CHAPITRE IV

L'axiome du choix

RÉSUMÉ. • *L'axiome du choix* AC *affirme qu'il est légitime de construire des objets mathématiques en répétant un nombre infini de fois l'opération de choisir un élément dans un ensemble non vide.*
• *Il existe de nombreux énoncés équivalents pour* AC *: existence d'une fonction de choix pour tout ensemble, d'une section pour toute surjection, d'un sélecteur pour toute relation, non vacuité du produit d'une famille d'ensembles non vides, lemme de Zorn affirmant que tout ensemble ordonné inductif a un élément maximal, théorème de Zermelo affirmant que tout ensemble est bien ordonnable.*
• *Deux formes faibles sont l'axiome du choix dénombrable* AC_ω *affirmant qu'un produit dénombrable d'ensembles non vides est non vide, et l'axiome des choix dépendants* ACD *affirmant l'existence de* $(x_n)_{n\in\omega}$ *vérifiant* $x_n\ R\ x_{n+1}$ *pour tout* n *dès que* $\forall x\,\exists y\,(x\ R\ y)$ *est vraie.*
• *En combinatoire,* AC_ω *entraîne qu'un ensemble* A *est infini si et seulement si il possède un sous-ensemble dénombrable si et seulement si il existe une injection non surjective de* A *dans lui-même ;* AC_ω *entraîne que toute réunion dénombrable d'ensembles dénombrables est dénombrable.*
• *En ce qui concerne les mathématiques discrètes,* ACD *entraîne qu'une relation est bien fondée si et seulement si elle ne possède pas de suite infinie décroissante ;* ACD *entraîne aussi le lemme de König : tout arbre infini dont les niveaux sont finis possède une branche infinie.*
• *En algèbre,* AC *entraîne l'inclusion de toute partie libre d'un espace vectoriel dans une base, celle de tout idéal d'un anneau dans un idéal maximal, celle de tout filtre dans un ultrafiltre.*
• *En analyse,* AC_ω *entraîne que* x *est dans l'adhérence de* A *si et seulement si on peut extraire dans* A *une suite convergeant vers* x*.*
• *En analyse,* AC *entraîne le théorème de Tychonoff (tout produit de compacts est compact) et le théorème de Hahn–Banach (prolongement des formes linéaires).*
• *En géométrie,* AC *donne lieu au paradoxe de Banach–Tarski : on peut découper la sphère en quatre pièces et les réarranger par rotation en deux copies de la sphère initiale.*
• *Adopter ou non* AC *dépend du type d'existence qu'on souhaite étudier pour des objets mathématiques ; dans la mesure où les deux options sont usuelles, il est commode de mentionner explicitement les appels à* AC*.*
• *Dans le contexte de la théorie des ensembles, il est usuel d'ajouter* AC *à* ZF*, obtenant ainsi* ZFC*.*

► Dans ce chapitre, on étudie un nouvel axiome affirmant l'existence d'ensembles, l'axiome du choix. La tâche principale est de montrer l'équivalence de diverses formes de cet axiome, et d'expliquer sur quelques exemples comment il est utilisé dans diverses branches des mathématiques.

Dans la première section de ce chapitre, on introduit l'axiome du choix AC et on établit en particulier l'équivalence de AC avec le théorème de Zermelo et le lemme de Zorn. Dans les sections suivantes, on établit quelques applications typiques de l'axiome du choix, successivement en théorie des ensembles, puis en combinatoire (existence d'ultrafiltres), en algèbre (existence de bases dans les espaces vectoriels), en topologie (théorème de Tychonoff sur les produits de compacts), en analyse (théorème de Hahn–Banach, existence de fonctions non mesurables), et en géométrie (décompositions paradoxales de la sphère). Dans la dernière section, on discute brièvement la validité de AC et l'opportunité de l'adopter ou de le rejeter, c'est-à-dire de l'inclure ou non dans le système des axiomes de base de la théorie des ensembles. ◀

▷ *Les axiomes de* ZF *permettent de légitimer l'existence d'un grand nombre d'ensembles, en particulier celle de suites définies récursivement. Cependant, certaines constructions ne relèvent pas, ou tout au moins ne semblent pas le faire, des principes de* ZF *: c'est en particulier le cas pour des constructions mettant en jeu une infinité de choix simultanés. On montre ici que de telles constructions dérivent toutes d'un unique principe, appelé axiome du choix et noté* AC*. Il n'est pas clair que* AC *dérive des axiomes de* ZF *— on esquissera au chapitre* **??** *une démonstration du fait qu'il n'en dérive pas — et, comme on l'a fait pour les axiomes de l'infini ou de remplacement, il faut envisager de l'ajouter aux axiomes de* ZF*. Néanmoins, comme* AC *affirme l'existence d'objets au sujet desquels l'intuition est incertaine, son usage n'est pas aussi systématiquement accepté que celui des autres axiomes de* ZF*, et il est d'usage de rendre ses utilisations aussi rares que possible, et d'en garder la trace.* ◁

1. L'axiome du choix

▶ L'axiome du choix affirme la possibilité de construire des ensembles en répétant une infinité de fois une opération de choix, même non spécifiée explicitement. Il existe de nombreuses formes équivalentes de l'axiome du choix. On part ici de la forme initiale, en termes de fonction de choix sur un ensemble, et on montre l'équivalence avec le théorème de Zermelo sur l'existence de bons ordres et avec le lemme de Zorn sur l'existence de chaines maximales. ◀

1.1. Fonctions de choix.

▶ Une fonction de choix sur un ensemble d'ensembles est une fonction qui sélectionne un élément distingué dans chacun de ses éléments non vides. L'axiome du choix affirme que tout ensemble d'ensembles possède une fonction de choix. ◀

Définition 1.1. (fonction de choix) Soit A un ensemble. On appelle *fonction de choix* sur A une application $F : A \setminus \{\emptyset\} \to \bigcup A$ vérifiant $F(x) \in x$ pour tout x non vide dans A.

Une fonction de choix est une fonction qui choisit un élément de chaque élément non vide de A, c'est-à-dire de chaque élément pour lequel un tel choix soit possible[1].

[1] Dans le contexte du système de Zermelo–Fraenkel choisi pour toute la suite, A désigne un ensemble pur, et par conséquent un ensemble d'ensembles. Si on souhaitait formuler AC dans un contexte qui ne soit pas celui des ensembles purs, il faudrait préciser ici que A est un ensemble (ou une famille) d'ensembles.

Exemple 1.2. (fonction de choix) Supposons $A = \{\{0\}, \{1,2\}, \{1,2,3\}\}$. Alors la fonction F définie par $F(\{0\}) = 0$, $F(\{1,2\}) = 1$ et $F(\{1,2,3\}) = 1$ est une fonction de choix sur A.

De nombreux ensembles possèdent des fonctions de choix.

Proposition 1.3. (choix fini) *Supposons que A est un ensemble fini, ou qu'il existe un bon ordre sur $\bigcup A$. Alors il existe une fonction de choix sur A.*

Démonstration. Les arguments sont différents dans les deux cas. Dans le cas d'un ensemble fini, on montre par induction sur l'entier n que, si A est en bijection avec $[0,n[$, alors A possède une fonction de choix. Pour $n = 0$, l'ensemble A est vide, et la fonction vide (c'est-à-dire l'ensemble vide) convient. Supposons $n > 0$, soit $n = m+1$. Soit h une bijection de $[0,n[$ sur A. Posons $A' = A \setminus \{h(m)\} = \mathrm{Im}(h\restriction_{[0,m[})$. Par hypothèse d'induction, il existe une fonction de choix F' sur A'. Deux cas sont possibles. Ou bien $h(m)$ est vide, et alors F' est, par définition, une fonction de choix sur A. Ou bien $h(m)$ n'est pas vide. Soit a un élément quelconque de $h(m)$. On pose $F = F' \cup \{(h(m), a)\}$, ce qui définit bien un ensemble par les axiomes de la paire et de l'union. Alors F est une fonction de choix sur A.

Supposons maintenant que $<$ est un bon ordre sur $\bigcup A$. On définit $F{:}A \to \bigcup A$ en posant $F(x) = \inf x$ pour x non vide, ce qui a un sens puisque, par définition, tout élément de A est une partie de $\bigcup A$. L'existence de F en tant qu'ensemble de couples est légitimée par l'irréprochable définition par séparation

$$F = \{(x,a) \in A \times \bigcup A\,;\, a \in x \text{ et } \forall b{\in}x\,(a = b \text{ ou } a < b)\}$$

où A et $<$ figurent comme paramètres. L'hypothèse que $<$ est un bon ordre garantit que F est partout définie, sauf, le cas échéant, en $\emptyset$. Par construction, $F(x) \in x$ est alors vérifiée pour tout x non vide. □

Corollaire 1.4. *Supposons qu'il existe une surjection d'un ordinal sur a. Alors il existe une fonction de choix sur $\mathfrak{P}(a)$.*

Démonstration. Pour tout ensemble a, on a l'égalité $\bigcup\mathfrak{P}(a) = a$. En effet, si b est dans a, alors on a $b \in \{b\} \in \mathfrak{P}(a)$, donc $b \in \bigcup\mathfrak{P}(a)$, d'où $a \subseteq \bigcup\mathfrak{P}(a)$. Inversement, $b \in \mathfrak{P}(a)$ entraîne $b \subseteq a$, d'où $\bigcup\mathfrak{P}(a) \subseteq a$.

Alors supposons que f est une surjection d'un ordinal α sur a. On obtient un bon ordre $\prec$ sur a en posant

$$x \prec y \;\Leftrightarrow\; \min f^{-1}(x) < \min f^{-1}(y).$$

On applique alors la proposition 1.3 à $A = \mathfrak{P}(a)$, qui est tel que $\bigcup A$ est bien ordonnable. □

On est donc mené naturellement au problème suivant :

Question 1.5. *Existe-t-il une fonction de choix sur tout ensemble ?*

▷ *L'argument inductif montrant qu'il existe une fonction de choix sur tout ensemble fini ne s'étend a priori pas au cas d'un ensemble quelconque : le problème, si on suppose par exemple A dénombrable, c'est-à-dire qu'on suppose l'existence d'une bijection $f : \omega \to A$, est que, même si, pour chaque entier n, on a une fonction de choix F_n sur le sous-ensemble $\{f(k)\,;\, k < n\}$ de A, rien ne garantit que la suite des fonctions $(F_n)_{n\in\omega}$ existe, c'est-à-dire qu'il existe un ensemble dont les éléments soient exactement tous les couples (n, F_n) avec n entier. Lorsque c'est le cas, et à supposer que, de surcroît, les fonctions F_n soient deux à deux compatibles, on obtient une fonction de choix sur A entier en prenant l'union de l'ensemble $(F_n)_{n\in\omega}$, c'est-à-dire l'ensemble de tous les couples $(x, F_n(x))$ pour n entier et x dans le domaine de F_n — c'est précisément ce qu'on fera dans la démonstration de la proposition 1.9.*

Constatant l'absence à ce point d'une démonstration d'existence à partir des axiomes précédemment formulés[2]*, on peut envisager d'introduire un nouvel axiome. Pour cela, il convient d'une part d'interroger l'intuition sur le caractère évident ou plausible d'une réponse positive à la question 1.5, et, d'autre part, d'examiner les conséquences d'un éventuel nouvel axiome pour évaluer sa cohérence et juger de l'opportunité de son introduction. Le premier point est affaire d'opinion, et on y reviendra plus loin ; le second est affaire de mathématiques, et c'est lui qu'on va développer en premier*[3]*. On introduit donc à titre d'hypothèse de travail l'énoncé suivant.* ◁

Définition 1.6. (axiome du choix) On appelle *axiome du choix* l'énoncé « Tout ensemble possède une fonction de choix », c'est-à-dire

$$(\mathsf{AC}) \qquad \forall A\, \exists F\colon A \setminus \{\emptyset\} \to \bigcup A\ \forall x{\in}A \setminus \{\emptyset\}\, (F(x) \in x).$$

Quitte à remplacer A par $A \setminus \{\emptyset\}$, on peut se restreindre au cas $\emptyset \notin A$, moyennant quoi les fonctions de choix sont partout définies, ce qui donne pour AC l'énoncé équivalent

$$\forall A\, (\emptyset \notin A \Rightarrow \exists F\colon A \to \bigcup A\ \forall x{\in}A\ (F(x) \in x)).$$

On considèrera également deux formes apparemment plus faibles de choix.

Définition 1.7. (axiome des choix dépendants) On appelle *axiome des choix dépendants* l'énoncé, noté ACD, « Si R est une relation binaire sur A vérifiant $\forall x\, \exists y\, (x\ R\ y)$, alors il existe une suite $(x_n)_{n\in\omega}$ d'éléments de A vérifiant $x_n\ R\ x_{n+1}$ pour tout entier n ».

Définition 1.8. (axiome du choix dénombrable) On appelle *axiome du choix dénombrable* l'énoncé, noté AC_ω, « Tout ensemble dénombrable possède une fonction de choix ».

Proposition 1.9. (formes faibles) *L'axiome du choix* AC *entraîne l'axiome des choix dépendants* ACD*, qui lui-même entraîne l'axiome du choix dénombrable* AC_ω*.*

Démonstration. (i) Soit R une relation binaire sur A, c'est-à-dire une partie de $A \times A$, vérifiant $\forall x\ \exists y\ (x\ R\ y)$. Supposons que F est une fonction de choix sur $\mathfrak{P}(A)$. Alors, par la proposition III.3.2 (définition récursive), il existe une suite $(x_n)_{n\in\omega}$ d'éléments de A vérifiant

$$x_0 = F(A) \quad \text{et} \quad x_{n+1} = F(\{x \in A\,;\, x_n\ R\ x\}),$$

puisque, par hypothèse, $\{x \in A\,;\, x_n\ R\ x\}$ n'est jamais vide. On a alors $x_n\ R\ x_{n+1}$ pour tout n, et, par conséquent, AC entraîne ACD.

(ii) Soit A un ensemble dénombrable ne contenant pas $\emptyset$, et f une bijection de ω sur A. Pour chaque entier n, on a vu qu'il existe une fonction de choix sur chaque sous-ensemble fini de A, donc en particulier sur chaque sous-ensemble du type $\{f(0), \ldots, f(n-1)\}$. Si on peut choisir une suite cohérente de telles fonctions de choix partielles, alors, par réunion, on obtiendra une fonction de choix sur tout A. C'est là que l'axiome des choix dépendants est utile. Notons C l'ensemble de toutes les fonctions de choix sur un sous-ensemble fini de A de

[2] absence qui, comme toujours, ne prouve pas qu'une telle démonstration n'existe pas

[3] ce qui est naturel : l'intuition mathématique semble en général se forger *a posteriori* par l'usage technique bien plus qu'*a priori* par l'exercice d'on ne sait trop quelles expériences de pensée

la forme $\{f(0), \dots, f(n-1)\}$, et R la relation d'inclusion stricte sur C : dire que F_1, F_2 sont des éléments de C vérifiant $F_1 \; R \; F_2$ signifie qu'il existe deux entiers n_1, n_2 avec $n_1 < n_2$ tels que le domaine de F_i est $\{f(0), \dots, f(n_i - 1)\}$ et F_2 est compatible avec F_1 sur le domaine de celle-ci. L'argument de la démonstration de la proposition 1.3 montre que, pour tout F dans C, on peut toujours ajouter un point au domaine de F et, par conséquent, il existe F' vérifiant $F \; R \; F'$. L'axiome des choix dépendants garantit alors l'existence d'une suite $(F_n)_{n\in\omega}$ telle qu'on ait $F_n \; R \; F_{n+1}$ pour tout n. Soit $F := \bigcup_{n\in\omega} F_n$, c'est-à-dire $F := \bigcup S$, où S est la suite $(F_n)_{n\in\omega}$ vue comme fonction, c'est-à-dire ensemble de couples. Par construction, le domaine de F contient $f(0)$ et, s'il contient $f(n)$, il contient aussi $f(n+1)$, donc il contient $f(n)$ pour tout n, et c'est donc A. Par ailleurs F est fonctionnelle, car les fonctions F_n sont cohérentes entre elles. Donc F est une fonction de choix sur A, et, par conséquent, ACD entraîne AC_ω. □

1.2. Variantes.

► On introduit ici quelques variantes de l'axiome du choix, où l'énoncé original est reformulé en des termes légèrement différents, respectivement en termes de sélecteur pour une relation binaire, de sections pour une surjection, et de produits d'ensembles. ◄

Si R est un ensemble de couples (c'est-à-dire une relation binaire), on note $\mathrm{pr}_1(R)$ la première projection de R, c'est-à-dire l'ensemble des x tels qu'il existe au moins un y vérifiant $(x, y) \in R$.

Définition 1.10. (sélecteur) Supposons $R \subseteq A \times A$. Un *sélecteur* pour R est une application $f : \mathrm{pr}_1(R) \to A$ vérifiant $(x, f(x)) \in R$ pour tout x dans $\mathrm{pr}_1(R)$.

Comme le nom le suggère, un sélecteur choisit pour chaque x un élément particulier y en relation avec x dès qu'un tel élément existe (Figure 1).

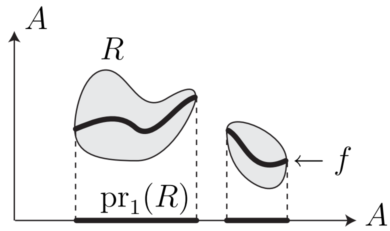

Figure 1. Sélecteur pour R : une fonction qui choisit une valeur dès que cela est possible, à savoir sur $\mathrm{pr}_1(R)$

Proposition 1.11. (sélecteur) *L'axiome du choix est équivalent à l'énoncé: « Tout ensemble de couples* [4] *admet un sélecteur ».*

Démonstration. Soit $R \subseteq A \times A$. Supposons que F est une fonction de choix sur $\mathfrak{P}(A)$. On définit un sélecteur f pour R en posant $f(x) := F(\{y\,;\,(x, y) \in R\})$.

Inversement, soit A un ensemble quelconque. On pose $R := \{(x, y)\,;\, x \in A \setminus \{\emptyset\} \text{ et } y \in x\}$. Par construction, R est un ensemble de couples dans $A \times \bigcup A$, et $\mathrm{pr}_1(R)$ est $A \setminus \{\emptyset\}$. Un sélecteur pour R est exactement une fonction de choix pour A. □

[4] ou, si on préfère, toute relation binaire qui est un ensemble

Corollaire 1.12. (AC) *Si R est une relation d'équivalence sur un ensemble A, il existe un sous-ensemble B de A qui rencontre chaque classe d'équivalence en exactement un point.*

Démonstration. Soit R' l'ensemble des couples de $\mathfrak{P}(A) \times \mathfrak{P}(A)$ de la forme $(X, \{x\})$ où X est la R-classe d'équivalence de x. Supposons que f est un sélecteur pour R', et posons $B = \bigcup \mathrm{pr}_2(f)$. Par construction, $\mathrm{pr}_1(f)$ est l'ensemble des classes d'équivalence de R, et $\mathrm{pr}_2(f)$ est un ensemble de singletons qui contient exactement un singleton $\{x\}$ avec $x \in C$ pour chaque classe C. L'union de ces singletons contient donc un point par classe d'équivalence de R. □

On rappelle que, si f est une surjection d'un ensemble A sur un ensemble B, on appelle *section* de f toute application $g : B \to A$ telle que $f \circ g$ soit l'identité.

Proposition 1.13. (section) *L'axiome du choix est équivalent à l'énoncé « Toute surjection admet une section »* [5].

Démonstration. Soit $f : A \to B$ une surjection. Supposons que F est une fonction de choix sur $\mathfrak{P}(A)$. On définit une section g pour f en posant $g(y) = F(f^{-1}(y))$ pour y dans B, c'est-à-dire en utilisant F pour choisir l'un parmi les antécédents de y.

Inversement, soit A un ensemble quelconque. On pose $B = \{(x, a) \in A \times \bigcup A \,;\, a \in x\}$, et on définit $f : B \to A \setminus \{\emptyset\}$ par $f((x, a)) = x$. Par construction, f est une surjection. Supposons que g est une section pour f. On définit $F : A \setminus \{\emptyset\} \to \bigcup A$ par la condition que $F(x)$ est la seconde composante du couple $g(x)$. L'hypothèse que $f \circ g$ est l'identité de $A \setminus \{\emptyset\}$ signifie que, pour tout x non vide dans A, l'élément $g(x)$ est un couple de la forme (x, a) avec $a \in x$, et, par construction, ce couple est $(x, F(x))$: donc $F(x)$ appartient à x, et, par conséquent, F est une fonction de choix sur A. □

Comme au chapitre III, si f est une application d'un ensemble I dans un ensemble X prenant la valeur x_i en i, on utilise $(x_i)_{i \in I}$ comme une notation alternative de f, alors appelée *suite* des x_i indexée par I.

Définition 1.14. (produit) Pour $(A_i)_{i \in I}$ suite d'ensembles indexée par I, on appelle *produit* des A_i, et on note $\prod_{i \in I} A_i$, l'ensemble des suites $(a_i)_{i \in I}$ vérifiant $a_i \in A_i$ pour tout i dans I.

▷ *En termes ensemblistes, une suite S étant vue comme un ensemble de couples, le produit $\prod S$ admet la définition*

$$\prod S = \{f : \mathrm{Dom} S \to \bigcup \mathrm{Im} S \,;\, \forall x {\in} \mathrm{Dom} S \, (f(x) \in S(x))\},$$

ce qui, en particulier, en garantit l'existence à partir de celle de S. Noter qu'en associant à tout couple (a_1, a_2) l'application qui envoie 1 sur a_1 et 2 sur a_2, on établit une bijection de $A_1 \times A_2$ sur $\prod_{i \in \{1,2\}} A_i$, et de même pour chaque entier intuitif n. *Il y a donc compatibilité entre le produit cartésien de deux ensembles et le produit généralisé au sens ci-dessus.* ◁

[5] Le résultat est un élément de réponse à la question 1.3 : faute d'y répondre pleinement, il relie le problème à l'axiome du choix et, de là, à tous les développements ultérieurs sur la question. On rappelle que l'énoncé symétrique « toute injection admet une rétraction », c'est-à-dire « pour toute injection $f : A \to B$, il existe $g : B \to A$ telle que $g \circ f$ soit l'identité de A ne requiert pas l'axiome du choix : a étant un élément quelconque de A, on définit g par $g(y) := x$ pour $y = f(x)$ et $g(y) = a$ pour $y \notin \mathrm{Im} f$.

PROPOSITION 1.15. (produit) *L'axiome du choix est équivalent à l'énoncé : « Tout produit d'ensembles non vides est non vide ». L'axiome du choix dénombrable est équivalent à l'énoncé : « Tout produit dénombrable d'ensembles non vides est non vide ».*

DÉMONSTRATION. Soit $(A_i)_{i \in I}$ une suite d'ensembles non vides. On introduit A comme l'ensemble $\{\{i\} \times A_i\,;\, i \in I\}$. L'existence de A est garantie par séparation puisque, si S est la suite $(A_i)_{i \in I}$ vue comme ensemble de couples, on a

$$A = \{z \in \mathfrak{P}(\mathrm{Dom}S) \times \mathrm{Im}S\,;\, \exists i{\in}\mathrm{Dom}S\,(z = \{i\} \times S(i))\}.$$

Supposons que F est une fonction de choix sur A. On définit une suite $(a_i)_{i \in I}$ en posant que a_i est la seconde composante de $F(\{i\} \times A_i)$. L'existence de cette suite s est garantie par la définition par séparation

$$s = \{(t, a) \in I \times \bigcup\nolimits_{i \in I} A_i\,;\, a \in F(\{i\} \times A_i)\}\,^{6},$$

son caractère fonctionnel par celui de F, et le fait qu'elle soit définie pout tout i par l'hypothèse que chacun des A_i est non vide. Alors, par construction, s est dans $\prod_{i \in I} A_i$, qui n'est donc pas vide. Par conséquent, AC entraîne que tout produit d'ensembles non vides est non vide.

Inversement, soit A un ensemble ne contenant pas $\emptyset$. Par hypothèse, $\prod_{x \in A} x$ est le produit d'une suite d'ensembles non vides. Si s est un élément de ce produit, alors, par définition, s est une application de A dans A vérifiant $s(x) \in x$ pour tout x non vide dans A : autrement dit, s est une fonction de choix sur A. Par conséquent, l'hypothèse que tout produit d'ensembles non vides est non vide entraîne AC.

L'argument est identique pour le cas dénombrable. Dans un sens, si $(A_n)_{n \in \omega}$ est une suite dénombrable d'ensembles non vides, l'ensemble $\{\{n\} \times A_n\,;\, n \in \omega\}$ est dénombrable, et AC_ω est suffisant pour garantir l'existence d'une fonction de choix sur cet ensemble. Dans l'autre sens, si A est dénombrable, alors $\prod_{x \in A} x$ est un produit dénombrable. □

1.3. Le théorème de Zermelo.

► Une question laissée en suspens au chapitre II est l'existence d'un bon ordre sur tout ensemble. On montre ici qu'une réponse positive est équivalente à l'axiome du choix. ◄

LEMME 1.16. *Soit A un ensemble quelconque. Alors il y a équivalence entre :*
(i) il existe une fonction de choix sur $\mathfrak{P}(A)$;
(ii) il existe un bon ordre sur A.

DÉMONSTRATION. (figure 2) L'équivalence est triviale si A est vide, et on suppose désormais A non vide. Supposons que F est une fonction de choix sur $\mathfrak{P}(A)$. Soit a un ensemble quelconque n'appartenant pas à A. Alors, par la proposition III.3.4 (récursion ordinale généralisée), il existe une suite $(x_\alpha)_{\alpha \in \mathbf{Ord}}$ d'éléments de $A \cup \{a\}$ indexée par les ordinaux et vérifiant

$$x_\alpha = \begin{cases} F(\{x \in A\,;\, \forall \beta{<}\, \alpha(x_\beta \neq x)\}) & \text{s'il existe au moins un } x \text{ dans } A \text{ distinct} \\ & \text{de tous les } x_\beta \text{ pour } \beta < \alpha, \\ a & \text{sinon.} \end{cases}$$

Deux cas sont *a priori* possibles, suivant que a apparaît ou non dans la suite $(x_\alpha)_{\alpha \in \mathbf{Ord}}$. Supposons d'abord qu'on a $x_\alpha \neq a$ pour tout ordinal α. Par construction, la suite $(x_\alpha)_{\alpha \in \mathbf{Ord}}$, qui est une certaine classe fonctionnelle $\mathbf{S}$, est injective : on peut montrer à partir des axiomes

[6] soit encore $s = \{(i, a) \in \mathrm{Dom}S \times \bigcup \mathrm{Im}S\,;\, a \in F(\{i\} \times S(i))\})$

de ZF que $\alpha \notin \beta$ entraîne $x_\alpha \neq x_\beta$. Il en résulte que la classe $\mathbf{S}^{-1}$, c'est-à-dire la classe des ensembles z vérifiant $\exists x, y (z = (y, x)$ et $(x, y) \in \mathbf{S})$, est elle-même fonctionnelle. Or le domaine de $\mathbf{S}^{-1}$ est l'ensemble A, donc, par remplacement, son image est un ensemble, et ne peut donc être la classe **Ord** tout entière.

On est donc dans le second cas, c'est-à-dire celui où a est dans l'image de la suite $(x_\alpha)_{\alpha \in \mathbf{Ord}}$. Il existe alors un plus petit ordinal θ vérifiant $x_\theta = a$. Alors, par construction, la suite $(x_\alpha)_{\alpha<\theta}$ réalise une bijection de θ sur A. On définit alors un bon ordre sur A en transportant le bon ordre des ordinaux, c'est-à-dire en proclamant que $x_\alpha \prec x_\beta$ est vrai si et seulement si $\alpha < \beta$ l'est.

Inversement, si $\prec$ est un bon ordre sur A, alors $F(x) := \inf_\prec(x)$ définit une fonction de choix sur $\mathfrak{P}(A)$, ainsi qu'on l'a déjà vu avec la proposition 1.3. $\square$

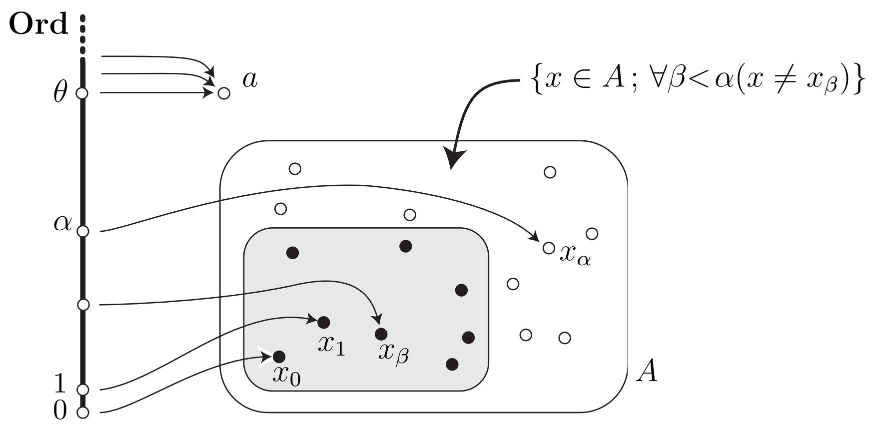

FIGURE 2. Démonstration du théorème de Zermelo : on définit récursivement une suite $(x_\alpha)_{\alpha<\theta}$ énumérant A en choisissant x_α parmi les éléments non encore atteints (en blanc), tant qu'il en existe ; par remplacement, il doit exister θ tel que la construction s'arrête.

On en déduit immédiatement :

PROPOSITION 1.17. (théorème de Zermelo) *L'axiome du choix équivaut à l'énoncé, appelé théorème de Zermelo, « tout ensemble est bien ordonnable ».*

1.4. Lemme de Zorn.

▶ Un autre énoncé équivalent à l'axiome du choix est le lemme de Zorn, qui affirme l'existence d'éléments extrémaux dans certains ensembles ordonnés, et qui est à la base de nombreuses applications, notamment en algèbre. ◀

Si $(A, \prec)$ est un ensemble ordonné, on appelle *chaîne* de $(A, \prec)$ une partie C de A dont les éléments sont deux à deux $\preccurlyeq$-comparables, c'est-à-dire telle que, pour tous x, y dans C, on a $x \preccurlyeq y$ ou $y \preccurlyeq x$; en d'autres termes, une chaîne est un sous-ensemble totalement ordonné de $(A, \prec)$.

DÉFINITION 1.18. (inductif) On dit qu'un ensemble ordonné $(A, \prec)$ est *inductif* si toute chaîne C de $(A, \prec)$ possède un majorant, c'est-à-dire s'il existe dans A un élément a vérifiant $x \preccurlyeq a$ pour tout x dans C.

Proposition 1.19. (lemme de Zorn) *L'axiome du choix est équivalent à l'énoncé, appelé lemme de Zorn : « Tout ensemble ordonné inductif a un élément maximal ».*

Démonstration. (figure 3) Soit $(A, \prec)$ un ensemble ordonné inductif. Supposons que F est une fonction de choix sur $\mathfrak{P}(A)$. Comme dans la démonstration de la proposition 1.17, la proposition III.3.4 (définition récursive généralisée) garantit l'existence d'une suite $(x_\alpha)_{\alpha\in\mathbf{Ord}}$ vérifiant

$$x_\alpha = \begin{cases} F(\{x \in A\,;\, \forall\beta{<}\,\alpha(x_\beta \prec x)\}) & \text{s'il existe au moins un } x \text{ dans } A \text{ plus grand} \\ & \qquad\text{que tous les } x_\beta \text{ pour } \beta < \alpha, \\ a & \text{sinon,} \end{cases}$$

où a est quelconque hors de A, et le même argument de remplacement montre qu'il doit exister un ordinal θ tel que la suite $(x_\alpha)_{\alpha<\theta}$ est à valeurs dans A et qu'il n'existe aucun élément x de A plus grand que tous les x_α pour $\alpha < \theta$. Par construction, l'ensemble $\{x_\alpha\,;\, \alpha < \theta\}$, c'est-à-dire l'image de la suite $(x_\alpha)_{\alpha<\theta}$, est une chaîne de $(A, \prec)$. L'hypothèse que A est inductif entraîne que cette chaîne possède un majorant x, autrement dit on a $x_\alpha \preccurlyeq x$ pour tout α dans θ. Comme il n'existe aucun élément strictement plus grand que tous les x_α, l'élément x est nécessairement maximal dans $(A, \prec)$[7].

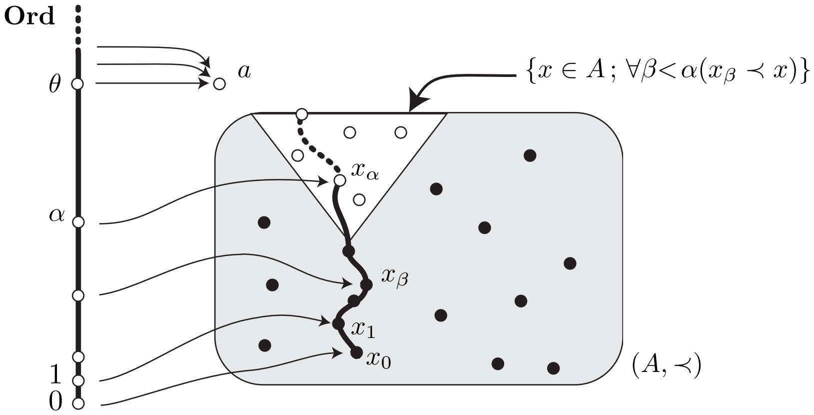

Figure 3. Démonstration du lemme de Zorn : on définit récursivement une suite croissante $(x_\alpha)_{\alpha<\theta}$ en choisissant x_α parmi les majorants des éléments précédemment atteints (en blanc), tant qu'il en existe ; par remplacement, il doit exister θ tel que la construction s'arrête.

Inversement, soit A un ensemble ne contenant pas $\emptyset$. Soit F l'ensemble de toutes les fonctions de choix sur une partie de A, et soit $\subseteq$ la relation de prolongement entre fonctions. Alors l'ensemble $(\mathsf{F}, \subseteq)$ est un ensemble ordonné, et il est inductif. En effet, si C est une chaîne de $(\mathsf{F}, \subseteq)$, l'ensemble $\bigcup C$ est une fonction de choix dont le domaine est l'union des domaines des éléments de C : l'hypothèse que C est une chaîne vis-à-vis de $\subseteq$ garantit que deux éléments de C prennent la même valeur en tout point où ils sont tous deux définis, et donc que $\bigcup C$ est fonctionnelle. Or soit F un élément de $(\mathsf{F}, \subseteq)$ dont le domaine est un sous-ensemble strict de A. Soit x un élément de A n'appartenant pas à $\mathrm{Dom}F$, et a un élément de x. Alors $F \cup \{(x, a)\}$ est un majorant strict de F dans $(\mathsf{F}, \subseteq)$, et donc F n'est pas maximal. Par conséquent, le domaine d'un élément maximal dans $(\mathsf{F}, \subseteq)$ est nécessairement A entier, donc un tel élément maximal

[7] et, de plus, x est un des x_α, donc nécessairement le plus grand ; par conséquent, θ est un ordinal successeur et x est $x_{\theta-1}$

est une fonction de choix sur A. Si l'existence d'un tel élément maximal est garantie, il existe une fonction de choix sur A.[8] □

2. Applications de l'axiome du choix

► Les applications de l'axiome du choix, c'est-à-dire les résultats mathématiques dont la démonstration requiert une des formes de l'axiome du choix, forment une famille beaucoup trop vaste pour qu'on puisse espérer être exhaustif. On se borne ici à mentionner quelques résultats typiques dans des domaines variés. ◄

2.1. Dénombrement.

[8]On peut démontrer le lemme de Zorn à partir de l'axiome du choix sans passer par une induction ordinale, mais au prix de l'argument suivant qui est moins lumineux : on mesurera sur cet exemple le bénéfice de disposer des définitions récursives. On part toujours de $(A, \prec)$ ordonné inductif, et on suppose que F est une fonction de choix sur $\mathfrak{P}(A)$. Pour $X \subseteq A$ et $a \in A$, on note $X \preccurlyeq a$ pour $\forall x{\in}X\,(x \preccurlyeq a)$, et de même avec $\prec$. Pour toute chaîne C, on pose

$$C^+ = \begin{cases} C \cup \{F(\{a\,;\, C \prec a\})\} & \text{s'il existe } a \text{ vérifiant } C \prec a,\\ C & \text{sinon.}\end{cases}$$

Soit C une chaîne quelconque. L'ensemble $(A, \preccurlyeq)$ étant inductif, il existe a vérifiant $C \preccurlyeq a$. Si a n'est pas maximal dans $(A, \preccurlyeq)$, il existe b vérifiant $a \prec b$, et donc $C \prec b$. Par conséquent, si C est une chaîne telle que $\{a\,;\, C \prec a\}$ soit vide, c'est-à-dire vérifiant $C^+ = C$, il existe un élément maximal dans A. On va construire une telle chaîne.

Pour cela, appelons *close* toute famille K de chaînes de $(A, \prec)$ telle que $C \in K$ entraîne $C^+ \in K$, et que, si J est une partie de K formée de chaînes deux à deux comparables pour l'inclusion, alors $\bigcup J$ appartienne à K. La famille de toutes les chaînes de $(A, \preccurlyeq)$ est close, et toute intersection de familles closes est close. Il existe donc une plus petite famille close K, à savoir l'intersection de toutes les familles closes. Posons

$$K' = \{C \in K\,;\, \forall D{\in}K\,(C \subseteq D \text{ ou } D \subseteq C\}.$$

On va montrer $K' = K$, c'est-à-dire que K est composée de chaînes deux à deux comparables pour $\subseteq$. Supposant ceci démontré, posons $C = \bigcup K$. Par définition, on a $C \in K$, et donc $C^+ \in K$. Or, par construction, on a $D \subseteq \bigcup K = C$ pour toute chaîne D dans K. Donc, en particulier, on a $C^+ \subseteq C$, donc $C^+ = C$, comme souhaité.

Puisque K est la plus petite famille close, et qu'on a $K' \subseteq K$, il suffit, pour montrer $K' = K$, de montrer que K' est close. Soit $C \in K'$. On veut montrer $C^+ \in K'$. Or posons $K_C = \{D \in K\,;\, D \subseteq C \text{ ou } C^+ \subseteq D\}$, et supposons $D \in K_C$. Si on a $C^+ \subseteq D$, on a *a fortiori* $C^+ \subseteq D^+$. Pour $C = D$, on a trivialement $C^+ = D^+$. Supposons $D \subsetneqq C$. Par hypothèse, D^+ est dans K, et C dans K', donc on a $D^+ \subseteq C$ ou $C \subsetneqq D^+$. Le second cas est incompatible avec $D \subsetneqq C$ puisque $D^+ \setminus D$ est un singleton. Donc, dans tous les cas, $D \in K_C$ entraîne $D^+ \in K_C$. Supposons maintenant que J est un sous-ensemble de K_C formé de chaînes deux à deux comparables pour l'inclusion. Ou bien on a $D \subseteq C$ pour tout D dans J, et alors on a $\bigcup J \subseteq C$, ou bien il existe D dans J vérifiant $C^+ \subseteq D$, et alors on a $C^+ \subseteq \bigcup J$: dans les deux cas, $\bigcup J$ est dans K_C. Donc K_C est une famille close, donc K_C est égale à K, ce qui montre que C^+ est dans K' dès que C y est.

Finalement, supposons que J est un sous-ensemble de K' formé de chaînes deux à deux comparables pour l'inclusion. Soit D une chaîne quelconque dans K. Ou bien on a $C \subseteq D$ pour toute chaîne C dans J, et alors on déduit $\bigcup J \subseteq D$, ou bien il existe C dans J vérifiant $D \subseteq C$, et on déduit $D \subseteq \bigcup J$. Donc, dans tous les cas, $\bigcup J$ est dans K'. Il en résulte que K' est close, et donc on a $K' = K$.

▶ L'axiome du choix est crucial dans plusieurs résultats de base concernant l'équipotence et la théorie des cardinaux, qui sera développée au chapitre V. On mentionne ici quelques points élémentaires sur les ensembles finis et dénombrables. ◀

LEMME 2.1. (AC) *Quels que soient A et B, il y a équivalence entre*
(i) *il existe une injection de A dans B;*
(ii) *il existe une surjection de B sur A.*

DÉMONSTRATION. L'implication $(i)\Rightarrow(ii)$ est toujours vraie (dans le contexte de Z). Inversement, si f est une surjection de B sur A, alors, par la proposition 1.13, AC garantit l'existence d'une section g pour f, laquelle est une injection de A dans B puisque, par définition, $g(x) = g(y)$ entraîne $x = f(g(x)) = f(g(y)) = y$. □

On a vu que AC implique que tout ensemble est bien ordonnable (théorème de Zermelo). Par le théorème de représentation, on déduit :

PROPOSITION 2.2. (ordinal) (AC) *Tout ensemble est en bijection avec un ordinal.*

COROLLAIRE 2.3. (AC) *Quels que soient les ensembles A, B, il existe une injection de A dans B ou une injection de B dans A.*

DÉMONSTRATION. Si α et β sont deux ordinaux, alors on a toujours $\alpha \subseteq \beta$ ou $\beta \subseteq \alpha$, donc l'application identité est une injection de α dans β ou une injection de β dans α. Si tout ensemble est en bijection avec un ordinal, le résultat s'exporte aux ensembles quelconques. □

L'axiome du choix permet également de clore certaines des questions concernant les ensembles dénombrables soulevées au chapitre I. On notera que les résultats suivants n'utilisent que l'axiome du choix dénombrable, qui est une hypothèse *a priori* plus faible que l'axiome du choix général.

PROPOSITION 2.4. (réunion dénombrable) (AC_ω) *Toute réunion dénombrable d'ensembles dénombrables est dénombrable.*

DÉMONSTRATION. Soit $(A_n)_{n\in\omega}$ une famille dénombrable d'ensembles dénombrables. Par hypothèse, il existe pour chaque entier n une bijection de ω sur A_n, mais, *a priori*, ce n'est pas la même chose que d'avoir une famille $(f_n)_{n\in\omega}$ telle que f_n soit une bijection de ω sur A_n : la différence consiste précisément à choisir une bijection pour chaque n. Or notons B_n l'ensemble des bijections de ω sur A_n. Par hypothèse, B_n est non vide pour chaque n, et la suite $(B_n)_{n\in\omega}$ est, elle, bien définissable à partir de la suite $(A_n)_{n\in\omega}$ qui existe par hypothèse. En effet, en notant $S = (A_n)_{n\in\omega}$, on a

$$(B_n)_{n\in\omega} = \{(n, f) \in \omega \times (\bigcup \mathrm{Im} S)^\omega \,;\, \text{« } f \text{ est une bijection de } \omega \text{ sur } S(n) \text{ »}\}.$$

Alors AC_ω garantit que le produit $\prod_{n\in\omega} B_n$ est non vide, c'est-à-dire qu'il existe une suite $(f_n)_{n\in\omega}$ telle que f_n est une bijection de ω sur A_n pour chaque n. Soient g et h les deux composantes d'une bijection de ω sur $\omega\times\omega$. Alors l'application $p \mapsto f_{g(p)}(h(p))$ est une surjection de ω sur $\bigcup_{n\in\omega} A_n$. □

PROPOSITION 2.5. (partie dénombrable) (AC_ω) *Tout ensemble infini possède un sous-ensemble dénombrable.*

Démonstration. Soit A un ensemble infini. Pour chaque entier n, notons I_n l'ensemble des injections de $[0, n[$ dans A. L'ensemble I_0 n'est pas vide, puisqu'il contient l'application vide (!). Ensuite, supposons I_n non vide. Cela signifie qu'il existe une injection f de $[0, n[$ dans A. Comme A est supposé infini, f n'est pas surjective, et il existe donc x dans $A \setminus \mathrm{Im} f$. Alors $f \cup \{(n, x)\}$ est une injection de $[0, n+1[$ dans A, et I_{n+1} n'est pas vide. Par conséquent, par induction, I_n est non vide pour tout entier n.

L'existence de la suite $(I_n)_{n \in \omega}$ est garantie par séparation à l'intérieur de $(A^\omega)^\omega$. Supposons que F est une fonction de choix sur l'ensemble dénombrable $\{I_n ; n \in \omega\}$, qui est l'image de la suite précédente considérée comme fonction de ω dans A^ω. On définit une fonction f de $\omega \times \omega$ dans A par $f(n, p) := F(I_n)(p)$ pour $p < n$: ainsi $f(n, p)$ est la valeur en p de l'injection choisie de n dans A. Soit B l'image de f. Pour tout n, l'ensemble B inclut l'image de l'injection $F(I_n)$, donc il ne peut être fini. Par ailleurs, f est une surjection de $\omega \times \omega$ sur B. En composant par une surjection de ω sur $\omega \times \omega$, on obtient une surjection g de ω sur B, puis une bijection h de ω sur B en définissant récursivement $h(q) := g(p)$ où p est le plus petit entier vérifiant $\forall r < q(g(p) \neq h(r))$. Par conséquent B est un sous-ensemble dénombrable de A. □

Corollaire 2.6. (AC_ω) *Un ensemble A est fini si et seulement si il n'existe pas d'injection non surjective de A dans lui-même.*

2.2. Ultrafiltres.

► Une conséquence directe de l'axiome du choix, cruciale pour les développements ultérieurs en topologie, est l'existence d'ultrafiltres et la possibilité de prolonger tout filtre en un ultrafiltre. ◄

On rappelle la notion de filtre. L'intuition, dans toute la suite, est que les éléments d'un filtre sur un ensemble I sont des parties de I qui sont grosses.

Définition 2.7. (filtre, ultrafiltre) Soit I un ensemble. Un ensemble non vide F de parties de I est appelé *filtre* sur I si (i) $\emptyset$ n'est pas dans F; (ii) si X est dans F et Y inclut X, alors Y est dans F; (iii) si X et Y sont dans F, alors $X \cap Y$ est dans F. Un filtre est appelé *ultrafiltre* si, de plus, (iv) si X n'est pas dans F, alors $I \setminus X$ est dans F.

Exemple 2.8. (filtre) Si I est un ensemble infini, les parties co-finies de I, c'est-à-dire celles dont le complémentaire est fini, forment un filtre. Ce filtre n'est pas un ultrafiltre puisqu'il existe des parties de I qui ne sont ni finies, ni co-finies.

Si F est un filtre sur I, alors I est nécessairement dans F par (ii). Si X est dans F, alors $I \setminus X$ n'est jamais dans F puisque, sinon, $X \cap (I \setminus X)$ y serait aussi, contredisant (i).

Proposition 2.9. (ultrafiltre) (AC) *Tout filtre peut être prolongé en un ultrafiltre.*

Démonstration. Soit F un filtre sur I. Il s'agit de compléter F en choisissant de façon cohérente un des deux éléments X, $I \setminus X$ pour chaque X telle que ni X, ni $I \setminus X$ n'est dans F. Pour ce faire, on fait appel au lemme de Zorn.

Soit A_F la famille de tous les filtres sur I qui incluent F. D'abord, $(A_F, \subseteq)$ est un ensemble ordonné inductif. En effet, soit $\{F_i ; i \in I\}$ une chaîne dans A_F, c'est-à-dire une famille de filtres sur I incluant F et deux à deux comparables vis-à-vis de l'inclusion. Soit $F' := \bigcup_i F_i$:

c'est un majorant de chacun des F_i, incluant F par construction. Si on montre que F' est un filtre, on déduit que la chaîne $\{F_i\,;\, i\in I\}$ a un majorant, et donc que $(A_F,\subseteq)$ est inductif.

Or, $\emptyset$ n'est dans aucun des F_i, donc il n'est pas dans leur réunion F', et F' satisfait (i). Ensuite, supposons $X \in F'$ et $X \subseteq Y$: cela signifie qu'il existe i tel que X est dans F_i, et, puisque F_i est un filtre, Y est aussi dans F_i, donc dans F', et F' satisfait (ii). Enfin, supposons $X, Y \in F'$. Il existe i, j tels que X est dans F_i et Y dans F_j. Par hypothèse, on a $F_i \subseteq F_j$ ou $F_j \subseteq F_i$. Dans le premier cas, on a $X \in F_j$ et $Y \in F_j$, donc $X \cap Y \in F_j$ puisque F_j est un filtre, et de là $X \cap Y \in F'$; de même dans le second cas avec F_i. Donc F' satisfait (iii), et est un filtre.

Par le lemme de Zorn, $(A_F,\subseteq)$ possède un élément maximal, qui, par construction, est un filtre sur I incluant F, et il ne reste qu'à montrer qu'un tel filtre maximal est un ultrafiltre. Pour cela, il suffit de montrer que, si F' est dans A_F et que J est une partie de I telle que ni J, ni $I \setminus J$ ne soit dans F', alors F' n'est pas maximal dans A_F. On remarque d'abord que, quel que soit Y dans F', on a $Y \cap J \neq \emptyset$. En effet, sinon, on aurait $Y \subseteq I \setminus J$, d'où $I \setminus J \in F'$, contrairement à l'hypothèse. Posons alors

$$F'' = \{X \subseteq I\,;\, \exists Y{\in}F'\,(Y \cap J \subseteq X)\}.$$

Si Y est dans F', on a $Y \cap J \subseteq Y$, donc Y est dans F'', et, par conséquent, F'' inclut F'. D'autre part, I est dans F' et on a $I \cap J \subseteq J$, donc J est dans F'', et par conséquent F'' inclut F' strictement. Reste à vérifier que F'' est un filtre. D'abord, F'' ne contient pas $\emptyset$ puisque, par hypothèse, $I \setminus J$ n'est pas dans F'. Donc F'' satisfait (i). Ensuite, $Y \cap J \subseteq X \subseteq X'$ implique $Y \cap J \subseteq X'$, donc F'' satisfait (ii). Enfin, supposons $X_1, X_2 \in F''$. Par hypothèse, il existe Y_1, Y_2 dans F' vérifiant $Y_i \cap J \subseteq X_i$ pour $i = 1, 2$. On déduit $(Y_1 \cap Y_2) \cap J \subseteq X_1 \cap X_2$, donc $X_1 \cap X_2 \in F''$ puisque $Y_1 \cap Y_2$ est dans F'. Donc F'' satisfait (iii). Par conséquent F'' est dans A_F, et F' n'est pas maximal dans A_F. □

2.3. Ordres.

▶ Dans le contexte des ordres, l'axiome des choix dépendants, donc *a fortiori* l'axiome du choix, permet de caractériser la propriété de bon ordre en termes de non-existence de suites descendantes, et garantit l'existence de branches dans certains arbres (théorème de König). ◀

On a vu au chapitre II qu'une relation bien fondée, donc en particulier un bon ordre, n'admet pas de suite infinie strictement décroissante. En présence de ACD, la condition est suffisante :

Proposition 2.10. (bon ordre) (ACD) *Supposons que R est une relation binaire bien fondée sur A n'admettant pas de suite infinie décroissante. Alors R est bien fondée. En particulier, si R est de surcroît un ordre total strict, c'est un bon ordre.*

Démonstration. Supposons que X est une partie de A ne possédant pas d'élément R-minimal. Cela signifie que, pour tout x dans A, il existe y dans A vérifiant $y\,R\,x$. Alors ACD montre l'existence d'une suite $(x_n)_{n\in\omega}$ d'éléments de X, donc de A, vérifiant $x_{n+1}\,R\,x_n$ pour tout n, c'est-à-dire l'existence d'une suite infinie décroissante pour R. □

Une autre conséquence de ACD est le théorème de König sur l'existence de branches dans les arbres infinis.

Définition 2.11. (arbre, niveau) On appelle *arbre* un ensemble ordonné $(T,<)$ possédant un plus petit élément et tel que les prédécesseurs de tout élément forment une chaîne finie. On dit alors qu'un élément x de T est *de niveau* n s'il existe une bijection de $[0, n[$ sur les prédécesseurs de x.

▷ *Un ensemble fini totalement ordonné, ou encore* $(\omega, <)$*, sont des arbres, mais il existe de nombreux arbres qui ne sont pas totalement ordonnés : par exemple, si* A *est un ensemble quelconque, alors l'ensemble* A^* *des suites finies d'éléments de* A *muni de l'ordre préfixe* $\sqsubset$*, c'est-à-dire* $\bigcup_{n\in\omega} A^n$ *ordonné par la relation « être un début de », est un arbre qui, si* A *a au moins deux éléments, n'est pas une chaîne : pour* $a \neq b$*, les suites* (a, a) *et* (a, b) *ne sont pas comparables, mais, par contre, les préfixes d'une suite donnée sont deux à deux comparables.* ◁

Proposition 2.12. (lemme de König) (ACD) *Supposons que* $(T, \prec)$ *est un arbre infini tel que, pour tout entier* n*, l'ensemble des éléments de niveau* n *est fini. Alors il existe une chaîne infinie dans* $(T, \prec)$*.*

Démonstration. Pour chaque élément x de T, on pose $T_x = \{y \in T\,;\, x \preccurlyeq y\}$. Alors $(T_x, \prec\restriction_{T_x})$ est un sous-arbre de $(T, \prec)$ dont le plus petit élément est x. On note T' l'ensemble des éléments x de T tels que T_x est infini, et on définit R par

$$R := \{(x, y) \in T' \times T'\,;\, « x \text{ est } \prec\text{-prédécesseur immédiat de } y »\}.$$

On note d'abord que T' n'est pas vide, puisqu'il contient au moins le plus petit élément de T. Ensuite tout élément non maximal x dans T possède des plus petits successeurs car, sinon, il existerait un majorant de x avec une infinité de prédécesseurs. Or, si x est de niveau n, tous les successeurs immédiats de x sont de niveau $n + 1$: donc, si nous supposons qu'il n'existe qu'un nombre fini d'éléments de T de niveau donné, nous déduisons que x n'a qu'un nombre fini de successeurs immédiats, soit $x_1, \ldots, x_k$. Par construction, on a $T_x = \{x\} \cup T_{x_1} \cup \cdots \cup T_{x_k}$, et donc, si T_x est infini, l'un au moins des ensembles T_{x_i} l'est aussi. Ceci est dire que, pour tout x dans T', il existe au moins un élément y dans T' vérifiant $(x, y) \in R$. L'axiome des choix dépendants, garantit alors l'existence d'une suite $(x_n)_{n\in\omega}$ telle que x_n est prédécesseur immédiat de x_{n+1} dans T pour tout entier n : une telle suite est en particulier une chaîne infinie dans T. □

2.4. Algèbre.

► On utilise le lemme de Zorn pour montrer l'existence de bases dans les espaces vectoriels et celle d'idéaux maximaux dans les anneaux. ◄

▷ *Le lemme de Zorn est le point de départ de la plupart des applications de l'axiome du choix en algèbre. Dans tous les cas, le principe est le même : pour montrer qu'il existe un certain objet avec telle ou telle propriété, on introduit une famille dans laquelle l'objet cherché soit un élément maximal : si la famille est inductive, ce qui est souvent facile à vérifier si l'ordre est une notion convenable d'inclusion, alors le lemme de Zorn permet d'affirmer l'existence d'un élément maximal.* ◁

Proposition 2.13. (base) (AC) *Toute partie libre d'un espace vectoriel peut être complétée en une base.*

Démonstration. Soit V un espace vectoriel et L une partie libre de V. Soit A_L l'ensemble des parties libres de V qui incluent L. Alors l'ensemble ordonné $(A_L, \subseteq)$ est inductif : il s'agit de montrer que, si $(L_i)_{i\in I}$ est une famille de parties libres de V incluant L et telle que, pour tous i, j dans I, on a $L_i \subseteq L_j$ ou $L_j \subseteq L_i$, alors il existe une partie libre L' de V incluant L et chaque L_i. Soit $L' = \bigcup_{i\in I} L_i$. Alors L' est libre. En effet, soit $\{e_i\,;\, i < n\}$ une famille finie d'éléments de L'. Supposons $\sum_{i<n} \lambda_i e_i = 0$. Pour chaque k, il existe i_k dans I tel que e_k est dans L_{i_k}. Comme les ensembles L_{i_k} sont deux à deux comparables, il existe i tel que L_i inclut

chacun des L_{i_k}. Alors tous les e_k sont dans L_i, et, comme L_i est libre, on déduit $\lambda_k = 0$ pour tout k. Donc L' est libre, c'est un majorant pour la chaîne $(L_i)_{i\in I}$, et $(A_L, \subseteq)$ est inductif.

Soit alors L' un élément maximal dans $(A_L, \subseteq)$, qui existe par le lemme de Zorn. Par construction, L' est une partie libre de V incluant L et telle qu'aucune partie incluant L' strictement ne soit libre. Soit e non nul dans V n'appartenant pas à L'. Alors, par hypothèse, $L' \cup \{e\}$ n'est pas libre, donc il existe une sous-famille finie $\{e_i\,;\, i < n\}$ de B et des scalaires $\{\lambda_i\,;\, i < n\}$ vérifiant $\lambda e + \sum_{i<n} \lambda_i e_i = 0$. Puisque L' est libre, λ est non nul, et on déduit $e = -\sum_{i<n} \lambda^{-1}\lambda_i e_i$. Par conséquent L' engendre V, dont il est une base. □

PROPOSITION 2.14. (idéal maximal) (AC) *Tout idéal propre d'un anneau est inclus dans un idéal propre maximal.*

DÉMONSTRATION. Soit R un anneau et I un idéal propre de R. Soit A_I la famille de tous les idéaux propres (c'est-à-dire ne contenant pas 1) incluant I. Alors $(A_I, \subseteq)$ est un ensemble ordonné inductif. En effet, si $(I_k)_{k\in K}$ est une chaîne dans $(A_I, \subseteq)$, posons $I' = \bigcup_{k\in K} I_k$. Alors I' est un idéal de R, car, pour a, b dans I', il existe k, j tels que a est dans I_k et b dans I_j ; par hypothèse, on a $I_k \subseteq I_j$, ou $I_j \subseteq I_k$; dans le premier cas, a et b sont dans I_j, et on a donc $a - b \in I_j \subseteq I'$, dans le second cas, on a de même $a - b \in I_k \subseteq I'$, et, dans tous les cas, I' est un sous-groupe du groupe additif de R. L'argument est similaire (et plus facile) pour la multiplication. De plus, I' est propre car, si 1 était dans I', il serait dans un des idéaux I_k. Alors I' est un majorant pour la chaîne $(I_k)_{k\in K}$, et $(A_I, \subseteq)$ est inductif.

Soit alors I' un élément maximal dans $(A_I, \subseteq)$, qui existe par le lemme de Zorn. Par construction, I' est un idéal propre de R incluant I et inclus dans aucun idéal propre de R, c'est donc un idéal propre maximal de R. □

▷ *On notera la grande similarité de toutes les démonstrations basées sur le lemme de Zorn. On peut extraire un principe commun, qui est que les diverses propriétés « être un filtre », « être une partie libre », « être un idéal » sont de caractère finitiste, ceci signifiant que la vérification de la propriété ne met en jeu que des sous-ensembles finis de l'ensemble considéré. Dans toute situation de ce type, le lemme de Zorn affirme l'existence d'objets maximaux.* ◁

2.5. Topologie.

► Les conséquences en topologie de l'axiome du choix et de ses formes faibles sont nombreuses. On montre ici comment utiliser AC_ω pour extraire des suites, puis on démontre le théorème de Tychonoff sur les produits d'espaces compacts. ◄

PROPOSITION 2.15. (clôture) (AC_ω) *Soit A un sous-ensemble d'un espace métrique X. Alors un élément a de X appartient à la clôture de A si et seulement si il existe une suite de points de A convergeant vers a.*

DÉMONSTRATION. Si on a $a = \lim a_n$ avec $a_n \in A$ pour tout n, alors a appartient à $\overline{A}$ de façon standard. C'est pour la réciproque qu'une forme d'axiome du choix est utilisée. Supposons $a \in \overline{A}$, et soit D_n le disque de centre a et de rayon $1/n$. Par hypothèse $D_n \cap A$ est une partie non vide de X pour tout n. Supposons que F est une fonction de choix sur l'ensemble dénombrable $\{D_n \cap A\,;\, n \in \omega\}$. Par construction, la suite $(F(D_n \cap A))_{n\in\omega}$ est une suite de points de A convergeant vers a. □

COROLLAIRE 2.16. (AC_ω) *Une fonction f d'un espace métrique dans un espace métrique est continue en a si et seulement si, pour toute suite $(a_n)_{n\in\omega}$ convergeant vers a, la suite des $f(a_n)$ converge vers $f(a)$.*

PROPOSITION 2.17. (théorème de Tychonoff) (AC) *Tout produit d'espaces compacts est compact.*

DÉMONSTRATION. Soit $(K_t)_{t\in T}$ une famille d'espaces compacts, et $K := \prod_{t\in T} K_t$. On note $\vec{x}$ un élément typique de K, et x_t est alors $\mathrm{pr}_t(\vec{x})$, c'est-à-dire la composante de $\vec{x}$ sur K_t.

Que K soit séparé ne pose pas de problème. Si $\vec{x}, \vec{y}$ sont des éléments de K distincts, il existe un indice t pour lequel x_t et y_t sont distincts. Comme K_t est séparé, il existe dans K_t des voisinages ouverts disjoints O_x de x_t et O_y de y_t, et alors $\mathrm{pr}_t^{-1}[O_x]$ et $\mathrm{pr}_t^{-1}[O_y]$ sont dans K des voisinages disjoints de $\vec{x}$ et $\vec{y}$.

Soit maintenant $\{F_i\,;\, i \in I\}$ une famille de fermés de K. Pour $\emptyset \neq J \subseteq I$, on pose $F_J := \bigcap_{i\in J} F_i$ [9]. On fait l'hypothèse $F_J \neq \emptyset$ pour J fini, et on veut montrer $F_I \neq \emptyset$.

Soit V l'ensemble des parties de K qui incluent (au moins) un ensemble F_J avec J fini non vide inclus dans I. Alors V est un filtre sur K. En effet, par hypothèse, $\emptyset$ n'est pas dans V puisque, par hypothèse, on a $F_J \neq \emptyset$, donc $F_J \not\subseteq \emptyset$ pour J fini non vide; ensuite $X' \supseteq X \supseteq F_J$ entraîne $X' \supseteq F_J$; enfin la conjonction de $X \supseteq F_J$ et $Y \supseteq F_K$ entraîne $X \cap Y \supseteq F_{J\cup K}$. Par la proposition 2.9, il existe un ultrafiltre U sur K étendant V.

Pour chaque t dans T, on pose $L_t := \bigcap_{X\in U} \overline{\mathrm{pr}_t[X]}$. On va montrer que L_t est un singleton. On montre d'abord que L_t est non vide. Puisque K_t est compact, il suffit de montrer que, pour toute partie finie U_0 de U, l'ensemble $\bigcap_{X\in U_0} \overline{\mathrm{pr}_t[X]}$ est non vide : or cet ensemble inclut $\mathrm{pr}_t[\bigcap_{X\in U_0} X]$, lequel n'est pas vide puisque U est un filtre. Reste à montrer que L_t contient au plus un élément. Pour commencer, on remarque que, pour x dans K_t,

(2.1) si x a un voisinage ouvert O vérifiant $\mathrm{pr}_t^{-1}[O] \notin U$, alors x n'est pas dans L_t.

En effet, dans ce cas, si on pose $X := \mathrm{pr}_t^{-1}[K_t \setminus O]$, alors X est dans U puisqu'il inclut $K \setminus \mathrm{pr}_t^{-1}[O]$; par ailleurs, puisque $K_t \setminus O$ est fermé, on a $\overline{\mathrm{pr}_t[X]} \subseteq \overline{K_t \setminus O} = K_t \setminus O$, d'où $x \notin \overline{\mathrm{pr}_t[X]}$, et $x \notin L_t$. On en déduit que L_t contient au plus un élément. En effet, puisque K_t est un espace séparé, si x, y sont des éléments de K_t distincts, ils admettent des voisinages ouverts disjoints O_x, O_y. Comme $\mathrm{pr}_t^{-1}[O_x]$ et $\mathrm{pr}_t^{-1}[O_y]$ sont disjoints, l'un au plus est dans U, donc, par (2.1), l'un au plus de x, y est dans L_t.

Finalement, soit $\vec{x}$ un élément de K vérifiant $\vec{x} \notin F_i$ pour un certain i. Alors on a $x_t \notin L_t$ pour au moins un t. En effet, F_i est un fermé, donc il existe un voisinage ouvert élémentaire O de $\vec{x}$ disjoint de F_i. Par définition, O est de la forme $\bigcap_{t\in T_0} \mathrm{pr}_t^{-1}[O_t]$, avec T_0 sous-ensemble fini de T et O_t ouvert de K_t pour $t \in T_0$. Par construction, F_i est dans V, donc dans U, donc, puisque O est disjoint de F_i, on a $\mathrm{pr}_t^{-1}[O_t] \notin U$ pour (au moins) un t dans T_0 ; comme on a $x_t \in O_t$ par construction, on déduit $x_t \notin L_t$ par (2.1).

La démonstration est terminée : soit $\vec{x}$ l'élément de K tel que x_t est l'unique élément de L_t pour chaque t. D'après ce qui précède, $\vec{x}$ est dans chacun des F_i, donc dans F_I, qui n'est non vide. Donc K est compact. □

2.6. Analyse.

► On établit deux applications de l'axiome du choix en analyse, à savoir le théorème de Hahn–Banach sur le prolongement des formes linéaires, et l'existence d'ensembles non mesurables pour la mesure de Lebesgue. ◄

[9] on a donc en particulier $F_{\{i\}} = F_i$

Proposition 2.18. (théorème de Hahn–Banach) *Soient E un $\mathbb{R}$-espace vectoriel, F un sous-espace de E, N une sous-norme sur E* [10], *et f une forme linéaire sur F tels qu'on ait $f(x) \leqslant N(x)$ pour tout x dans F. Alors il existe une forme linéaire $\widehat{f}$ définie sur E, prolongeant f, et telle qu'on ait $\widehat{f}(x) \leqslant N(x)$ pour tout x dans E.*

Démonstration. Soit A l'ensemble de toutes les formes linéaires g prolongeant f, définies sur un sous-espace de E et telles qu'on ait $g(x) \leqslant N(x)$ pour tout x dans le domaine de g. On ordonne A par inclusion.

On montre d'abord que $(A, \subseteq)$ est inductif. Pour cela, soit C une chaîne de A. On pose $g = \bigcup C$. Comme C est composé de fonctions deux à deux compatibles, g est fonctionnelle. Supposons que x et y appartiennent au domaine de g. Alors il existe g_x et g_y dans C tels que x est dans le domaine de g_x et y dans celui de g_y. L'hypothèse que C est une chaîne entraîne que g_x prolonge g_y, ou g_y prolonge g_x. Suppossons par exemple le premier cas. Alors y est dans le domaine de g_x, et donc, pour tous λ, μ dans $\mathbb{R}$, il en est de même de $\lambda x+\mu y$. Donc le domaine de g est un sous-espace de E, et on a $g(\lambda x + \mu y) = g_x(\lambda x + \mu y) = \lambda g_x(x) + \mu g_x(y) = \lambda g(x) + \mu g(y)$, donc g est une forme linéaire. Par ailleurs, on a $g(x) = g_x(x) \leqslant N(x)$. Donc g appartient à A.

On suppose maintenant que g est un élément de A dont le domaine G est un sous-espace propre de E. Soit z un vecteur de E n'appartenant pas à G. Soit G' le sous-espace engendré par G et z. On va montrer que g peut être prolongée en un élément g' de A dont le domaine est G', ce qui permettra de conclure que g n'est pas maximal dans A. Par conséquent, un élément maximal, dont l'existence est garantie par le lemme de Zorn, est une forme linéaire définie sur E entier et satisfaisant toutes les conditions souhaitées.

Les éléments de G' s'écrivent de façon unique $x + \lambda z$ avec x dans G, et, si g' prolonge g, on doit avoir $g'(x + \lambda z) = g(x) + \lambda g'(z)$. Il s'agit de choisir le réel $g'(z)$ de façon à ce qu'on ait $g(x) + \lambda g'(z) \leqslant N(x + \lambda z)$ pour tout x dans G et tout λ dans $\mathbb{R}$. L'inégalité étant satisfaite pour $\lambda = 0$ par hypothèse, il s'agit de la satisfaire à la fois pour $\lambda > 0$ et pour $\lambda < 0$, c'est-à-dire de garantir, pour tous x, y dans G et tout λ strictement positif, à la fois $g(x)+\lambda g'(z) \leqslant N(x+\lambda z)$ et $g(y) + (-\lambda)g'(z) \leqslant N(y + (-\lambda)z)$, soit

$$g(\lambda^{-1}y) - N(\lambda^{-1}y - z) \leqslant g'(z) \leqslant N(\lambda^{-1}x + z) - g(\lambda^{-1}x). \tag{2.2}$$

Pour tous u, v dans G, on a par hypothèse $g(u)+g(v) = g(u+v) \leqslant N(u+v) \leqslant N(u-z)+N(v+z)$, donc

$$g(u) - N(u - z) \leqslant N(v + z) - g(v). \tag{2.3}$$

Posons $c_- := \sup\{g(u) - N(u - z) \,;\, u \in G\}$ et $c_+ := \inf\{N(v + z) - g(v) \,;\, v \in G\}$. Alors (2.3) garantit $c_- \leqslant c_+$, et tout choix de $g'(z)$ entre c_- et c_+ satisfait (2.2). □

Proposition 2.19. (non mesurable) (AC) *Il existe un sous-ensemble de $\mathbb{R}$ non mesurable pour la mesure de Lebesgue.*

Démonstration. Notons μ la mesure de Lebesgue. Pour x, y dans $\mathbb{R}$, notons $x \equiv y$ pour $x - y \in \mathbb{Q}$. Alors $\equiv$ est une relation d'équivalence sur $\mathbb{R}$. Par le corollaire 1.12, il existe un sous-ensemble A de $\mathbb{R}$ qui contient un point exactement dans chaque classe d'équivalence.

Supposons A Lebesgue-mesurable. Par construction, $\mathbb{R}$ est réunion disjointe des ensembles $A + q$ pour q dans $\mathbb{Q}$. Puisque $\mathbb{Q}$ est dénombrable, ceci entraîne $\mu(A) > 0$. Comme A est la réunion des $A \cap [k, k + 1]$ pour $k \in \mathbb{Z}$, il existe n vérifiant $\mu(A \cap [n, n + 1]) > 0$. Posons

[10] c'est-à-dire une application $N : E \to \mathbb{R}_+$ vérifiant $N(\lambda x) = \lambda N(x)$ pour $\lambda \geqslant 0$ et $N(x + y) \leqslant N(x) + N(y)$ pour tous x, y

$B = A \cap [n, n+1]$. Alors $\bigcup_{q \in \mathbb{Q} \cap [0,1]} (B+q)$ a une mesure infinie, alors qu'il est inclus dans $[n, n+2]$, dont la mesure est 2. L'hypothèse que A est mesurable est donc à rejeter. □

2.7. Géométrie.

▶ On établit une application surprenante de l'axiome du choix en géométrie, à savoir le paradoxe de Banach–Tarski affirmant la possibilité de découper une sphère de $\mathbb{R}^3$ en un nombre fini de pièces et de réarranger celles-ci par des rotations de façon à obtenir deux copies de la sphère initiale. ◀

Dans ce qui suit, on écrit $A = A_1 \sqcup ... \sqcup A_n$ pour « A est union disjointe des A_i », c'est-à-dire si $\{A_1, ..., A_n\}$ est une partition de A.

Proposition 2.20. (paradoxe de Banach–Tarski) (AC) *Il existe une décomposition $S = S_1 \sqcup \ldots \sqcup S_6$ de la sphère-unité de $\mathbb{R}^3$ en six pièces disjointes et six rotations $h_1, ..., h_6$ de $\mathbb{R}^3$ telles qu'on ait*

$$S = h_1[S_1] \sqcup h_2[S_2] \sqcup h_3[S_3] = h_4[S_4] \sqcup h_5[S_5] \sqcup h_6[S_6]. \tag{2.4}$$

Le même résultat existe avec une décomposition en quatre pièces, mais la démonstration est plus longue. Dans tous les cas, le résultat provient des propriétés d'autosimilarité du groupe libre de rang 2.

Définition 2.21. (groupe libre, mot réduit) On note F un groupe *libre* de base $\{a, b\}$, réalisé comme l'ensemble des *mots réduits* formés sur les quatre lettres a, b, a^{-1}, et b^{-1}, ces derniers étant définis comme les mots ne contenant aucun motif du type xx^{-1} ou $x^{-1}x$; le *produit* de deux mots réduits u, v est le mot obtenu en concaténant u et v et supprimant itérativement les motifs xx^{-1} et $x^{-1}x$ jusqu'à obtention d'un mot réduit.

Par exemple, le produit de aba^{-1} et $ab^{-1}a$ est aa. On note 1 pour le mot vide, élément neutre de F.

Lemme 2.22. *Il existe une partition $\{F_1, ..., F_4\}$ de F vérifiant*

$$F = F_1 \sqcup F_2 a = F_3 \sqcup F_4 b. \tag{2.5}$$

Démonstration. Posons
$F_1 := \{u \in F\,;\, u$ se termine par a ou $u = 1$ ou $u = a^{-n}$ avec $n \geqslant 1\}$,
$F_2 := \{u \in F\,;\, u$ se termine par a^{-1} et n'est pas de la forme a^{-n} avec $n \geqslant 1\}$,
$F_3 := \{u \in F\,;\, u$ se termine par $b\}$,
$F_4 := \{u \in F\,;\, u$ se termine par $b^{-1}\}$.
Que $F_1, ..., F_4$ forment une partition de F est clair. Ensuite, si u n'est pas dans F_3, alors ub^{-1} est réduit. On a donc $u = ub^{-1} \cdot b$, d'où $F \setminus F_3 \subseteq F_4 b$. Inversement, l'avant-dernière lettre d'un mot de F_4, si elle existe, ne peut être un b (sinon ce mot ne serait pas réduit), donc un mot dans $F_4 b$ ne saurait se terminer par b, et on a donc $F_4 b \subseteq F \setminus F_3$, d'où $F_4 b = F \setminus F_3$, soit $F = F_3 \sqcup F_4 b$.

L'argument pour F_1 et F_2 est similaire. Si u n'est pas dans F_1, alors ua^{-1} est réduit et on a $u = ua^{-1} \cdot a$. L'hypothèse que u n'est pas dans F_1 entraîne que ua^{-1} n'est pas de la forme a^{-n} avec $n \geqslant 1$, donc ua^{-1} est dans F_2, et on a $F \setminus F_1 \subseteq F_2 a$. Inversement, l'avant-dernière lettre d'un mot de F_2, si elle existe, ne peut être un a, donc un mot de $F_2 a$ ne peut se terminer par a. Par ailleurs, un mot de F_2 ne peut être de la forme a^{-n} avec $n \geqslant 1$. Donc un mot de F_2a ne peut être ni 1, ni un mot de la forme a^{-n} avec $n \geqslant 1$. Autrement dit, on a $F_2a \subseteq F \setminus F_1$, d'où $F_2a = F \setminus F_1$, soit $F = F_1 \sqcup F_2a$. $\square$

LEMME 2.23. *Soient f, g les rotations de $\mathbb{R}^3$ d'angle* $\arccos(3/5)$ *et d'axes respectifs Oz et Ox. Le sous-groupe de* SO_3 *engendré par f et g est libre de rang* 2.

DÉMONSTRATION. Les rotations f et g sont représentées dans la base canonique par les matrices $\begin{pmatrix} 3/5 & 4/5 & 0 \\ -4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 4/5 \\ 0 & -4/5 & 3/5 \end{pmatrix}$. Soit φ le morphisme du groupe libre F dans SO_3 qui envoie a sur f et b sur g. Il s'agit de montrer que, si u est un mot non vide, alors la matrice M_u de $\varphi(u)$ n'est pas la matrice identité. Si u est un mot de longueur ℓ, les coefficients de M_u peuvent s'écrire $x/5^\ell$ avec x entier. Soit $(p_u/5^\ell, q_u/5^\ell, r_u/5^\ell)$ la première ligne de M_u. Les relations d'induction sont $p_\varepsilon = 1$, $q_\varepsilon = r_\varepsilon = 0$ et

$$\begin{aligned} p_{ua^{\pm1}} &= 3p_u \mp 4q_u, & q_{ua^{\pm1}} &= \pm 4p_u + 3q_u, & r_{ua^{\pm1}} &= 5r_u, \\ p_{ub^{\pm1}} &= 5p_u, & q_{ub^{\pm1}} &= 3q_u \mp 4r_u, & r_{ub^{\pm1}} &= \pm 4q_u + 3r_u. \end{aligned}$$

Par construction, 5 divise r_u si u se termine par $a^{\pm1}$, et 5 divise p_u si u se termine par $b^{\pm1}$. Par ailleurs, un calcul direct donne $q_{uxx} = 6q_{ux} - 25q_u$ pour $x = a, a^{-1}, b$ et b^{-1}. On montre alors par induction sur la longueur de u que, si u n'est pas de la forme b^n, alors 5 ne divise pas q_u. Le résultat est clair pour u de longueur 1. Supposons u de longueur au moins 2, soit $u = vxy$. On considère $y = a^{\pm1}$. Pour $x = b^{\pm1}$, on a $q_u = \mp 4p_{vx} + 3q_{vx}$. Si v ne contient que des $b^{\pm1}$, alors on a $p_{vx} = 1$, et $q_{vx} = 0$, donc le résultat est vrai. Sinon, on sait que 5 divise p_{vx}, et l'hypothèse d'induction est que 5 ne divise pas q_{vx}, donc 5 ne divise pas q_u. Pour $x = a^{\pm1}$ maintenant, on a $q_u = 6q_{vx} - 25q_v$, et l'hypothèse d'induction est que 5 ne divise pas q_{vx}, donc à nouveau 5 ne divise pas q_u. Les cas $y = b^{\pm1}$ sont traités de même. La conclusion est que $\varphi(u)$ n'est pas l'identité quand u n'est pas de la forme b^n avec $n \in \mathbb{Z}$. Puisque $\arccos(3/5)$ n'est pas commensurable avec π, il en est de même pour u de la forme b^n avec $n \neq 0$. $\square$

On peut maintenant montrer un résultat proche de la proposition 2.20.

LEMME 2.24. (AC) *Il existe une décomposition $S = D \sqcup X_1 \sqcup \ldots \sqcup X_4$ avec D dénombrable et deux rotations f, g de $\mathbb{R}^3$ vérifiant*

$$S \setminus D = X_1 \sqcup f[X_2] = X_3 \sqcup g[X_4]. \tag{2.6}$$

DÉMONSTRATION. Soient f, g les rotations du lemme 2.23. On définit une action (à droite) du groupe F sur la sphère S en posant $P{*}a := f(P)$ et $P{*}b := g(P)$ [11]. Soit D l'ensemble des points de S laissés fixes par au moins un mot non vide de F : comme le groupe F est dénombrable, et que chaque rotation laisse exactement deux points de S fixes, à savoir les intersections de son axe avec S, l'ensemble D est dénombrable.

Soit P un point de $S \setminus D$. Alors le stabilisateur de P est trivial, c'est-à-dire que l'application de F dans S qui envoie u sur $P{*}u$ est injective. Autrement dit, l'orbite $P{*}F$ de P est une copie de F (figure 4). On peut alors utiliser le lemme 2.22 pour décomposer cette orbite, en quatre parties disjointes $P{*}F_1, ..., P{*}F_4$ vérifiant

$$P{*}F = P{*}F_1 \sqcup P{*}F_2a = P{*}F_3 \sqcup P{*}F_4b. \tag{2.7}$$

[11] Le changement de côté est là pour tenir compte de ce que, dans $f \circ g$, on fait d'abord g.

Maintenant, pour tout mot réduit u, on a $P{*}ua = (P{*}u){*}a = f(P{*}u)$, donc l'ensemble $P{*}F_2a$ est l'image de $P{*}F_2$ par la rotation f et, de même, $F_4b{*}P$ est l'image de $F_4{*}P$ par la rotation g. Par conséquent, (2.7) se récrit

$$P{*}F = P{*}F_1 \sqcup f[P{*}F_2] = F_3{*}P \sqcup g[P{*}F_4]. \tag{2.8}$$

Utilisons alors AC pour extraire de $S \setminus D$ un sous-ensemble X contenant un point exactement dans chaque orbite, et posons $X_i := \bigcup_{P\in X} F_i{*}P$ pour $i = 1, ..., 4$. Par construction, les ensembles X_i, $i = 1, ..., 4$, forment une partition de $S \setminus D$, et (2.8) donne directement (2.6). $\square$

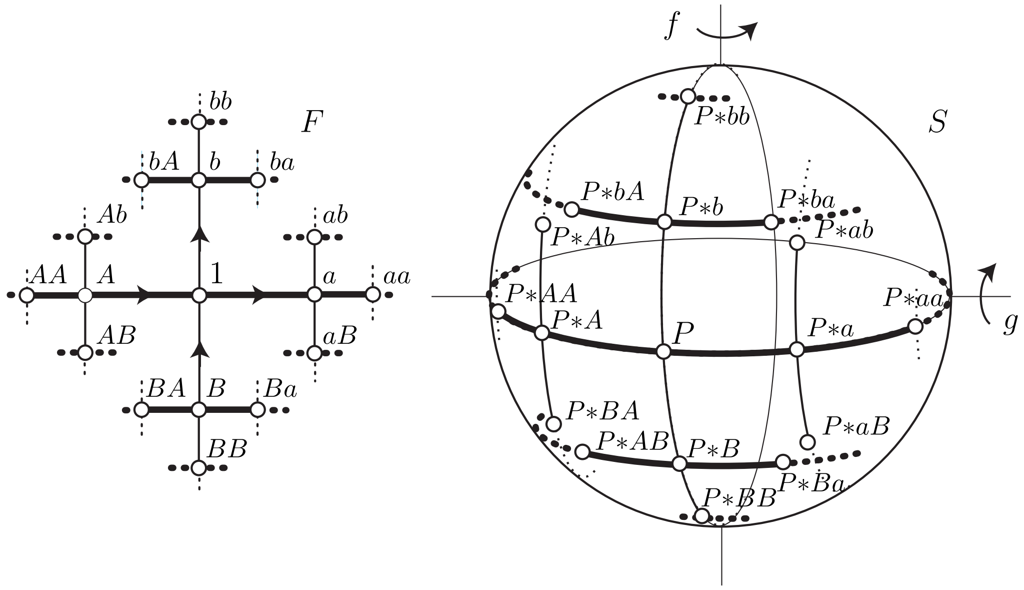

Figure 4. Décomposition d'une orbite : la décomposition du groupe F, ici représenté par son graphe de Cayley — on a écrit A pour a^{-1} et B pour b^{-1} — induit une décomposition analogue de chaque orbite non dégénérée sur S ; par construction, les translations par a et b sur F correspondent aux rotations f et g sur S.

On peut maintenant conclure.

Démonstration de la proposition 2.20. On reprend les notations du lemme 2.24. Il s'agit d'éliminer l'ensemble dénombrable D, ce qui est facile, quitte à augmenter le nombre de pièces. Puisque D est dénombrable, on peut trouver une rotation h ne laissant fixe aucun point de D. Posons $D^* := \bigcup_{n\geqslant 0} h^n[D]$, puis $Y_1 := h[D^*]$ et $Y_2 := S \setminus D^*$. Par construction on a

$$S \setminus D = Y_1 \sqcup Y_2 \quad \text{et} \quad S = h^{-1}[Y_1] \sqcup Y_2.$$

On a ainsi deux décompositions de $S \setminus D$, l'une en $X_1 \sqcup \ldots \sqcup X_4$ et l'autre en $Y_1 \sqcup Y_2$. On peut raffiner ces deux partitions en une partition unique en posant $Z_{i,j} := X_i \cap Y_j$ pour $1 \leqslant i \leqslant 4$ et $1 \leqslant j \leqslant 2$. Les pièces $Z_{i,j}$ peuvent se réarranger d'une part en S, d'autre part en deux copies de $S \setminus D$. On en déduit facilement un réarrangement de S en deux copies de S. Posons (*cf.* figure 5) $S_1 := h[Z_{1,1}] \sqcup Z_{1,2}$, $S_2 := h[Z_{2,1}]$, $S_3 := Z_{2,2}$, $S_4 := h[Z_{3,1}] \sqcup Z_{3,2}$, $S_5 := h[Z_{4,1}]$, et $S_6 := Z_{4,2}$. Par construction, les pièces S_i constituent une partition de S. La relation $f[X_2] = X_2 \sqcup X_3 \sqcup X_4$ implique $f[Z_{2,j}] = Z_{2,j} \sqcup Z_{3,j} \sqcup Z_{4,j}$ pour $j = 1, 2$, d'où

$$S_1 \sqcup hfh^{-1}[S_2] \sqcup f[S_3] = h[Z_{1,1}] \sqcup hf[Z_{2,1}] \sqcup Z_{1,2} \sqcup f[Z_{2,2}]$$

$$= h[Z_{1,1} \sqcup \ldots \sqcup Z_{4,1}] \sqcup (Z_{2,1} \sqcup \ldots \sqcup Z_{4,2}) = h[Y_1] \sqcup Y_2 = S.$$

On a donc le résultat escompté en posant $h_1 := \mathrm{id}$, $h_2 := hfh^{-1}$, et $h_3 := f$, puis, de la même façon, $h_4 := \mathrm{id}$, $h_5 := hgh^{-1}$, et $h_6 := g$. □

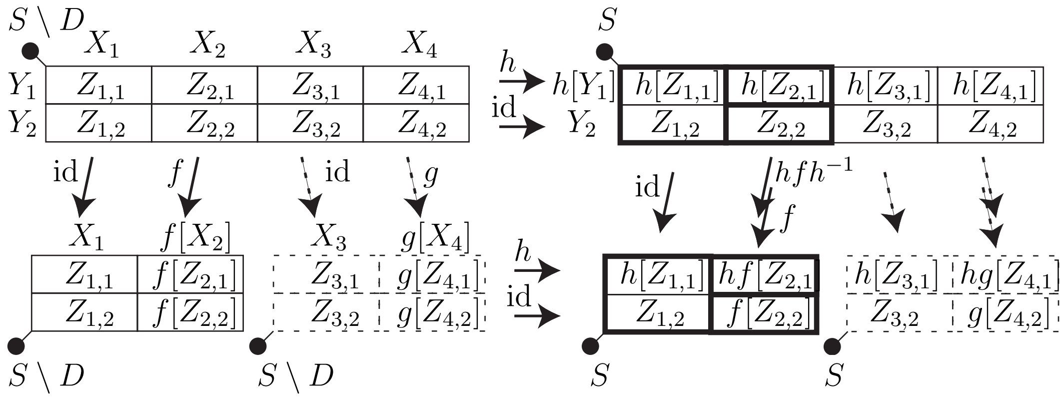

FIGURE 5. Transitivité de l'équidécomposabilité : on peut décomposer $S \setminus D$ en quatre pièces qui se réarrangent en deux copies de $S \setminus D$, et en deux pièces qui se réarrangent en S ; en considérant les huit intersections, on peut réarranger S en deux copies de $S \setminus D$, puis même de S ; on peut finalement regrouper des pièces associées aux mêmes rotations pour n'avoir que six pièces en tout.

▷ *On peut obtenir une décomposition à quatre pièces en partant du lemme 2.24 et en traitant directement toutes les orbites y compris celles qui ont un stabilisateur non trivial. Les détails n'importent pas ici, puisqu'on veut mettre l'accent sur l'utilisation de* AC. *Noter que les pièces d'une décomposition telle que mentionnée dans la proposition 2.20 sont composées de points épars et, par exemple, ne peuvent être mesurables pour la mesure de Lebesgue puisque, sinon, celle-ci étant préservée par rotation, on aurait* $1 = 1 + 1$. *En un sens, le résultat ne dit pas beaucoup plus que le fait (trivial) qu'il existe une bijection entre* S *et* $S \sqcup S$, *si ce n'est qu'ici la bijection est construite comme collage de six rotations.*

Le résultat précédent est anecdotique. Par contre, il a mené à de vastes développements en théorie de la mesure. R. Dougherty et M. Foreman ont montré en 1991 que, dans la proposition 2.20, on peut supposer que les pièces ont la propriété de Baire, c'est-à-dire ne diffèrent d'un ouvert que par un ensemble maigre, c'est-à-dire une réunion dénombrable d'ensembles nulle part denses. Un corollaire remarquable de leur construction (extrêmement sophistiquée) est qu'on peut obtenir sans axiome du choix une décomposition approchée dont le défaut est un ensemble maigre. ◁

3. L'axiome du choix est-il vrai ?

► On discute briévement le statut de l'axiome du choix et l'opportunité ou non de l'inclure dans les axiomes de base de la théorie des ensembles. ◄

▷ *Dégagé et énoncé explicitement au début du XXe siècle, l'axiome du choix a suscité des discussions innombrables où partisans et adversaires se sont parfois violemment opposés. Aujourd'hui, les positions sont plus nuancées : on sait que* AC *n'est ni prouvable ni réfutable à partir des axiomes de* ZF, *mais ceci ne dit rien quant à l'opportunité de l'adopter ou non comme axiome additionnel. Ce dernier point dépend surtout du type d'objet, effectif ou non, auquel on s'intéresse.* ◁

3.1. Une question ambiguë.

▶ La question brutale « l'axiome du choix est-il vrai ? » n'est pas bien posée, et elle recouvre deux questions de natures complètement différentes, l'une objet d'une possible démonstration, l'autre seulement objet d'un possible consensus. ◀

▷ *Avant de pouvoir — éventuellement — répondre à la question de la vérité de l'axiome du choix, il convient de s'entendre sur ce que signifie la vérité d'un enoncé mathématique. Un énoncé peut certainement être tenu pour vrai s'il en existe une démonstration convaincante. Le problème est qu'une démonstration ne procède pas ex nihilo, mais consiste à dériver de nouvelles propriétés à partir d'autres qui soit ont été démontrées antérieurement, soit ont fait l'objet d'un consensus pour être prises comme point de départ axiomatique. Dès lors que le système* ZF *constitue une base possible pour la théorie des ensembles, deux questions distinctes se posent donc en présence d'un énoncé tel que l'axiome du choix :* ◁

Question 3.1. *L'axiome du choix — ou sa négation — est-il conséquence des autres principes de base, donc, dans le contexte présent, des axiomes du système* ZF *?*

Question 3.2. *Est-il opportun — ou, au contraire, déconseillé — d'inclure l'axiome du choix dans les principes de base des mathématiques ?*

3.2. L'axiome du choix est-il conséquence de ZF ?

▶ Ce sera un des buts de la suite de ce texte de démontrer que la réponse à cette question est négative, dans les deux directions — ou au moins d'en esquisser une démonstration. ◀

Théorème 3.3. (Gödel, 1938) *S'il est cohérent, le système* ZF *ne réfute pas l'axiome du choix, c'est-à-dire qu'il n'existe pas de preuve de la négation de* AC *à partir des axiomes du système* ZF.

Théorème 3.4. (Cohen, 1963) *S'il est cohérent, le système* ZF *ne démontre pas l'axiome du choix, c'est-à-dire qu'il n'existe pas de preuve de* AC *à partir des axiomes du système* ZF.

3.3. Est-il opportun d'adopter l'axiome du choix ?

▶ Les objets construits avec l'aide de l'axiome du choix n'ont pas la même qualité d'existence que ceux qui ont été définis explicitement. Sans s'interdire d'utiliser AC, il est naturel de chercher à en éviter l'usage quand on le peut, et, suivant les contextes, de mentionner explicitement les usages qu'on en fait. ◀

▷ *Les résultats précédents ne répondent en rien à la question de l'opportunité d'adopter ou non l'axiome du choix : tout au plus garantissent-ils que l'introduction de l'axiome du choix ou de sa négation ne risquent pas d'introduire de contradiction à la façon dont le ferait par exemple postuler l'existence d'un ensemble de tous les ensembles.*

Ne reste alors qu'à discuter de l'opportunité d'accepter ou de refuser l'axiome du choix, notamment en interrogeant l'intuition et la pratique des mathématiciens. Le cœur de la discussion concerne la nature des objets mathématiques qu'on souhaite étudier. Ainsi qu'on l'a vu dans ce chapitre, l'intérêt de l'axiome du choix est de permettre sans difficulté la construction d'objets qu'il serait difficile ou impossible de spécifier complètement. On sent bien qu'une base du $\mathbb{Q}$*-espace vectoriel* $\mathbb{R}$ *est un objet très bizarre, et d'un tout autre type que l'ensemble* $\{1, 2, 3\}$

ou le nombre π. Est-il légitime de considérer qu'on a défini un objet mathématique lorsqu'on a répété (par la pensée) une infinité de fois une opération de choix qui, prise individuellement, ne pose pas de problème ? Pour qui pense que les objets mathématiques doivent être explicitement définis, à la façon des objets qu'un ordinateur peut manipuler, alors la réponse sera probablement négative. Pour celui au contraire qui est prêt à considérer globalement tout objet que la pensée peut saisir même si une description exhaustive et effective fait défaut, alors la réponse sera tout aussi probablement positive. Les points de vue les plus divers ont été développés, reflétant les positions de leurs auteurs quant à ce que sont les objets mathématiques, à leur contingence ou à leur nécessité, à leur caractère absolu ou relatif, objectif ou subjectif. Ce débat-là n'est pas tranché, et ne le sera probablement jamais. Au demeurant, il est facile d'éviter les querelles en renonçant à l'espoir d'un système universel et en développant pragmatiquement plusieurs systèmes reposant sur des bases différentes. Dans le cas de l'axiome du choix, cette solution est très facile : elle correspond à mentionner explicitement tous les usages de l'axiome, ainsi que nous l'avons fait dans ce chapitre. De la sorte, le lecteur a le choix des principes de base qu'il souhaite adopter : s'il accepte l'axiome du choix, tous les théorèmes sont démontrés, s'il le refuse, il ne doit tenir pour démontrés que les théorèmes où l'axiome du choix n'est pas utilisé. L'un des intérêts des controverses sur l'axiome du choix aura été de pousser les mathématiciens à limiter, autant que faire se peut, les appels à un axiome moins indiscutable que les autres axiomes de ZF.

On notera que l'utilisation, à un moment donné, de AC *comme outil de démonstration pour un certain résultat n'entraîne pas que cet axiome soit indispensable : il peut exister une autre démonstration du même résultat, ou d'une variante de celui-ci, qui n'utilise pas* AC. *Il serait donc regrettable et imprudent, même pour un mathématicien (ou un informaticien) exclusivement intéressé aux objets effectifs, de rejeter en bloc tous les résultats établis à partir de* AC. *Un exemple typique est la riche collection de résultats de théorie de la mesure issus du paradoxe de Banach–Tarski (section 2.7), en particulier le théorème de Dougherty–Foreman dont on a dit que certaines versions ne font absolument pas appel à l'axiome du choix, lesquelles n'auraient probablement jamais été découvertes si on avait rejeté* AC *d'emblée.* ◁

3.4. Le système ZFC.

► Il est usuel d'adjoindre AC au système ZF, parvenant ainsi au système ZFC qui est le cadre standard de la théorie des ensembles. ◄

▷ *Le point de vue de la théorie des ensembles est d'explorer une notion d'objet mathématique aussi large que possible, et en particulier de ne pas requérir de condition d'effectivité préalable. Il est donc cohérent d'adjoindre l'axiome du choix aux principes de base. Cette position correspond à l'option de construire un cadre très général à l'intérieur duquel des théories plus spécifiques peuvent être développées le cas échéant. Dans une telle approche, il est naturel de considérer la forme la plus libérale, la moins contraignante d'existence pour les objets mathématiques. C'est ainsi qu'on a, avec l'axiome des parties, posé pour chaque ensemble A l'existence d'un ensemble de tous les sous-ensembles de A, sans préciser comment ces sous-ensembles sont délimités, quitte à introduire ultérieurement un cadre plus restrictif où une spécification plus précise des sous-ensembles est requise.* ◁

Définition 3.5. (système ZFC) On note ZFC le système axiomatique obtenu en ajoutant au système de Zermelo–Fraenkel ZF l'axiome du choix AC.

▷ *Sauf mention du contraire, le système* ZFC *est le cadre formel pour toute la suite de ce texte, c'est-à-dire qu'on s'autorise à utiliser tous les axiomes de* ZFC *pour justifier les constructions effectuées, mais seulement ceux-là : toute construction faisant appel à d'autres hypothèses de base que des énoncés démontrables à partir de* ZFC *sera mentionnée explicitement.* ◁

Exercices

Exercice 1. (AC_ω) Montrer que l'axiome du choix dénombrable AC_ω est équivalent à l'énoncé: « Si S est un sous-ensemble dénombrable de $\mathfrak{P}(A) \setminus \{\emptyset\}$, alors il existe $F : S \to A$ vérifiant $F(x) \in x$ pour tout x ».

Exercice 2. (bons ordres) Montrer par un argument direct que le lemme de Zorn entraîne le théorème de Zermelo. [Considérer l'ensemble W des bons ordres définis sur un sous-ensemble de l'ensemble A à bien ordonner, ordonner W par la relation « être un segment initial de », et montrer qu'un élément maximal est un bon ordre défini sur A entier.]

Exercice 3. (clôture algébrique) Montrer, avec l'axiome du choix, que tout corps possède une clôture algébrique.

Exercice 4. (Hahn–Banach) Démontrer la forme géométrique suivante du théorème de Hahn–Banach : Si E est un espace vectoriel topologique, O un ouvert convexe non vide, et M un sous-espace affine non vide ne rencontrant pas O, alors il existe au moins un hyperplan fermé affine H contenant M et ne rencontrant pas O.

Exercice 5. (mesure) Montrer, avec AC, qu'il n'existe pas de mesure additive invariante sur $\mathbb{R}$ étendant la mesure de Lebesgue et partout définie sur $\mathfrak{P}(\mathbb{R})$.