

Choix, Zorn et Zermelo

Marc Lorenzi

3 juillet 2022

Résumé

Une activité fréquente du mathématicien est, lorsqu'il a en sa possession des ensembles non vides, d'avoir à choisir un élément dans chacun des ensembles. Cette activité apparemment anodine doit pourtant être justifiée. Nous allons voir dans cet article que si le pouvoir de choisir implique des théorèmes d'une grande puissance. Ceci justifie la nécessité d'introduire un axiome disant au mathématicien que de tels choix sont possibles.

1 L'axiome du choix

1.1 Introduction

Si E est un ensemble d'ensembles, nous noterons $\bigcup E$ l'ensemble des éléments des éléments de E . Une notation plus courante est

$$\bigcup E = \bigcup_{A \in E} A$$

Ainsi, $x \in \bigcup E$ si et seulement si il existe $A \in E$ tel que $x \in A$.

Définition 1. Soit E un ensemble d'ensembles non vides. Une *fonction de choix* pour E est une fonction $f : E \longrightarrow \bigcup E$ telle que pour tout $A \in E$, $f(A) \in A$.

Le terme *fonction de choix* vient de ce qu'une telle fonction permet de choisir, pour chaque ensemble $A \in E$, un élément de A , à savoir $f(A)$. Pour certains ensembles E , l'existence d'une fonction de choix peut être montrée de façon élémentaire. Passons quelques exemples en revue.

Proposition 1. Soit E un ensemble de parties non vides de $\mathbb{N}$. Alors E possède une fonction de choix.

Démonstration. Toute partie non vide de $\mathbb{N}$ possède un plus petit élément. On peut donc poser pour tout $A \in E$, $f(A) = \min A$. La fonction f est une fonction de choix pour E . $\square$

Proposition 2. Soient X et Y deux ensembles en bijection. Si tout ensemble de parties non vides de X possède une fonction de choix, alors tout ensemble de parties non vides de Y possède une fonction de choix.

Démonstration. Soit $\varphi : X \longrightarrow Y$ une bijection. Soit F un ensemble de parties non vides de Y . L'ensemble

$$E = \{\varphi^{-1}(B) : B \in F\}$$

est un ensemble de parties non vides de X , il possède donc une fonction de choix f . L'application $g : F \longrightarrow \bigcup F$ définie par

$$g(B) = \varphi(f(\varphi^{-1}(B)))$$

est une fonction de choix sur E . $\square$

Voici un dernier exemple.

Proposition 3. *Soit E un ensemble fini d'ensembles non vides. Alors, E possède une fonction de choix.*

Démonstration. Montrons ce résultat par récurrence sur le cardinal de E . C'est clair si $E = \emptyset$. Soit $n \in \mathbb{N}$. Supposons que tout ensemble fini de cardinal n dont les éléments sont des ensembles non vides possède une fonction de choix. Soit E un ensemble d'ensembles non vides. Supposons que E est fini de cardinal $n + 1$. Écrivons

$$E = E' \cup \{A_0\}$$

où A_0 est un ensemble non vide et $E' = E \setminus \{A_0\}$ est fini, de cardinal n . Soit $a_0 \in A_0$. Par l'hypothèse de récurrence, E' possède une fonction de choix f' . La fonction $f : E \longrightarrow \bigcup E$ définie par $f(A) = f'(A)$ si $A \in E'$ et $f(A_0) = a_0$ est alors une fonction de choix sur E . $\square$

Pour d'autres ensembles E , l'existence d'une fonction de choix pour E est plus problématique. Par exemple, si E est un ensemble de parties non vides de $\mathbb{R}$, il n'est pas évident de *choisir* pour chaque ensemble $A \in E$ un élément de A particulier. Prenons pour exemple un théorème classique d'Analyse dans lequel un tel problème de choix survient.

Proposition 4. *Soit A une partie de $\mathbb{R}$ non vide et majorée. Il existe alors une suite d'éléments de A qui converge vers $\sup A$.*

Démonstration. Soit $M = \sup A$. Pour tout $n \geq 1$, le réel $M - \frac{1}{n}$ ne majore pas A . Il existe donc $x_n \in A$ tel que

$$M - \frac{1}{n} < x_n \leq M$$

Par le théorème d'encadrement des limites, x_n tend vers M lorsque n tend vers l'infini. $\square$

Nous avons dans cette démonstration considéré l'ensemble

$$E = \left\{ \left[M - \frac{1}{n}, M \right] \cap A : n \in \mathbb{N}^* \right\}$$

dont les éléments sont des ensembles non vides, et nous avons « choisi » dans chacun des ensembles de E un élément x_n .

L'axiome du choix est précisément là pour nous dire que ce genre de raisonnement est pertinent.

Remarque. Les puristes objecteront à notre exemple que nous avons seulement utilisé ce que l'on appelle l'axiome du choix dénombrable. Nous n'entrerons pas dans ces détails dans cet article, qui n'est qu'une introduction.

1.2 L'énoncé

Voici l'axiome du choix, ou du moins une version simple et très courante de celui-ci.

AXIOME DU CHOIX

(AC) *Tout ensemble d'ensembles non vides possède une fonction de choix.*

Nous allons maintenant énoncer quelques propriétés qui sont « élémentairement » équivalentes à l'axiome du choix. Entendons par là que la démonstration de l'équivalence est courte et sans malice (ou presque).

1.3 Quelques propriétés équivalentes à AC

Proposition 5. *AC est équivalent à*

(AC₁) Tout ensemble d'ensembles non vides et disjoints possède une fonction de choix.

Démonstration. Évidemment, AC implique AC₁.

Supposons AC₁. Soit E un ensemble d'ensembles non vides. Soit

$$F = \{A \times \{A\} : A \in E\}$$

L'ensemble F est un ensemble d'ensembles non vides et disjoints. Par AC₁, il existe une fonction $f : F \rightarrow \bigcup F$ telle que pour tout $B \in F$, $f(B) \in B$. Dit autrement, pour tout $A \in E$,

$$f(A \times \{A\}) \in A \times \{A\}$$

Pour tout $A \in E$, on a $f(A \times \{A\}) = (x, A)$ où $x \in A$. Posons $g(A) = x$. La fonction g est alors une fonction de choix pour E . $\square$

Proposition 6. *AC est équivalent à*

(AC₂) Pour tout ensemble non vide E , l'ensemble $\mathcal{P}(E) \setminus \{\emptyset\}$ possède une fonction de choix.

Démonstration. AC implique clairement AC₂ puisque $\mathcal{P}(E) \setminus \{\emptyset\}$ est un ensemble d'ensembles non vides.

Supposons AC₂. Soit E un ensemble d'ensembles non vides. Soit $A = \bigcup E$. Par AC₂, l'ensemble $\mathcal{P}(A) \setminus \{\emptyset\}$ possède une fonction de choix f . La restriction de f à E est alors une fonction de choix pour E . $\square$

Proposition 7. *AC est équivalent à*

(AC₃) Toute surjection possède une inverse à droite.

Démonstration. Supposons AC. Soit $f : E \rightarrow F$ une surjection. Pour tout $y \in F$, l'ensemble $f^{-1}(\{y\})$ est non vide. Posons

$$\mathcal{E} = \{f^{-1}(\{y\}) : y \in F\}$$

Remarquons que

$$\bigcup_{y \in F} \mathcal{E} = \bigcup_{y \in F} f^{-1}(\{y\}) = f^{-1}\left(\bigcup_{y \in F} \{y\}\right) = f^{-1}(F) = E$$

Par AC, il existe une fonction $g : \mathcal{E} \rightarrow E$ telle que pour tout $A \in \mathcal{E}$, $g(A) \in A$. Soit $h : F \rightarrow E$ définie par

$$h(y) = g(f^{-1}(\{y\}))$$

On a pour tout $y \in E$,

$$(f \circ h)(y) = f(g(f^{-1}(\{y\})))$$

Or, $g(f^{-1}(\{y\})) \in f^{-1}(\{y\})$, donc

$$f(g(f^{-1}(\{y\}))) = y$$

Ainsi, $f \circ h = \text{id}_F$.

Montrons maintenant que AC_3 implique AC_1 . Supposons que toute surjection admet un inverse à droite. Soit E un ensemble d'ensembles non vides disjoints. Soit $f : \bigcup E \rightarrow E$ définie pour tout $x \in F$ par $f(x) =$ l'unique ensemble $A \in E$ tel que $x \in A$. L'application f est surjective, elle admet donc un inverse à droite $g : E \rightarrow \bigcup E$. Pour tout $A \in E$, on a

$$f(g(A)) = A$$

et donc $g(A) \in A$. Ainsi, g est une fonction de choix pour E . Nous avons donc AC_1 , et donc aussi AC . $\square$

Terminons par l'équivalence de AC avec un résultat qui semble « évident » ... et qui ne l'est donc pas.

Proposition 8. *AC est équivalent à*

(AC_4) Un produit cartésien d'ensembles non vides est non vide.

Démonstration. Supposons AC . Soit $(E_i)_{i \in I}$ une famille d'ensembles non vides. Soit $E = \{E_i : i \in I\}$. Le produit cartésien des E_i est

$$\prod_{i \in I} E_i = \{f : I \rightarrow \bigcup E : \forall j \in I, f(j) \in E_j\}$$

Par AC , l'ensemble E possède une fonction de choix φ . Soit $f : I \rightarrow E$ définie pour tout $j \in J$ par $f(j) = \varphi(E_j)$. Alors, f appartient au produit cartésien des E_i , qui est donc non vide.

Supposons maintenant AC_4 . Soit E un ensemble d'ensembles non vides. Pour tout $A \in E$ posons assez pompeusement $E_A = A$. Considérons la famille d'ensembles $(E_A)_{A \in E}$. Par AC_4 , le produit cartésien $\prod_{A \in E} E_A$ est non vide. Il existe donc une fonction $f : E \rightarrow \bigcup E$ telle que pour tout $A \in E$, $f(A) \in E_A = A$. La fonction f est donc une fonction de choix pour E . $\square$

Les preuves de l'équivalence de AC avec AC_1 , AC_2 , AC_3 et AC_4 n'étaient pas difficiles. Nous allons maintenant aborder une propriété qui elle aussi est équivalente à AC . La preuve de l'équivalence sera tout sauf facile.

2 Le lemme de Zorn

2.1 Introduction

Commençons par introduire un certain nombre de notions qui nous éviteront d'avoir à faire de longues périphrases.

Tout d'abord, définissons le mot « chaîne » qui est un terme consacré dans la littérature.

Définition 2. Soit $(E, \leq)$ un ensemble ordonné. Une *chaîne* de E est une partie de E totalement ordonnée.

Remarquons que $\emptyset$ et les singletons sont des chaînes. Si $(E, \leq)$ est un ensemble totalement ordonné, la notion de chaîne n'a aucun intérêt puisque dans ce cas toutes les parties de E , y compris E , sont des chaînes.

Définition 3. Soit $(E, \leq)$ un ensemble ordonné. La relation $\leq$ est un *bon ordre* si toute partie non vide de E possède un plus petit élément.

Remarquons qu'un ensemble bien ordonné est automatiquement totalement ordonné. En effet, si x et y sont deux éléments de E , l'ensemble $\{x, y\}$ possède un plus petit élément.

La notion de bon ordre est une notion capitale de la théorie des ensembles. Elle est en particulier à la source de la théorie des ordinaux transfinis. Nous ne l'utiliserons dans cet article que par nécessité.

Définition 4. Soit $(E, \leq)$ un ensemble bien ordonné. Pour tout $x \in E$, le *segment initial* de E défini par x est l'ensemble

$$\{y \in E : y < x\}$$

Terminons par ce que nous appellerons la *propriété de Zorn*.

Définition 5. Soit $(E, \leq)$ un ensemble ordonné. Nous dirons que E vérifie la *propriété de Zorn* lorsque toute chaîne de E est majorée.

Venons-en enfin au lemme de Zorn.

LEMME DE ZORN

Soit $(E, \leq)$ un ensemble ordonné non vide vérifiant la propriété de Zorn. Alors, E possède un élément maximal.

Cette propriété n'a d'intérêt que lorsque l'ordre $\leq$ est un ordre partiel. En effet, si l'ordre est total, E lui-même est une chaîne. Si E vérifie la propriété de Zorn, alors E est majoré. Si a est un majorant de E , comme on a aussi $a \in E$, a est le plus grand élément de E .

Montrons sur un exemple simple comment le lemme de Zorn est utilisé en pratique. Ce lemme permet de montrer l'existence d'un objet vérifiant certaines propriétés.

Proposition 9. On suppose le lemme de Zorn vérifié. Alors, tout espace vectoriel sur un corps $\mathbb{K}$ admet des bases.

Démonstration. Soit E un $\mathbb{K}$ -espace vectoriel.

1. On introduit un ensemble E judicieux et on le munit d'une relation $\leq$ tout aussi judicieuse. Et non, dans le cas qui nous intéresse, E n'est pas l'espace vectoriel E . D'où l'appellation « judicieux ».

Pour notre propos, point n'est besoin de chercher très loin. Soit $\mathcal{L}$ l'ensemble des ensembles libres de vecteurs de E . Munissons $\mathcal{L}$ de la relation d'inclusion $\subseteq$.

2. On montre que la relation $\subseteq$ est une relation d'ordre sur $\mathcal{L}$.

Dans le cas qui nous intéresse, c'est immédiat puisque la relation d'inclusion est une relation d'ordre.

3. On montre que $(\mathcal{L}, \subseteq)$ vérifie la propriété de Zorn.

N'oublions pas de vérifier que $\mathcal{L}$ est non vide. En effet, l'ensemble vide ne risque pas d'avoir un élément maximal, au même titre qu'il n'a pas d'élément qui va à la pêche le dimanche. Ici, c'est évident, car l'ensemble vide est libre et appartient donc à $\mathcal{L}$.

Soit $\mathcal{C}$ une chaîne de $\mathcal{L}$. Considérons l'ensemble

$$M = \bigcup_{L \in \mathcal{C}} L$$

Montrons que M est libre, c'est à dire que toute partie finie de M est libre. Soient $e_1, \dots, e_n \in M$. Pour tout $i \in \llbracket 1, n \rrbracket$, il existe $L_i \in \mathcal{C}$ tel que $e_i \in L_i$. La chaîne $\mathcal{C}$ est totalement ordonnée. Comme toute partie finie d'un ensemble totalement ordonné admet un plus grand élément, il existe $j \in \llbracket 1, n \rrbracket$ tel que pour tout $i \in \llbracket 1, n \rrbracket$, $L_i \subseteq L_j$. Les vecteurs $e_1, \dots, e_n$ appartiennent ainsi tous à l'ensemble L_j qui est un ensemble libre. Ces vecteurs sont donc libres.

Clairement, M majore $\mathcal{C}$. Nous avons donc montré que $\mathcal{L}$ vérifie la propriété de Zorn.

4. *On applique le lemme de Zorn et on explore les conséquences. Très souvent, l'objet dont l'existence est assurée par le lemme de Zorn est l'objet dont on cherche à prouver l'existence. Parfois, un peu de travail en plus est nécessaire.*

L'ensemble $\mathcal{L}$ vérifie la propriété de Zorn. Il possède donc un élément maximal, que nous appellerons B . B est donc un ensemble libre de vecteurs de E maximal au sens de l'inclusion. Montrons que B est une base de E . Comme B est libre, il suffit de prouver que cet ensemble est générateur.

Soit $x \in E$. Si x est l'un des vecteurs de B , x est clairement combinaison linéaire de B . Sinon, on a

$$B \subsetneq B \cup \{x\}$$

Par la maximalité de B , l'ensemble $B \cup \{x\}$ est lié. Facilement, comme B est libre, cela implique que x est combinaison linéaire des vecteurs de B .

□

Le lemme de Zorn est très puissant. *Trop* puissant, diront certains. Pour comprendre sa puissance, nous allons d'abord montrer que celui-ci implique l'axiome du choix. La démonstration n'est pas difficile, ce qui prouve bien la puissance du lemme de Zorn.

2.2 Zorn implique choix

Proposition 10. *Le lemme de Zorn implique AC.*

Démonstration. Supposons le lemme de Zorn vérifié.

Soit E un ensemble d'ensembles non vides. Reprenons les étapes que nous avons décrites dans notre exemple d'utilisation du lemme de Zorn.

1. Considérons

$$\mathcal{E} = \{(A, f) : A \subseteq E, f : A \longrightarrow \bigcup E, \forall X \in A, f(X) \in X\}$$

Définissons sur $\mathcal{E}$ une relation $\preceq$ en posant $(A, f) \preceq (B, g)$ si $A \subseteq B$ et pour tout $X \in A$, $g(X) = f(X)$.

2. On vérifie facilement que $\preceq$ est une relation d'ordre sur $\mathcal{E}$.
3. Le couple $(\emptyset, f)$, où f est l'application triviale, appartient à $\mathcal{E}$. Ainsi, $\mathcal{E}$ n'est pas vide.

Soit $\mathcal{C}$ une chaîne de $\mathcal{E}$. Soit

$$M = \bigcup_{(A, f) \in \mathcal{C}} A$$

Soit $X \in M$. Supposons qu'il existe (A, f) et (A', f') appartenant à $\mathcal{C}$ tels que $X \in A$ et $X \in A'$. Comme l'ensemble $\mathcal{C}$ est totalement ordonné, on a par exemple $(A, f) \preceq (A', f')$, c'est à dire $A \subseteq A'$ et aussi $f'(X) = f(X)$. On définit ainsi une application $g : M \longrightarrow \bigcup E$ en posant pour tout $(A, f) \in \mathcal{C}$ et tout $X \in A$, $g(X) = f(X)$.

Montrons que g appartient à $\mathcal{E}$. Soit $X \in M$. Soit $(A, f) \in \mathcal{C}$ tel que $X \in A$. On a alors $g(X) = f(X) \in X$ puisque $(A, f) \in \mathcal{E}$.

Montrons maintenant que g majore $\mathcal{C}$. Soit $(A, f) \in \mathcal{C}$. Par définition même de M , on a $A \subseteq M$. De plus, pour tout $X \in A$ on a, par définition de g , $g(X) = f(X)$.

Ainsi, $\mathcal{E}$ satisfait la propriété de Zorn.

4. L'ensemble $\mathcal{E}$ possède donc un élément maximal que nous noterons (U, m) .

Supposons que $U \neq E$. Soit $B \in E \setminus U$. L'ensemble B est non vide, il contient donc un élément X_0 . Définissons

$$h : V = U \cup \{B\} \longrightarrow \bigcup E$$

en posant $h(X) = m(X)$ si $X \in U$ et $h(B) = X_0$. On a alors $h \in \mathcal{E}$, $(U, m) \preceq (V, h)$ et $(U, m) \neq (V, h)$, ce qui contredit la maximalité de (U, m) .

Ainsi, $U = E$ et donc $m : E \longrightarrow \bigcup E$ est une fonction de choix pour E . $\square$

2.3 Choix implique Zorn

L'axiome du choix se démontre à partir du Lemme de Zorn. Le lemme de Zorn est donc, d'une certaine manière, plus « puissant » que l'axiome du choix. Nous devrions plutôt dire plus « évolué ». Car nous allons maintenant prouver la réciproque. Le lecteur doit être prévenu : la preuve est *très* compliquée. Il existe des preuves plus simples que celle que nous allons donner, mais elles utilisent des concepts sophistiqués, comme par exemple des démonstrations par induction sur les ordinaux transfinis. Notre preuve sera quant à elle « élémentaire ». Élémentaire, mais compliquée.

Proposition 11. *L'axiome du choix implique le lemme de Zorn.*

Que le lecteur soit prévenu, nous voici partis pour une démonstration de 4 pages.

Supposons l'axiome du choix vérifié.

Soit $(E, \leq)$ un ensemble ordonné dont toute chaîne possède un majorant. Nous allons raisonner par l'absurde en supposant que E n'a pas d'élément maximal.

On commence fort. Une longue démonstration par l'absurde se profile à l'horizon.

Appliquons tout d'abord AC_2 . Il existe une fonction

$$\varphi : \mathcal{P}(E) \setminus \{\emptyset\} \longrightarrow E$$

telle que pour toute partie non vide A de E , on ait $\varphi(A) \in A$.

Soit C une chaîne de E . Par la propriété de Zorn, C est majorée. En d'autres termes, l'ensemble

$$U'_C = \{y \in E : \forall x \in C, x \leq y\}$$

est une partie non vide de E . Soit $y \in U'_C$. Comme E ne possède pas d'élément maximal, il existe $z \in E$ tel que $y < z$. On a alors aussi $z \in U'_C$, et de plus $z \notin C$ puisque pour tout $x \in C$, $x \leq y < z$. Ainsi, l'ensemble

$$U_C = \{y \in E : \forall x \in C, x < y\}$$

est une partie non vide de E . Comme $\varphi(U_C) \in U_C$, nous disposons d'une fonction u qui à toute chaîne C de E associe $u(C) = \varphi(U_C)$, un majorant strict de C dans E . Retenons donc que pour toute chaîne C ,

- $u(C)$ majore C .
- $u(C) \notin C$.

Nous voici donc en possession de la baguette magique u qui, pour chaque chaîne C , permet de créer un majorant strict de C .

Considérons l'ensemble Γ des parties C de E vérifiant les propriétés suivantes :

- C est bien ordonnée.
- Pour tout $x \in C$, $x = u(\{y \in C : y < x\})$.

Comme tout ensemble bien ordonné est totalement ordonné, les éléments de Γ sont donc des chaînes.

Rappelons qu'une chaîne est une partie totalement ordonnée. Toute partie d'une chaîne est donc aussi totalement ordonnée, et est donc aussi une chaîne. La seconde propriété nous dit que si $x \in C$, le segment initial de C déterminé par x (qui est une chaîne) a précisément x pour image par u .

Dit autrement, l'application u détermine de façon unique un élément x de C si on connaît les prédécesseurs de x .

Clairement, $\emptyset \in \Gamma$ et donc Γ n'est pas vide. Si C est une chaîne non vide de Γ , elle possède un plus petit élément x , puisqu'elle est bien ordonnée. Que vaut x ? Il nous suffit de remarquer que $\{y \in C : y < x\} = \emptyset$ et donc

$$x = u(\{y \in C : y < x\}) = u(\emptyset)$$

Ainsi, $u(\emptyset)$ est le plus petit élément de toutes les chaînes non vides de Γ .

Si C est une chaîne de Γ ayant au moins deux éléments, quel est son deuxième plus petit élément ? Eh bien, si nous notons x l'élément en question, alors

$$x = u(\{y \in C : y < x\}) = u(\{u(\emptyset)\})$$

De même, si C est une chaîne de Γ ayant au moins trois éléments, le troisième plus petit élément de C est

$$u(\{\emptyset, \{u(\emptyset)\}\})$$

Il semblerait que les éléments d'une chaîne ne « dépendent pas » de la chaîne. C'est effectivement ce que nous aurons bientôt à montrer, et la preuve de ce résultat nous donnera beaucoup de mal.

Lemme 12. Pour toute chaîne $C \in \Gamma$, $C \cup \{u(C)\} \in \Gamma$.

Démonstration. Soit $C \in \Gamma$. Soit $C' = C \cup \{u(C)\}$.

Soit A une partie non vide de C' . Si $A = \{u(C)\}$, alors A a un plus petit élément, qui est $u(C)$. Sinon, soit $B = A \cap C$. L'ensemble B est une partie non vide de C , B a donc un plus petit élément x . Comme $u(C)$ majore strictement C , on a $x < u(C)$, et donc x est aussi le plus petit élément de A .

Soit $x \in C'$. Si $x \neq u(C)$, alors

$$\{y \in C' : y < x\} = \{y \in C : y < x\}$$

et donc, comme $C \in \Gamma$,

$$u(\{y \in C' : y < x\}) = \{y \in C : y < x\} = x$$

Si, au contraire, $x = u(C)$, alors

$$\{y \in C' : y < x\} = C$$

et donc

$$u(\{y \in C' : y < x\}) = u(C) = x$$

Ainsi, $C' \in \Gamma$. $\square$

Passons à la partie ardue de notre preuve.

Lemme 13. Soient C_1 et C_2 deux chaînes distinctes éléments de Γ . Alors, l'une de ces deux chaînes est un segment initial de l'autre.

Démonstration. Si C_1 ou C_2 est vide, c'est évident. Supposons donc ces deux chaînes non vides.

Supposons que $C_1 \not\subseteq C_2$. L'ensemble $C_1 \setminus C_2$ est alors une partie non vide de C_1 qui est un ensemble bien ordonné. Cet ensemble a donc un plus petit élément. Posons

$$a = \min(C_1 \setminus C_2)$$

Par minimalité de a , tout $x \in C_2$ tel que $x < a$ est aussi dans C_1 . Ainsi,

$$\{x \in C_2 : x < a\} \subseteq \{x \in C_1 : x < a\}$$

En fait, nous avons égalité. En effet, soit $x \in C_1$ tel que $x < a$. Par minimalité de a , $x \notin C_1 \setminus C_2$, c'est à dire $x \notin C_1$ ou $x \in C_2$. Mais $x \in C_1$. On a donc $x \in C_2$, et comme $x < a$, on a $x \in \{x \in C_2 : x < a\}$. Retenons donc que

$$(1) \{x \in C_2 : x < a\} = \{x \in C_1 : x < a\}$$

Considérons l'ensemble

$$D = C_2 \setminus \{x \in C_2 : x < a\}$$

Si $D = \emptyset$ alors, par (1), $C_2 \subseteq C_1$. Sinon, comme C_2 est bien ordonnée, D admet un plus petit élément c . On a alors

$$C_2 \setminus \{x \in C_2 : x < a\} = \{x \in C_2 : c \leq x\}$$

et donc, puisque C_2 est totalement ordonnée,

$$\{x \in C_2 : x < a\} = \{x \in C_2 : x < c\}$$

On a donc, en reportant dans (1),

$$(1) \{x \in C_2 : x < c\} = \{x \in C_1 : x < a\}$$

En appliquant la fonction u à cette égalité, il vient

$$c = u(\{x \in C_2 : x < c\}) = u(\{x \in C_1 : x < a\}) = a$$

contradiction, puisque $c \in C_2$ et $a \in C_1 \setminus C_2$. Nous avons ainsi montré que $C_2 \subseteq C_1$. Comme C_1 et C_2 sont distinctes, on a même l'inclusion stricte $C_2 \subsetneq C_1$.

Il reste à montrer que C_2 est un segment initial de C_1 . Soit $m = \min(C_1 \setminus C_2)$. Soit $x \in C_2$. Alors, $x \in C_1$ mais $x \notin C_1 \setminus C_2$, donc $m \not\leq x$. Comme C_1 est totalement ordonnée, $x < m$. Inversement, soit $x \in C_1$ tel que $x < m$. Alors, $x \notin C_1 \setminus C_2$, donc $x \in C_2$. Ainsi,

$$C_2 = \{x \in C_1 : x < m\}$$

est un segment initial de C_1 . $\square$

Le plus dur est passé. Il nous reste maintenant à étudier $\bigcup \Gamma$ qui, intuitivement, est une « méga-chaîne ».

Lemme 14. $\bigcup \Gamma$ est une chaîne bien ordonnée de E .

Démonstration. Soit A une partie non vide de $\bigcup \Gamma$. Soit $x \in A$. Il existe $C \in \Gamma$ tel que $x \in C$. L'ensemble $A \cap C$ est une partie non vide de l'ensemble bien ordonné C , il possède donc un plus petit élément a . Soit $y \in A$. Soit $C' \in \Gamma$ tel que $y \in C'$.

- Si $C' = C$ ou C' est un segment initial de C , alors $y \in C$ et donc $a \leq y$.
- Si C est un segment initial de C' , soit $b \in C'$ tel que

$$C = \{z \in C' : z < b\}$$

Si $y < b$ alors $y \in C$ et donc $a \leq y$. Si $b \leq y$ alors $a < b \leq y$.

Ainsi, $a \in A$ et pour tout $y \in A$, $a \leq y$. L'ensemble A possède un plus petit élément.

□

En fait, $\bigcup \Gamma$ vérifie aussi la propriété magique numéro 2 des chaînes de Γ .

Proposition 15. $\bigcup \Gamma \in \Gamma$.

Démonstration. Nous avons déjà vu que $\bigcup \Gamma$ est une chaîne bien ordonnée de E . Il nous reste à montrer que pour tout $x \in \bigcup \Gamma$,

$$u(\{y \in \bigcup \Gamma : y < x\}) = x$$

Soit donc $x \in \bigcup \Gamma$. Il existe une chaîne $C \in \Gamma$ telle que $x \in C$. Soit $y \in \Gamma$ tel que $y < x$. Tout comme pour x , il existe une chaîne $C' \in \Gamma$ telle que $y \in C'$. Si $C' = C$ ou si C' est un segment initial de C , alors $C' \subseteq C$ et donc $y \in C$. Si, au contraire, C est un segment initial de C' alors, comme $x \in C$, on a aussi $y \in C$. Ainsi,

$$\{y \in \bigcup \Gamma : y < x\} = \{y \in C : y < x\}$$

et donc, comme $C \in \Gamma$,

$$u(\{y \in \bigcup \Gamma : y < x\}) = u(\{y \in C : y < x\}) = x$$

□

Nous y sommes. Une contradiction va bientôt apparaître.

Comme $\bigcup \Gamma \in \Gamma$, on a

$$(\star) \quad \bigcup \Gamma \cup \{u(\bigcup \Gamma)\} \in \Gamma$$

- D'un côté, par la définition de u , $u(\bigcup \Gamma)$ est strictement plus grand que tous les éléments de $\bigcup \Gamma$.
- D'un autre côté, par $(\star)$, $u(\bigcup \Gamma)$ appartient à $\bigcup \Gamma$, puisque c'est un élément d'un élément de Γ .

Contradiction. Ainsi, l'hypothèse que nous avons faite au tout début de la démonstration est fausse : l'ensemble E possède un élément maximal.

2.4 Conclusion

Le lemme de Zorn est ainsi équivalent à l'axiome du choix. S'il fallait faire une analogie hasardeuse, ce pourrait être la suivante. L'axiome du choix et le lemme de Zorn, c'est un peu comme la programmation en langage machine et la programmation en langage Python. L'axiome du choix est une propriété « atomique », alors que le lemme de Zorn est une propriété « évoluée ». Cela dit, ces deux propriétés sont équivalentes et ont donc toutes les deux exactement les mêmes conséquences. Pour rester dans notre analogie, on peut *a priori* écrire les mêmes programmes en langage machine et en Python.

Nous allons maintenant nous attaquer à un second théorème, lui aussi très évolué et équivalent à l'axiome du choix : le théorème de Zermelo.

3 Le théorème de Zermelo

3.1 L'énoncé

THÉORÈME DE ZERMELO

Tout ensemble peut être bien ordonné.

Un énoncé aussi simple nous fait suspecter que prouver que AC implique Zermelo sera extrêmement compliqué, et que la réciproque sera facile à montrer. C'est effectivement le cas. En fait, pour prouver l'implication difficile, nous passerons par le lemme de Zorn.

Commençons par le plus facile.

3.2 Zermelo implique Choix

Proposition 16. *Le théorème de Zermelo implique l'axiome du choix.*

Démonstration. Supposons le théorème de Zermelo vérifié. Soit E un ensemble. Munissons $\mathcal{P}(E)$ d'un bon ordre $\leq$. La fonction

$$f : \mathcal{P}(E) \setminus \{\emptyset\} \longrightarrow E$$

définie par $f(A) = \min A$ est alors une fonction de choix sur $\mathcal{P}(E) \setminus \{\emptyset\}$. Nous avons donc prouvé AC_2 . $\square$

Passons à l'implication « difficile ». Nous sommes ici heureux d'avoir prouvé que AC équivaut au lemme de Zorn : ceci nous permettra de nous reposer sur Zorn pour prouver cette implication.

3.3 Zorn implique Zermelo

Proposition 17. *Le lemme de Zorn implique le théorème de Zermelo.*

Démonstration. Supposons le lemme de Zorn vérifié. Soit E un ensemble. Il s'agit de prouver que E peut être bien ordonné. L'idée est de considérer tous les bons ordres définis sur des parties de E et d'exhiber *via* le lemme de Zorn un bon ordre maximal (en un certain sens à préciser). Puis de constater que ce bon ordre maximal est en fait un bon ordre sur E .

Soit

$$\mathcal{E} = \{(A, \mathcal{R}) : A \subseteq E, \mathcal{R} \text{ est un bon ordre sur } A\}$$

Munissons $\mathcal{E}$ d'une relation $\preceq$ en posant $(A, \mathcal{R}) \preceq (B, \mathcal{S})$ lorsque

- $A = B$, ou bien A est un segment initial de B pour la relation $\mathcal{S}$.
- Pour tous $x, y \in A$, $x\mathcal{R}y \iff x\mathcal{S}y$.

Remarquons que la première propriété entraîne que $A \subseteq B$.

La relation $\preceq$ est clairement une relation d'ordre sur $\mathcal{E}$.

L'ensemble $\mathcal{E}$ est non vide car il contient le couple $(\emptyset, \mathcal{R})$ où $\mathcal{R}$ est la relation triviale sur l'ensemble vide.

Soit $\mathcal{C}$ une chaîne de $\mathcal{E}$. Soit

$$M = \bigcup_{(A, \mathcal{R}) \in \mathcal{C}} A$$

Soient x et y deux éléments de M . Définissons une relation $\mathcal{T}$ sur M en posant pour tous $x, y \in M$,

$$x\mathcal{T}y \iff \exists (A, \mathcal{R}) \in \mathcal{C}, x, y \in A \text{ et } x\mathcal{R}y$$

Montrons que $\mathcal{T}$ est un ordre sur M .

- Soit $x \in M$. Il existe $(A, \mathcal{R}) \in \mathcal{C}$ tel que $x \in A$. Comme $\mathcal{R}$ est un ordre sur A , $x\mathcal{R}x$ et donc $x\mathcal{T}x$.
- Soient $x, y \in M$. Supposons que $x\mathcal{T}y$. Il existe $(A, \mathcal{R}) \in \mathcal{C}$ tel que $x, y \in A$ et $x\mathcal{R}y$. Comme $\mathcal{R}$ est un ordre sur A , $y\mathcal{R}x$ et donc $y\mathcal{T}x$.
- Soient $x, y, z \in M$. Supposons que $x\mathcal{T}y$ et $y\mathcal{T}z$. Il existe $(A, \mathcal{R}) \in \mathcal{C}$ tel que $x, y \in A$ et $x\mathcal{R}y$. Il existe également $(A', \mathcal{R}') \in \mathcal{C}$ tel que $y, z \in A'$ et $y\mathcal{R}'z$. Comme l'ensemble $\mathcal{C}$ est totalement ordonné, on a par exemple $(A, \mathcal{R}) \preceq (A', \mathcal{R}')$. De là, $x, y \in A'$ et $x\mathcal{R}'y$. Comme $\mathcal{R}'$ est un ordre sur A' , $x\mathcal{R}'z$ et donc $x\mathcal{T}z$.

Montrons enfin que M est bien ordonné par $\mathcal{T}$. Soit D une partie non vide de M . Soit $x \in D$. Il existe $(A, \mathcal{R}) \in \mathcal{C}$ tel que $x \in A$. L'ensemble $A \cap D$ est une partie non vide de A , qui est bien ordonné par $\mathcal{R}$. Il existe donc $a \in A \cap D$ tel que pour tout $y \in A \cap D$, $a\mathcal{R}y$. Soit $y \in D$. Il existe $(A', \mathcal{R}') \in \mathcal{C}$ tel que $y \in A'$. L'ensemble $\mathcal{C}$ est totalement ordonné. On a donc deux possibilités.

- Cas 1, $(A, \mathcal{R}) \preceq (A', \mathcal{R}')$. $A = A'$ ou alors A est un segment initial de A' pour la relation $\mathcal{R}'$, donc $a\mathcal{R}'y$, d'où $a\mathcal{T}y$.
- Cas 2, $(A', \mathcal{R}') \preceq (A, \mathcal{R})$. On a alors $y \in A$, donc $y \in A \cap D$, et $a\mathcal{R}y$ d'où $a\mathcal{T}y$.

Ainsi, a est le plus petit élément de D pour la relation $\mathcal{T}$.

Nous avons donc prouvé que $(M, \mathcal{T}) \in \mathcal{E}$. Clairement, pour tout $(A, \mathcal{R}) \in \mathcal{C}$, on a $(A, \mathcal{R}) \leq (M, \mathcal{T})$. Ainsi, $\mathcal{C}$ est majoré.

Par le lemme de Zorn, $\mathcal{E}$ possède un élément maximal. Appelons-le $(A, \mathcal{R})$. Supposons un instant que $A \neq E$. Soit $a \in E \setminus A$. Soit $B = A \cup \{a\}$. Soit $\mathcal{S}$ le prolongement de $\mathcal{R}$ à B défini en posant pour tout $x \in B$, $x\mathcal{S}a$. Facilement, $(B, \mathcal{S})$ est un bon ordre sur E et $(A, \mathcal{R}) \preceq (B, \mathcal{S})$. Cependant, $(A, \mathcal{R}) \neq (B, \mathcal{S})$ ce qui contredit la maximalité de $(A, \mathcal{R})$. Ainsi, $A = E$ et $\mathcal{R}$ est un bon ordre sur E . $\square$