Chapitre 1

Ensembles quotients

Dans ce court chapitre préliminaire, nous rappelons quelques notions générales de théorie des ensembles. La notion de *relation d'équivalence* sur un ensemble X est l'outil qui permet en pratique d'identifier les éléments de X partageant une certaine propriété (les rendant "équivalents"). Nous rappelons comment la donnée d'une telle relation sur X définit une partition naturelle de X en *classes d'équivalence*. L'ensemble de ces classes, un sous-ensemble de l'ensemble de toutes les parties de X, est appelé *ensemble quotient* de X par R, et il est noté X/R. Sa propriété principale (dite "universelle") est qu'une application $f : X \to Y$ constante sur les classes d'équivalence de R se factorise canoniquement en une application $\overline{f} : X/R \to Y$, appelé *passage au quotient* de f. La notion de *système de représentants* de R, qui consiste à choisir un élément par classe, conduit naturellement à discuter l'*axiome du choix*, un énoncé bien intuitif mais qui a joué un rôle historiquement important dans les fondements des mathématiques (il ne jouera que peu de rôle dans ce cours). Le chapitre culmine avec la discussion de plusieurs énoncés surprenants entrainés par (en fait équivalents à) l'axiome du choix, comme le *lemme de Zorn* ou le *théorème de Zermelo*. Dans la démonstration du Lemme de Zorn, donnée en complément, nous nous approcherons sans la définir, de la notion d'ordinal, et nous renvoyons au cours de logique pour des développements sur ce sujet. Les exercices contiennent notamment un fascicule de résultats classiques de théorie des ensembles (dénombrabilité, cardinalité) utiles à tout mathématicien.

Quelques références : L'appendice 2 de *Algebra*, 3ème ed. (S. Lang), le premier chapitre de *Algèbres et théories galoisiennes* (R. & A. Douady).

1. Partitions et relations d'équivalence

Définition 1.1. *Soit X un ensemble. Une partition de X est la donnée d'un ensemble $\{X_i\}_{i\in I}$ de parties non vides X_i de X tel que X est la réunion disjointe des X_i, i.e. $X = \cup_{i\in I} X_i$ et $X_i \cap X_j = \emptyset$ pour $i \neq j$. On note alors simplement*

$$X = \coprod_{i\in I} X_i.$$

Par exemple, l'ensemble $X = \{1, 2, 3\}$ possède exactement 5 partitions, à savoir $\{\{1, 2, 3\}\}$, $\{\{1\}, \{2, 3\}\}$, $\{\{2\}, \{1, 3\}\}$, $\{\{3\}, \{1, 2\}\}$ et $\{\{1\}, \{2\}, \{3\}\}$.

Exemple 1.2. (*Partition en fibres*) *Soit $f : X \to Y$ une application et $y \in Y$. On appelle fibre de f en y l'ensemble*[1]

$$f^{-1}(\{y\}) = \{x \in X \mid f(x) = y\}$$

1. Le $(\{\})$ est lourd, et on la note donc parfois $f^{-1}(y)$, même si cette notation désigne aussi l'unique antécédant de y par f quand f est bijective.

des antécédants de y par f. Lorsque f est surjective, ses fibres $X_y := f^{-1}(\{y\})$, avec y parcourant X, sont non vides et forment une partition de X indexée par Y. Toutes les partitions de X s'obtiennent en fait ainsi : si l'on a $X = \coprod_{i\in I} X_i$ avec $X_i \neq \emptyset$ pour tout i, alors l'application $f : X \to I$, associant à $x \in X$ l'unique $i \in I$ vérifiant $x \in X_i$, est surjective et vérifie $f^{-1}(\{i\}) = X_i$.

Une *relation* sur un ensemble X est la donnée d'une partie R de $X \times X$. On note suggestivement « $x\,R\,y$ » pour « $(x, y) \in R$ ».

Définition 1.3. *Une relation R sur un ensemble X est une* relation d'équivalence *si elle vérifie :*

(i) $x\,R\,x$, pour tout x dans X (réflexivité),

(ii) $x\,R\,y \Rightarrow y\,R\,x$, pour tout $x, y \in X$ (symétrie), et

(iii) $x\,R\,y$ et $y\,R\,z \Rightarrow x\,R\,z$, pour tout x, y, z dans X (transitivité).

Des notations standards pour les relations d'équivalence sont $\sim$, $\simeq$ ou $\equiv$. Soient R une relation d'équivalence sur X et $x \in X$. La *classe de R-équivalence* de x est par définition $[x]_R = \{y \in X \mid y\,R\,x\}$. On la note aussi $x \bmod R$. Lorsque R est sous-entendue, on parle simplement de classe d'équivalence de x et on la note en général $[x]$ ou $\overline{x}$ au lieu de $[x]_R$. On a $x \in [x]_R$ (réflexivité) ainsi que

$$\forall\, x, y \in X,\ y \in [x]_R \implies [y]_R = [x]_R. \tag{1}$$

En effet, $z\,R\,y$ implique $z\,R\,x$ (transitivité), puis $[y]_R \subset [x]_R$. On conclut par symmétrie car $y \in [x]_R$ implique $y \in [x]_R$.

Proposition 1.4. *Si R est une relation d'équivalence sur X, ses classes d'équivalence forment une partition de X.*

Démonstration — On a $x \in [x]_R$ par (i), donc X est réunion des classes d'équivalence, et ces dernières sont non vides. On conclut car d'après (1), deux classes $[x]_R$ et $[y]_R$, avec $x, y \in X$, sont égales ou disjointes. □

Remarque 1.5. *Réciproquement, toute partition $X = \coprod_{i\in I} X_i$ est la partition en classes d'équivalence d'une unique relation d'équivalence sur X. En effet, on constate que $x\,R\,y \Leftrightarrow \exists i \in I,\ x \in X_i$ et $y \in X_i$ est une relation d'équivalence sur X (et même manifestement l'unique) dont les classes sont les X_i.*

On vérifie aisément qu'une intersection de relations d'équivalence sur X est d'équivalence. En revanche, c'est faux en général pour une réunion (pourquoi ?). Si X est un ensemble, on note $\mathrm{P}(X)$ l'ensemble des parties de X.

Définition 1.6. *Si R est une relation d'équivalence sur X, le sous-ensemble de $\mathrm{P}(X)$ constitué des classes de R-équivalence est appelé* ensemble quotient *de X par R, et il est noté X/R. L'application $\pi_R : X \to X/R$, $x \mapsto [x]_R$, est appelée* projection canonique *associée à R. C'est une surjection dont les fibres sont par définition les classes d'équivalence de R.*

EXEMPLE 1.7. (*L'ensemble* $\mathbb{Z}/N\mathbb{Z}$) Soit $N \in \mathbb{Z}$. On définit une relation d'équivalence sur $\mathbb{Z}$ en posant $a\,R\,b \Leftrightarrow N|a-b$ (le vérifier). Depuis Gauss, on note en général $a \equiv b \bmod N$ pour $a\,R\,b$. L'ensemble $\mathbb{Z}/R$ est le familier $\mathbb{Z}/N\mathbb{Z}$ de l'arithmétique modulaire. La classe de $a \in \mathbb{Z}$ est le sous-ensemble $a + N\mathbb{Z} = \{a + Nm \mid m \in \mathbb{Z}\}$ de $\mathbb{Z}$, aussi noté $a \bmod N$ ou simplement $\overline{a}$.

EXEMPLE 1.8. *(Décomposition en "translations" et "cycles" d'une bijection)* Soient X un ensemble et $f : X \to X$ une bijection. On pose [2]

$$x\,R\,y \;\Leftrightarrow\; \exists\, i \in \mathbb{Z},\quad y = f^i(x).$$

C'est une relation d'équivalence sur X : on a $x = f^0(x)$, $y = f^i(x) \Leftrightarrow x = f^{-i}(y)$ et $f^i(f^j(x)) = f^{i+j}(x)$. On a en outre ici $x\,R\,f(x)$, donc f préserve les classes. Quelle est la classe de $x \in X$? Il y a deux cas, soit tous les $f^i(x)$, $i \in \mathbb{Z}$, sont distincts, auquel cas $\mathbb{Z} \to [x]$, $i \mapsto f^i(x)$ est bijective, et identifie f à la translation $i \mapsto i+1$. Soit il existe $i < j$, avec disons $d := j - i$ minimal, tel que $f^j(x) = f^i(x)$. Appliquant f^{-i} on a alors $f^d(x) = x$, puis que $a \equiv b \bmod d$ implique $f^a(x) = f^b(x)$. On en déduit

$$|[x]| = d \text{ et } [x] = \{x, f(x), f^2(x), \ldots, f^{d-1}(x)\},$$

par minimalité de d. Dans ce cas, f préserve, et permute circulairement, les d éléments de $[x]$.

L'Exemple 1.8 est en fait un cas particulier d'action de groupes, que nous introduirons plus tard (action du groupe $\mathbb{Z}$ sur X). L'Exemple 1.7 est aussi un cas particulier de l'Exemple 1.8 : considérer $X = \mathbb{Z}$ et $f : \mathbb{Z} \to \mathbb{Z}, x \mapsto x + N$. Une conséquence de l'Exemple 1.8 est la suivante :

COROLLAIRE 1.9. *Soient X un ensemble fini et $f : X \to X$ une application telle que $f^p = \mathrm{id}_X$ avec p premier. Soit $\mathrm{Fix}\,X = \{x \in X \mid f(x) = x\}$ l'ensemble des points fixes de f. Alors on a $|X| \equiv |\mathrm{Fix}\,X| \bmod p$.*

DÉMONSTRATION — On est dans la situation de l'Exemple 1.8, car f est bijective d'inverse f^{p-1}. Soient $x \in X$ et $d = |[x]|$. Montrons $d = 1$ (*i.e.* $x \in \mathrm{Fix}\,X$) ou $d = p$. On a $f^d(x) = x$ et $f^p(x) = x$ (donc $d \leq p$ par minimalité de d). On en déduit $f^i(x) = x$ pour tout i dans $d\mathbb{Z} + p\mathbb{Z}$. Si on a $d < p$ alors $1 \in d\mathbb{Z} + p\mathbb{Z}$ par Bezout (car p est premier), et donc $f(x) = x$, *i.e.* $d = 1$. On conclut en écrivant $|X| = \sum_{C \in X/R} |C|$ (partition en classes, Proposition 1.4). □

Ce corollaire peut être utilisé pour prouver l'existence de points fixes. En voici une application particulièrement amusante due à Zagier, dans le cas $p = 2$ (une application $f : X \to X$ telle que $f^2 = \mathrm{id}_X$ s'appelle une *involution*.) Il s'agit d'une démonstration « en une seule phrase » du fait que tout nombre premier $q \equiv 1 \bmod 4$ est somme de deux carrés d'entiers (un résultat fondateur de la théorie des nombres dû à Fermat, dont redonnerons une démonstration plus conceptuelle due à Dedekind un peu plus loin dans le cours).

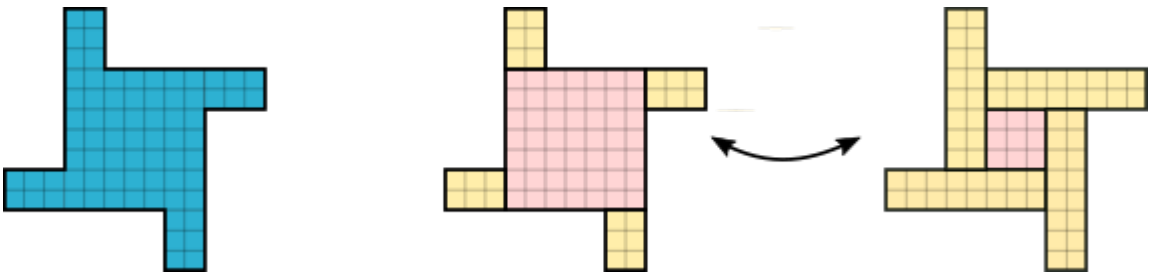
2. On rappelle que f^i désigne la composée $f \circ f \circ \cdots \circ f$ (i fois) pour $i > 0$, $(f^{-1})^{-i}$ pour $i < 0$, et la convention $f^0 = 1_X$. On a alors $f^{i+j} = f^i \circ f^j$ pour tout i, j dans $\mathbb{Z}$.

EXEMPLE 1.10. (Zagier, Amer. Math. Monthly 97 (1990), no. 2, 144) « *The involution on the finite set* $S = \{(x, y, z) \in \mathbb{N}^3 \mid x^2 + 4yz = q\}$ *defined by*

$$(x,\, y,\, z) \mapsto \begin{cases} (x + 2z,\, z,\, y - x - z) & \text{if} \quad x < y - z, \\ (2y - x,\, y,\, x - y + z) & \text{if} \quad y - z < x < 2y, \\ (x - 2y,\, x - y + z,\, y) & \text{if} \quad x > 2y. \end{cases} \tag{2}$$

has only one fixed point, so $|S|$ *is odd and the involution defined by* $(x, y, z) \mapsto (x, z, y)$ *also has a fixed point.* »

Pour la petite histoire, le fait que l'application ci-dessus est une involution de S est laissée au lecteur par Zagier (et en effet, facile à vérifier), tout comme le fait que son unique point fixe est $(1, 1, \frac{q-1}{4})$ (c'est ici que l'on utilise que q est premier). Zagier n'explique pas d'où vient la formule donnée : on en trouvera sur la chaine Youtube de Mathologer https://www.youtube.com/watch?v=DjI1NICfjOk une explication géométrique limpide (les « moulins »), que l'on peut résumer en :



2. Passage au quotient

Examinons comment définir une application dont la source est un ensemble quotient. Observons que si l'on dispose d'une application $g : X/R \to Y$, alors l'application qui s'en déduit $f := g \circ \pi_R : X \to Y$ est manifestement constante sur chaque classe de R-équivalence. C'est la situation générale :

PROPOSITION 2.1. *(Propriété universelle du quotient) Soient* $f : X \to Y$ *une application et* R *une relation d'équivalence sur* X*. On suppose que* f *est constante sur chaque classe d'équivalence : pour tout* $x, y \in X$*,* $x\,R\,y \Rightarrow f(x) = f(y)$*. Alors existe une unique application* $g : X/R \to Y$ *telle que* $g([x]_R) = f(x)$ *pour tout* $x \in X$*, ou ce qui revient au même, vérifiant* $g \circ \pi_R = f$*.*

DÉMONSTRATION — L'unicité de g découle de la surjectivité de π_R. Pour son existence, on observe que si C est une classe de R-équivalence, il y a un sens à poser $g(C) = f(x)$ où x est un élément quelconque[3] de C, car C est non vide et f est constante sur C par hypothèses. Pour tout $x \in X$ on a alors $f(x) = g([x]_R)$, *i.e.* $f = g \circ \pi_R$. □

3. Dit comme ça, on a l'impression que l'on a utilisé l'axiome du choix (cf. plus bas). On peut faire sans : par hypothèse sur f, pour C une classe d'équivalence alors l'ensemble $f(C)$ est un singleton, et on définit alors $g(C)$ comme étant son unique élément.

L'application g de l'énoncé est appelée « passage au quotient de f », on la note souvent $\overline{f}$. Essentiellement, nous avons simplement vérifié *une fois pour toutes* que l'application $X/R \to Y, [x]_R \mapsto f(x)$, est bien définie... ce qui est assez trivial ! L'identité $f = \overline{f} \circ \pi_R$ est appelée « factorisation canonique de f ». À moins de bien savoir ce que l'on fait, on procèdera toujours de cette manière pour définir une application dont la source est un quotient. On retiendra le slogan : « *c'est la même chose de se donner une application $X/R \to Y$ et se donner une application $X \to Y$ constante sur les classes de R-équivalence* » (les bijections réciproques étant $g \mapsto g \circ \pi_R$ et $f \mapsto \overline{f}$).

EXEMPLE 2.2. (*Contraction d'une partie*) Soit A une partie de X. La relation $x\,R\,y \Leftrightarrow x = y$ ou ($x \in A$ et $y \in A$), est une relation déquivalence sur X dont les classes d'équivalence sont A et les $\{x\}$ avec $x \in X \smallsetminus A$. Ainsi, la projection $\pi_R : X \to X/R$ est la « *contraction de A en un point* ». Considérons par exemple le cas $X = [0,1]$ et $A = \{0,1\}$. Alors X/R s'identifie naturellement au cercle unité S^1. En effet, l'application $f : x \mapsto e^{2i\pi x}, [0,1] \to \mathrm{S}^1$, est constante sur les classes de R car on a $f(0) = f(1)$. Elle se factorise donc en une application $\overline{f} : [0,1]/R \to \mathrm{S}^1, \overline{x} \mapsto e^{2i\pi x}$, qui est manifestement bijective. La notion de topologie quotient (voir le cours d'analyse) permettrait même de voir $\overline{f}$ comme un homéomorphisme.

3. Sections et systèmes de représentants

DÉFINITION 3.1. *Soit $f : X \to Y$ une application. Une section de f (ou "inverse à droite") est une application $s : Y \to X$ vérifiant $f \circ s = \mathrm{id}_Y$.*

Autrement dit, une application $s : Y \to X$ est une section de f si et seulement si pour tout $y \in Y$, $s(y)$ est un élément de la fibre $f^{-1}(\{y\})$. Une section est toujours injective, et uniquement déterminée par son image $s(Y)$, qui est une partie de X rencontrant chaque fibre de f en un et un seul point (*transversale de f*).

Si f possède une section, alors f est surjective. Réciproquement, intuitivement, toute surjection $f : X \to Y$ admet une section : il suffit de *choisir*, pour chaque $y \in Y$, un élément arbitraire de la fibre $f^{-1}(\{y\})$, et de l'appeler $s(y)$. L'axiome autorisant cette construction en théorie des ensemble s'appelle l'*axiome du choix*.

(AC) *Pour tout ensemble X, il existe une application $\tau : \mathrm{P}(X) - \{\emptyset\} \to X$ telle que $\tau(E) \in E$ pour toute partie non vide E de X.*

Une telle fonction τ est appelée *fonction de choix* sur X.

PROPOSITION 3.2. *Les énoncés suivants sont équivalents à l'axiome du choix :*

(i) Toute surjection admet une section.

(ii) Pour toute famille $\{X_i\}_{i \in I}$ de sous-ensembles non vides X_i d'un ensemble X, le produit $\prod_{i \in I} X_i$ est non vide.

DÉMONSTRATION — Si $f : X \to Y$ est une surjection, et si τ est une fonction de choix sur X, alors $s(y) := \tau(f^{-1}(\{y\}))$ est une section de f (noter ($f^{-1}(\{y\}) \neq \emptyset$ car f est surjective). Donc AC $\Rightarrow$ (i).

Soient $\{X_i\}_{i\in I}$ des ensembles non vides comme au (ii). Posons $X = \coprod_{i\in I} X_i$ leur réunion disjointe externe [4] On dispose d'une application $f : X \to I$ vérifiant $f^{-1}(i) = X_i$ pour $i \in I$ (déjà vue dans l'Exemple 1.2). Pour toute section s de f, on a $(s(i))_{i\in I} \in \prod_{i\in I} X_i$: cela montre (i) $\Rightarrow$ (ii).

Enfin, appliquant (ii) à l'ensemble de toutes les parties non vides de X, on en déduit que $\prod_{E\in \mathrm{P}(X)-\{\emptyset\}} E$ est non vide : un élément quelconque $\tau := (\tau(E))_E$ est une fonction choix sur X. □

Bien que très intuitif, AC a aussi des conséquences surprenantes, comme le *paradoxe de Banach-Tarski*, ou plus simplement les théorèmes de Zorn et de Zermelo (voir §4). Pour de nombreux exemples concrets d'ensembles X, on peut construire une fonction de choix sur X sans appel à AC : « *Pour choisir une chaussette plutôt que l'autre pour chaque paire d'une collection infinie, on a besoin de l'axiome du choix. Mais pour les chaussures, ce n'est pas la peine* » (Russel).

Remarque 3.3. (Culturelle) On sait depuis Gödel et Cohen que si l'on rajoute l'axiome du choix, ou son contraire, à la théorie axiomatique des ensembles de Zermelo et Fraenkel (ZF), et si l'on suppose cette dernière cohérente, alors on obtient une théorie cohérente. Nous renvoyons au cours de logique pour comprendre le sens de cette affirmation !

Définition 3.4. *Soient X un ensemble et R une relation d'équivalence sur X. Un* représentant *d'une classe d'équivalence est la donnée d'un élément de cette classe. Un* système de représentants *de (X, R) est la donnée d'un sous-ensemble de X contenant un et un seul représentant de chaque classe d'équivalence ; autrement dit, c'est l'image d'une section de π_R.*

Le fait que toute relation d'équivalence possède un système de représentants est donc une autre formulation de AC. Dans le cas de $\mathbb{Z}/n\mathbb{Z}$ avec $n \geq 1$, l'unicité de la division euclidienne montre bien sûr que $\{0, \ldots, n-1\}$ est un système de représentants (et donc $|\mathbb{Z}/n\mathbb{Z}| = n$), mais il y en a bien d'autres ! (une infinité dénombrable).

Exemple 3.5. (*Infinite prisonners wearing hats, without hearing*) Nous renvoyons à https://risingentropy.com/axiom-of-choice-and-hats/ pour une explication d'un puzzle surprenant montrant comment l'axiome du choix peut ... sauver des vies ! Il est basé sur le choix d'un système de représentants de la relation d'équivalence sur $\{0,1\}^{\mathbb{N}}$ définie par $(x_n) \sim (y_n) \Leftrightarrow \exists N \geq 1,\ x_n = y_n$ pour $n \geq N$.

4. Le lemme de Zorn

Nous profitons de cette petite discussion sur l'axiome du choix pour discuter du Lemme de Zorn. C'est un énoncé moins intuitif qui se déduit de AC et qui a de nombreuses applications : existence d'une base dans un espace vectoriel général, existence des clôtures algébriques, théorème de prolongement de Hahn-Banach, théorème de Tychonoff, existence d'un idéal maximal dans un anneau commutatif non nul... Il nous faut d'abord faire quelques rappels sur les relations d'ordre.

4. Formellement, on identifie naturellement pour tout i l'ensemble X_i au sous-ensemble $Y_i = X_i \times \{i\}$ de $E \times I$, on observe que les Y_i y sont deux à deux disjoints, et on pose $X = \bigcup_{i\in I} Y_i$.

DÉFINITION 4.1. *Une relation d'ordre sur un ensemble X est une relation R sur X supposée réflexive, transitive et vérifiant en outre $x\,R\,y$* et *$y\,R\,x \Rightarrow x = y$, pour tout $x, y \in X$ (antisymétrie).*

Une relation d'ordre est en général notée $\leq$. On note alors aussi $x \geq y$ pour $y \leq x$, $x < y$ pour "$x \leq y$ et $x \neq y$", et $x > y$ pour $y < x$. L'ordre $\leq$ est dit *total* si deux éléments quelconques de X sont comparables : pour tout $x, y \in X$ on a $x \leq y$ ou $y \leq x$. Un ensemble muni d'une relation d'ordre est appelé *ensemble ordonné*.

Soit $(X, \leq)$ un ensemble ordonné. Toute partie Y de X est naturellement ordonnée par *l'ordre induit* $\leq \cap\,(Y \times Y)$, encore noté $\leq$ en général. De plus, on appelle *majorant* (resp. *majorant strict*) de Y tout élément $x \in X$ vérifiant $y \leq x$ (resp. $y < x$) pour tout $y \in Y$.

DÉFINITION 4.2. *Soient $(X, \leq)$ un ensemble ordonné et $x \in X$. On dit que x est un élément maximal si le seul élément $y \in X$ avec $x \leq y$ est $y = x$. On dit que x est un plus grand élément si c'est un majorant de X,* i.e. *si on a $y \leq x$ pour tout $y \in X$. Un plus grand élément est nécessairement maximal, et unique s'il existe, auquel cas on le note* $\max X$.

Par symétrie, on dispose aussi de notions de minorants, d'élément minimal (un $x \in X$ tel que $y \leq x \Rightarrow y = x$) et de plus petit élément (un $x \in X$ tel que pour tout $y \in X$ on a $x \leq y$) et on note $\min X$ l'unique plus petit élément de X s'il existe.

REMARQUE 4.3. Il faut bien noter que quand $\leq$ n'est pas total, un élément maximal n'est pas nécessairement un plus grand élément, et n'est pas nécessairement unique. Par exemple, soient n un entier ≥ 2 et X l'ensemble des parties à $< n$ éléments de $\{1, \ldots, n\}$ muni de l'ordre $X \leq Y \Leftrightarrow X \subset Y$. Alors $(X, \leq)$ n'a pas de plus grand élément, ses éléments maximaux sont les parties à $n-1$ éléments de $\{1, \ldots, n\}$, et on a $\min X = \emptyset$.

Nous pouvons désormais énoncer le *lemme de Zorn*. Un ensemble ordonné est dit *inductif* si tout sous-ensemble totalement ordonné admet un majorant. Par exemple, un ensemble totalement ordonné est inductif si, et seulement si, il admet un plus grand élément.

THÉORÈME 4.4. (Zorn) *Tout ensemble ordonné inductif possède au moins un élément maximal.*

(Un ensemble ordonné inductif est non vide, comme on le voit en considérant un majorant de sa partie vide...) Une application typique de cet énoncé est la suivante.

COROLLAIRE 4.5. *Tout espace vectoriel possède une base.*

On rappelle qu'une partie X d'un k-espace vectoriel V est dite *libre*, si pour toute famille *finie* non vide $\{v_j\}_{j \in J}$ d'éléments distincts de X, la relation $\sum_{i \in J} \lambda_j v_j = 0$ avec les $\lambda_j \in k$ entraîne $\lambda_j = 0$ pour tout $j \in J$. De plus, X est dite *génératrice* si tout élément de V est combinaison linéaire *finie*[5] d'éléments de X . Enfin, X est une *base* si elle est à la fois libre et génératrice.

5. Bien noter que les combinaisons linéaires infinies n'ont pas de sens en général.

Démonstration — Soient k un corps et V un k-espace vectoriel. Soit $\mathscr{L} \subset \mathrm{P}(V)$ le sous-ensemble des parties libres de V (noter $\emptyset \subset \mathscr{L}$!). On le munit de la relation d'inclusion $\subset$ induite de $\mathrm{P}(V)$. Montrons que $(\mathscr{L}, \subset)$ est inductif. Soit $\{L_i\}_{i\in I} \subset \mathscr{L}$ une famille totalement ordonnée de parties libres de V. Alors leur réunion $L' = \cup_{i\in I} L_i$ est encore une partie libre, car toute famille finie d'éléments de L' appartient à un même L_i, et L' contient aussi chaque L_i : c'est un majorant de $\{L_i \mid i \in I\}$ dans $\mathscr{L}$. D'après Zorn, il existe un élément maximal de $\mathscr{L}$, notons-le B.

La fin de l'argument est comme en dimension finie. Supposons B non génératrice : il existe $b \in V$ non dans $\mathrm{Vect}_k B$ (en particulier, $b \notin B$). Alors $B \coprod \{b\}$ contient strictement B et elle est libre : si on a $\lambda b + \sum_{j\in J} \lambda_j b_j = 0$, avec les $\{b_j\}_{j\in J}$ distincts dans B, et λ et les λ_j dans k, on a $\lambda \neq 0$ car B est libre, donc λ inversible car k est un corps, puis $b = -\lambda^{-1} \sum_{j\in J} \lambda_j b_j$. On a contredit la maximalité de B. □

Remarque 4.6. Une modification simple de la démonstration montre que de toute famille génératrice on peut extraire une base. De plus, comme en dimension finie, on peut montrer que si $\{e_i\}_{i\in I}$ et $\{f_j\}_{j\in J}$ sont des bases d'un même espace vectoriel, alors I est en bijection avec J (voir l'Exercice1.17).