récurrence suivants

Initialisation : $n_0 \in \mathcal{A}$

Hérédité : $n \in \mathcal{A} \Rightarrow n+1 \in \mathcal{A}$.

Alors $\mathcal{A}$ contient la **section finissante** $[\![n_0;+\infty[\![$. En particulier si $a_0 = 0$, alors $\mathcal{A} = \mathbf{N}$.

♠

L'axiomatisation (hors-programme) de **N** par Richard Dedekind (1888) et Giuseppe Peano (1889) est fondée sur l'existence d'une application successeur, injective et n'ayant pas 0 dans son image, et sur l'axiome de récurrence. John von Neumann (1923) quant à lui définit l'entier 0 comme l'ensemble vide et le successeur d'un entier a comme $a \cup \{a\}$. Ainsi on a $2 = \{\emptyset, \{\emptyset\}\} = \{0,1\}$, ensemble qui a manifestement deux éléments ! Plus généralement $n = \{x \in \mathbf{N} \mid x < n\}$.

L'emploi le plus fréquent du théorème de récurrence est dans la démonstration d'une propriété $\mathcal{P}$ dont l'énoncé dépend d'un entier n (on parle de prédicat sur les entiers). On peut énoncer les trois corollaires usuels suivants :

Corollaire 1 - 1

Récurrence simple

Soit $\mathcal{P}$ un prédicat sur **N**. Si, pour un certain entier naturel n_0, $\mathcal{P}(n_0)$ est vrai et que, de plus,

$$\forall n \geq n_0\,, \qquad \mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$$

alors $\mathcal{P}(n)$ est vrai pour tout entier n supérieur (ou égal) à n_0.

Corollaire 1 - 2

Récurrence double

Soit $\mathcal{P}$ un prédicat sur **N**. Si, pour un certain entier naturel n_0, $\mathcal{P}(n_0)$ et $\mathcal{P}(n_0+1)$ sont vrais et que, de plus,

$$\forall n \geq n_0+1\,, \qquad (\mathcal{P}(n-1) \wedge \mathcal{P}(n)) \Rightarrow \mathcal{P}(n+1)$$

alors $\mathcal{P}(n)$ est vrai pour tout entier n supérieur (ou égal) à n_0.

Corollaire 1 - 3

Récurrence forte

Soit $\mathcal{P}$ un prédicat sur $\mathbf{N}$. Si, pour un certain entier naturel n_0, $\mathcal{P}(n_0)$ est vrai et que, de plus,

$$\forall n \geq n_0\,, \qquad (\forall k \in [\![n_0; n]\!]\ ,\ \mathcal{P}(k)) \Rightarrow \mathcal{P}(n+1)$$

alors $\mathcal{P}(n)$ est vrai pour tout entier n supérieur (ou égal) à n_0.

♥

Rédaction des récurrences

On commence par définir un prédicat $\mathcal{P}$ en précisant son ensemble de définition, que l'on munit d'un (bon) ordre, i.e. de sorte qu'il y ait un élément minimal (disons n_0) et une application successeur (en général notée $n \mapsto n+1$). L'initialisation consiste à démontrer $\mathcal{P}(n_0)$ et l'hérédité consiste à déduire $\mathcal{P}(n+1)$ à partir de $\mathcal{P}(n)$.

♠

On peut également raisonner par récurrence noethérienne (ou bien fondée) en montrant $\forall n \in \mathbf{N}\ (\forall k \in \mathbf{N}\ (k < n \Longrightarrow \mathcal{P}(k))) \Longrightarrow \mathcal{P}(n)$.

Exemple 1 - 14

Blaise Pascal (1664–1662)

Soit $(\mathbf{H}_n)$ le prédicat sur $\mathbf{N}^*$ défini par $(\mathbf{H}_n)$: « Pour tout entier naturel p inférieur à n, $\binom{n}{p} = \dfrac{n-p+1}{p}\binom{n}{p-1}$ » où, comme il est d'usage en français, « inférieur » signifie « inférieur ou égal ».

Remarque 1 - 6

À noter également la récurrence finie et la récurrence descendante. On s'intéresse alors à des parties finies de $\mathbf{N}$, i.e. de la forme $[\![p; q]\!]$ et l'hérédité peut prendre la forme usuelle $(\mathbf{H}_n) \Rightarrow (\mathbf{H}_{n+1})$ ou au contraire la forme descendante $(\mathbf{H}_{n+1}) \Rightarrow (\mathbf{H}_n)$. Dans le premier cas on démontre $(\mathbf{H}_p)$, dans le second $(\mathbf{H}_q)$ et dans tous les cas on en déduit que $(\mathbf{H}_n)$ est vraie pour tout entier n dans $[\![p; q]\!]$.

Définition 1 - 21

Soit E un ensemble, on appelle suite d'éléments de E toute famille de E indexée par $\mathbf{N}$, i.e. une application de $\mathbf{N}$ dans E. On note une telle suite $(u_n)_{n\in\mathbf{N}}$.

On admet les théorèmes suivants beaucoup moins banals qu'il n'y paraît :

Théorème 1 - 3

Soit $f : E \to E$ une application d'un ensemble dans lui-même, et a un élément de E. Alors il existe une unique suite $(u_n)_{n\in\mathbf{N}}$ dans $E^{\mathbf{N}}$ vérifiant

1. $u_0 = a$
2. $\forall n \in \mathbf{N}$, $u_{n+1} = f(u_n)$.

Théorème 1 - 4

Soit $f : E \times E \to E$ et a et b deux éléments de E. Alors il existe une unique suite $(u_n)_{n\in\mathbf{N}}$ dans $E^{\mathbf{N}}$ vérifiant

1. $u_0 = a$ et $u_1 = b$
2. $\forall n \in \mathbf{N}$, $u_{n+2} = f(u_{n+1}, u_n)$.

8 Ensembles finis

Rappel

On dit que deux ensembles E et F sont équipotents s'il existe une bijection de E sur F.

Théorème 1 - 5

Soit p et n des entiers naturels.

1. S'il existe une injection de $[\![1;p]\!]$ dans $[\![1;n]\!]$ alors $p \leq n$.
2. S'il existe une surjection de $[\![1;p]\!]$ dans $[\![1;n]\!]$ alors $p \geq n$.
3. S'il existe une bijection de $[\![1;p]\!]$ dans $[\![1;n]\!]$ alors $p = n$.

Définition 1 - 22

On dit qu'un ensemble E est fini non vide si il existe un entier p strictement positif tel que E soit équipotent à $[\![1;p]\!]$. L'entier p est alors unique et est appelé cardinal de E et noté $\text{Card}(E)$, $|E|$ ou $\sharp E$.

Convention

On convient de dire que l'ensemble vide est fini de cardinal 0.

Définition 1 - 23

Tout ensemble non fini est dit infini.

Théorème 1 - 6

Les parties finies non vides de $\mathbf{N}$ sont les parties non vides majorées.

Démonstration.

a. Condition nécessaire. On travaille par récurrence sur le cardinal p d'une partie non vide finie. Soit $\mathcal{P}$ le prédicat sur $\mathbf{N}^*$ donné par $\mathcal{P}(n)$: toute partie de $\mathbf{N}$ de cardinal n est majorée.

a. Initialisation : $\mathcal{P}(1)$ est immédiat puisque qu'une partie de cardinal 1 est majorée par son unique élément.

b. Hérédité. Soit p dans $\mathbf{N}^*$ tel que $\mathcal{P}(p)$ soit vrai et P une partie finie de $\mathbf{N}$ de cardinal $p+1$ éléments. Soit enfin f une bijection de $[\![1;p+1]\!]$ sur P.

On note $P_1 = P \setminus \{f(p+1)\}$. La birestriction $g = f_{|[\![1;p]\!]}^{|P_1}$ de f est alors une bijection de $[\![1;p]\!]$ sur P_1. Ceci fait de P_1 un ensemble fini de cardinal p et donc majoré. Soit M_1 un de ses majorants. Alors $M = \max(M_1, f(p+1))$ est un majorant de P qui est donc bien majoré.

b. Condition suffisante. On travaille par récurrence forte sur le plus grand élément M d'une partie non vide majorée de $\mathbf{N}$. Soit $\mathcal{P}$ le prédicat sur $\mathbf{N}$ donné par $\mathcal{P}(n)$: toute partie de $\mathbf{N}$ de plus grand élément n est finie.

Soit P une partie non vide et majorée de $\mathbf{N}$ et soit M son plus grand élément. On montre par récurrence forte sur M que P est fini.

a. Initialisation : $\mathcal{P}(0)$ est immédiat puisque qu'une partie de plus grand élément nul est égale à $\{0\}$ et est donc de cardinal fini égal à 1.

b. Hérédité. Soit M dans $\mathbf{N}$ tel que $\mathcal{P}(n)$ soit vrai pour $n \leq M$ et P une partie non vide majorée de plus grand élément $M+1$. Soit alors $P_1 = P \setminus \{M+1\}$.

Si P_1 est vide, alors P est de cardinal 1. Sinon P_1 est non vide et majoré par M. Son plus grand élément est donc inférieur à M. Par hypothèse de récurrence, P_1 est donc fini. Soit alors p son cardinal ; il existe une bijection f de $[\![1;p]\!]$ sur P_1 que l'on peut étendre à $[\![1;p+1]\!]$ en posant $\overline{f}(p+1) = M+1$. L'extension $\overline{f}$ de f ainsi construite est alors une bijection de $[\![1;p+1]\!]$ sur P qui est bien fini.

□

Théorème 1 - 7

Soit P une partie finie non vide de $\mathbf{N}$ de cardinal p. Alors il existe une unique bijection (strictement) croissante de $[\![1;p]\!]$ sur P.

Démonstration. Par récurrence sur p dans $\mathbf{N}^*$.

a. Initialisation : pour $p = 1$, toute bijection est croissante.

b. Hérédité. On suppose la propriété satisfaite au rang $p-1$, avec $p \geq 2$. Soit alors une partie P finie non vide de cardinal p. Elle est majorée et admet donc un plus grand élément α. Soit $P_1 = P \setminus \{\alpha\}$. Alors P_1 est une partie finie de cardinal $p-1$ et, par hypothèse de récurrence, il existe une unique bijection croissante f de $[\![1;p-1]\!]$ sur P_1. Le prolongement $\overline{f}$ de f obtenu en posant $\overline{f}(p) = \alpha$ réalise une bijection strictement croissante de $[\![1;p]\!]$ sur P. Si g est une autre bijection croissante de $[\![1;p]\!]$ sur P, alors $g(p) \geq g(k)$ pour tout $k \leq p$. Donc $g(p)$ est le plus grand élément de P : $g(p) = \alpha$. La birestriction de g à $[\![1;p-1]\!]$ au départ et à P_1 à l'arrivée est une bijection strictement croissante de $[\![1;p-1]\!]$ sur P_1 et donc coïncide avec f par hypothèse de récurrence, si bien que $g = \overline{f}$.

□

Les propriétés suivantes résultent directement des théorèmes précédents 1 - 5, 1 - 6 et 1 - 7. C'est un bon exercice de les démontrer !

Propriétés 1 - 9

— Soit E et F deux ensembles équipotents. Si l'un est fini, l'autre l'est aussi et ils ont même cardinal.

— Toute partie E' d'un ensemble fini E est elle-même finie et $\text{Card}(E') \leq \text{Card}(E)$ avec égalité si et seulement si $E' = E$.

— Soit E et F deux ensembles finis de même cardinal et $f : E \to F$. Alors l'injectivité, la surjectivité et la bijectivité de f sont équivalentes.

— Soit $f : E \to F$ où E est un ensemble fini. Alors $f(E)$ est une partie finie de F et on a $\text{Card}(f(E)) \leq \text{Card}\,(E)$ avec égalité si et seulement si f est injective.

— Toute intersection d'ensembles finis est finie.

— Toute réunion finie d'ensembles finis est finie de cardinal inférieur ou égal à la somme des cardinaux.

9 Sommes finies et produits finis

Dans un monoïde additif[a], on définit par récurrence la notation $\sum_{k=1}^{n} a_k$ par $\sum_{k=1}^{0} a_k = 0$ et $\sum_{k=1}^{n+1} a_k = \left(\sum_{k=1}^{n} a_k\right) + a_{n+1}$.

On peut alors définir, pour tout ensemble fini d'indices I, la notation $\sum_{i\in I} a_i$ à l'aide d'une bijection f de $[\![1;n]\!]$ sur I par

$$\sum_{i\in I} a_i = \sum_{k=1}^{n} a_{f(k)} \ .$$

Cette formule, de par la commutativité de l'addition, ne dépend pas de la bijection f choisie.

En particulier, pour n entier et a dans le monoïde, les notations na et n correspondent respectivement à $a_k = a$ ou $a_k = 1$ pour tout k dans $[\![1;n]\!]$. De plus

Convention

$$\sum_{i\in\emptyset} a_i = 0$$

Dans un monoïde multiplicatif[b], on définit par récurrence la notation $\prod_{k=1}^{n} a_k$ par $\prod_{k=1}^{0} a_k = 1$ et $\prod_{k=1}^{n+1} a_k = \left(\prod_{k=1}^{n} a_k\right) \times a_{n+1}$.

Dans un monoïde multiplicatif **commutatif** on peut alors définir pour tout ensemble fini d'indices I la notation $\prod_{i\in I} a_i$ à l'aide d'une bijection f de $[\![1;n]\!]$ sur I par

$$\prod_{i\in I} a_i = \prod_{k=1}^{n} a_{f(k)} \ .$$

Cette formule, de par la commutativité de la multiplication, ne dépend pas de la bijection f choisie.

En particulier, pour n entier et a dans le monoïde, les notations a^n et $n!$ (lire « factorielle n » – la notation de factorielle vient des travaux de Christian Kramp, 1760–1826) correspondent respectivement à $a_k = a$ ou $a_k = k$ pour tout k dans $[\![1;n]\!]$. Par convention $0! = 1$ et plus généralement

Convention

$$\prod_{i\in\emptyset} a_i = 1$$

a. un tel monoïde est commutatif
b. un tel monoïde n'est pas nécessairement commutatif

10 Dénombrements, analyse combinatoire

Théorème 1 - 8

Principe des bergers

Soit $f : E \to F$ une surjection de E sur un ensemble fini F. On suppose qu'il existe un entier naturel non nul p vérifiant

$$\forall y \in F \qquad \mathrm{Card}(f^{-1}(y)) = p \ .$$

Alors E est fini et on a

$$\mathrm{Card}(E) = p \times \mathrm{Card}(F) \ .$$

Ce théorème résulte de l'exemple 1 - 1 et de l'énoncé plus général suivant

Théorème 1 - 9

Pour toute partition $(A_i)_{i\in I}$ d'un ensemble fini E, l'ensemble I est lui-même fini et, de plus,

$$\mathrm{Card}(E) = \sum_{i\in I} \mathrm{Card}(A_i) \ .$$

Un problème de dénombrement consiste à démontrer qu'un ensemble est fini et à en déterminer le cardinal. Les théorèmes essentiels auxquels on se réfère sont le théorème d'équipotence et le principe des bergers, avec pour but de se ramener à des ensembles finis de cardinal connu. Cependant le plus souvent, les démonstrations seront rédigées de façon plus « littéraire », ces théorèmes apparaissant alors en filigrane. Voici quelques illustrations ces méthodes.

Théorème 1 - 10

Si E et F sont des ensembles finis alors $E \times F$ est fini et

$$\mathrm{Card}(E \times F) = \mathrm{Card}(E) \times \mathrm{Card}(F) \ .$$

Démonstration. Il suffit d'appliquer le principe des bergers à l'application $f : E \times F \to F$ qui à (x, y) associe y (i.e. $f = pr_2$). Les images réciproques $f^{-1}(y)$ sont toutes équipotentes à E ce qui permet de conclure. □

Théorème 1 - 11

Si E et F sont des ensembles finis alors $\mathscr{F}(E, F)$ est fini et

$$|\mathscr{F}(E, F)| = \left|F^E\right| = |F|^{|E|} \ .$$

Idée

Le principe est le suivant : pour construire une application de E dans F il suffit de savoir construire une application de $E_1 = E \setminus \{a\}$ dans F et de définir $f(a)$. Ainsi la démonstration sera faite par récurrence sur $p = \mathrm{Card}(E)$. Pour $f_{|E_1}$ on a n^{p-1} choix (avec $n = \mathrm{Card}(F)$) et pour chacun d'eux on a n choix possibles pour $f(a)$. La mise en forme d'une telle démonstration relève donc du principe des bergers.

Démonstration.

a. L'initialisation à $p = 1$ est claire puisqu'alors $F^E \simeq F$.

b. Hérédité. On suppose la propriété établie pour $p-1$, avec $p \in \mathbf{N}^*$. Soit alors E fini de cardinal p, a dans E fixé quelconque et $E_1 = E \setminus \{a\}$.

On applique le principe des bergers à l'application

$$\varphi_a \ : \quad \mathscr{F}(E,F) \to \mathscr{F}(E_1,F)$$

qui à f associe $f_{|E_1}$. En effet :

a. le principe de prolongement assure le caractère surjectif de φ_a ;

b. l'hypothèse de récurrence assure l'équipotence des images réciproques des éléments de F avec pour cardinal commun $\mathrm{Card}(F)$.

Et donc le principe des bergers assure bien l'hérédité.

□

Théorème 1 - 12

Soit E un ensemble fini de cardinal p, alors $\mathscr{P}(E)$ est fini et

$$\mathrm{Card}\left(\mathscr{P}(E)\right) = 2^p \ .$$

Démonstration. Il suffit de constater que $\mathscr{P}(E)$ et $\mathscr{F}(E,\{0,1\})$ sont équipotents via la bijection qui à une partie A de E associe sa fonction caractéristique $\mathbb{1}_A$. □

Remarque 1 - 7

La fonction indicatrice intervient souvent en dénombrement, notamment pour la raison suivante : si E est un ensemble fini et $A \subset E$, alors

$$\sum_{x \in E} \mathbb{1}_A(x) = \mathrm{Card}(A) \ .$$

comme on le voit en prenant une bijection de $[\![1;\mathrm{Card}(A)]\!]$ sur A que l'on prolonge en une bijection de $[\![1;\mathrm{Card}(E)]\!]$ sur E.

Théorème 1 - 13

Formule de Poincaré, dite du crible ou Principe d'inclusion-exclusion

Soit A et B deux parties finies d'un même ensemble E. On a

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B) \ .$$

Henri Poincaré (1854–1912).

Démonstration. Cela résulte de l'identité entre fonctions indicatrices

$$\mathbb{1}_{A\cup B} + \mathbb{1}_{A\cap B} = \mathbb{1}_A + \mathbb{1}_B \ ,$$

que l'on peut vérifier pour tout x dans E dans les quatre cas exclusifs $x \in A \cap B$, $x \in A \setminus A \cap B$, $x \in B \setminus A \cap B$ et $x \notin A \cup B$, et de la remarque précédente appliquée à $E = A \cup B$ et à ses parties A, B, $A \cap B$ et $A \cup B$. □

Formule de Poincaré

Si A, B et C sont des parties finies d'un même ensemble, alors on a

$$\begin{aligned}\mathrm{Card}(A\cup B\cup C) &= \mathrm{Card}(A)+\mathrm{Card}(B)+\mathrm{Card}(C)\\ &\quad -[\mathrm{Card}(A\cap B)+\mathrm{Card}(B\cap C)+\mathrm{Card}(C\cap A)]\\ &\quad +\mathrm{Card}(A\cap B\cap C)\,.\end{aligned}$$

♠

Plus généralement si $(A_i)_{1\leq i\leq n}$ est une famille de parties finies d'un même ensemble, on a la formule

$$\mathrm{Card}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{k=1}^{n}(-1)^{k+1}\sum_{1\leq i_1<\cdots<i_k\leq n}\mathrm{Card}\left(\bigcap_{j=1}^{k}A_{i_j}\right)$$

Théorème 1 - 14

Soit E et F deux ensembles finis de cardinaux respectifs p et n. Alors l'ensemble $\mathcal{I}(E,F)$ des injections de E dans F est fini de cardinal A_n^p, avec $A_n^p=\prod_{k=0}^{p-1}(n-k)$, i.e.

$$A_n^p=\begin{cases}0 & \text{si } n<p\\ \dfrac{n!}{(n-p)!} & \text{sinon.}\end{cases}$$

Idée

Pour construire une injection de E dans F il suffit de savoir construire une injection de $E\backslash\{a\}$ (que l'on notera E_1) dans F et de choisir $f(a)$ parmi les images possibles restantes. La démonstration s'obtient par récurrence sur $\mathrm{Card}(E)$ (que l'on notera p).

En notant $n=\mathrm{Card}(F)$, on a A_n^{p-1} choix pour $f_{|E_1}$ et, pour chacun d'eux, on a $n-p+1$ choix possibles si $n\geq p$ et aucun sinon. D'où, pour $n\geq p$, $A_n^p=(n-p+1)A_n^{p-1}$ et donc, compte tenu de l'initialisation $A_n^1=n$, il vient $A_n^p=\dfrac{n!}{(n-p)!}$.

Corollaire 1 - 4

Soit E un ensemble fini de cardinal n, S_E l'ensemble des permutations de E (i.e. les bijections de E dans E) est fini de cardinal $n!$.

Définition 1 - 24

On appelle p-liste d'un ensemble E tout p-uplet $(a_1,a_2,\ldots,a_p)$ constitué d'éléments de E. En particulier l'unique 0-liste est la liste vide.

Théorème 1 - 15

Le nombre de p-listes d'éléments d'un ensemble fini E de cardinal n est n^p.

Démonstration. Pour $p\geq 1$, se donner une p-liste revient à se donner une application de $[\![1;p]\!]$ dans E. Pour $p=0$, l'unique 0-liste est la liste vide. □

Définition 1 - 25

On appelle arrangement de p éléments de E toute p-liste de E constituée d'éléments deux à deux distincts.

Théorème 1 - 16

Le nombre d'arrangements de p éléments parmi n est A_n^p.

Démonstration. Pour $p \geq 1$, se donner un arrangement de p éléments de E ensemble à n éléments, revient à se donner une injection de $[\![1;p]\!]$ dans E. Pour $p = 0$, l'unique 0-liste est la liste vide. □

Définition 1 - 26

On appelle combinaison de p éléments d'un ensemble E toute partie finie de E de cardinal p. En particulier l'unique combinaison à 0 élément de E est la combinaison vide.

Théorème 1 - 17

Si E est un ensemble fini de cardinal n, l'ensemble des combinaisons de p éléments de E est fini, de cardinal noté $\binom{n}{p}$ (ou parfois C_n^p) et on a

$$\binom{n}{p} = \frac{n!}{p!\ (n-p)!}$$

pour $0 \leq p \leq n$ et $\binom{n}{p} = 0$ si $n < p$. En particulier $\binom{n}{0} = 1$.

Idée

À tout arrangement à p éléments, qui est une p-liste, on associe son support, i.e. l'ensemble des éléments constituant la p-liste. Comme c'est un arrangement, cet ensemble a p éléments, i.e. est une combinaison à p éléments. Réciproquement à toute combinaison à p éléments de E, on peut associer autant d'arrangements à p éléments qu'on a de possibilités de ranger p éléments, à savoir $p!$. Par conséquent $A_n^p = p!\binom{n}{p}$, d'où la propriété.

Théorème 1 - 18

Cas particulier de la formule du binôme de Newton

Pour tout entier naturel n, on a

$$2^n = \sum_{p=0}^{n} \binom{n}{p}.$$

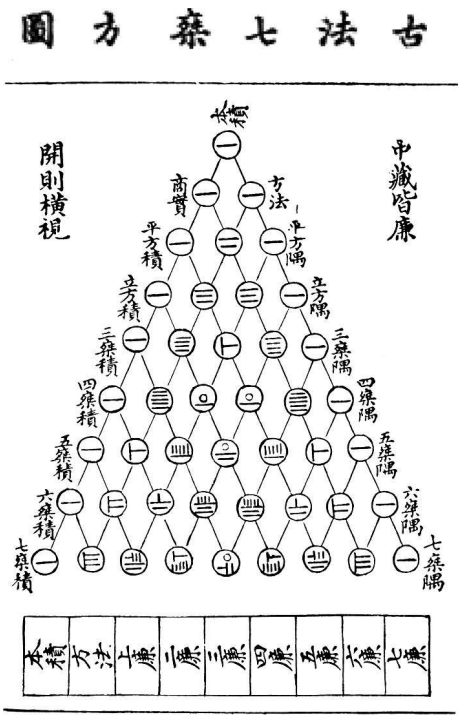
Isaac Newton (1642–1727).

Démonstration. Pour p entier vérifiant $0 \leq p \leq n$, on note $\mathcal{P}_p(E)$ l'ensemble des p-combinaisons d'éléments de E. Alors la famille $(\mathcal{P}_p(E))_{0\leq p\leq n}$ réalise une partition de $\mathscr{P}(E)$. En prenant les cardinaux on trouve l'égalité recherchée. □

Définition 1 - 27

Les nombres $\binom{n}{p}$ s'appellent coefficients binomiaux.

On peut les représenter sur un triangle, dit triangle de Pascal (Blaise Pascal, 1623-1662).



Aparté

Ce triangle était connu du mathématicien perse Abu Bekr ibn Muhammad ibn al-Husayn Al-Karaji (953–1029) et est connu en Chine sous le nom de triangle de Yang Hui (1238-1298). Dans ses écrits, il l'attribue à un autre mathématicien chinois, de deux siècles son prédecesseur : Jia Xian (ca. 1010-1070), dont les travaux ont été perdus. Il était utilisé pour extraire des racines carrées ou cubiques. Le triangle apparaît également sous le nom *meruprastāra* dans un commentaire du X^{e} siècle des travaux du mathématicien indien Acharya Pingala (ca. -300/-200) sur la prosodie. Dans ces travaux on trouve le 0, la notation binaire et la suite de Fibonacci !

Le terme $\binom{n}{k}$ est donc défini comme le nombre de façons de choisir k éléments dans un ensemble de n éléments, notamment k termes dans un produit de n termes. On en déduit directement la formule du binôme

♥

Formule du binôme de Newton

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

formule valide pour a et b dans un anneau général, avec $[a,b]=0$, i.e. a et b commutant entre eux (où $[a,b]$ désigne, comme souvent, le commutateur de a et b, i.e. $[a,b]=ab-ba$).

Une propriété élémentaire est la formule, qui permet de calculer les coefficents binomiaux de façon récursive, $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$. Elle s'interprète combinatoirement en comptant de façon séparée les choix pour lesquels le $n+1^{e}$ terme choisi est a ou b.

On peut également construire directement la n^{e} ligne du triangle de PASCAL en partant de 1 et en multipliant successivement par des fractions de numérateur décroissant et dont le numérateur et le dénominateur ont pour somme $n+1$: 1, $1\times\dfrac{n}{1}$, $n\times\dfrac{n-1}{2}$, $\dfrac{n(n-1)}{2}\times\dfrac{n-2}{3}$ etc. En particulier, on retrouve la formule

$$\binom{n}{k}=\frac{n(n-1)\cdots(n-k+1)}{1.2.\cdots.k}=\frac{n!}{k!\,n-k!}\ .$$

Théorème 1 - 19

Les coefficients binomiaux satisfont aux formules suivantes valables pour des entiers vérifiant $n\geq p\geq 0$:

Symétrie $\binom{n}{p}=\binom{n}{n-p}$

n^{e} ligne $\binom{n}{p+1}=\dfrac{n-p}{p+1}\times\binom{n}{p}$

Diagonales $\binom{n+1}{p+1}=\dfrac{n+1}{p+1}\times\binom{n}{p}$

Relation de PASCAL $\binom{n+1}{p+1}=\binom{n}{p+1}+\binom{n}{p}$.

Démonstration. Les vérifications sont triviales à partir de la formule $\binom{n}{p}=\dfrac{n!}{p!\ (n-p)!}$, mais on peut également les interpréter de façon ensembliste.

La première formule correspond à l'échange entre a et b dans la formule du binôme, c'est-à-dire à l'application bijective qui à une partie A associe son complémentaire, échangeant ainsi les parties de cardinal p avec celles de cardinal $n-p$.

La seconde provient de la remarque sur le calcul direct sur la n^{e} ligne. Elle exprime que pour construire une partie à $p+1$ éléments à partir d'une partie à p éléments, il faut lui adjoindre un élément parmi les $n-p$ de son complémentaire, et réciproquement il faut supprimer un des $p+1$ éléments d'une partie à $p+1$ éléments pour en obtenir une à p éléments. On peut l'écrire directement sous forme d'égalité d'une somme double en utilisant une bijection entre les couples (A,x), formés d'une partie A à $p+1$ éléments et d'un élément x de A, et les couples (B,y), formés d'une partie B à p éléments et d'un élément y de $\overline{B}$. La bijection envoie (A,x) sur $(A\setminus\{x\},x)$ dont la bijection réciproque envoie (B,y) sur $(B\cup\{y\},y)$. Il vient

$$(p+1)\binom{n}{p+1}=\sum_{|A|=p+1}\sum_{x\in A}1=\sum_{|B|=p}\sum_{y\notin B}1=(n-p)\binom{n}{p}\ .$$

La troisième provient d'une interversion de deux signes somme ou encore de la bijection associant à (A,x), avec A une partie à $p+1$ éléments de $[\![1;n+1]\!]$ et x dans A, le couple $(x,A\setminus\{x\})$, formé d'un élément de $[\![1;n+1]\!]$ et d'une partie à p éléments ne le contenant pas. De façon plus parlante, on peut l'écrire

$$(p+1)\binom{n+1}{p+1}=\sum_{|A|=p+1}\sum_{x\in A}1=\sum_{1\leq x\leq n+1}\sum_{A\ni x}1=\sum_{1\leq x\leq n+1}\sum_{|B|=p,x\notin B}1=(n+1)\binom{n}{p}\ .$$

On a déjà expliqué la dernière : une partie A à $p+1$ éléments de $[\![1;n+1]\!]$ contient ou non $n+1$ et on peut lui associer une partie de $[\![1;n]\!]$ soit à p éléments, soit à $p+1$

éléments. De façon légèrement abusive cela peut se sythétiser ainsi :

$$\sum_{|A|=p+1} 1 = \sum_{n+1\in A} 1 + \sum_{n+1\notin A} 1\,.$$

□

L'égalité $\binom{n}{k} = \dfrac{n(n-1)\cdots(n-k+1)}{1.2.\cdots.k}$ est définie a priori pour n et k entiers naturels. On l'étend directement au cas où n est dans un anneau contenant $\mathbf{Z}$ et où tous les entiers non nuls sont inversibles, comme $\mathbf{R}$, $\mathbf{C}$, $\mathbf{R}[X]$ ou $\mathscr{M}_n(\mathbf{R})$ par exemple. On note U_n les polynômes de Newton définis par $U_n = X(X-1)\cdots(X-n+1)$ et T_n les polynômes de Hilbert (aussi appelés polynômes factoriels) définis par $T_n = \dfrac{U_n}{n!} = \binom{X}{n}$. Ils jouent un rôle important dans les questions d'interpolation, avec l'interpolation de Newton.

Elle consiste à approcher une fonction définie sur $[\![0;n]\!]$ par un polynôme P_n de degré au plus n prenant les mêmes valeurs que la fonction, de façon incrémentale : plus le nombre de points d'approximation augmente, plus on élève le degré. Par opposition à l'interpolation de Lagrange (Joseph-Louis Lagrange, 1736–1813), à chaque étape on ne rajoute qu'un seul terme. On étudie l'application linéaire $P \mapsto (P(0),\dots,P(n))$, dont le noyau est formé des multiples de U_{n+1}. On remarque, en notant τ et e_x les applications linéaires $P \mapsto P(X+1)$ et $P \mapsto P(x)$, qu'on a $e_k(P) = e_0(\tau^k(P))$. La matrice de τ dans la base canonique étant triangulaire supérieure à diagonale formée de 1, on étudie $\tau - \mathrm{Id}$, noté Δ et on remarque $\Delta(T_k) = T_{k-1}$ (pour $k \geq 1$) et $\Delta(T_0) = 0$. Comme $\deg(T_k) = k$, on peut écrire $P = \sum_{k=0}^{n} a_k T_k$ et alors $e_0(\Delta^k(P)) = a_k$ puisque $e_0(T_k) = \delta_{k,0}$.

On peut le dire en termes de suites. Si $(u_k)_{k\in\mathbf{N}}$ est une suite à valeurs réelles ou complexes, on lui associe (et ceci donne naissance à un opérateur linéaire sur l'espace vectoriel des suites) $(\Delta u_k)_{k\in\mathbf{N}}$ définie par $\Delta u_k = u_{k+1} - u_k$. Le polynôme d'approximation de Newton d'ordre n, P_n, coïncidant avec $(u_k)_{k\in\mathbf{N}}$ sur ses n premiers termes s'obtient en calculant successivement $P_0 = u_0$, $P_1 = P_0 + \Delta u_0 X$, $\cdots$, $P_n = P_{n-1} + (\Delta^n u)_0 \dfrac{X(X-1)\cdots(X-n+1)}{n!}$, avec $\Delta^n = \Delta\circ\cdots\circ\Delta$, i.e.

$$P_n = \sum_{k=0}^{n} \frac{(\Delta^k u)_0}{k!} U_k = \sum_{k=0}^{n} (\Delta^k u)_0 T_k = \sum_{k=0}^{n} (\Delta^k u)_0 \binom{X}{k}\,.$$

Aparté

La méthode se généralise avec des pas constants (i.e. en remplaçant $0, 1, \cdots, n$ par $x_0, x_0+h, \cdots, x_0+nh$) ou non. On parle de méthode des différences divisées.

L'opérateur Δ est appelé opérateur de dérivation discrète et est lié au triangle de Pascal :

$$(\Delta^n u)_0 = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} u_k \quad \text{et donc aussi} \quad u_n = \sum_{k=0}^{n} \binom{n}{k} (\Delta^k u)_0$$

et on peut le définir également comme application linéaire sur l'espace des polynômes [c] par $\Delta(P) = P(X+1) - P$. Dans ce cadre les polynômes de Newton et Hilbert constituent une base de triangulation de l'opérateur Δ : $\Delta(T_{n+1}) = T_n$ ou encore $\Delta(U_{n+1}) = (n+1)U_n$.

c. On rappelle que les polyômes sont, techniquement, des suites, mais attention les deux définitions de Δ qui en résultent **ne coïncident pas**.

Opérateur Δ – dérivation discrète

L'opérateur Δ permet de calculer les sommes du type $\sum_{n=1}^{N} n^k$ pour k entier naturel. En effet, si P est un polynôme tel que $\Delta(P) = X^k$, alors la somme précédente est une somme télescopique et donc, en supposant $P(0) = 0$, ce qui est toujours possible, elle est égale à $P(N+1)$.

Par exemple, on a $\Delta(T_2) = T_1 = X$ et donc $\sum_{n=1}^{N} n = T_2(N+1) = \dfrac{N(N+1)}{2}$.

Pour la somme des cubes, on écrit d'abord X^3 dans la base des polynômes de Hilbert. Puisqu'un polynôme de degré 3 est approché de façon exacte par l'approximation de degré 3, on a

Remarque 1 - 8

$$X^3 = P_3 = \sum_{k=0}^{3} (\Delta^k u)_0 T_k \quad \text{avec } (u_k) = (k^3) \ .$$

On calcule aisément l'effet successif de Δ en retranchant le terme courant du suivant :

k	u_k	Δu_k	$\Delta^2 u_k$	$\Delta^3 u_k$
0	0	1	6	6
1	1	7	12	
2	8	19		
3	27			

et donc $X^3 = 6T_3 + 6T_2 + T_1$. Un antécédent de X^3 par Δ est donc $6T_4 + 6T_3 + T_2$, i.e. $\Delta\left(\dfrac{1}{4}X^2(X-1)^2\right) = X^3$, et donc $\sum_{n=1}^{N} n^3 = \dfrac{N^2(N+1)^2}{4}$.

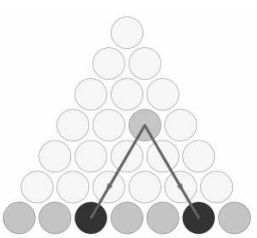
Aparté

On a donc $\Delta(T_2) = T_1$ et $\Delta(T_2^2) = T_1^3$. C'est remarquable, mais exceptionnel. En fait Δ n'est pas tout à fait une dérivation. On a, pour P et Q deux polynômes,

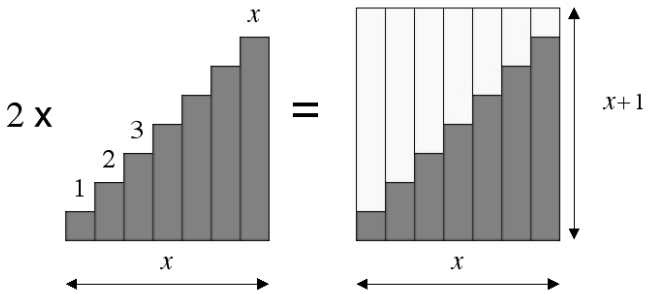
$$\Delta(PQ) = \begin{vmatrix} P(X+1) & P \\ Q & Q(X+1) \end{vmatrix} = \begin{vmatrix} \Delta(P) & P \\ -\Delta(Q) & Q(X+1) \end{vmatrix} = \begin{vmatrix} \Delta(P) & P \\ -\Delta(Q) & Q+\Delta(Q) \end{vmatrix}$$

et donc $\Delta(PQ)$ est la somme du terme attendu $\Delta(P)Q + P\Delta(Q)$, correspondant à une dérivation, et du terme du second ordre $\Delta(P)\Delta(Q)$. Il vient $\Delta(T_2^2) = 2T_1T_2 + T_1^2 = T_1(T_1 + 2T_2)$.

L'égalité fondamentale utilisé pour X^3, donne avec X^2 : $T_1 + 2T_2 = X^2 = T_1^2$ et on retrouve $\Delta(T_2^2) = T_1^3 = \Delta(T_2)^3$.



Le dessin précédent offre une démonstration sans mots de l'égalité $\sum_{n=1}^{N} n = \binom{N+1}{2} = T_2(N+1)$. En le joignant à la classique méthode de Gauss de sommation par complément,



$$2(1+2+3+\ldots+x) = x(x+1)$$

on en déduit $T_2(N+1) = \sum_{n=1}^{N} n = \dfrac{N(N+1)}{2}$. S'il était besoin on pourrait aussi en déduire $T_2 = \dfrac{1}{2}X(X-1)$. On termine ce paragraphe avec une identité remarquable, souvent associée à la formule du binôme du moins lorsque $n = 2$:

♥

Formule de (Jakob) Bernoulli

Pour n entier strictement positif, la formule

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}$$

est valide dans un anneau général à condition qu'on ait $[a, b] = 0$.

Jakob Bernoulli, 1654–1705.

11 Ensembles dénombrables.

Idée

Les ensembles infinis sont les ensembles non finis, mais y-a-t-il plusieurs infinis ? En d'autres termes deux ensembles infinis sont-ils toujours équipotents ? La réponse est non, comme le prouvent les recherches de Georg Cantor (1845–1918) et notamment le théorème hors-programme suivant.

Théorème 1 - 20

Théorème de Cantor ♠

Soit E un ensemble quelconque, il n'existe pas de surjection (et donc a fortiori de bijection) de E sur $\mathscr{P}(E)$.

Démonstration. On effectue une démonstration par l'absurde.

Soit φ une telle surjection, $\mathcal{A}$ l'ensemble $\{x \in E \,|\, x \notin \varphi(x)\}$ et a un antécédent de $\mathcal{A}$.

a. Si on suppose $a \in \mathcal{A}$, alors $a \notin \varphi(a) = \mathcal{A}$, ce qui est contradictoire.

b. Si on suppose $a \notin \mathcal{A}$, alors $a \in \varphi(a) = \mathcal{A}$, ce qui est aussi contradictoire.

Ainsi une telle application φ ne saurait exister. □

En particulier il y a « plusieurs infinis » ! Tout comme on a classé les ensembles finis par taille, on peut « classer » les ensembles infinis. Ainsi les ensembles équipotents à $\mathbf{N}$ sont dits de cardinal $\aleph_0$ (lire aleph zéro, comme la lettre de l'alphabet hébreu). Ceux, comme $\mathbf{R}$, qui sont équipotents à $\mathscr{P}(\mathbf{N})$ sont dits de cardinal $2^{\aleph_0}$.

Pour aller plus loin

L'hypothèse du continu affirme $2^{\aleph_0} = \aleph_1$, autrement dit qu'il n'existe pas d'ensemble « compris entre $\mathbf{N}$ et $\mathbf{R}$ ».

Ces questions sont au cœur des travaux de Georg Cantor, 1845–1918.

Définition 1 - 28

Un ensemble E est dit

— dénombrable s'il est équipotent à $\mathbf{N}$.

— au plus dénombrable s'il est fini ou dénombrable.

Théorème 1 - 21

Toute partie de $\mathbf{N}$ est au plus dénombrable. Autrement dit, toute partie infinie de $\mathbf{N}$ est dénombrable.

Plus précisément, pour toute partie infinie P de $\mathbf{N}$, il existe une bijection strictement croissante et une seule de $\mathbf{N}$ sur P.

Démonstration. Soit P une partie de $\mathbf{N}$.

Si P est fini, il est au plus dénombrable. Si P est infini, il est non vide. Soit alors a_0 son plus petit élément et $P_1 = P \setminus \{a_0\}$. L'ensemble P_1 est alors non vide car P est infini, et on peut poser $a_1 = \min P_1$.

Supposons avoir construit les p premiers éléments $(a_0, a_1, \ldots, a_{p-1})$ d'une suite vérifiant

— $a_0 < a_1 < \cdots < a_{p-1}$;

— $\forall a \in P \setminus \{a_0, a_1, \ldots, a_{p-1}\}$, $a > a_{p-1}$.

On pose alors $P_p = P \setminus \{a_0, a_1, \ldots, a_{p-1}\}$, de sorte que P_p est également une partie de $\mathbf{N}$ non vide. On peut poser $a_p = \min P_p$ et la suite $(a_n)_{n\in\mathbf{N}}$ ainsi construite définit une injection strictement croissante de $\mathbf{N}$ dans P.

Il reste à vérifier qu'elle est surjective. Par une récurrence immédiate, pour tout entier n on a $a_n \geq n$. Soit donc $q \in P$. On a $a_q \geq q$ et, par construction de la suite, on a $q \notin P_q$. Par conséquent q est l'un des éléments a_0, a_1, …, a_{q-1} et l'application $\varphi : n \mapsto a_n$ est bien surjective. Ainsi c'est une bijection croissante de $\mathbf{N}$ sur P, et P est dénombrable.

Soit ψ est une bijection croissante de $\mathbf{N}$ sur P. Elle vérifie par monotonie $\psi(0) \leq \psi(n)$ pour tout entier naturel n et donc $\psi(0) = \min \psi(\mathbf{N}) = \min P = \varphi(0)$.

Soit $n \in \mathbf{N}^*$ tel que, pour tout k inférieur à $n-1$, on ait $\psi(k) = \varphi(k)$. Par monotonie on a

$$\begin{aligned}\psi(n) = \min \psi(\mathbf{N} \setminus [\![0; n-1]\!]) &= \min\left(P \setminus \{\psi(0), \ldots, \psi(n-1)\}\right) \\ &= \min\left(P \setminus \{\varphi(0), \ldots, \varphi(n-1)\}\right) = \varphi(n).\end{aligned}$$

et donc $\psi = \varphi$. □

Exemple 1 - 15

Pour tout $p \geq 2$ l'ensemble $p\mathbf{N}$ est dénombrable (bijection $n \mapsto pn$).

Corollaire 1 - 5

Un ensemble est fini ou dénombrable si et seulement s'il est en bijection avec une partie de $\mathbf{N}$. Autrement dit un ensemble E est fini ou dénombrable si et seulement s'il existe une injection de E dans $\mathbf{N}$, ou encore si et seulement s'il existe une surjection de $\mathbf{N}$ dans E.

Démonstration. Le sens direct résulte de la définition d'ensemble fini ou d'ensemble dénombrable. Le sens réciproque est une conséquence du théorème précédent et du fait que l'équipotence est une relation d'équivalence.

Pour la réinterprétation on suppose E non vide, sinon c'est immédiat.

Soit f une bijection de E sur une partie de $\mathbf{N}$, notée F. L'injection canonique ι de F dans $\mathbf{N}$ permet de construire une injection $\iota \circ f$ de E dans $\mathbf{N}$. Un prolongement à $\mathbf{N}$ quelconque (par exemple en envoyant $\overline{F}$ sur $\min(F)$) de l'identité de F, noté p_F permet de construire une surjection $f^{-1} \circ p_F$ de $\mathbf{N}$ sur E.

S'il existe une injection de E dans $\mathbf{N}$, alors E est en bijection avec son image, qui est une partie de $\mathbf{N}$. S'il existe une surjection f de $\mathbf{N}$ dans E, alors l'application $g : x \mapsto \min(f^{-1}(x))$ est bien définie et est une injection de E dans $\mathbf{N}$, puisque $f \circ g = \mathrm{Id}_E$. □

Théorème 1 - 22

L'ensemble $\mathbf{N} \times \mathbf{N}$ est dénombrable.

Démonstration. Tout entier naturel non nul n admet une unique écriture du type $n = 2^p(2q+1)$ avec p et q entiers naturels. L'application $(p,q) \mapsto 2^p(2q+1) - 1$ est donc une bijection de $\mathbf{N} \times \mathbf{N}$ sur $\mathbf{N}$. □

Théorème 1 - 23

Si k est un entier naturel non nul, $\mathbf{N}^k$ est dénombrable.

Plus généralement un produit cartésien (fini) d'ensembles dénombrables est dénombrable.

Démonstration (non exigible). L'ensemble des nombres premiers étant infini, on note $(p_i)_{i \in \mathbf{N}^*}$ la suite strictement croissante des nombres premiers. L'application de $\mathbf{N}^k$ dans $\mathbf{N}$ donnée par $(n_i)_{1 \leq i \leq k} \mapsto \prod_{i=1}^{k} p_i^{n_i}$ est alors injective, d'après le théorème de factorisation des entiers, et donc $\mathbf{N}^k$ est au plus dénombrable. Comme il n'est pas fini puisqu'il contient une partie équipotente à $\mathbf{N}$ (par exemple celle dont toutes les coordonnées sont nulles sauf peut-être la première), il est donc dénombrable. On note φ une bijection de $\mathbf{N}^k$ sur $\mathbf{N}$.

Soit maintenant $(E_i)_{1 \leq i \leq k}$ des ensembles dénombrables et $(f_i)_{1 \leq i \leq k}$ des bijections $f_i : E_i \simeq \mathbf{N}$. Alors $(x_i)_{1 \leq i \leq k} \mapsto \varphi(f_1(x_1), \ldots, f_k(x_k))$ est une bijection de $E_1 \times \cdots \times E_k$ sur $\mathbf{N}$, d'où le résultat annoncé. □

Corollaire 1 - 6

Les ensembles $\mathbf{Z}$ et $\mathbf{Q}$ sont dénombrables.

Démonstration. L'application $(a,b) \mapsto a - b$ fournit une surjection de $\mathbf{N}^2$ dans $\mathbf{Z}$ et donc, par composition avec une bijection de $\mathbf{N}$ sur $\mathbf{N}^2$, une surjection de $\mathbf{N}$ dans $\mathbf{Z}$.

L'application $(a,b,c) \mapsto \dfrac{a-b}{c+1}$ fournit une surjection de $\mathbf{N}^3$ dans $\mathbf{Q}$ et, comme précédemment, une surjection de $\mathbf{N}$ dans $\mathbf{Q}$. □

Définition 1 - 29

Limite monotone d'ensemble ♠

Soit $(E_n)_{n\in\mathbf{N}}$ une suite croissante d'ensembles. On appelle **limite** de cette suite, et on la note $\lim\uparrow E_n$, l'ensemble défini par

$$\lim\uparrow E_n = \bigcup_{n\in\mathbf{N}} E_n\ .$$

On l'appelle aussi **borne supérieure** de la suite $(E_n)_{n\in\mathbf{N}}$.

De même pour une suite $(E_n)_{n\in\mathbf{N}}$ décroissante d'ensemble, on appelle **limite** ou **borne inférieure** de cette suite, et on la note $\lim\downarrow E_n$, l'ensemble défini par

$$\lim\downarrow E_n = \bigcap_{n\in\mathbf{N}} E_n\ .$$

Théorème 1 - 24

Caractérisation des ensembles dénombrables par limite croissante ♠

Un ensemble I est au plus dénombrable si et seulement s'il existe une suite croissante $(I_n)_{n\in\mathbf{N}}$ de parties finies de I dont la réunion est égale à I, i.e. telles que $I = \lim\uparrow I_n$.

Démonstration.

a. Condition nécessaire. Si I est fini la propriété est claire en prenant une suite constante égale à I. S'il est dénombrable, il existe une bijection φ de $\mathbf{N}$ sur I. Comme $\mathbf{N} = \bigcup_{n\in\mathbf{N}} [\![0;n]\!]$, il vient

$$I = \bigcup_{n\in\mathbf{N}} \varphi([\![0;n]\!]) = \bigcup_{n\in\mathbf{N}} I_n$$

en posant $I_n = \varphi([\![0;n]\!])$, et (I_n) est bien une suite croissante de parties finies de I dont la limite est I.

b. Condition suffisante. Soit n dans $\mathbf{N}$. Comme I_n est fini on dispose d'une injection j_n de I_n dans $\mathbf{N}$. Soit maintenant $j : I \to \mathbf{N}\times\mathbf{N}$ l'application qui à x dans I associe $(n, j_n(x))$ avec $n = \min\{k\in\mathbf{N}\,|\,x\in I_k\}$. Cette application est injective et fournit, par composition d'une bijection de $\mathbf{N}\times\mathbf{N}$ sur $\mathbf{N}$, une injection de I dans $\mathbf{N}$. On en déduit que I un ensemble au plus dénombrable.

□

Théorème 1 - 25

Toute réunion au plus dénombrable d'ensembles eux-mêmes au plus dénombrables est au plus dénombrable.

Démonstration (non exigible). C'est une conséquence du théorème précédent. Soit I un ensemble au plus dénombrable et $(E_i)_{i\in I}$ une famille d'ensembles au plus dénombrables. On dispose de (I_n) et $(E_{i,n})$ des parties finies telles que $I = \lim\uparrow I_n$ et, pour i dans I, $E_i = \lim\uparrow E_{i,n}$.

On pose alors, pour n dans $\mathbf{N}$, $J_n = \bigcup_{i\in I_n} E_{i,n}$. C'est une suite croissante de parties finies incluses dans $\bigcup_{i\in I} E_i$ par construction. Réciproquement pour x dans cette réunion, on dispose de i dans I tel que $x\in E_i$ et donc de j dans $\mathbf{N}$ tel que $i\in I_j$ et de k dans $\mathbf{N}$ tel que $x\in E_{i,k}$. Par croissance, en posant $n = \max(j,k)$, il vient $x\in E_{i,n}$ et $i\in I_n$, donc $x\in J_n$. Il vient $\bigcup_{i\in I} E_i = \lim\uparrow J_n$ et le résultat s'ensuit. □

Théorème 1 - 26

Les ensembles $\mathbf{R}$, $]0;1[$, $[0;1[$, $]0;1]$ et $[0;1]$ ne sont pas dénombrables.

Démonstration (non exigible). On commence par démontrer la non dénombrabilité de $[0;1[$.

a. Soit $\varphi : [0;1[\to \{0,1\}^{\mathbf{N}}$ l'application qui à un élément associe la suite des termes de son développement dyadique propre.

On a en particulier, en prenant x dans $[0;1]$ et en posant $\varphi(x) = (a_n)_{n\in\mathbf{N}}$,

$$x = \sum_{n=0}^{\infty} \frac{a_n}{2^{n+1}} \ .$$

Il en résulte que φ est injective.

b. Si $(a_n)_{n\in\mathbf{N}}$ est dans $\{0,1\}^{\mathbf{N}}$ et n'est pas stationnaire égale à 1, alors en posant $x = \sum_{n=0}^{\infty} \frac{a_n}{2^{n+1}}$, on a $\varphi(x) = (a_n)_{n\in\mathbf{N}}$. Il en résulte que l'image de φ est constituée des suites qui ne sont pas stationnaires égales à 1.

c. Soit I_n l'ensemble des suites stationnaires égales à 1 à partir du rang n. On a affaire à une suite croissante d'ensembles de cardinaux donnés par $\mathrm{Card}(I_n) = 2^n$ et donc I donné par $I = \lim \uparrow I_n$ est (au plus) dénombrable.

d. Si $[0;1[$ était au plus dénombrable, alors son image par φ le serait aussi (puisqu'elle lui est équipotente) et donc $\{0,1\}^{\mathbf{N}}$ serait réunion de deux ensembles au plus dénombrables. Comme $\{0,1\}^{\mathbf{N}}$ est équipotent à $\mathscr{P}(\mathbf{N})$, par la bijection qui à une partie associe sa fonction caractéristique, d'après le théorème de Cantor (théorème 1 - 20), $\{0,1\}^{\mathbf{N}}$ n'est pas dénombrable et cette contradiction montre que $[0;1[$ n'est pas dénombrable.

Comme une partie d'un ensemble dénombrable est au plus dénombrable, il en résulte par contraposée que $\mathbf{R}$ et $[0;1]$ ne sont pas dénombrables.

Comme th est une bijection de $\mathbf{R}$ sur $]0;1[$, ce dernier intervalle n'est pas dénombrable et l'argument précédent permet de conclure que $]0;1]$ ne l'est pas non plus. □

Exercice

Expliciter des bijections entre les cinq ensembles précédents.

Pour aller plus loin

L'ensemble des nombres **algébriques**, i.e. des racines d'une équation polynomiale à coefficients entiers, est dénombrable. Par contre $\mathbf{R}$ n'est pas dénombrable. Ainsi les nombres **transcendants** (i.e. non algébriques) sont beaucoup plus nombreux que les nombres algébriques.

Et pourtant les nombres transcendants sont difficiles à exhiber ! On sait grâce à Charles Hermite (1822–1901) et Ferdinand von Lindemann (1852–1939) que e et π sont transcendants, grâce à des théorèmes démontrés en 1873 et 1882 respectivement. Celle de la constante de Gelfond, e^{π}, a fallu attendre 1934 et les travaux d'Alexandre Ossipovitch Gelfond (1906–1968) : elle résulte de la formule d'Euler et du théorème de Gelfond-Schneider en écrivant $e^{\pi} = (-1)^{i}$. Un autre exemple, constante de Gelfond-Schneider, est $2^{\sqrt{2}}$. Par contre, on ne sait pas à l'heure actuelle si π^e ou $e+\pi$ sont rationnels, algébriques ou transcendants.

12 Espaces probabilisés

La notion de hasard ne s'intègre pas totalement dans le formalisme des structures mères, d'ailleurs N. Bourbaki n'a essentiellement jamais traité des probabilités dans ses ouvrages. Elle est étroitement liée aux jeux de dés : *az-zahr* signifie *dé* en arabe, tout comme *aléa* en latin.

Un phénomène aléatoire est une expérience que l'on peut renouveler et dont le résultat échappe à toute prédiction absolue, comme un lancer de dé. Il est nécessaire de pouvoir renouveler l'expérience un grand nombre de fois (dans des conditions réputées identiques), fut-ce par la pensée, pour pouvoir parler d'aléa. L'aléatoire ne s'oppose pas nécessairement au déterminisme physique : le dé obéit aux lois de la physique qui sont, au moins à son échelle, déterministes, mais la complexité du phénomène fait que le résultat auquel on s'intéresse n'est pas prédictible. On peut prédire de façon raisonnable la vitesse, la durée du mouvement, la hauteur du rebond etc. mais pas l'équilibre stable sur lequel le dé finit sa trajectoire, i.e. la face sur laquelle il repose *in fine*.

Pour modéliser un phénomène aléatoire, on se donne un univers qui est simplement l'ensemble de tous les résultats possibles. Il peut-être plus ou moins détaillé : l'ensemble $[\![1;6]\!]$, le produit cartésien $[\![1;6]\!]\times\mathbf{R}_+$ donnant la face d'arrêt et la durée du mouvement (ou la hauteur du rebond) ou encore l'ensemble des couples formés de la face d'arrêt et de la position de la Lune à ce moment-là, voire la position de toutes les particules de l'univers pendant le trajet ! Qu'importe ce choix, qui appartient à la personne qui modélise, au final l'univers est un ensemble. Pour faire des mathématiques, il est néanmoins nécessaire que cet ensemble en soit un au sens des mathématiques !

Définition 1 - 30

Un **univers**, le plus souvent noté Ω, est un ensemble. Ses éléments sont le plus souvent notés ω et sont appelés indifféremment **épreuve**, aléa, résultat de l'expérience, événement élémentaire, événement atomique, réalisation du hasard ...

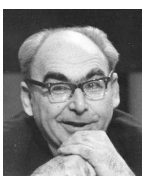
En fait un événement au sens probabiliste ne se réduit pas à un seul résultat possible, c'est une partie de Ω, autrement dit un ensemble de résultats possibles.

Définition 1 - 31

Un **événement** est une partie A de l'univers Ω. Si une épreuve ω appartient à A, on dit que A se réalise dans l'épreuve ω.

On identifie souvent ω au singleton $\{\omega\}$, mais un événement **élémentaire** est *stricto sensu* le singleton $\{\omega\}$.

Deux événements A et B sont dits **incompatibles** lorsqu'ils sont disjoints, i.e. $A\cap B=\emptyset$.



Alfréd Rényi

Pour aller plus loin

Le point de départ de la théorie des probabilités se confond donc avec la théorie des ensembles. Il existe des approches différentes, plus générales, de celle que l'on va maintenant développer et qui résulte des travaux d'Andreï Nikolaïevitch Kolmogorov (1903–1987), notamment la notion de probabilité subjective ou encore d'espace de probabilités conditionnelles introduite par Alfréd Rényi (1921–1970).

Comme par exemple la notation de limite pour une suite monotone d'ensembles, les probabilités utilisent de nombreuses conventions de notations bien utiles, mais qui ne sont pas au programme.

Somme (disjointe) ♠

Soit A et B deux parties **disjointes** d'un ensemble. On note $A \coprod B$ ou encore $A + B$ leur réunion.

Plus généralement si $(A_i)_{i\in I}$ est une famille de parties **deux à deux disjointes** d'un ensemble, on note $\coprod_{i\in I} A_i$ ou encore $\sum_{i\in I} A_i$ leur réunion.

Notation

La notation $\coprod$ est un produit renversé, ou coproduit. Cette notion a un sens en elle-même, largement en dehors du cadre du programme.

Ce coproduit ressemble parfois à une somme, comme dans le cas de l'algèbre $\mathscr{P}(\Omega)$, en ce sens que le produit est distributif par rapport à la somme. C'est évident ici puisque la réunion est la première loi de l'algèbre $\mathscr{P}(\Omega)$ et que la seconde est l'intersection. C'est d'ailleurs pourquoi certains, comme Rényi, écrivent AB au lieu de $A \cap B$. Avec ces notations les lois de de Morgan prennent une forme très simple, comme $A(B+C) = AB+AC$, ce qui fait partie de la définition d'algèbre. On dit que la catégorie des ensembles est distributive.

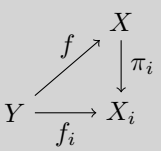
Pour aller plus loin

Mais le coproduit ressemble parfois à un produit, jusqu'à lui être isomorphe. C'est le cas pour les espaces vectoriels : on a un isomorphisme canonique entre $E \oplus F$ et $E \times F$. En fait, pour être plus précis, l'application somme de $E \times F$ dans $E + F$ est toujours surjective par définition, et n'est injective que lorsque la somme est directe. L'utilisation de cette application est d'ailleurs à la source d'une démonstration de la formule de Grassmann, qui n'est autre que la formule de Poincaré mais vue pour les espaces vectoriels ! On dit que la catégorie des espaces vectoriels est linéaire.

Qu'est-ce qu'un produit du point de vue des applications ? Si on se donne une famille $(X_i)_{i\in I}$ d'ensembles et $(f_i)_{i\in I}$ une famille d'applications ayant même espace de départ Y et telles que l'ensemble d'arrivée de f_i est X_i, pour i dans I, alors on peut fabriquer un couple (X, f) formé d'un ensemble X et d'une application f, tous deux construits à partir des données précédentes de façon « universelle ». On prend pour X le produit cartésien $\prod_{i\in I} X_i$, f l'application définie par $y \mapsto (f_i(y))_{i\in I}$ et on a la propriété fondamentale : $\forall i \in I \ f_i = \pi_i \circ f$ où π_i désigne la projection de X sur sa coordonnée i.

Remarque 1 - 9

Cette construction a un sens même si I est quelconque. On parle de solution au problème universel représenté par le dessin suivant

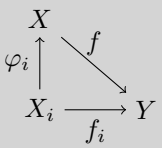


En d'autres termes on a obtenu un objet X tel que $\prod_{i\in I} X_i^Y = X^Y$ pour tout ensemble Y.

Pour aller plus loin

Recherche

On veut maintenant répondre à la question d'écrire $\prod_{i\in I} Y^{X_i} = Y^X$ pour un certain objet X, appelé coproduit. Autrement dit on veut résoudre le problème universel représenté par le dessin



où φ_i désigne une injection de X_i dans X.

Quand on dispose d'un objet plus grand (un ensemble contenant tous les X_i, ou un espace vectoriel etc.), la réponse est donnée par l'union disjointe ou la somme directe. Quand les unions ne sont pas disjointes ou les sommes non directes, on peut effectuer une construction visant à oublier qu'elles ne le sont pas.

Danger

Tout comme la notation $E \oplus F$, la notation $A + B$ est porteuse d'un double sens : elle commence par réquérir (ou affirmer) que A et B sont disjoints, puis elle désigne leur réunion.

Pour calculer des probabilités, on a besoin de savoir ce qu'on peut calculer et ce qu'on ne peut pas calculer. Ce qui est calculable peut varier en fonction des besoins mais se doit de ressembler à une algèbre. Comme $\mathscr{P}(\Omega)$ est une algèbre, on peut commencer caractériser les sous-algèbres. Il suffit de trois propriétés :

— La première propriété est le fait que l'élément neutre pour la multiplication (intersection), i.e. Ω, appartient à la sous-algèbre.

— La seconde est que l'opposé pour l'addition (union), i.e. le complémentaire, aussi.

— La dernière est la stabilité par l'addition (union).

En effet, grâce aux lois de DE MORGAN la stabilité pour la réunion et par passage au complémentaire entraîne celle par intersection. Le passage au complémentaire assure que l'élément neutre pour l'addition appartient aussi à la sous-algèbre.

Si Ω est fini, cela suffit, mais dès qu'il faudra calculer la probabilité d'un événement infini à partir des probabilités élémentaires, associées aux événements élémentaires, on ne pourra se contenter de manipuler des réunions et des intersections finies. Une tribu est définie comme une sous-algèbre avec une propriété en plus : la stabilité par réunion (ou intersection) dénombrable.

Définition 1 - 32

On appelle **tribu** sur Ω une partie $\mathcal{A}$ de $\mathscr{P}(\Omega)$ qui possède les propriétés suivantes :

Élément neutre : $\Omega \in \mathcal{A}$;

Stabilité par passage au complémentaire : $\forall A \in \mathcal{A} \quad \overline{A} \in \mathcal{A}$;

σ-additivité : si $(A_i)_{i\in I}$ est une famille au plus dénombrable d'éléments de $\mathcal{A}$, alors $\bigcup_{i\in I} A_i \in \mathcal{A}$.

Un couple $(\Omega, \mathcal{A})$ formé d'un univers et d'une tribu sur cet univers est appelé espace probabilisable. Les éléments de $\mathcal{A}$ sont appelés événements **observables**.

Danger

On prendra garde au fait que $\mathcal{A}$ ne contient pas nécessairement les singletons !

Proposition 1 - 3

Une tribu est stable par intersection dénombrable, i.e. si $(A_i)_{i\in I}$ est une famille d'éléments de $\mathcal{A}$, alors $\bigcap_{i\in I} A_i \in \mathcal{A}$.

Démonstration. Il suffit d'appliquer la σ-additivité à la famille $(\overline{A_i})_{i\in I}$, ce qui est licite par stabilité par passage au complémentaire, et passer au complémentaire. On obtient

$$\overline{\bigcup_{i\in I} \overline{A_i}} \in \mathcal{A}\,,$$

ce qui est exactement l'assertion recherchée en vertu des lois de DE MORGAN. □

Définition 1 - 33

Si $\mathcal{A}$ est une tribu sur Ω, une **probabilité** sur $(\Omega, \mathcal{A})$ est une application $\mathbf{P}$ définie sur $\mathcal{A}$, à valeurs dans $[\,0;1\,]$, telle que $\mathbf{P}(\Omega) = 1$ et, pour toute suite $(A_n)_{n\geq 0}$ d'événements deux à deux disjoints dans $\mathcal{A}$, on ait :

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n)\ .$$

Autrement dit

$$\mathbf{P}\left(\sum_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n)\ .$$

On dit que $\mathbf{P}$ est σ-additive et que $(\Omega, \mathcal{A}, \mathbf{P})$ est un espace probabilisé.

Remarque 1 - 10

▶

Si Ω est fini ou dénombrable et si $\mathcal{A} = \mathscr{P}(\Omega)$, une probabilité $\mathbf{P}$ sur $(\Omega, \mathcal{A})$ s'identifie, via la formule

$$\mathbf{P}(\{\omega\}) = p_\omega\,,$$

à une famille $(p_\omega)_{\omega\in\Omega}$ de réels positifs, sommable de somme 1. En particulier si $\Omega = \mathbf{N}$ alors $\lim p_n = 0$.

L'ensemble $\mathscr{P}(\Omega)$ est bien entendu une tribu sur Ω. Cette tribu est amplement suffisante pour faire des probabilités lorsque Ω est fini ou dénombrable. Mais même dans ce cas on a parfois besoin de restreindre ce qui est mesurable dès qu'on considère des modélisations un tant soit peu évoluées : avec deux ou plus phénomènes modélisés en même temps, on a besoin de choisir un univers qui contient toutes les informations et résultats des phénomènes, et plusieurs tribus permettent de rendre compte du fait que l'on sait des choses sur l'un ou l'autre des phénomènes mais pas sur les autres.

Aparté

Lorsque l'univers n'est pas fini ni dénombrable, se posent de nombreuses questions. Il résulte en particulier des travaux de Stanislaw ULAM que si Ω peut s'envoyer surjectivement sur $\mathbf{R}$ (on dit qu'il a la puissance du continu), alors les seules probabilités que l'on peut définir sur $\mathscr{P}(\Omega)$ sont celles qui sont à support sur une partie dénombrable D de Ω. Autant dire que cela revient à travailler sur D, et donc sur $\mathscr{P}(D)$, avec D dénombrable.

Exemples 1 - 16

— Soit Ω un ensemble fini, $\mathcal{A} = \mathscr{P}(\Omega)$. La **probabilité uniforme** sur Ω est la fonction caractérisée sur les événements élémentaires par $\mathbf{P}(\{\omega\}) = \dfrac{1}{|\Omega|}$ pour tout ω dans Ω. Autrement dit $\mathbf{P}(A) = \dfrac{|A|}{|\Omega|}$ pour toute partie A de Ω.

— Soit Ω un univers quelconque et a un élément de Ω. La **masse de Dirac** (Paul Adrien Maurice Dirac, 1902–1984) centrée en a est définie par $\delta_a(A) = \mathbb{1}_A(a)$, i.e. cette probabilité vaut 1 sur les ensembles contenant a et 0 ailleurs.

— On appelle **probabilité finie** (ou, plus exactement, à support fini) une combinaison convexe de masses de Dirac, i.e. la donnée d'épreuves $(a_i)_{1\leq i\leq n}$ et de réels positifs $(p_i)_{1\leq i\leq n}$ de somme 1 vérifiant $\mathbf{P} = \sum_{i=1}^{n} p_i \delta_{a_i}$.

Propriétés 1 - 10

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé et A et B dans $\mathcal{A}$. On a

1. $\mathbf{P}(\emptyset) = 0$;
2. $\mathbf{P}$ est additive, i.e. $\mathbf{P}(A+B) = \mathbf{P}(A) + \mathbf{P}(B)$ – on prendra garde que cela requiert que A et B soient incompatibles ;
3. $\mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$ et, en particulier, si $B \subset A$, alors $\mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(B)$.
4. $\mathbf{P}$ est croissante, i.e. si $A \subset B$, alors $\mathbf{P}(A) \leq \mathbf{P}(B)$;
5. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$;
6. $\mathbf{P}\left(\overline{A}\right) = 1 - \mathbf{P}(A)$.
   Plus généralement si $(A_i)_{1\leq i\leq n}$ est une famille finie d'événements observables, alors
7. $\mathbf{P}$ est additive, i.e. $\mathbf{P}(A_1 + \cdots + A_n) = \mathbf{P}(A_1) + \cdots + \mathbf{P}(A_n)$ – on prendra garde que cela requiert que les événements $(A_i)_{1\leq i\leq n}$ soient deux à deux incompatibles ;
8. $\mathbf{P}$ est sous-additive, i.e.

$$\mathbf{P}(A_1 \cup A_2 \cup \cdots \cup A_n) \leq \mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_n) \ .$$

Démonstration.

1. On applique la σ-additivité à la suite dont le premier élément est Ω et les autres sont tous égaux à $\emptyset$. Ce sont bien des éléments de $\mathcal{A}$ incompatibles deux à deux. On en déduit que la série de terme constant égal à $\mathbf{P}(\emptyset)$ est convergente, i.e. que ce terme est nul.
2. On applique la σ-additivité à la suite dont les deux premiers termes sont A et B et dont les autres sont tous égaux à $\emptyset$. Le résultat précédent entraîne le résultat recherché.
3. Puisque $A \setminus B = A \cap \overline{B}$, on a bien affaire à des éléments de $\mathcal{A}$. De plus l'additivité appliquée aux événements disjoints $A \cap B$ et $A \setminus B$ permet d'obtenir $\mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) = \mathbf{P}(A)$, ce qui est essentiellement l'assertion voulue. Le cas particulier résulte du fait qu'alors $A \cap B = B$.
4. Le résultat précédent, couplé au fait que $\mathbf{P}(A \setminus B)$ est positif puisque $\mathbf{P}$ est à valeurs positives, donne $\mathbf{P}(A) - \mathbf{P}(B) \geq 0$ si $A \supset B$, d'où le résultat.

5. On applique l'additivité aux événements incompatibles B et $A \setminus B$. Le point 3 permet alors de conclure.
6. On applique le point 3 à $A = \Omega$ et $B = A$.
7. On applique la σ-additivité à la suite dont les n premiers termes sont A_1, …, A_n et dont les autres sont tous égaux à $\emptyset$.
8. D'après le point 5, on a $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$, par positivité de $\mathbf{P}$, et l'assertion en résulte par récurrence immédiate.

□

Remarque 1 - 11

En vertu de la propriété 3, on peut avoir envie de noter $A - B$ l'ensemble $A \setminus B$ lorsque $B \subset A$.

Théorème 1 - 27

Continuité monotone

Soit $(A_n)_{n\in\mathbf{N}}$ une suite d'événements observables, monotone pour l'inclusion. Alors la suite $(\mathbf{P}(A_n))_{n\in\mathbf{N}}$ est monotone de même sens et, de plus, on a :

— si la suite est croissante, alors

$$\lim \mathbf{P}(A_n) = \mathbf{P}(\lim\uparrow A_n) \ ;$$

— si elle est décroissante, alors

$$\lim \mathbf{P}(A_n) = \mathbf{P}(\lim\downarrow A_n) \ .$$

Démonstration. L'assertion sur la monotonie de $(\mathbf{P}(A_n))_{n\in\mathbf{N}}$ résulte directement de la croissance de $\mathbf{P}$.

Dans le cas croissant, on pose $B_0 = A_0$ et, pour $n \geq 1$, $B_n = A_n \setminus A_{n-1}$. Alors la suite $(B_n)_{n\in\mathbf{N}}$ est à valeurs dans $\mathcal{A}$ et est formée d'événements incompatibles car, si $m < n$, $B_m \subset A_m \subset A_{n-1}$ et $B_n \cap A_{n-1} = \emptyset$. De plus, par construction, on a pour tout entier naturel n

$$A_n = \coprod_{m=0}^{n} B_m \ .$$

Il vient alors, par σ-additivité,

$$\mathbf{P}(\lim\uparrow A_n) = \mathbf{P}\left(\coprod_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(B_n) \ .$$

Or, pour tout entier naturel N, on a

$$\sum_{n=0}^{N} \mathbf{P}(B_n) = \mathbf{P}\left(\coprod_{n=0}^{N} B_n\right) = \mathbf{P}(A_N) \ .$$

Il en résulte que ces quantités ont une limite puis, par passage à la limite, on en déduit la propriété recherchée.

Dans le cas décroissant, on applique le résultat précédent à la suite $(\overline{A_n})_{n\in\mathbf{N}}$ qui est bien à valeurs dans $\mathcal{A}$ et croissante. Or

$$\lim\uparrow \overline{A_n} = \bigcup_{n=0}^{+\infty} \overline{A_n} = \overline{\bigcap_{n=0}^{+\infty} A_n} = \overline{\lim\downarrow A_n}$$

et le résultat en découle par linéarité de la limite, puisqu'il vient

$$1 - \mathbf{P}\left(\lim \downarrow A_n\right) = \lim\left(1 - \mathbf{P}\left(A_n\right)\right) \ .$$

□

Théorème 1 - 28

Inégalité de Boole – σ-sous-additivité

Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements observables, alors :

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbf{P}\left(A_n\right) \ ,$$

où la somme de la série s'entend comme la limite dans $\mathbf{R}_+ \cup \{+\infty\}$ de la suite, croissante, des sommes partielles.

Démonstration. Pour n dans $\mathbf{N}$, on pose $B_n = \bigcup_{m=0}^{n} A_m$, de sorte que (B_n) est une suite croissante d'événements de $\mathcal{A}$. De plus, par construction, pour tout entier naturel n, on a $\bigcup_{m=0}^{n} A_m = \bigcup_{m=0}^{n} B_m$ et donc aussi

$$\bigcup_{m=0}^{+\infty} A_m = \bigcup_{m=0}^{+\infty} B_m \ .$$

Par continuité croissante, on a donc $\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim \mathbf{P}\left(B_n\right)$.

Or la suite $(\mathbf{P}\left(B_n\right))$ est majorée par la suite $(\mathbf{P}\left(A_1\right) + \cdots + \mathbf{P}\left(A_n\right))$, par sous-additivité et donc, par croissance de la limite, il en est de même des limites de ces deux suites (éventuellement infinie pour la seconde). L'inégalité de Boole en résulte. □

Définition 1 - 34

On appelle événement **négligeable** tout élément A de $\mathcal{A}$ de probabilité nulle, i.e. vérifiant $\mathbf{P}\left(A\right) = 0$.

A contrario, on appelle événement **presque sûr** tout élément A de $\mathcal{A}$ de probabilité 1, i.e. vérifiant $\mathbf{P}\left(A\right) = 1$.

Une propriété est dite presque sûre quand l'événement qui lui correspond l'est.

Remarque 1 - 12

La notion de propriété presque sûre n'a réellement d'intérêt que lorsqu'on travaille avec des univers non-dénombrables.

On choisit $\Omega = \mathbf{R}$, muni d'une tribu adéquate, dite tribu des boréliens, et de la probabilité donnée par la formule suivante $\mathbf{P}\left(A\right) = \displaystyle\int_0^1 \mathbb{1}_A(t)\,\mathrm{d}t$. L'objectif ici n'est pas d'expliciter la tribu, ni de donner un sens à l'intégrale quand $\mathbb{1}_A$ n'est pas continue par morceaux. Le support de la probabilité est manifestement $[\,0;1\,]$, c'est-à-dire qu'en fait les résultats de l'expérience sont des réels de cet intervalle, et ils apparaissent avec ce qu'on peut appeler une probabilité uniforme.

Néanmoins, individuellement, aucun n'a réellement de probabilité de se réaliser car $\mathbf{P}\left(\{\omega\}\right) = 0$ pour tout ω dans $\mathbf{R}$! On sent pourtant qu'il y a une différence entre les

événements $[\,-2;-1\,]$ et $\mathbf{R}_-$. Les deux sont négligeables et le premier semble bien impossible : on aurait pu (dû ?) prendre $\Omega = [\,0;1\,]$ et alors on aurait bien au affaire à $\emptyset$ puisque $[\,-2;-1\,] \cap [\,0;1\,] = \emptyset$. Mais le second se serait alors restreint à $\{0\}$, qui n'est pas vide.

A contrario il y a également une différence entre $\mathbf{R}$ et $\mathbf{R}_+^*$. Le premier est presque sûr et on peut même dire qu'il est certain, par définition, puisqu'aucune épreuve ne peut exister en dehors de l'univers de modélisation ! Le second est aussi presque sûr, mais on peut garder un petit pincement au cœur en se demandant si, vraiment, 0 ne sortira jamais ...

Théorème 1 - 29

Un événement observable qui est inclus dans un événement négligeable l'est également.

Une réunion finie ou dénombrable d'événements négligeables est négligeable.

Démonstration. Le premier point résulte directement de la croissance de $\mathbf{P}$.

Soit $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements négligeables. Par croissance et σ-sous-additivité, il vient pour tout entier naturel N

$$\mathbf{P}\left(\bigcup_{n=0}^{N} A_n\right) \leq \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbf{P}(A_n) = 0\,.$$

Le théorème en résulte par positivité de $\mathbf{P}$. □

Définition 1 - 35

Soit A un événement observable non négligeable. Pour tout événement observable B, on définit la **probabilité conditionnelle de B sachant A**, notée $\mathbf{P}(B \mid A)$ ou parfois $\mathbf{P}_A(B)$, par la formule

$$\mathbf{P}(B \mid A) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}\,.$$

La notation $\mathbf{P}_A$ se lit *probabilité issue de A*.

Proposition 1 - 4

Soit A un événement observable non négligeable. La fonction $B \mapsto \mathbf{P}_A(B)$ est une probabilité sur la tribu $\mathcal{A}$.

Démonstration. Puisque $\mathcal{A}$ est stable par intersection, $\mathbf{P}_A$ est défini sur $\mathcal{A}$ et, par croissance et positivité de $\mathbf{P}$, $\mathbf{P}_A$ est à valeurs dans $[\,0;1\,]$.

On a $\mathbf{P}_A(\Omega) = \dfrac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1$ et, d'après les lois de DE MORGAN, pour toute famille dénombrable d'événements observables $(B_i)_{i\in\mathbf{N}}$ incompatibles deux à deux,

$$\left(\coprod_{i\in\mathbf{N}} B_i\right) \cap A = \coprod_{i\in\mathbf{N}} B_i \cap A$$

et donc, par σ-additivité de $\mathbf{P}$,

$$\mathbf{P}\left(\coprod_{i\in\mathbf{N}} B_i \;\middle|\; A\right) = \sum_{i=0}^{+\infty} \mathbf{P}(B_i \mid A)\,.$$

Par conséquent $\mathbf{P}_A$ est bien une probabilité sur $\mathcal{A}$. □

On déduit de la définition la

Théorème 1 - 30

Formule des probabilités composées

Si A est un événement observable non négligeable, alors pour tout événement observable B, on a

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B \mid A) \ .$$

Plus généralement, si $(A_i)_{1\leq i\leq n}$ est une famille d'événements observables et que l'intersection des $n-1$ premiers n'est pas négligeable, alors

$$\mathbf{P}(A_1 \cap \cdots \cap A_n) = \mathbf{P}(A_1) \cdot \mathbf{P}(A_2 \mid A_1) \cdots \mathbf{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}) \ .$$

Démonstration. Puisque $A_1 \cap \cdots \cap A_{n-1}$ n'est pas négligeable, aucun événement observable qui le contient ne l'est et donc toutes les probabilités conditionnelles apparaissant dans la formule recherchée sont bien définie. Cette dernière résulte alors d'une récurrence immédiate. □

Théorème 1 - 31

Formule des probabilités totales

Soit $(A_i)_{i\in I}$ une famille finie ou dénombrable d'événements observables non négligeables et qui est également une partition de Ω (on dit que c'est un **système complet d'événements**), et B un événement observable. On a

$$\mathbf{P}(B) = \sum_{i\in I} \mathbf{P}(A_i)\,\mathbf{P}(B \mid A_i) \ .$$

Démonstration. Puisqu'on a affaire à une partition, les lois de DE MORGAN entraînent

$$B = \coprod_{i\in I} B \cap A_i$$

et le résultat en découle par σ-additivité et la formule des probabilités composées. □

Théorème 1 - 32

Formules de BAYES

Soit A, B et $(A_i)_{i\in I}$ des événements observables non négligeables, avec I fini ou dénombrable, et tels que $(A_i)_{i\in I}$ soit une partition de Ω.

On a

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A)\,\mathbf{P}(B \mid A)}{\mathbf{P}(B)}$$

et, pour tout i dans I,

$$\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\,\mathbf{P}(B \mid A_i)}{\sum_{j\in I} \mathbf{P}(A_j)\,\mathbf{P}(B \mid A_j)} \ .$$

Thomas BAYES, 1702 – 1761.

Démonstration. La première formule résulte de la symétrisation de la formule des probabilités composées :

$$\mathbf{P}(A)\,\mathbf{P}(B \mid A) = \mathbf{P}(A \cap B) = \mathbf{P}(B)\,\mathbf{P}(A \mid B)$$

et la seconde s'obtient à partir de la première en transformant le dénominateur grâce à la formule des probabilités totales. □

Définition 1 - 36

Deux événements observables A et B sont dits **indépendants** si $\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B)$.

Remarque 1 - 13

Autrement dit, à condition d'avoir affaire à des événements non négligeables, l'indépendance se traduit par le fait que la probabilité conditionnelle et la probabilité usuelle se confondent : $\mathbf{P} = \mathbf{P}_B$, ou encore, pour tout événement A, $\mathbf{P}(A) = \mathbf{P}(A \,|\, B)$.

En d'autres termes le fait que B soit réalisé n'apporte rien quant à la réalisation ou non de A.

Définition 1 - 37

Soit $(A_i)_{i\in I}$ une famille quelconque d'événements observables. On dit qu'ils sont **mutuellement indépendants** si, pour toute partie finie J de I, on a

$$\mathbf{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbf{P}(A_i) \ .$$

Remarque 1 - 14

Cette propriété est plus forte que de simplement demander que les événements soit **deux à deux indépendants**.

Exercice

Un test sanguin est positif avec probabilité $0,95$ lorsque la personne est malade, négatif avec probabilité $0,9$ lorsqu'elle ne l'est pas. La maladie a une prévalence de 20%. Quelle est la probabilité que le test se trompe ?

Exercice

Un(e) interne arrive souvent en retard au self le matin. Son retard est estimé avec probabilité $0,4$, sauf en cas de retard la veille auquel cas la probabilité n'est que de $0,1$. On note p_n la probabilité de son retard lors du n^{e} jour. Calculer $\lim p_n$.

Exercice

On choisit au hasard un entier X dans $[\![1;2n]\!]$, puis un autre, Y, dans $[\![1;X]\!]$. On pose $p_n = \mathbf{P}(Y \leq n \leq X \,\&\, X - Y \leq n)$.

a. Calculer p_n.

b. Étudier $\lim p_n$.

Exercice

Trois étudiant(e)s de MP* jouent à shifumi en s'affrontant deux par deux. La personne qui vient de gagner affronte toujours celle qui n'a pas joué, le but étant de gagner deux parties de suites. Il n'y a pas d'ex-aequo et la probabilité de gain est réputée identique pour chacun(e).

a. Montrer que le jeu s'arrête presque sûrement. Plus précisément donner la probabilité qu'il dure plus longtemps qu'un cours de maths.

b. A-t-on intérêt à commencer à jouer ou à laisser les deux autres commencer ? Quantifier.

13 Variable aléatoire discrète

Dans tout ce chapitre $(\Omega, \mathcal{A}, \mathbf{P})$ désigne un espace probabilisé.

En général, pour décrire des phénomènes aléatoires de masse, on ne se contente pas de savoir si un événement a lieu ou pas : on mesure le phénomène grâce à une ou plusieurs grandeurs. C'est ce qui donne naissance à la notion de variable aléatoire.

Définition 1 - 38

Soit E un ensemble. Une variable aléatoire discrète définie sur Ω et à valeurs dans E est une application X de Ω dans E telle que $X(\Omega)$ soit au plus dénombrable et, pour tout x dans E, $X^{-1}(x) \in \mathcal{A}$.

Lorsque $E = \mathbf{R}$, la variable aléatoire est dite réelle.

L'ensemble E, ou $X(\Omega)$, est appelé univers-image, ensemble des valeurs prises par X ou encore ensemble des réalisations de X.

Remarque 1 - 15

La condition $X^{-1}(x) \in \mathcal{A}$ assure que les événements associés aux réalisations de X sont observables. Néanmoins on peut toujours remplacer l'espace probabilisé par un ensemble fini ou dénombrable et donc prendre $\mathcal{A} = \mathscr{P}(\Omega)$, du moins lorsqu'on n'étudie qu'une seule variable aléatoire.

L'ensemble $X(\Omega)$ est muni de la topologie discrète, d'où le nom. Attention ! une variable à valeurs dans **Q est discrète**. La topologie sur $X(\Omega)$ n'est pas celle héritée de $\mathbf{R}$.

Dans la suite X désigne une variable aléatoire discrète à valeurs dans E.

Notations

Pour x dans E, on note $(X = x)$ l'ensemble $X^{-1}(x)$.

L'ensemble $\{(X = x) \,|\, x \in X(\Omega)\}$ est appelé système complet d'événements induit par X. Plus généralement si A est une partie de E, on note $(X \in A)$ l'ensemble $X^{-1}(A)$.

Lorsque X est réelle, on note $(X \leq x)$, $(X < x)$, $(X > x)$ et $(X \geq x)$ les ensembles donnés par

$$(X \leq x) = X^{-1}\left(]-\infty; x]\right), (X < x) = X^{-1}\left(]-\infty; x[\right)$$

$$(X \geq x) = X^{-1}\left([x; +\infty[\right), (X > x) = X^{-1}\left(]x; +\infty[\right).$$

Définition 1 - 39

On note $X(\Omega) = \{x_i \,|\, i \in I\}$, avec $I \subset \mathbf{N}$. La loi de la variable aléatoire X est l'application de $\mathscr{P}(E)$ dans $\mathbf{R}_+$ donnée par

$$\mathbf{P}_X(A) = \mathbf{P}(X \in A) = \mathbf{P}\left(X \in \{x_i \,|\, i \in I \text{ et } x_i \in A\}\right) = \sum_{i \,|\, x_i \in A} \mathbf{P}(X = x_i).$$

Remarque 1 - 16

La loi de X est entièrement caractérisée par les probabilités atomiques $(\mathbf{P}_X(x_i))_{i \in I}$ puisque, pour toute partie A de E, $X^{-1}(A)$ est réunion disjointe d'événements appartenant au système complet d'événements induit par X.

Dans la suite on notera $p_i = \mathbf{P}_X(x_i)$. On a donc

$$\mathbf{P}_X(A) = \sum_{i \,|\, x_i \in A} p_i \text{ et en particulier } \mathbf{P}_X(E) = \sum_{i \in I} p_i = 1.$$

Notation

Si $\mathcal{L}$ est une application de $\mathscr{P}(E)$ dans $\mathbf{R}_+$, on dit que X suit la loi $\mathcal{L}$, et on note $X \sim \mathcal{L}$, si $\mathbf{P}_X = \mathcal{L}$.

Si Y est une seconde variable aléatoire définie sur un espace probabilisé éventuellement différent et à valeurs dans E, on dit que X et Y ont même loi, et on note $X \sim Y$, si $\mathbf{P}_X = \mathbf{P}'_Y$.

Pour aller plus loin

Fonction de répartition

Lorsque X est réelle, on appelle fonction de répartition de X la fonction donnée par $F(x) = \mathbf{P}(X \leq x)$.

C'est une fonction en escalier croissante, de limite nulle en $-\infty$ et 1 en $+\infty$. Ses points de discontinuité sont les éléments de $X(\Omega)$ et si x est un tel point, on a

$$F(x) - F(x^-) = F(x) - \lim_{y \to x^-} F(y) = \mathbf{P}_X(x) \ .$$

14 Lois usuelles

Définition 1 - 40

Loi uniforme

Soit $[\![a;b]\!]$ un intervalle entier, avec a et b dans $\mathbf{Z}$ vérifiant $a \leq b$. On dit que X suit une loi uniforme sur $[\![a;b]\!]$, et on note $X \sim \mathcal{U}([\![a;b]\!])$, si $X(\Omega) = [\![a;b]\!]$ et si, pour $a \leq k \leq b$, on a

$$\mathbf{P}_X(k) = \frac{1}{b-a+1} \ .$$

Autrement dit tous les entiers entre a et b ont même probabilité.

Définition 1 - 41

Loi de Bernoulli

Soit p dans $]0;1[$. On dit que X suit une loi de Bernoulli de paramètre p, et on note $X \sim \mathcal{B}(p)$, si $X(\Omega) = \{0,1\}$ et $\mathbf{P}_X(1) = p$. On a alors $\mathbf{P}_X(0) = 1-p$.

On modélise ainsi une expérience n'ayant que deux issues possibles, appelées succès et échec. Le succès est l'événement $(X = 1)$ tandis que l'échec est l'événement $(X = 0)$.

Définition 1 - 42

Loi binomiale

Soit n dans $\mathbf{N}^*$ et p dans $]0;1[$. On dit que X suit une loi binomiale de paramètres n et p, et on note $X \sim \mathcal{B}(n,p)$, si $X(\Omega) = [\![0;n]\!]$ et, pour $0 \leq k \leq n$,

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

Autrement dit X modélise le nombre de succès lors de la réalisation de n épreuves de Bernoulli de paramètre p, indépendantes les unes des autres.

On considère maintenant un événement très rare, i.e. p proche de 0, mais que l'on observe durant une grande période de temps, i.e. n très grand. On peut supposer par exemple que la probabilité que l'événement se produise est constante, ce qui est le cas si

on s'intéresse à un événement comme la désintégration d'atomes au sein d'un gramme d'atomes instables. Le nombre d'atomes est tellement grand que la désintégration de l'un d'entre eux n'affecte pas vraiment la probabilité d'oberver une désintégration. De même la probabilité de désintégration est tellement petite que la probabilité d'observer deux désintégrations simultanément est négligeable. On modélise le phénomène en posant que la probabilité d'observer une désintégration durant un petit intervalle de temps Δt est proportionnelle à la longueur cet intervalle, par un facteur α. On découpe alors un intervalle de temps en n petits intervalles de taille Δt, de sorte que le nombre total de désintégrations observées durant ce laps de temps suit une loi binomiale de paramètres n et $\alpha\Delta t$. On remarque de plus que, dans cette modélisation, $n\Delta t$ représente le laps de temps d'observation et est donc une constante. Pour étudier ce phénomène, on doit donc s'intéresser aux lois binomiales de paramètres n et p_n avec np_n constant et n tendant vers l'infini.

Théorème 1 - 33

Approximation d'une loi binomiale par une loi de Poisson

Soit (X_n) des variables aléatoires réelles suivant des lois binomiales. On suppose $X_n \sim \mathcal{B}(n, p_n)$ et $\lim np_n = \lambda$. Alors, pour tout entier k dans $\mathbf{N}$, on a

$$\lim_{n\to+\infty} \mathbf{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}\ .$$

Démonstration. Pour k et n entiers, on a

$$\begin{aligned}
\mathbf{P}(X_n = k) &= \binom{n}{k} p_n^k (1-p_n)^{n-k} \\
&= \frac{\prod_{j=0}^{k-1}(n-j)}{k!} p_n^k \exp\left((n-k)\ln(1-p_n)\right) \\
&= \frac{\prod_{j=0}^{k-1}[(n-j)p_n]}{k!} \exp(-np_n + o(np_n)) \\
&= \frac{\prod_{j=0}^{k-1}(np_n + o(1))}{k!} \exp(-np_n + o(1)) \\
&= e^{-\lambda}\frac{\lambda^k}{k!} + o(1)
\end{aligned}$$

et l'assertion en découle. □

Pour aller plus loin

On dit que la suite de variables aléatoires (X_n) converge en loi vers une loi de Poisson, du nom de Siméon Denis Poisson, 1781–1840.

Définition 1 - 43

Loi de Poisson

Soit λ un réel strictement positif. On dit que X suit une loi de Poisson de paramètre λ, et on note $X \sim \mathcal{P}(\lambda)$, si $X(\Omega) = \mathbf{N}$ et, pour k dans $\mathbf{N}$, $\mathbf{P}_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$.

Une loi de Poisson modélise les événements rares, c'est une loi des petits nombres.

Remarque 1 - 17

Une somme d'un grand nombre n de variables de BERNOULLI indépendantes et de petit paramètre p suit approximativement une loi de POISSON de paramètre np. On verra que np représente l'espérance mathématique de cette somme de variables aléatoires. Dans la pratique on préconise $n \geq 30$, $p \leq 0,1$ et $np \leq 15$.

Exemples 1 - 17

La loi de POISSON est utilisée dans de nombreux modèles comme

— nombre d'accidents dus par ruade de cheval dans la cavalerie prussienne,

— nombre de suicides d'enfants,

— nombre d'arrivées de bâteaux au port,

— nombre de clients se présentant à une caisse ou un guichet,

— nombre de passagers à un arrêt de bus,

— nombre de connexions à un serveur web,

— nombre de communications,

— nombre de mutations en biologie,

— nombre de désintégrations,

etc. durant une période de temps fixée.

Remarque 1 - 18

Une loi de POISSON fournit une approximation d'une loi binomiale, plutôt que le contraire.

Définition 1 - 44

Loi géométrique

Soit p dans $]0;1[$. On dit que X suit une loi géométrique de paramètre p, et on note $X \sim \mathcal{G}(p)$, si $X(\Omega) = \mathbf{N}^*$ et, pour k dans $\mathbf{N}^*$, $\mathbf{P}_X(k) = p(1-p)^{k-1}$.

Une loi géométrique modélise l'instant de premier succès lors de la réalisation d'une infinité dénombrable d'épreuves de BERNOULLI.

Remarque 1 - 19

La loi géométrique est, par convention, la loi du premier succès. On pourrait néanmoins s'intéresser à la loi du nombre d'échecs avant le premier succès. C'est également une loi géométrique, mais elle est définie sur $\mathbf{N}$. On a donc $\mathbf{P}_X(n) = p(1-p)^n$ et on note $X \sim \mathcal{G}_{\mathbf{N}}(p)$. En cas de besoin la loi géométrique du premier succès se note $\mathcal{G}_{\mathbf{N}^*}(p)$ pour la distinguer de la précédente.

Danger

L'instant de premier succès n'est pas vraiment une variable aléatoire au sens strict.

En effet il se peut qu'il n'y ait jamais de succès. On choisit $\Omega = \{0,1\}^{\mathbf{N}^*}$ et, pour ω dans Ω avec $\omega = (\omega_i)_{i \in \mathbf{N}^*}$, on définit $X(\omega)$ par $X(\omega) = \min\{i \in \mathbf{N}^* \,|\, \omega_i = 1\}$, ce qui a un sens sauf si l'ensemble sur lequel on prend le minimum est vide, i.e. ω est la suite nulle. Par conséquent X est défini sur Ω' avec $\Omega' = \Omega \setminus \{(0)\}$. Il faudrait encore définir $\mathcal{A}$ et $\mathbf{P}$, mais quoiqu'il en soit, si ces objets existent, puisque

$$\mathbf{P}(\Omega') = \mathbf{P}_X(\mathbf{N}^*) = \sum_{k=1}^{+\infty} p(1-p)^{k-1} = \frac{p}{1-(1-p)} = 1\,,$$

on aura $\mathbf{P}(\Omega') = 1$, ce qui revient à dire d'une part que $\{(0)\}$ est négligeable et d'autre part que X est défini presque partout.

Pour aller plus loin

On appelle cylindre élémentaire une partie de Ω déterminée par un nombre fini de coordonnées, i.e. l'intersection d'un nombre fini d'images réciproques par les projections canoniques de Ω sur $\{0,1\}$. Plus généralement on appelle cylindre une partie C de la forme

$$C = \{\omega \in \Omega \,|\, (\omega_1, \ldots, \omega_r) \in K\}$$

avec $r \in \mathbf{N}^*$ et $K \subset \{0,1\}^r$. La tribu $\mathcal{A}$ est alors la tribu engendrée par les cylindres, i.e. la plus petite tribu les contenant, et on définit $\mathbf{P}$ sur les cylindres par

$$\mathbf{P}(C) = \frac{1}{2^r} \sum_{(\omega_1,\ldots,\omega_r)\in K} p^{\sum \omega_i}(1-p)^{\sum(1-\omega_i)},$$

autrement dit en considérant les éléments de K comme des réalisations d'une succession de variables de Bernoulli de paramètre p. C'est un théorème important et difficile de Constantin Carathéodory qui permet d'affirmer que la probabilité $\mathbf{P}$ existe, et est entièrement déterminée par ses valeurs sur les cylindres.

Proposition 1 - 5

Une variable aléatoire géométrique est sans mémoire, c'est-à-dire que pour k et ℓ entiers, on a

$$\mathbf{P}\left(X > k + \ell \,|\, X > \ell\right) = \mathbf{P}(X > k)\,.$$

De plus $\mathbf{P}(X > k) = (1-p)^k$.

Démonstration. Soit k et ℓ des entiers, il vient

$$\mathbf{P}(X > k) = \sum_{n=k+1}^{+\infty} p(1-p)^{n-1} = \frac{p(1-p)^k}{1-(1-p)} = (1-p)^k$$

et

$$\begin{aligned}
\mathbf{P}\left(X > k + \ell \,|\, X > \ell\right) &= \frac{\mathbf{P}(X > k+\ell, X > \ell)}{\mathbf{P}(X > \ell)} \\
&= \frac{\mathbf{P}(X > k+\ell)}{\mathbf{P}(X > \ell)} \\
&= \frac{(1-p)^{k+\ell}}{(1-p)^{\ell}} \\
&= (1-p)^k \\
&= \mathbf{P}(X > k)\,.
\end{aligned}$$

□

Autrement dit la probabilité d'attendre au moins dix lancers avant d'obtenir un *pile* est la même que celle d'attendre vingt-cinq lancers sachant que les quinze premiers étaient des *face*.

Aparté

On parle de loi sans mémoire ou sans vieillissement. Ces lois permettent de modéliser des phénomènes où le vieillissement n'intervient pas.

Théorème 1 - 34

Lois sans mémoire

Soit X une variable aléatoire à valeurs dans $\mathbf{N}^*$ telle que, pour tous entiers k et ℓ, on ait $\mathbf{P}(X = k+1) > 0$ et $\mathbf{P}(X > k+\ell \,|\, X > \ell) = \mathbf{P}(X > k)$. Alors X suit une loi géométrique de paramètre $\mathbf{P}_X(1)$.

Démonstration. On pose $p = \mathbf{P}(X = 1)$ et $q = 1-p$, de sorte qu'on a $q = \mathbf{P}(X > 1)$. Pour ℓ dans $\mathbf{N}^*$, on a

$$\frac{\mathbf{P}(X > \ell+1)}{\mathbf{P}(X > \ell)} = \mathbf{P}(X > \ell+1 \,|\, X > \ell) = \mathbf{P}(X > 1) = q$$

et donc $\mathbf{P}(X > \ell) = q^\ell$. Il en résulte

$$\mathbf{P}(X = \ell) = \mathbf{P}(X > \ell-1) - \mathbf{P}(X > \ell) = q^{\ell-1}(1-q) = p(1-p)^{\ell-1},$$

i.e. $X \sim \mathcal{G}(p)$. □

15 Algèbre linéaire

Cette section reprend le cours de MPSI et en étend brièvement le formalisme, dans l'esprit des travaux fondateurs de Grassmann et Peano.

Définition 1 - 45

Sous-espace vectoriel

Un **sous-espace vectoriel** de E est une partie F de E qui, munie des bi-resctrictions des lois de $(E, +, \star)$, est un $\mathbf{K}$-espace vectoriel.

Définition 1 - 46

Application linéaire

Une **application linéaire** est un morphisme d'espaces vectoriels, i.e. une application respectant la structure de $\mathbf{K}$-espace vectoriel.

Exemples 1 - 18

Espaces vectoriels de référence

La notion de corps inclut celle d'espace vectoriel et ainsi $\mathbf{K}$ est un $\mathbf{K}$-espace vectoriel. Plus généralement un produit cartésien est muni d'une structure vectorielle par multiplication coordonnée par coordonnée. Ainsi $\mathbf{K}^n$, $\mathscr{M}_{n,p}(\mathbf{K})$, $\mathbf{K}[X]$, $\mathbf{K}^{\mathbf{N}}$ sont des $\mathbf{K}$-espaces vectoriels. Il en va de même pour $\mathbf{K}^I$ si I est un ensemble quelconque.

Le sous-espace $\{0\}$ est un sous-espace vectoriel minimal de E : il est inclus dans tout sous-espace vectoriel de E. De même E est un sous-espace vectoriel maximal de E. Tout sous-espace vectoriel distinct de $\{0\}$ et E est dit non-trivial.

Exemples 1 - 19

Sous-espaces vectoriels

Les sous-espaces vectoriels non-triviaux du plan $\mathbf{K}^2$ sont des droites vectorielles. Ceux de l'espace $\mathbf{K}^3$ sont des droites et des plans vectoriels. Les parties $\mathbf{K}_n[X]$ et $\mathbf{K}^{(\mathbf{N})}$ sont des sous-espaces vectoriels respectivement de $\mathbf{K}[X]$ et $\mathbf{K}^{\mathbf{N}}$. Si I est un intervalle de $\mathbf{R}$, $C^0(I, \mathbf{K})$ est un sous-espace vectoriel de $\mathbf{K}^I$.

Exemples 1 - 20

Sous-espaces vectoriels et applications linéaires

Si E et F sont deux $\mathbf{K}$-espaces vectoriels, l'ensemble $\mathscr{L}(E,F)$ des applications linéaires de E dans F est également un $\mathbf{K}$-espace vectoriel. Si u appartient à $\mathscr{L}(E,F)$ on note $\mathrm{Ker}(u)$ le **noyau** de u, i.e. $u^{-1}(0)$, et $\mathrm{Im}(u)$ son **image**, i.e. $u(E)$. Si E' et F' sont des sous-espaces vectoriels respectivement de E et F, $u(E')$ et $u^{-1}(F')$ sont des espaces vectoriels. Si $F=\mathbf{K}$, u est appelée **forme linéaire** et si elle est non nulle son noyau est appelé **hyperplan** de E.

Remarque 1 - 20

Composition des applications linéaires

La composée de deux applications linéaires en est une, de même que la réciproque d'un isomorphisme (i.e. une application linéaire bijective). Une application linéaire u est injective si et seulement si son noyau $\mathrm{Ker}(u)$ est réduit à $\{0\}$.

Remarque 1 - 21

Produit de composition et produit matriciel

L'application $(u,v)\mapsto u\circ v$ est bilinéaire. La linéarité à gauche est tautologique tandis que la linéarité à droite résulte de la définition d'application linéaire.

Pour A dans $\mathscr{M}_{n,p}(\mathbf{K})$ l'application $X\mapsto AX$ de $\mathscr{M}_{p,1}(\mathbf{K})$ dans $\mathscr{M}_{n,1}(\mathbf{K})$ est linéaire. La composée de deux telles applications associées respectivement à A et B est donnée par le produit matriciel, qui est donc bilinéaire et associatif.

On étend aux familles les objets introduits pour les ensembles finies : intersection quelconque, combinaison linéaire.

Remarque 1 - 22

Intersection

Une intersection quelconque de sous-espaces vectoriels de E en est un.

Définition 1 - 47

Soit I un ensemble quelconque d'indices, on appelle **famille à support fini** de scalaires indexée par I toute famille presque nulle $(\lambda_i)_{i\in I}$, i.e. telle que $\{i\in I\,|\,\lambda_i\neq 0\}$ est fini. Cet ensemble est le support de la famille $(\lambda_i)_{i\in I}$. On note $\mathbf{K}^{(I)}$ leur ensemble.

Propriété 1 - 11

Combinaison linéaire

Soit $(\lambda_i)_{i\in I}\in\mathbf{K}^{(I)}$ une famille de scalaires de support fini S et $(x_i)_{i\in I}\in E^I$ une famille de vecteurs de E. Alors pour toute partie finie J de I contenant S, on a $\displaystyle\sum_{i\in J}\lambda_i x_i=\sum_{i\in S}\lambda_i x_i$, i.e. la quantité $\displaystyle\sum_{i\in J}\lambda_i x_i$ ne dépend pas de J. On la note $\displaystyle\sum_{i\in I}\lambda_i x_i$ et on dit que c'est une combinaison linéaire de la famille $(x_i)_{i\in I}$.

Pour toute famille $(x_i)_{i\in I}$ de E^I, on a une application linéaire canonique de $\mathbf{K}^{(I)}$ dans E donnée par $(\lambda_i)_{i\in I}\mapsto\displaystyle\sum_{i\in I}\lambda_i x_i$.

Propriété 1 - 12

Caractérisation des applications linéaires

Une application entre $\mathbf{K}$-espaces vectoriels est linéaire si et seulement si elle préserve les combinaisons linéaires, ce qui peut se réduire à l'une des deux assertions suivantes : $\forall\lambda\in\mathbf{K},\forall(x,y)\in E^2,(u(\lambda x)=\lambda u(x))\wedge(u(x+y)=u(x)+u(y))$ ou encore $\forall(\lambda,\mu)\in\mathbf{K}^2,\forall(x,y)\in E^2,u(\lambda x+\mu y)=\lambda u(x)+\mu u(y)$.

La notion de combinaison linéaire est la notion clef en algèbre linéaire. Elle permet notamment de décrirer les sous-espaces vectoriels.

Propriété 1 - 13

Espace engendré

Soit $A \in \mathscr{P}(E)$ où E est un $\mathbf{K}$-espace vectoriel. L'intersection de tous les sous-espaces vectoriels de E contenant A est un sous-espace vectoriel de E contenant A et c'est le plus petit (au sens de l'inclusion) sous-espace vectoriel de E contenant A. On le note indifféremment $\mathrm{V}ect(A)$, $\langle A \rangle$ ou $\mathrm{V}ect(a)_{a\in A}$ et, si $A = \{x_i \,|\, i \in I\}$, on le note également $\mathrm{V}ect(x_i)_{i\in I}$. On a

$$x \in \mathrm{V}ect(A) \Leftrightarrow \exists (\lambda_a)_{a\in A} \in \mathbf{K}^{(A)}, \quad x = \sum_{a\in A} \lambda_a a \,.$$

On en déduit quatre caractérisations (au moins) des sous-espaces vectoriels.

Propriété 1 - 14

Une partie F de E en est un **sous-espace vectoriel** si et seulement si $F = \mathrm{V}ect(F)$. On en déduit les caractérisations suivantes :

- $\forall (\lambda_x)_{x\in F} \in \mathbf{K}^{(F)}$, $\sum_{x\in F} \lambda_x x \in F$ (avec la convention qu'une somme vide est égale à 0) ;
- $0 \in F$ et $\forall \lambda \in \mathbf{K}$, $\forall (x,y) \in F^2$, $\lambda x \in F \wedge x + y \in F$;
- $0 \in F$ et $\forall \lambda \in \mathbf{K}$, $\forall (x,y) \in F^2$, $\lambda x + y \in F$;
- $0 \in F$ et $\forall (\lambda,\mu) \in \mathbf{K}^2$, $\forall (x,y) \in F^2$, $\lambda x + \mu y \in F$.

La notion de combinaison linéaire, une famille étant fixée, permet de décrire des vecteurs de E à partir de familles de scalaires. La nature de cette description, selon qu'elle permet de décrire tous les vecteurs, de façon unique ou non, est une notion cruciale et amène à la définition des propriétés les plus importantes des familles de vecteurs.

Définition 1 - 48

Soit $(x_i)_{i\in I}$ une famille de vecteurs de E. Si l'application canonique de $\mathbf{K}^{(I)}$ dans E donnée par $(\lambda_i)_{i\in I} \mapsto \sum_{i\in I} \lambda_i x_i$ est injective, surjective, bijective, on dit qu'on a affaire à une **famille libre**, une **famille génératrice** ou une **base** respectivement.

Propriété 1 - 15

Image d'une famille

L'image d'une base de E, ou plus généralement d'une famille génératrice, par une application linéaire u dans $\mathscr{L}(E,F)$ engendre $\mathrm{Im}(u)$. De plus u est un isomorphisme si et seulement si l'image d'une base de E est une base de F.

Réciproquement une application d'une base de E dans F étant donnée, il existe une et une seule application linéaire dans $\mathscr{L}(E,F)$ la prolongeant.

Exemples 1 - 21

Base canonique

L'espace nul admet $\emptyset$ comme base et $\mathbf{K}$ admet $\{1\}$ comme base en tant que $\mathbf{K}$-espace vectoriel. Elle est appelée base canonique. Plus généralement la base canonique d'un espace produit est obtenu par produit cartésien. Ainsi les bases canoniques de $\mathbf{K}^n$, $\mathscr{M}_{n,p}(\mathbf{K})$, $\mathbf{K}_n[X]$ et $\mathbf{K}[X]$ sont respectivement $((\delta_{ij})_{1\leq j\leq n})_{1\leq i\leq n}$, $(E_{i,j})_{(i,j)\in[\![1;n]\!]\times[\![1;p]\!]}$, $(X^k)_{0\leq k\leq n}$ et $(X^k)_{k\in\mathbf{N}}$.

Définition 1 - 49

Formes coordonnées

Si E est un $\mathbf{K}$-espace vectoriel et $(e_i)_{i\in I}$ en est une base. Les formes coordonnées associées à cette base sont les formes linéaires $(e_i^*)_{i\in I}$ uniquement déterminées par les conditions $e_i^*(e_j) = \delta_{ij}$. Les formes linéaires $(e_i^*)_{i\in I}$ constituent une base de $\mathscr{L}(E, \mathbf{K})$.

♠

L'espace vectoriel $\mathscr{L}(E, \mathbf{K})$, appelé dual de E, est aussi noté E^*.

Définition 1 - 50

Hyperplans et codimension

Si E est un espace de dimension finie n et H un hyperplan obtenu comme $\mathrm{Ker}(u)$, où u est une forme linéaire, une décomposition $\sum_{i=1}^{n} a_i e_i^*$ de u correspond à une équation $\sum_{i=1}^{n} a_i x_i = 0$ de H, relativement à la base $(e_i)_{1\leq i\leq n}$. Pour H donné, il n'y a pas unicité de u et donc pas non plus de l'équation, mais il y a unicité à multiplication par un scalaire non nul près.

Un hyperplan H de E est un sous-espace vectoriel de codimension 1, i.e. $\dim(E) - \dim(H) = 1$. Plus généralement si $\dim(E) - \dim(F) = m$, on dit que F est de codimension m.

L'intersection de m hyperplans est un espace de codimension au plus m et, réciproquement, tout espace de codimension m est intersection de m hyperplans. Un tel espace est donc défini par un système de m équations (non unique).

♠

La codimension de F peut se définir même si E est de dimension infinie, c'est alors la dimension de l'espace vectoriel quotient E/F. Elle correspond au nombre d'équations nécessaires pour définir F.

Définition 1 - 51

Applications linéaires, bases et matrices

Étant donné deux $\mathbf{K}$-espaces vectoriels de dimensions finies E et F et deux bases (e) et (f) de E et F respectivement, avec $(e) = (e_j)_{1\leq j\leq q}$ et $(f) = (f_i)_{1\leq i\leq p}$ une application linéaire entre E et F correspond de façon unique à une matrice $(a_{ij})_{(i,j)\in[\![1;p]\!]\times[\![1;q]\!]}$ dans $\mathscr{M}_{p,q}(\mathbf{K})$ telle que, pour tout x dans E on ait $u(x) = \sum_{i=1}^{p}\sum_{j=1}^{q} a_{ij} e_j^*(x) f_i$.

La matrice précédente est appelée matrice de u relativement aux bases (e) et (f) et notée $\mathrm{Mat}_{(e),(f)}(u)$. Ainsi le vecteur de coordonnées $(x_i)_{1\leq i\leq q}$ dans (e) a pour image le vecteur de coordonnées $(\sum_{j=1}^{q} a_{ij}x_j)_{1\leq i\leq p}$ dans (f).

L'application $u \mapsto \mathrm{Mat}_{(e),(f)}(u)$ est un isomorphisme de $\mathbf{K}$-espaces vectoriels.

Remarque 1 - 23

La composition entre applications linéaires correspond à la multiplication matricielle. Si $E = F$ et $(e) = (f)$, on écrit $\mathrm{Mat}_{(e)}(u)$ et l'application $u \mapsto \mathrm{Mat}_{(e)}(u)$ est un isomorphisme de $\mathbf{K}$-algèbres entre $\mathrm{End}(E)$ et $\mathscr{M}_n(\mathbf{K})$.

Aparté

On écrit aussi $\mathrm{Mat}(u;(e),(f))$ ou $\mathrm{Mat}_{(f)}^{(e)}(u)$, ainsi que $u = \sum_{i=1}^{p}\sum_{j=1}^{q} a_{ij} e_j^* \otimes f_i$ et on a $\mathscr{L}(E,F) \simeq E^* \otimes F$. Les trois dernières écritures rappellent que la dépendance en E est **contravariante** alors que celle en F est **covariante**. La matrice de $e_j^* \otimes f_i$ est la matrice E_{ij} de la base canonique, i.e. $a_{ij} = E_{ij}^*(\mathrm{Mat}_{(e),(f)}(u))$.

Définition 1 - 52

Application linéaire canoniquement associée à une matrice

À M dans $\mathscr{M}_{p,q}(\mathbf{K})$ on associe canoniquement l'application linéaire de $\mathscr{M}_{q,1}(\mathbf{K})$ dans $\mathscr{M}_{p,1}$ donnée par $X \mapsto MX$. Sa matrice relativement aux bases canoniques est alors M. Le noyau et l'image de M sont ceux de l'application linéaire associée.

On peut également identifier canoniquement $\mathscr{M}_{q,1}(\mathbf{K})$ à $\mathbf{K}^q$ et $\mathscr{M}_{p,1}$ à $\mathbf{K}^p$, puisque ce sont des espaces de mêmes dimensions et admettant chacun une base canonique. Ainsi M est également canoniquement associée à une application linéaire de $\mathbf{K}^q$ dans $\mathbf{K}^p$.

La notion de base est fondamentale en algèbre linéaire. Le travail de Grassmann a justement été de permettre de ne pas travailler avec une seule base, i.e. de se détacher du modèle de $\mathbf{K}^n$, pour pouvoir *choisir* des bases adaptées à chaque problème.

Théorème 1 - 35

Lemme fondamental de la théorie de la dimension

Soit $(x_i)_{i\in I}$ une famille finie de vecteurs de E et $(y_j)_{j\in J}$ une famille finie de vecteurs, tous combinaisons linéaires des (x_i) avec $\mathrm{Card}\, J > \mathrm{Card}(I)$. Alors la famille (y_j) est liée.

Théorème 1 - 36

Dimension

Un espace vectoriel est dit de **dimension finie** s'il admet une famille génératrice finie. Tout tel espace admet une base et toutes ses bases ont un même cardinal fini.

Plus précisément on a le résultat suivant, valide en toute dimension, et conséquence des mêmes arguments.

Théorème 1 - 37

Base incomplète et base extraite

Soit E un espace vectoriel de dimension finie, L une famille libre dans E et G une famille génératrice de E, vérifiant $L \subset G$. Alors il existe une base B de E vérifiant $L \subset B \subset G$. En particulier pour $G = E$, cela permet de compléter toute famille libre en une base et, pour $L = \emptyset$, cela permet d'extraire une base de toute famille génératrice.

Stricto sensu le théorème de la base incomplète est le cas $G = E$, celui de la base extraite est le cas $L = \emptyset$ et celui de l'existence de base le cas $L = \emptyset$ et $G = E$.

♠

Un espace vectoriel qui n'est pas de dimension finie est dit de dimension infinie. Un tel espace admet également une base et toutes ses bases sont de cardinal infini. Le théorème précédent est encore vrai en dimension infinie.

Le résultat en dimension infinie résulte, comme en dimension finie, de l'existence d'élément maximal parmi les familles libres ou, comme on voudra, de l'existence d'élément minimal parmi les familles génératrices. En dimension finie la maximalité ou la minimalité s'expriment en termes de cardinaux (finis) et l'existence de base résulte du lemme fondamental de la théorie de la dimension.

Dimensions de référence

La dimension d'un produit d'espaces vectoriels (de dimensions finies) est la somme de leurs dimensions. Ainsi les espaces $\mathbf{K}^n$, $\mathscr{M}_n(\mathbf{K})$, $\mathscr{M}_{n,p}(\mathbf{K})$ et $\mathbf{K}_n[X]$ sont de dimensions respectives n, n^2, np et $n+1$.

Exemples 1 - 22

L'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre n (sous forme résolue) est un espace vectoriel de dimension n. Il en va de même pour l'ensemble des suites satisfaisant une relation de récurrence linéaire homogène (à coefficients constants) d'ordre n.

Rang d'une famille de vecteurs

La dimension de l'espace engendré par une famille de vecteurs est appelé rang de la famille. Le rang est inférieur à la fois au cardinal de la famille et à la dimension de l'espace ambiant.

Définition 1 - 53

Rang d'une application linéaire ou d'une matrice

Le rang d'une application linéaire u, noté $\mathrm{rg}(u)$ est la dimension de $\mathrm{Im}(u)$, i.e. le rang de l'image d'une base par u. Le rang est invariant par composition (à gauche comme à droite) par un isomorphisme.

Définition 1 - 54

Le rang d'une matrice est le rang de l'application linéaire canoniquement associée ou, ce qui revient au même, au rang de la famille de ses vecteurs colonnes.

Le rang est en particulier invariant par pré-composition ou post-composition par une application linéaire bijective. Matriciellement il est donc invariant par multiplication par une matrice inversible.

Rang de la transposée

C'est un fait non totalement évident que le rang d'une matrice est *aussi* celui de ses vecteurs lignes. Autrement dit le rang d'une matrice est invariant par transposition.

Remarque 1 - 24

Caractérisation par la dimension

Dans un espace vectoriel E de dimension finie n, une famille de cardinal n est une base si et seulement si elle est libre ou génératrice. De plus tout sous-espace vectoriel de E est de dimension inférieure à n et E est le seul sous-espace de dimension exactement n.

Soit u une application linéaire entre deux espaces vectoriels de même dimension finie. Alors u est bijective si et seulement si elle est injective (i.e. $\mathrm{Ker}(u) = \{0\}$) et si et seulement si elle est surjective. En particulier un endomorphisme d'un espace vectoriel de dimension finie est inversible à droite si et seulement s'il l'est à gauche ou encore est inversible.

Propriété 1 - 16

Il existe un unique $\mathbf{K}$-espace vectoriel de dimension finie n donnée, à isomorphisme près.

L'intersection est, on l'a vu, compatible à la notion de sous-espace vectoriel : une intersection quelconque en est un. Il n'en va pas de même pour la réunion et ni non plus pour le complémentaire. On y supplée en considérant le sous-espace engendré par la réunion. En termes de familles génératrices il s'agit donc bien d'une réunion : si (f) et (g) engendrent respectivement les sous-espaces vectoriels F et G, alors $F+G$ est l'espace engendré par $(f)\cup(g)$.

Le travail avec des bases n'étant pas intrinsèque, la définition de la somme peut se faire en considérant les sommes de vecteurs. Il est alors bien plus commode de raisonner avec les applications linéaires naturelles associées.

Somme de sous-espaces vectoriels

Soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels d'un même espace vectoriel E. On a une application linéaire canonique

$$\begin{array}{rccc}\sigma: & \displaystyle\prod_{i\in I} E_i & \to & E \\ & (x_i)_{i\in I} & \mapsto & \displaystyle\sum_{i\in I} x_i\end{array}.$$

Remarque 1 - 25

Elle est surjective de sorte qu'on peut définir la somme $\displaystyle\sum_{i\in I} E_i$ comme l'image de cette application linéaire : $\displaystyle\sum_{i\in I} E_i = \mathrm{Im}(\sigma)$.

Somme directe

Soit I un ensemble fini et $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. On dit que la somme $\displaystyle\sum_{i\in I} E_i$ est directe si l'application

$$\begin{array}{rccc}\sigma: & \displaystyle\prod_{i\in I} E_i & \to & \displaystyle\sum_{i\in I} E_i \\ & (x_i)_{i\in I} & \mapsto & \displaystyle\sum_{i\in I} x_i\end{array}$$

Définition 1 - 55

est bijective.

Par construction σ est surjective. Elle est donc bijective si et seulement si elle est injective, ou encore si et seulement si son noyau est réduit à $(0)_{i\in I}$.

Autrement dit, la somme est directe si et seulement si l'écriture de tout élément de $\displaystyle\sum_{i\in I} E_i$ comme somme d'éléments des $(E_i)_{i\in I}$ est unique. Quand la somme est directe, on écrit $\oplus_{i\in I} E_i$ au lieu de $\displaystyle\sum_{i\in I} E_i$.

Danger

Tout comme la notation $\exists!$, la notation $\oplus$ est dangereuse. En effet c'est à la fois une notation et une assertion mathématique. On ne peut en particulier par l'utiliser pour définir un nouvel objet, i.e. on ne peut pas écrire *Soit E l'espace vectoriel défini par $E = F\oplus G$* sans avoir démontré avant que F et G sont en somme directe.

Remarque 1 - 26

Caractérisation de la somme directe

Soit E_1 et E_2 deux sous-espaces vectoriels de E. On a

$$E_1 + E_2 = E_1 \oplus E_2 \iff E_1 \cap E_2 = \{0\}$$
$$\iff (\forall (x,y) \in E_1 \times E_2 \quad x+y=0 \Longrightarrow x = y = 0) \ .$$

♥

La caractérisation précédente est un **faux ami**. En effet, plus généralement, si $(E_i)_{i\in I}$ est une famille de sous-espaces vectoriels de E, on a

$$\sum_{i\in I} E_i = \oplus_{i\in I} E_i \iff \forall i \in I \qquad E_i \cap \sum_{j\in I\setminus\{i\}} E_j = \{0\}$$
$$\iff \left(\forall (x_i)_{i\in I} \in \prod_{i\in I} E_i \ \left(\sum_{i\in I} x_i = 0 \Longrightarrow (\forall i \in I \, , x_i = 0) \right) \right) \ .$$

Remarque 1 - 27

Caractérisation de la somme directe par la dimension

Si E est de dimension finie on a $\dim(\sum_{i\in I} E_i) \leq \sum_{i\in I} \dim(E_i)$ avec égalité si et seulement si la somme est directe.

Propriété 1 - 17

Détermination d'une application linéaire

Par propriété universelle de la somme directe, si E et F sont des $\mathbf{K}$-espaces vectoriels et si $E = \oplus_{i\in I} E_i$, alors $\mathscr{L}(E,F) = \prod_{i\in I} \mathscr{L}(E_i, F)$ et plus précisément étant donné des applications linéaires u_i dans $\mathscr{L}(E_i,F)$, il existe une unique application linéaire les prolongeant simultanément à E.

Théorème 1 - 38

Existence de supplémentaire

Soit E un espace vectoriel de dimension finie et F un sous-espace de E. Alors il existe au moins un sous-espace G de E tel que $E = F \oplus G$. On dit alors que G est un supplémentaire de F dans E.

Deux sous-espaces vectoriels F et G de E sont supplémentaires si et seulement si $\dim(F) + \dim(G) = \dim(E)$ et $F \cap G = \{0\}$.

C'est une conséquence directe du théorème de la base incomplète. En particulier

♠

Ce résultat est encore valide en dimension infinie.

Théorème 1 - 39

Théorème du rang

Soit u dans $\mathscr{L}(E,F)$, avec E et F deux $\mathbf{K}$-espaces vectoriels, et E' un supplémentaire de $\mathrm{Ker}(u)$ dans E. Alors

$$u_{|E'} \ : \ E' \simeq \mathrm{Im}(u) \ .$$

En particulier, en dimension finie, on a $\mathrm{rg}(u) = \dim(E) - \dim(\mathrm{Ker}(u))$.

Exemple 1 - 23

Isomorphisme entre supplémentaires

Soit F un sous-espace vectoriel de E et G_1 et G_2 deux supplémentaires de F, de sorte qu'on a $E = F \oplus G_1$ et $E = F \oplus G_2$. Soit p le projecteur canonique sur G_1 associé à la décomposition $E = F \oplus G_1$, i.e. le projecteur sur G_1 parallèlement à F. Alors p induit, par restriction à G_2 un isomorphisme de G_2 sur G_1.

En effet $\mathrm{Ker}(p) = F$ et G_2 en est un supplémentaire, tandis que $\mathrm{Im}(u) = G_1$.

Danger

Ne pas confondre les mots supplémentaire et complémentaire : un supplémentaire est un espace vectoriel tandis que le complémentaire est un ensemble.

Il n'y a pas unicité d'un supplémentaire de F, par exemple parce qu'il n'y a pas unicité de la façon de compléter une base de F en une base de E.

Exemple 1 - 24

Formule de Grassmann

Soit E et F deux sous-espaces vectoriels d'un même espace vectoriel. On a $\dim(E \times F) = \dim(E \cap F) + \dim(E + F)$ ou encore

$$\dim(E + F) + \dim(E \cap F) = \dim(E) + \dim(F) \ .$$

Exemple 1 - 25

Endomorphismes remarquables

Si E est un $\mathbf{K}$-espace vectoriel $\mathrm{End}(E)$, i.e. $\mathscr{L}(E, E)$, est une $\mathbf{K}$-algèbre. Ses éléments sont appelés endomorphismes de E. Les endomorphismes scalaires sont les multiples de l'identité, elle-même notée Id_E. L'application $\lambda \mathrm{Id}_E$ est appelée **homothétie** de rapport λ. L'algèbre $\mathrm{End}(E)$ n'est pas commutative, sauf en dimension inférieure à 1, et son centre (i.e. l'ensemble des endomorphismes commutant à tous les autres) est formé des homothéties.

Si $E = F \oplus G$, le **projecteur** sur F parallèlement à G est l'unique endomorphisme dont la bi-restriction à F est l'identité et celle à G est l'endomorphisme nul. On le note p_F^G et on parle aussi de projection sur F parallèlement à G. Géométriquement, pour x dans E, le projeté $p_F^G(x)$ est l'unique point d'intersection de F avec l'espace affine passant par x de direction G, i.e. $(x+G) \cap F = \{p_F^G(x)\}$. Un endomorphisme p est un projecteur si et seulement si $p^2 = p$ et alors c'est un projecteur sur $\mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$.

La **symétrie** par rapport à F et parallèlement à G est l'isomorphisme s_F^G donné par $s_F^G = 2p_F^G - \mathrm{Id}_E$, i.e. sa birestriction à F est Id_F et celle à G est $-\mathrm{Id}_G$. Le symétrique $s_F^G(x)$ est l'unique point de E tel que le milieu de x et $s_F^G(x)$ soit $p_F^G(x)$. Un endomorphisme s est une symétrie si et seulement si c'est une involution, i.e. $s^2 = \mathrm{Id}_E$, et alors c'est un symétrie par rapport à $\mathrm{Ker}(s - \mathrm{Id}_E)$ et parallèlement à $\mathrm{Ker}(s + \mathrm{Id}_E)$.

Propriété 1 - 18

Projecteurs et décomposition de l'unité

Soit $(p_i)_{i \in I}$ la famille (finie) de projecteurs associée à une décomposition en somme directe $E = \oplus_{i \in I} E_i$. On a les relations suivantes :

1. Idempotence - $\forall i \in I,\ p_i^2 = p_i$
2. Orthogonalité - $\forall (i, j) \in I^2,\ i \neq j \Rightarrow p_i p_j = 0$
3. Décomposition de l'unité - $\displaystyle\sum_{i \in I} p_i = \mathrm{Id}_E$.

Propriété 1 - 19

Bases adaptées et dimensions

Soit F un sous-espace vectoriel de E. Tous les supplémentaires de F ont même dimension, appelée co-dimension de F et notée $\text{codim}(F)$. (♠) Ce résultat est encore valide en dimension infinie.

Une base $\mathcal{B}_E$ de E est dite adaptée à F si on peut en extraire une base $\mathcal{B}_F$ de F. Une telle base peut être obtenue en complétant une base de F en une base de E. Le complémentaire relativement à $\mathcal{B}_E$ de la base $\mathcal{B}_F$ de F ainsi obtenue est une base d'un supplémentaire de F.

Une base $\mathcal{B}_E$ de E est dite adaptée à une somme directe $\oplus_{i\in I}E_i$ si on peut en extraire une base de chacun des E_i. Si $\mathcal{B}_\rangle$ est une base de E_i ainsi obtenue, les $(\mathcal{B}_\rangle)_{i\in I}$ sont deux à deux disjoints. Si de plus $E = \oplus_{i\in I}E_i$, les $(\mathcal{B}_\rangle)_{i\in I}$ forment un partage de $\mathcal{B}_E$.

Exemple 1 - 26

Équations différentielles linéaires à coefficients constants

L'ensemble des solutions de l'équation $y' + ay = 0$ sur $\mathbf{R}$ est une droite vectorielle. Toutes ses bases sont donc proportionnelles. Si on rajoute un second membre, lui-même solution d'une telle équation, e.g. $y' + ay = e^{bx}$, on obtient une droite affine.

Si $a + b \neq 0$, on peut considérer le plan vectoriel engendré par les fonctions données par e^{-ax} et e^{bx}. L'ensemble des solutions est alors une droite affine contenue dans ce plan et de direction donnée par la première fonction (considérée comme un vecteur de ce plan). Toute droite horizontale coupant l'axe vertical, on en déduit qu'il existe une solution de cette équation différentielle proportionnelle à la seconde fonction. On peut ainsi complètement résoudre cette équation en appliquant une recette : chercher β tel que βe^{bx} soit solution, i.e. tel que $\beta(b + a)e^{bx} = e^{bx}$, puis rajouter un multiple quelconque de e^{-ax} pour obtenir la solution générale ou un multiple précis si on dispose d'une condition initiale. Ainsi la formule $y(x) = \alpha e^{-ax} + \beta\frac{e^{bx}-e^{-ax}}{a+b}$ définit l'unique solution du problème de Cauchy : $y' + ay = e^{bx}$ et $y(0) = \alpha$.

On voit ici l'intérêt d'une base. L'espace ambient est un espace de dimension infinie, celui des fonctions de classe C^∞ de $\mathbf{R}$ dans $\mathbf{C}$. On en extrait un sous-espace vectoriel (de dimension 2), puis on choisit une base de ce sous-espace.

Exemple 1 - 27

Ordre 2 homogène

Les équations du type $y'' + ay' + b = 0$ se résolvent en considérant l'équation polynomiale associée $X^2+aX+b = 0$. Si cette équation a deux solutions distinctes, i.e. si $a^2 + 4b \neq 0$, les solutions à valeurs complexes sont de la forme $\alpha y_1 + \beta y_2$ où $y_k(x) = e^{r_k x}$ en notant r_1 et r_2 sont les deux racines de $X^2 + aX + b$.

La base (y_1, y_2) n'est toutefois pas la plus adaptée si a et b sont réels alors que r_1 et r_2 ne le sont pas. On peut préférer une base prenant des valeurs réelles, e.g. $(\frac{1}{2}(y_1 + y_2), \frac{1}{2i}(y_1 - y_2))$. Si $r_k = \lambda \pm i\varphi$, il s'agit de $e^{\lambda x}\cos(\varphi x)$ et $e^{\lambda x}\sin(\varphi x)$.

Si $a^2 + 4b = 0$ et r est racine double de $X^2 + aX + b$, une base de l'espace des solutions est donnée par e^{rx} et xe^{rx}.

En fonction du problème considérée la première base trouvée n'est donc pas toujours la bonne et tout l'objet de l'algèbre linéaire est de comprendre comment trouver des bases adaptées et calculer avec.

Aparté

On verra que d'une façon générale une équation différentielle linéaire à coefficients constants d'ordre n dont le second membre est lui-même solution d'une telle équation d'ordre m, admet un espace de solutions qui est naturellement un sous-espace affine de dimension n d'un espace vectoriel de dimension $n+m$.

Exemple 1 - 28

Suites récurrentes linéaires d'ordre 2

Le même principe prévaut dans l'étude des suites complexes vérifiant $\forall n \in \mathbf{N}$ $u_{n+2}+au_{n+1}+bu_n = 0$. On résout la même équation polynomiale $X^2+aX+b=0$ et on dispose d'une base de l'espace vectoriel des solutions donnée par les suites (r_1^n) et (r_2^n), si $a^2+4b \neq 0$, ou (r^n) et (nr^n) sinon. L'analogie est frappante.

Pour voir que les espaces de solutions, dans le cas homogène, sont des espaces vectoriels, il suffit d'introduire une application linéaire dont c'est le noyau. Ici il s'agit de $y \mapsto y''+ay'+by$, d'un côté, et de $(u_n)_{n\in\mathbf{N}} \mapsto (u_{n+2}+au_{n+1}+bu_n)_{n\in\mathbf{N}}$, de l'autre. Ce sont des endomorphismes l'une de $C^\infty(\mathbf{R},\mathbf{C})$, l'autre de $\mathbf{C}^\mathbf{N}$.

Aparté

On verra qu'on les construits de la même façon à partir de la dérivation ou de l'opération de décalage respectivement, i.e. de $y \mapsto y'$ et $(u_n)_{n\in\mathbf{N}} \mapsto (u_{n+1})_{n\in\mathbf{N}}$ respectivement. En notant u cette application (linéaire), on étudie le noyau de $u\circ u + au + b\mathrm{Id}$, ce que l'on pourrait noter $u^2+au^1+bu^0$ en prenant soin d'interpréter les exposants relativement au produit de composition. Autrement dit, en notant $P = X^2+aX+b$, le noyau de $P(u)$ s'étudie en trouvant les racines de $P\ldots$ ce qui n'est plus très surprenant.

Pour aller plus loin

La bonne analogie entre continu et discret est donnée par l'opérateur de dérivation discrète, i.e. $(u_n)_{n\in\mathbf{N}} \mapsto (u_{n+1}-u_n)_{n\in\mathbf{N}}$. Le noyau de cet opérateur est formé des suites constantes, tandis que celui de la dérivation est formé des fonctions constantes.

Pour clore ces rappels, on donne quelques propriétés du rang.

Proposition 1 - 6

Composition et rang

Soit $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$ des applications linéaires de rangs finis. On a

— $\mathrm{rg}(v\circ u) \leq \min(\mathrm{rg}(u),\mathrm{rg}(v))$

— si v est injective, $\mathrm{rg}(v\circ u) = \mathrm{rg}(u)$

— si u est surjective, $\mathrm{rg}(v\circ u) = \mathrm{rg}(v)$.

En particulier la composition à gauche ou à droite par une application linéaire bijective ne modifie pas le rang.

Exercices

Ensembles, applications, lois

1 - 1 Ⓢ ★

Soit E un ensemble. On définit une loi interne sur $\mathscr{P}(E)$ par $A\sharp B = E \setminus (A \cap B)$.

Montrer que l'on peut définir les lois $\cap$, $\cup$ et Δ ainsi que le passage au complémentaire rien qu'en utilisant la loi $\sharp$.

1 - 2 Ⓢ ★

Soit E un ensemble, A et B des parties de E telles qu'il existe une partie X de E telle que $A \cap X \subset B \cap X$ et $A \cup X \subset B \cup X$.

Montrer $A \subset B$.

1 - 3 Ⓢ ★

Soit E un ensemble et A une partie de E. On appelle fonction caractéristique de A l'application χ_A de E dans $\mathbf{F}_2$ définie par $\chi_A(x) = 1$ si et seulement si $x \in A$.

a. Pour A et B des parties de E, on note $A\Delta B$ leur différence symétrique, i.e. $A\Delta B = A \cup B \setminus A \cap B$. Montrer que $\chi_A + \chi_B$ est la fonction caractéristique de $A\Delta B$.

b. Montrer que $\chi_A\chi_B$ est la fonction caractéristique de $A \cap B$.

c. En déduire une structure de $\mathbf{F}_2$-algèbre sur $\mathscr{P}(E)$, muni des opérations Δ et $\cap$;

d. Montrer que tout élément de $\mathscr{P}(E)$ est égal à son carré.

On dit que $\mathscr{P}(E)$ est une algèbre de Boole, du nom de George Boole (1815–1864). Voir exercice 1 - 11

1 - 4 Ⓢ ★★

Soit $f : X \to Y$, $g : Y \to Z$ et $h : Z \to X$.

a. On suppose que parmi les trois applications $h \circ g \circ f$, $g \circ f \circ h$ et $f \circ h \circ g$ deux sont surjectives et la dernière injective. Montrer que f, g et h sont bijectives.

b. Même conclusion en échangeant surjectif et injectif.

1 - 5 Ⓢ ★★ Théorème de factorisation

Soit X, Y et Z trois ensembles non vides.

a. Soit $f : X \to Y$ et $h : X \to Z$. Montrer : $(\exists g : Y \to Z, h = g \circ f) \Leftrightarrow (\forall (x, x') \in X^2, f(x) = f(x') \Rightarrow h(x) = h(x'))$. Justifier l'appellation du théorème. À quelle condition g est-elle uniquement déterminée ?

b. ♠ Soit $g : Y \to Z$ et $h : X \to Z$. Montrer : $(\exists f : X \to Y,\ h = g \circ f) \Leftrightarrow (\forall x \in X,\ \exists y \in Y,\ h(x) = g(y))$.

Indication : on utilisera l'axiome du choix.

1 - 6 Ⓢ ★★ Idempotents

Soit E un ensemble fini non vide muni d'une loi de composition interne associative notée $\intercal$.

Montrer qu'il existe e dans E tel que $e \intercal e = e$.

1 - 7 Ⓢ C ★★ Anneaux et idempotents

Soit A un anneau tel que $x^3 = x$ pour tout x dans A.

a. Déterminer les éléments nilpotents de A.

b. Soit $e \in A$ tel que $e^2 = e$. On pose $b = ea(1-e)$ et $c = (1-e)ae$. Calculer b^2 et c^2. En déduire que $ae = ea$.

c. En considérant les carrés de A, montrer que A est commutatif.

1 - 8 Ⓢ U 2013 ★★ Associaèdre de Stasheff

Soit $(A, \times)$ un magma associatif.

Démontrer la proposition suivante généralement admise : pour tout entier n strictement positif, tout n-uplet $(a_1, \ldots, a_n)$ d'éléments de A et tout parenthésage « admissible » de la multiplication $a_1 \times \cdots \times a_n$ (par exemple, pour $n = 4$, $(a_1 \times a_2) \times (a_3 \times a_4)$, $a_1 \times (a_2 \times (a_3 \times a_4))$, $(a_1 \times (a_2 \times a_3)) \times a_4$ sont des parenthésages admissibles), le résultat de la multiplication est le même.

Relations d'équivalence et relations d'ordre

1 - 9 Ⓢ ★★

Soit E un ensemble ordonné et f, $g : E \to E$, deux applications croissantes qui vérifient $f \circ f = f$, $g \circ g = g$ (f et g sont idempotentes) et, $\forall x \in E$, $f(x) \leq x \leq g(x)$.

Montrer que $f \circ g$ et $g \circ f$ sont idempotentes.

1 - 10 Ⓢ ★★★ Point fixe

Soit E un ensemble non vide ordonné dans lequel toute partie non vide possède une borne inférieure et une borne supérieure, et soit $f : E \to E$ croissante.

Montrer que f possède un point fixe.

1 - 11 Ⓢ ★★★ Anneaux de Boole

Soit $(A, +, \cdot)$ un anneau de Boole, i.e. tel que tout élément x de A vérifie $x^2 = x$.

a. Montrer $\forall x \in A$, $2x = 0$. (On rappelle qu'on a, par définition, $2x = x + x$.)

b. Montrer que A est commutatif.

c. ►Calculer $xy(x+y)$ pour x et y dans A. En déduire que si A possède strictement plus de deux éléments, il n'est pas intègre (i.e. on peut trouver deux éléments non nuls dont le produit est nul).

d. Montrer par un exemple que A peut être de cardinal 2. Peut-il être de cardinal 3 ?

e. Soit E un ensemble non vide. Montrer que $(\mathscr{P}(E), \Delta, \cap)$ est un anneau de BOOLE.

f. Soit $\mathcal{R}$ la relation dans A définie par $x\mathcal{R}y \equiv xy = x$. Montrer que $\mathcal{R}$ est une relation d'ordre. On note $\leq$ cet ordre. Est-il compatible à l'addition et à la multiplication, i.e. $\forall (x,y,z) \in A^3$, $(x \leq y \Rightarrow x+z \leq y+z) \wedge ((0 \leq x \wedge 0 \leq y) \Rightarrow 0 \leq xy)$?

1 - 12 ★★★ Ordre noethérien

Soit $\mathcal{R}$ une relation d'ordre sur un ensemble E. Pour x et y dans E, on note $x < y$ si $x\mathcal{R}y$ et $x \neq y$. Autrement dit $<$ est l'ordre strict associé à $\mathcal{R}$. Montrer que les deux conditions suivantes sont équivalentes :

1. Toute partie non vide X de E admet un élément minimal, i.e. $\exists x \in X\ \forall y \in E\ y < x \Longrightarrow y \notin X$.
2. il n'existe pas de suite infinie (x_n) d'éléments de E telle qu'on ait $\forall n \in \mathbf{N}\ x_{n+1} < x_n$.

On dira que $\mathcal{R}$ est un ordre noethérien, du nom d'Emmy NOETHER (1882-1935).

1 - 13 ★★★ Récurrence noethérienne

On utilise la notion d'ordre noethérien définie à l'exercice 1 - 12.

a. Vérifier que $\leq$ est un ordre noethérien sur $\mathbf{N}$.

b. Soit $E = \mathbf{N} \setminus \{0,1\}$. Vérifier que la relation $a \mid b$ définit un ordre noethérien sur E. Quels sont les éléments minimaux de E ?

c. Vérifier que l'ordre lexicographique est un ordre noethérien sur $\mathbf{N}^2$. Quels sont les éléments minimaux de $\mathbf{N}^2$?

d. En déduire un ordre noethérien sur $\mathbf{Q}_+$.

e. Soit $\mathcal{R}$ un ordre noethérien sur un ensemble E. On notera que $\mathcal{R}$ n'est pas nécessairement un ordre total. On note $<$ l'ordre strict associé. Montrer le principe de récurrence noethérienne : soit X une partie de E vérifiant

$$\forall x \in E,\ (\forall y \in E\ y < x \Longrightarrow y \in X) \Longrightarrow x \in X$$

alors $X = E$.

Entiers naturels

1 - 14 Ⓢ ★★ Suite de FIBONACCI

La suite de FIBONACCI est donnée par $F_0 = 0$, $F_1 = 1$ et, $\forall n \in \mathbf{N}^*$, $F_{n+1} = F_n + F_{n-1}$.

a. Démontrer que pour tous entiers naturels n et m, on a $F_{n+m} = F_{n+1}F_m + F_nF_{m-1}$.

b. En déduire $F_{n+m} \wedge F_m = F_m \wedge F_n$ puis $F_m \wedge F_n = F_{m \wedge n}$ (où $a \wedge b$ désigne le PGCD de a et b).

c. Soit $\varphi = (1+\sqrt{5})/2$. Démontrer qu'on a $F_n = \dfrac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}}$ et que F_n est l'entier le plus proche de $\dfrac{\varphi^n}{\sqrt{5}}$.

1 - 15 Ⓢ ★★

Soit f une fonction strictement croissante de $\mathbf{N}$ dans lui-même et multiplicative (i.e. $\forall (m,n) \in \mathbf{N}^2$, $f(mn) = f(m)f(n)$).

Montrer que si f admet un point fixe supérieur (ou égal) à 2, c'est l'identité.

1 - 16 Ⓢ ★★★ Représentation de ZECKENDORFF

On note $(F_n)_{n\in\mathbf{N}}$ la suite de FIBONACCI privée de ses deux premiers termes ou, ce qui revient au même, la suite définie par $F_0 = 1$, $F_1 = 2$ et, $\forall n \in \mathbf{N}$, $F_{n+2} = F_n + F_{n+1}$. On définit pour $n \in \mathbf{N}$, σ_n et S_n par

$$\sigma_n = \sum_{0 \leq k \leq \frac{n}{2}} F_{n-2k} \quad \text{et} \quad S_n = \sum_{k=0}^{n} F_k\ .$$

Si $(a_k)_{0\leq k\leq n}$ est une suite finie de chiffres valant 0 ou 1, avec $a_n = 1$, on note $\overline{a_n a_{n-1} \cdots a_0}$ l'entier m défini par $m = \sum_{k=0}^{n} a_k F_k$. Une telle écriture est appelée représentation de FIBONACCI de m. Elle n'est a priori pas unique. Si, de plus, la suite ne prend pas la valeur 1 de façon consécutive (i.e. $a_k = 1 \Rightarrow a_{k+1} = 0$) on dit qu'on a affaire à une représentation de ZECKENDORFF. Par exemple $8 = \overline{1100} = \overline{10000}$, la première représentation n'étant pas une représentation de ZECKENDORFF, alors que la seconde l'est.

a. Donner une relation entre σ_n et F_{n+1}, ainsi qu'entre S_n et F_{n+2}.

b. Déterminer les représentations de FIBONACCI de 44 et préciser si ce sont des représentations de ZECKENDORFF.

c. On se donne une représentation de ZECKENDORFF $\overline{a_n a_{n-1} \cdots a_0}$ d'un entier m.

i. Montrer $m \leq \sigma_n$.

ii. En déduire que F_n est le plus grand des termes de $(F_k)_{k\in\mathbf{N}}$ majorés par m.

iii. Démontrer que tout entier admet exactement une représentation de ZECKENDORFF.

iv. Calculer la représentation de ZECKENDORFF de 444.

d. Déterminer les représentations de ZECKENDORFF de σ_n et S_n.

e. On s'intéresse au nombre $\delta(m)$ de représentations de FIBONACCI de m.

i. Montrer qu'on a : $\delta(m) = 1$ si et seulement si m est l'un des σ_n.

ii. Établir la relation de récurrence $\delta(F_n) = 1 + \delta(F_{n-2})$ et en déduire la valeur de $\delta(F_n)$.

f. Démontrer que, sur $[\![F_n - 1; F_{n+1} - 1]\!]$, le graphe de δ est symétrique.

1 - 17 Ⓢ ★★★ **Algorithme d'Euclide**

Soit (F_n) la suite de Fibonacci (voir 1 - 14).

a. Soit x et y des entiers tels que $0 < y < x$ et $d = x \wedge y$. Montrer que si l'algorithme d'Euclide appliqué à x et y admet n étapes, alors $x \geq dF_{n+2}$ et $y \geq dF_{n+1}$.

b. Montrer que le nombre d'étapes dans l'algorithme d'Euclide appliqué à x et y avec $0 \leq y \leq x$ est inférieur à 5 fois le nombre de chiffres de y en base 10.

c. Montrer que le nombre d'étapes précédent est au plus $\frac{3}{2}\log_2(y) + 1$.

Ensembles dénombrables

1 - 18 Ⓢ ★★★

On appelle nombre algébrique tout nombre complexe z tel qu'il existe un polynôme P dans $\mathbf{Z}[X]$ dont il soit racine, i.e. tel qu'il existe un entier naturel non nul n et des entiers relatifs $(a_0, \ldots, a_n)$ vérifiant $a_n \neq 0$ et $a_n z^n + \cdots + a_1 z + a_0 = 0$. On note $\overline{\mathbf{Q}}$ l'ensemble des nombres algébriques.

Montrer que $\overline{\mathbf{Q}}$ est dénombrable.

Dénombrements et probabilités élémentaires

1 - 19 Ⓢ ★

Soit E un ensemble fini de cardinal n.

Calculer le nombre de recouvrements de E du type (E_1, E_2) avec $E = E_1 \cup E_2$ et $\text{Card}(E_1 \cap E_2) = 1$.

1 - 20 ★

On lance une pièce et l'on obtient *pile* avec une probabilité p dans $]0;1[$.

Quelle est la probabilité, en effectuant autant de lancers que nécessaire, d'obtenir deux *piles* consécutifs sans avoir eu auparavant une séquence *pile-face* ?

1 - 21 Ⓢ ★★ **Chevalier de Méré**

Quel est le plus probable : jouant avec un dé, obtenir au moins une fois 6 en quatre coups, ou, jouant avec deux dés, obtenir au moins une fois deux 6 en vingt-quatre coups ?

On s'interdira un calcul instrumenté et on utilisera une inégalité de concavité pour conclure.

1 - 22 Ⓢ ★★

De combien de manières différentes peut-on placer p objets sur un damier $n \times n$ de sorte qu'il y ait au plus un objet par ligne et par colonne ?

1 - 23 Ⓢ ★★

Étant donné n points du plan affine réel, avec $n \geq 4$, tels que trois d'entre eux ne soient jamais alignés et les droites qu'ils définissent deux à deux soient deux à deux sécantes en des points distincts.

En combien de points ces droites se coupent-elles ?

1 - 24 Ⓢ ★★

Quel est le nombre de diagonales d'un polygone convexe à n côtés ?

1 - 25 ★★

Un employé New Yorkais habite au carrefour de 5^{th} street et 3^{rd} avenue. Il travaille au carrefour de 12^{th} street et 13^{th} avenue et décide de prendre un chemin différent tous les jours (mais de longueur minimale !) et ce, le plus longtemps possible.

Combien de jours tiendra-t-il ?

1 - 26 Ⓢ ★★

a. A quelle condition nécessaire et suffisante portant sur (Y, Z) dans $\mathscr{P}(X)^2$ l'application φ de $\mathscr{P}(X)$ dans $\mathscr{P}(Y) \times \mathscr{P}(Z)$ qui à A associe $(A \cap Y, A \cap Z)$ est-elle injective ? surjective ? bijective ?

b. Établir la relation valable pour tous entiers (p, q, r) vérifiant $r \leq p + q$

$$\binom{p+q}{r} = \sum_{\max(0,r-q)}^{\min(p,r)} \binom{p}{i}\binom{q}{r-i}.$$

1 - 27 Ⓢ ★★

Soit n et p des entiers naturels.

a. Démontrer, pour $p \leq n$, $\displaystyle\sum_{k=p}^{n} \binom{k}{p} = \binom{n+1}{p+1}$ et en déduire $\displaystyle\sum_{k=0}^{n} k^3$.

b. Calculer $\displaystyle\sum_{k=0}^{n} k\binom{n}{k}$.

1 - 28 Ⓢ **X 2002** ★★

a. Déterminer le nombre F_n de façons de recouvrir un damier de dimension $2 \times n$ par des dominos blancs de dimension 1×2.

b. Montrer que, si n est assez grand, alors F_n est la partie entière de $\frac{1}{2} + \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}$.

1 - 29 Ⓢ ★★★ **Inégalité triangulaire**

Pour A et B deux événements observables. On pose $d(A, B) = \mathbf{P}(A \Delta B)$. Si $A \cup B$ est négligeable, on pose $d'(A, B) = 0$ et sinon $d'(A, B) = \dfrac{\mathbf{P}(A \Delta B)}{\mathbf{P}(A \cup B)}$.

a. Démontrer $d(A, C) \leq d(A, B) + d(B, C)$, avec A, B, C des événements observables.

b. Même question pour d'.

1 - 30 Ⓢ **X 2007** ★★★ **Dérangements**

Un dérangement est une permutation sans point fixe.

a. †Calculer $\displaystyle\sum_{k=0}^{p}(-1)^k\binom{n}{k}\binom{n-k}{p-k}$ pour $p \leq n$.

b. Soit D_n le nombre de dérangements d'un ensemble E à n éléments. On pose $D_0 = 1$. Montrer $\displaystyle\sum_{k=0}^{n}\binom{n}{k}D_{n-k} = n!$.

c. Établir la formule $D_n = n!\displaystyle\sum_{k=0}^{n}\frac{(-1)^k}{k!}$.

d. Quelle est la limite de $D_n/n!$?

e. Quel est le nombre moyen de points fixes d'une permutation de S_n ?

1 - 31 **X 2013** ★★★

Soit E un ensemble non vide, $A_1, \cdots, A_n$ des parties de E et $P_E(A_1, \ldots, A_n)$ l'ensemble de toutes les parties de E que l'on peut former avec des opérations ensemblistes (union, intersection, complémentaire) en utilisant $A_1, \ldots, A_n$.

a. Déterminer $\max_{A_1,\ldots,A_n}|P_E(A_1, \ldots, A_n)|$ en fonction de n.

b. Déterminer quand ce maximum est atteint.

1 - 32 Ⓢ **X 2003** ★★★ **Formule de Legendre**

a. Soit p un nombre premier et n un entier naturel. Montrer

$$v_p(n!) = \sum_{k=1}^{+\infty}\left\lfloor\frac{n}{p^k}\right\rfloor$$

où $v_p(x)$ désigne la *valuation p-adique* d'un entier x. On montrera au préalable que la somme est en réalité finie.

b. En déduire que, pour m et n dans $\mathbf{N}$, $\binom{n+m}{n}$ divise $\binom{2n}{n}\binom{2m}{m}$.

Espaces probabilisés

1 - 33 ★ **Indépendance mutuelle**

Exhiber trois événements deux à deux indépendants qui ne sont pas mutuellement indépendants.

1 - 34 Ⓢ ★

Soit $(A_n)_{n\in\mathbf{N}}$ une suite d'événements deux à deux incompatibles. Montrer $\lim \mathbf{P}(A_n) = 0$.

1 - 35 Ⓢ ★

Soit A et B deux événements observables non négligeables. Montrer $\mathbf{P}(A \mid B) \geq \mathbf{P}(A) \iff \mathbf{P}(B \mid A) \geq \mathbf{P}(B)$.

1 - 36 Ⓢ ★

On lance deux dés, un vert et un bleu, et on introduit les événements

— A : *le dé vert tombe sur 1*,

— B : *l'un des deux dés tombe sur 1*

— C : *la somme des deux dés est 7*.

Quels sont les événements indépendants parmi eux ?

1 - 37 Ⓢ ★

On note $\mathcal{A}$ l'ensemble de tous les intervalles de $\mathbf{R}$ de la forme $]a;b]$ avec $-\infty \leq a \leq b \leq +\infty$ et les conventions $]a;a] = \emptyset$ et $]a;+\infty] =]a;+\infty[$. Est-ce une tribu ?

1 - 38 Ⓢ ★

Soit $(A_i)_{i\in I}$ une partition dénombrable de Ω. Décrire la plus petite tribu la contenant.

1 - 39 Ⓢ ★ **Anagrammes**

a. Combien y a-t-il d'anagrammes du mot *ananas* ?

b. En permutant au hasard ses lettres, quelle est la probabilité de retomber sur le mot *ananas* ?

1 - 40 Ⓢ **CCP 18** ★ **Covariance**

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisés et A et B deux observables.

a. Montrer, pour $x \in [0;1]$, $x(1-x) \leq \frac{1}{4}$.

b. Montrer, pour A et B incompatibles,

$$\mathbf{P}(A)\,\mathbf{P}(B) \leq \frac{1}{4}\,.$$

c. On pose $c = \mathbf{P}(A \cap B) - \mathbf{P}(A)\,\mathbf{P}(B)$.

i. Montrer

$$c = \mathbf{P}(A \cap B)\,\mathbf{P}\left(\overline{A}\right) - \mathbf{P}\left(\overline{A} \cap B\right)\mathbf{P}(A)\,.$$

ii. Montrer $|\mathbf{P}(A \cap B) - \mathbf{P}(A)\,\mathbf{P}(B)| \leq \frac{1}{4}$.

iii. Déterminer les cas d'égalité.

1 - 41 Ⓢ ★★ **Dé pipé**

On considère un dé vert équilibré et un dé bleu pipé. Ce dernier a une chance sur 3 de tomber sur un 6 et ses autres résultats sont équiprobables. Deux personnes s'affrontent en lançant le dé suivant la règle suivante : le plus haut score l'emporte, et en cas d'égalité c'est celui qui a lancé le dé vert qui gagne.

Quel dé vaut-il mieux prendre pour jouer ?

1 - 42 Ⓢ ★★ **Continuité monotone en $\emptyset$**

Soit $\mathbf{P}$ une application additive de $\mathcal{A}$ dans $[0;1]$ telle que $\mathbf{P}(\Omega) = 1$. On suppose que pour toute suite décroissante $(A_n)_{n\in\mathbf{N}}$ d'événements observables vérifiant $\lim\downarrow A_n = \emptyset$, on a $\lim \mathbf{P}(A_n) = 0$.

Montrer que $\mathbf{P}$ est une probabilité.

1 - 43 Ⓢ ★★★

Une limace se déplace sur les arêtes d'un cube posé au sol. Au moment où elle démarre son périple, elle est sur un des sommets, au sol. On place deux feuilles de laitues, l'une au sommet opposé à son sommet de départ et l'autre à la verticale de ce dernier (donc au sol). La limace commence alors à se déplacer et, à chaque sommet, elle emprunte au hasard une des arêtes de façon uniforme. Autrement dit elle peut revenir en arrière et chaque arête a une probabilité $1/3$ d'être choisie. Son périple se termine quand elle trouve une feuille de laitue : elle peut alors la manger !

a. On cherche la probabilité p que la limace mange la feuille au sol. En introduisant les probabilités que la limace mange cette même feuille **mais** en partant d'un autre sommet, établir un système de quatre équations à quatre inconnues, dont p.

b. Déterminer p.

c. Quelle est la probabilité que la limace mange la feuille en l'air ?

d. Quelle est la probabilité que la limace erre indéfiniment sans trouver de pitance ?

Variables aléatoires discrètes

1 - 44 Ⓢ ★

Donner la loi de X^2 si $X \sim \mathcal{B}(p)$ ou $X \sim \mathcal{U}(\llbracket -2;3 \rrbracket)$.

1 - 45 Ⓢ ★

Un produit est vendu par lots de 10 et la société qui les vend s'engage à les rembourser s'ils contiennent au moins deux produits défectueux.

Chacun des produits a une probabilité 1% d'être défectueux.

Quelle est la proportion moyenne de lots que la société devra rembourser ?

1 - 46 Ⓢ ★

Sachant qu'il y a en moyenne une faute toutes les deux pages dans un cours de six cents pages, quelle est la probabilité qu'il y ait effectivement une erreur dans une page donnée ?

1 - 47 Ⓢ ★

Sachant qu'il y a en moyenne cinq défauts sur cent mètres de tissus et que le tissu est débité en coupons de trois mètres, quelle est la proportion de coupons sans défaut que l'on peut espérer ?

1 - 48 Ⓢ ★

Par un soir d'été on observe en moyenne une étoile filante toutes les dix minutes.

Quelle est la probabilité d'en observer deux en un quart d'heure ?

1 - 49 Ⓢ ★

Comparer $\mathbf{P}(X \in 2\mathbf{N})$ et $\mathbf{P}(X \in 1+2\mathbf{N})$ quand X suit une loi $\mathcal{G}(p)$, $\mathcal{G}_{\mathbf{N}}(p)$ ou $\mathcal{P}(\lambda)$.

1 - 50 Ⓢ ★★ **Loi hypergéométrique**

a. Déterminer la loi que suit un tirage aléatoire de n boules parmi N, sachant que la répartition initiale est telle qu'il y a une proportion p de boules vertes et $1-p$ de boules rouges, et que l'on compte le nombre de boules vertes. On parle de la loi hypergéométrique de paramètres N, n et p, notée $\mathcal{H}(N,n,p)$.

b. En déduire la formule de Vandermonde $\displaystyle\sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}$.

c. On fixe n dans $\mathbf{N}$ et k dans $\llbracket 0;n \rrbracket$. Soit, pour N supérieur à n, X_N tel que $X_N \sim \mathcal{H}(N,n,p_N)$ avec $\lim p_N = p \in \left]0;1\right[$. Montrer $\lim \mathbf{P}(X_N = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Interpréter. On rappelle la formule de Stirling, pour n tendant vers l'infini on a $n! \sim \dfrac{\sqrt{2\pi} n^{n+1/2}}{e^n}$.

1 - 51 Ⓢ ★★ **Loi binomiale négative**

On se donne une suite $(X_k)_{k\in\mathbf{N}^*}$ de variables aléatoires indépendantes et identiquement distribuées suivant une loi $\mathcal{B}(p)$. On note, pour r dans $\mathbf{N}^*$,

$$X = \min\left\{ n \in \mathbf{N}^* \mid X_1 + \cdots + X_n = r \right\} .$$

a. Montrer $\mathbf{P}(X=n) = \binom{-r}{k} p^r (p-1)^k$ où $k = n-r$.

b. Montrer que X est une variable aléatoire presque sûrement finie.

c. En déduire une expression de la probabilité pour que le r^{e} succès intervienne avant le s^{e} échec dans la suite $(X_k)_{k\in\mathbf{N}^*}$.

d. Stefan Banach était un fumeur invétéré. Il avait une boîte d'alumettes dans chaque poche et prenait au hasard une boîte dans une de ses poches pour prendre une alumette. Au moment où il veut prendre une alumette dans une boîte et qu'il la découvre vide, quelle est la probabilité pour que l'autre boîte soit également vide ?

1 - 52 ★★ **Modes**

Soit X suivant une loi $\mathcal{P}(\lambda)$. Déterminer son ou ses modes, i.e. les entiers de probabilité maximale selon X

1 - 53 Ⓢ ★★ **Maximum d'une loi uniforme**

On tire au hasard des notes entre 0 et 20, i.e. on se donne $(X_k)_{1\leq k\leq n}$ des variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{U}(\llbracket 0;20 \rrbracket)$. On s'intéresse à la note maximum, i.e. $Y_n = \max_{1\leq k\leq n} X_k$.

a. Montrer que Y_n est une variable aléatoire.

b. Pour $n = 4$, calculer $\mathbf{P}(Y_4 \leq 10)$. Comparer avec le résultat obtenu si on tirait au hasard les notes avec des jetons numérotés, sans remise.

c. Calculer $\mathbf{P}(Y_4 < 10 \,|\, Y_4 \leq 10)$ et en déduire $\mathbf{P}(Y_4 = 10 \,|\, Y_4 \leq 10)$.

d. En déduire $\mathbf{P}(Y_4 = 10)$. Comparer avec le tirage sans remise dans le cas $n = 4$.

e. Montrer $\lim \mathbf{P}(Y_n < 20) = 0$.

1 - 54 Ⓢ ★★ Élections américaines

On définit la puissance électorale lors d'une élection par l'espérance du nombre de sièges que peut rapporter un vote individuel, autrement dit comme le produit de la probabilité pour que le vote fasse basculer l'élection par le nombre de sièges à pourvoir. Dans un État des USA, le nombre de sièges accordé après une élection est de la forme $pn+2$ où n est la taille de la population de l'État et p une constante de proportionnalité (le 2 correspond au Sénat).

a. On suppose que la taille de la population est impaire et on pose $n = 2k+1$. Montrer que la probabilité qu'un vote donné fasse basculer l'élection est approximativement égale à $(k\pi)^{-1/2}$. On pourra utiliser la formule de Stirling : pour n tendant vers l'infini on a $n! \sim \dfrac{\sqrt{2\pi}n^{n+1/2}}{e^n}$.

b. Qui a la plus grande puissance électorale entre les habitant(e)s d'un État fortement peuplé et ceux d'un État moins peuplé ?

1 - 55 Ⓢ ★★ Estimation binomiale

Soit p la probabilité d'un événement A. On effectue n épreuves indépendantes et on désigne par f la fréquence relative de A dans cette série d'épreuves. On rappelle que si X suit une loi binomiale $\mathcal{B}(n,p)$ et a est dans $\mathbf{R}_+^*$, l'inégalité de Bienaymé-Tchebychev fournit $\mathbf{P}(|X - np| > a) \leq np(1-p)a^{-2}$.

a. Soit $p = 0,375$. Combien d'épreuves suffit-il pour que la probabilité d'avoir $|f - p| \leq 0,01$ soit supérieure à 0,995 ?

b. Soit $p = \dfrac{2}{3}$ et $n = 1200$. Comment choisir ε pour que la probabilité d'avoir $|f - p| < \varepsilon$ soit supérieure à 0,985 ?

c. Soit $n = 14400$. Pour quelles valeurs de p la probabilité d'avoir $|f - p| < 0,01$ est-elle supérieure à 0,99 ?

1 - 56 Ⓢ ★★★ Approximation Gaussienne

Soit $X_n \sim \mathcal{B}(n,p)$. On suppose p constant et on fait tendre n vers l'infini. Soit (k_n) une suite d'entiers, avec k_n dans $[\![0;n]\!]$, telle qu'on ait $\lim \dfrac{(k_n - np)^3}{n^2} = 0$. On pose $q = 1-p$ et $p_n = \mathbf{P}(X_n = k_n)$.

a. En utilisant la formule de Stirling rappelée en 1 - 54, montrer $p_n \sim \dfrac{1}{\sqrt{2\pi npq}} \exp\left(-\dfrac{(k_n - np)^2}{2npq}\right)$. Interpréter.

b. Soit $\varepsilon > 0$ et F la fonction de répartition de la loi gaussienne centrée réduite (que l'on trouve dans des tables), i.e. pour x et y réels,

$$F(y) - F(x) = \int_x^y e^{-t^2/2} \frac{\mathrm{d}t}{\sqrt{2\pi}},$$

$\lim_{-\infty} F = 0$ et $\lim_{+\infty} F = 1$. Montrer $\mathbf{P}(|X_n - np| \leq n\varepsilon) \simeq 2F(\varepsilon\sqrt{n}/\sqrt{pq}) - 1$.

c. Reprendre l'exercice 1 - 55 et comparer les approximations.

d. Soit $p = 0,4$ et $n = 1500$. Quelle est la probabilité que f soit compris entre 0,40 et 0,44 ?

1 - 57 ★★★ Paradoxe des anniversaires

On se donne $(X_k)_{1\leq k\leq n}$ des variables aléatoires indépendantes et identiquement distribuées selon $\mathcal{U}([\![1;365]\!])$. On note T_n la variable égale à 0 si $\mathrm{Card}\{X_k \,|\, 1 \leq k \leq n\} = n$ et à 1 sinon.

a. Montrer que T_n est une variable aléatoire.

b. Montrer que T_n suit une loi de Bernoulli de paramètre p_n que l'on déterminera. Montrer en particulier $p_{23} > \dfrac{1}{2}$ et $p_{50} \sim 97\%$.

c. Pour $1 \leq k < \ell \leq n$, on note $Y_{k,\ell}$ la variable donnée par $\mathbb{1}_{X_k = X_\ell}$.

i. Quelle est la loi des variables $Y_{k,\ell}$?

ii. Les variables $Y_{k,\ell}$ sont-elles deux à deux indépendantes ? mutuellement indépendantes ?

iii. On note $Y_n = \displaystyle\sum_{1\leq k<\ell\leq n} Y_{k,\ell}$. Peut-on justifier qu'on a $\mathbf{P}(Y_n = 0) \approx \exp\left(-\dfrac{n(n-1)}{730}\right)$?

iv. En utilisant l'approximation précédente, trouver n tel que $\mathbf{P}(Y_n = 0) \leq \dfrac{1}{2}$.

d. Utiliser les approximations faites dans la question précédente pour montrer que, pour $n \geq 84$, la probabilité pour que la suite $(X_k)_{1\leq k\leq n}$ prenne trois fois la même valeur est supérieure à $\dfrac{1}{2}$.

1 - 58 ★★★ Théorème de Simmons

Soit $X \sim \mathcal{B}(n,p)$ avec $p < \dfrac{1}{2}$ et tel que np soit entier. On pose $m = np$ et on veut démontrer $\mathbf{P}(X \leq m) > \mathbf{P}(X > m)$.

a. On pose $B_r = \binom{n}{m-r} p^{m-r} q^{n-m+r}$ et $C_r = \binom{n}{m+r} p^{m+r} q^{n-m-r}$. Montrer qu'il existe k dans $[\![1;m]\!]$ tel que : $(B_r < C_r) \iff (k \leq r)$.

b. En déduire qu'on a toujours $(k-r-1)B_r \geq (k-r-1)C_r$ et donner les cas d'égalité.

c. Montrer $\displaystyle\sum_{r=0}^{m} rB_r = \sum_{r=0}^{n-m} rC_r$.

d. Conclure.

Algèbre linéaire

1 - 59 Ⓢ ★ Image réciproque

Soit $u \in \mathscr{L}(E,F)$ avec E de dimension finie. Soit H un sous-espace vectoriel de F. Établir $\dim\left(u^{-1}(H)\right) = \dim E - \mathrm{rg}(u) + \dim\left(H \cap \mathrm{Im}(u)\right)$.

1 - 60 Ⓢ ★ Valeurs absolues

Soit $(a_i)_{i\in I}$ une famille de réels deux à deux distincts et $(f_i)_{i\in I}$ la famille de fonctions de $C(\mathbf{R},\mathbf{R})$ définie par $f_i : x \to |x-a_i|$. Cette famille est-elle libre ?

1 - 61 Ⓢ ★★ Centre ♥

Soit E un $\mathbf{K}$-espace vectoriel de dimension supérieure à 2 et $\mathcal{A}$ partie de $\mathscr{L}(E)$. On appelle commutant de $\mathcal{A}$ l'ensemble donné par $c(\mathcal{A}) = \{u \in \mathscr{L}(E) \,|\, \forall a \in \mathcal{A}, [a,u]=0\}$ et centre de $\mathscr{L}(E)$ le commutant de $\mathscr{L}(E)$.

a. Montrer que le centre de $\mathscr{L}(E)$ est constitué des homothéties de E.

b. On suppose E euclidien. Déterminer le commutant de $\mathscr{O}(E)$.

c. On suppose de plus E orienté. Déterminer le commutant de $\mathrm{SO}(E)$.

1 - 62 Ⓢ X 2003 ★★ Endomorphisme nilpotent

Soit $f \in \mathscr{L}(E)$. Montrer qu'on a $f^2 = 0$ si et seulement si $\exists (g,h) \in \mathscr{L}(E)^2$, $f = g\circ h$ et $h \circ g = 0$.

1 - 63 Ⓢ ★★ Formule de Grassmann

Soit E et F deux sous-espaces vectoriels d'un même espace vectoriel de dimension finie. On écrit la suite d'applications linéaires

$$\{0\} \longrightarrow E\cap F \longrightarrow E\times F \longrightarrow E+F \longrightarrow \{0\}$$

donnée respectivement par l'application nulle, $x \mapsto (x,-x)$, $(x,y)\mapsto x+y$ et l'application nulle. On note $(u_k)_{0\leq k\leq 3}$ ces applications et $(E_k)_{0\leq k\leq 4}$ les espaces vectoriels considérés.

En utilisant le théorème du rang, montrer $\displaystyle\sum_{k=1}^{3}(-1)^{k+1}\dim(E_k) = 0$ et en déduire la formule de Grassmann.

1 - 64 Ⓢ ★★ Images et noyaux itérés ♥♥

Soit E un $\mathbf{K}$-espace vectoriel de dimension finie d, avec $d>0$ et u dans $\mathscr{L}(E)$.

a. Montrer que la suite $\left(\mathrm{Im}(u^n)\right)_{n\in\mathbf{N}}$ est décroissante stationnaire. On définit alors

$$p = \min\left\{k\in\mathbf{N} \;\middle|\; \mathrm{Im}(u^{k+1}) = \mathrm{Im}(u^k)\right\}.$$

b. Que dire de la suite $\left(\mathrm{Ker}(u^n)\right)_{n\in\mathbf{N}}$?

c. Montrer $p = \min\left\{k\in\mathbf{N} \;\middle|\; \mathrm{Ker}(u^{k+1}) = \mathrm{Ker}(u^k)\right\}$.

d. Montrer que, pour $n\geq p$, on a $E = \mathrm{Im}(u^n)\oplus\mathrm{Ker}(u^n)$.

1 - 65 Ⓢ ★★ Supplémentaire commun

Soit E un espace vectoriel de dimension finie. Montrer que deux sous-espaces vectoriels de E de même dimension admettent un supplémentaire commun.

Indication : On pourra procéder par récurrence.

1 - 66 Ⓢ ★★ Espace quotient ♠

Soit F un sous-espace vectoriel de E espace vectoriel de dimension quelconque. On munit E de la relation d'équivalence $\mathcal{R}$ définie par $x\mathcal{R}y \Leftrightarrow x-y\in F$.

a. On munit l'ensemble quotient, noté E/F, des « lois quotients » définies par

$$\overline{x}+\overline{y} = \overline{x+y} \quad\text{et}\quad \lambda.\overline{x} = \overline{\lambda.x}\,.$$

Montrer que E/F admet alors une structure d'espace vectoriel et que la projection canonique π de E dans E/F est une application linéaire surjective.

b. Montrer que, si F est de codimension finie, alors $\mathrm{codim}(F) = \dim(E/F)$.

Compléments

1 - 67 ★★ Formule du multinôme

Soit $p\geq 2$ et $(a_i)_{1\leq i\leq p}\in A^p$ une famille d'éléments commutants deux à deux.

Montrer que, pour tout entier naturel non nul n, on a

$$(a_1+a_2+\cdots+a_p)^n = \sum_{|\alpha|=n}\frac{n!}{\alpha_1!\ldots\alpha_p!}a_1^{\alpha_1}\cdots a_p^{\alpha_p}$$

où la somme est étendue sur les p-uplets α dans $\mathbf{N}^p$, $\alpha=(\alpha_1,\cdots,\alpha_p)$ avec $|\alpha| = \alpha_1+\cdots+\alpha_p = n$.

1 - 68 Ⓢ ★★ Formule d'Euler

Soit $\Omega = [\![1;n]\!]$ et, pour p premier, A_p l'événement p divise ω.

Montrer que les événements $(A_p)_{p|n}$ sont mutuellement indépendants et en déduire une formule pour le nombre d'entiers premiers à n dans Ω.

1 - 69 ★★★ Exponentiation rapide

Dans cet exercice on donne des règles de récriture : chaque ligne est une règle qui décrit une application.

Lorsque l'application n'est pas partout définie, son ensemble de définition est décrit par une condition entre crochets. Appliquer ces règles à un élément a signifie choisir, s'il en existe, une règle applicable à a, remplacer a par le résultat de la règle, disons b, et recommencer avec b, et ainsi de suite tant qu'on peut continuer. La procédure s'arrête quand on ne peut plus continuer. Commenter les algorithmes suivants.

a. On se donne u et v deux entiers naturels et on initialise un couple à (u, v). On se donne ensuite les règles suivantes, avec $q = \lfloor v/u \rfloor$:

$$\begin{array}{lll} [0 < u \leq v] & : & (u,v) \mapsto (u, v-u) \\ [0 < v \leq u] & : & (u,v) \mapsto (u-v, v) \\ [\ u = 0\ ] & : & (u,v) \mapsto v \\ [\ v = 0\ ] & : & (u,v) \mapsto u \end{array}$$

b. On se donne x et y deux entiers naturels et on initialise un sextuplet à $(x, 1, 0, y, 0, 1)$. On se donne ensuite les règles suivantes :

$$\begin{array}{lll} [u \neq 0] & : & (u,c,d,v,a,b) \mapsto (v-qu, a-qc, b-qd, u, c, d) \\ [u = 0] & : & (u,c,d,v,a,b) \mapsto (v, a, b) \end{array}$$

c. On se donne un élément a dans un anneau et m un entier strictement positif. On initialise un triplet à (e, a, m) avec e l'élément neutre pour la multiplication de l'anneau. On applique ensuite les règles :

$$\begin{array}{lll} [\ \ n \text{ pair} \ \ ] & : & (y,x,n) \mapsto (y, x^2, n/2) \\ [n \text{ impair} \neq 1] & : & (y,x,n) \mapsto (yx, x^2, (n-1)/2) \\ [\ \ n = 1 \ \ ] & : & (y,x,n) \mapsto yx \end{array}$$

Estimer le nombre de multiplications nécessitées par cet algorithme.

d. Montrer que les coefficients de la matrice $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n$ sont des termes de la suite de Fibonacci et les préciser. En déduire un algorithme rapide pour calculer cette suite et préciser la rapidité en la comparant à un algorithme naïf.

1 - 70 Ⓢ ★★★ Tribus

Soit Ω un univers et $\mathcal{T}$ l'ensemble des tribus sur Ω. On munit cet ensemble de l'ordre donné par l'inclusion.

a. L'ensemble $\mathcal{T}$ admet-il un plus petit élément ? un plus grand ?

b. Montrer que $\mathcal{T}$ est stable par intersection quelconque.

c. Soit F dans $\mathscr{P}(\mathscr{P}(\Omega))$, i.e. un ensemble de parties de Ω. Montrer qu'il existe une plus petite tribu dans $\mathcal{T}$ contenant F. On l'appelle la TRIBU ENGENDRÉE par F et on la note $\sigma(F)$.

d. Montrer que si une tribu contient une partie $(A_i)_{1 \leq i \leq n}$ d'éléments non vides et disjoints deux à deux, alors cette tribu contient au moins 2^n éléments.

e. En déduire qu'une tribu est soit finie, soit non dénombrable (et même admet la puissance du continu).

1 - 71 Ⓢ ★★★ Théorème de CANTOR-BERNSTEIN

Pour *classer les infinis*, l'idée est que s'il existe une injection de E dans F, alors F est plus gros que E. Pour que ce soit effectivement une bonne *notion de grandeur* (on parle en fait de *puissance d'un ensemble*), il convient de vérifier que si E est plus gros que F et F est plus gros que E alors E et F ont même taille : i.e. E et F sont équipotents. C'est l'objet du théorème de CANTOR-BERNSTEIN (Georg CANTOR, 1845–1918, et Félix BERNSTEIN, 1878–1956).

a. Soit $f : \mathscr{P}(E) \to \mathscr{P}(E)$, croissante au sens de l'inclusion. Montrer qu'elle admet un point fixe.

b. Soit E un ensemble non vide, $h : E \to E$, et D dans $\mathscr{P}(E)$. Montrer qu'il existe A dans $\mathscr{P}(E)$ tel que $A = h(A) \cup D$.

c. Soit E et F deux ensembles tels qu'il existe deux applications injectives $f : E \to F$ et $g : F \to E$. On pose $D = E \setminus g(F)$ et $h = g \circ f$. D'après **b.**, il existe alors A tel que $A = h(A) \cup D$. On pose $B = E \setminus A$.

i. Montrer qu'il existe une unique application $j : E \to F$ telle que
- $\forall x \in A,\ j(x) = f(x)$;
- $\forall x \in B,\ g(j(x)) = x$.

ii. Montrer que E et F sont équipotents.