Démonstration du lemme de Zorn.

Paris 7, L3 – théorie des ensembles
(Paul Rozière)

27 avril 2017

Soit $(E, \leqslant)$ un ensemble ordonné. On appelle *chaîne* de E un sous-ensemble totalement ordonné de $(E, \leqslant)$. En particulier l'ensemble vide est une chaîne de E. Tout singleton est une chaîne de E.
On appelle *segment initial* de $(E, \leqslant)$ un sous-ensemble I de $(E, \leqslant)$ tel que si $x \in I$, alors $\{y \in E \mid y < x\} \subset I$. Une réunion de segments initiaux est un segment initial (de même qu'une intersection mais on ne l'utilisera pas).
On rappelle que, quand $(C, \leqslant)$ est un ensemble bien ordonné, pour tout segment initial I de C différent de C, il existe un unique $x \in C$ tel que $I = \{y \in C \mid y < x\}$; il suffit de prendre pour x le plus petit élément de l'ensemble des majorants stricts de I qui est non vide car $I \neq C$.
Un élément m de $(E, \leqslant)$ est un *élément maximal* s'il ne possède aucun majorant strict : $\forall x \in E$ $(m \leqslant x \Rightarrow m = x)$.
Une chaîne qui n'est strictement contenue dans aucune chaîne est appelée *chaîne maximale*. Si une telle chaîne possède un majorant, celui-ci est alors un élément maximal de l'ensemble E.

Théorème 1 (Lemme de Zorn) *Soit $(E, \leqslant)$ un ensemble ordonné dont toute chaîne possède un majorant, alors $(E, \leqslant)$ possède un élément maximal.*

Une idée de démonstration serait de partir d'un élément quelconque — le singleton constitué de cet élément est alors une chaîne — puis, si celui-ci n'est pas maximal, de lui ajouter un majorant strict de façon à former une nouvelle chaîne, et ainsi de suite.

En général ce procédé ne s'arrête pas en un nombre fini d'étapes. Supposons que l'on ait construit de cette façon une suite de chaînes finies, chacune prolongeant la précédente en ajoutant un élément. La chaîne de longueur n est un segment initial des chaînes de longueur supérieure. La réunion de ces chaînes forme encore une chaîne (infinie cette fois-ci). Elle a donc un majorant et on peut l'ajouter à celle-ci de façon à obtenir une nouvelle chaîne et recommencer. Il faut faire un choix à chaque étape parmi les majorants stricts, et comme il peut y avoir une infinité d'étapes, on utilise l'axiome du choix.

Une chaîne ainsi construite est non seulement totalement ordonnée mais *bien ordonnée*, c'est-à-dire que toute partie non vide de la chaîne possède un plus petit élément. On peut en fait démontrer un énoncé renforcé du lemme de Zorn, dont le premier est une conséquence immédiate.

Théorème 2 (Lemme de Zorn pour les chaînes bien ordonnées) *Soit $(E, \leqslant)$ un ensemble ordonné dont toute chaîne bien ordonnée possède un majorant, alors $(E, \leqslant)$ possède un élément maximal.*

La démonstration demande de préciser la construction de la chaîne maximale esquissée ci-dessus, et de justifier qu'elle s'arrête (ce sera alors forcément sur un élément maximal). Il faudrait quelques outils supplémentaires de théorie des ensembles concernant les bons ordres (par exemple la théorie des ordinaux) pour rendre cette démarche rigoureuse, il s'agit essentiellement, en plus de l'utilisation de l'axiome du choix, d'un principe de définition par récurrence plus fort que celui sur les entiers (il faut pouvoir aller « bien au delà » des entiers).

La démonstration qui suit décrit directement cette construction : la chaîne est construite par « approximations successives ». C'est une adaptation (due à Helmut Kneser) de la première démonstration que tout ensemble peut être bien ordonné due à Ernst Zermelo en 1904.

Une démonstration différente, qui construit plutôt la chaîne par intersection, est donnée dans le polycopié de J.L. Krivine « Logique et théorie axiomatiques ».

Une chaîne bien ordonnée maximale peut être obtenue par réunion de chaînes bien ordonnées qui se « prolongent l'une l'autre ». Le lemme suivant donne une condition pour qu'une réunion de chaînes bien ordonnées soit encore bien ordonnée.

Lemme 1 *Soit $(E, \leqslant)$ un ensemble ordonné, et $\mathscr{S}$ un ensemble de chaînes bien ordonnées de E tel qu'étant données deux chaînes C et D dans $\mathscr{S}$, C est un segment initial de D ou D est un segment initial de C. Alors*

i. tout élément de $\mathscr{S}$ est un segment initial de $\bigcup \mathscr{S}$ (la réunion des chaînes de $\mathscr{S}$) ;

ii. $\bigcup \mathscr{S}$ est une chaîne bien ordonnée de E.

Démonstration.

i. Soit $C \in \mathscr{S}$. Soit $x \in C$. Soit $y \in \bigcup \mathscr{S}$ tel que $y < x$. Alors il existe D dans $\mathscr{S}$ tel que $y \in D$. Si $D \subset C$, $y \in C$. Sinon C est un segment initial de D et comme $y < x$, on a encore $y \in C$. On a montré que C est un segment initial de $\bigcup \mathscr{S}$.

ii. Soit X un sous-ensemble non vide de $\bigcup \mathscr{S}$. Alors il existe une chaîne C de $\mathscr{S}$ telle que $X \cap C \neq \emptyset$. Soit a le plus petit élément de $X \cap C$. Montrons que c'est aussi le plus petit élément de X. Soit $y \in X$. Si $y \in C$, alors $a \leqslant y$ par définition de a. Sinon, $y \notin C$, il existe $D \in \mathscr{S}$, tel que $y \in D$. Comme $D \not\subset C$, C est un segment initial de D, et donc comme $y \notin C$, on ne peut avoir $y < a$, d'où par totalité de l'ordre dans D, $a \leqslant y$.

   On a montré que $\bigcup \mathscr{S}$ est une chaîne bien ordonné par l'ordre de $(E, \leqslant)$. ■

On veut utiliser ce lemme pour construire une chaîne bien ordonnée maximale comme réunion de chaînes toutes comparables par segment initial. On utilise maintenant l'axiome du choix pour définir une fonction qui permet de prolonger de façon uniforme les chaînes bien ordonnées de $(E, \leqslant)$.

Lemme 2 (AC) *Soit un ensemble ordonné $(E, \leqslant)$, alors il existe une fonction g définie sur toutes les chaînes bien ordonnées de $(E, \leqslant)$ qui possèdent un majorant strict, et telle que $g(C)$ est un majorant strict de C.*

Démonstration. Soit f une fonction de choix sur $\mathscr{P}(E) \setminus \{\emptyset\}$. Soit $\mu(C)$ l'ensemble des majorants stricts de la chaîne bien ordonnée C. Si $\mu(C) \neq \emptyset$, on pose $g(C) = f(\mu(C))$. ■

Étant donnée une fontion g définie sur une partie de l'ensemble des chaînes bien ordonnées de $(E, \leqslant)$, on appelle *g-chaîne* une chaîne bien ordonnée C de $(E, \leqslant)$ vérifiant

$$\forall x \in C \; x = g(\{y \in C \;/\; y < x\}) \; .$$

Si g est une fonction telle que celle fournie par le lemme 2, c'est-à-dire définie sur toutes les chaînes bien ordonnées strictement majorées, alors :

— $g(\emptyset)$ est forcément défini (si $E \neq \emptyset$), et appartient donc à toute g-chaîne non vide ;

— si C est une g-chaîne qui possède des majorants stricts, alors $C \cup \{g(C)\}$ est encore une g-chaîne.

L'idée telle qu'exposée initialement était de construire une chaîne maximale en « itérant suffisamment » une telle fonction g, à partir de $g(\emptyset)$. On va la réaliser en construisant une chaîne maximale comme réunion de toutes les g-chaînes. Le lemme suivant assure que les g-chaînes se « recollent » comme il faut, c'est-à-dire qu'un ensemble de g-chaînes satisfait les hypothèses du lemme 1.

Lemme 3 *Soit $(E, \leqslant)$ un ensemble ordonné, g une fonction définie sur une partie de l'ensemble des chaînes bien ordonnées de E. Alors, étant données deux g-chaînes de $(E, \leqslant)$, l'une est segment initial de l'autre.*

Démonstration. Soient C et D deux g-chaînes. Soit $\mathscr{S}$ l'ensemble des chaînes bien ordonnées qui sont des segments initiaux à la fois de C et de D. L'ensemble $\bigcup \mathscr{S}$ est un sous-ensemble d'un ensemble bien ordonné (C ou D) donc c'est une chaîne bien ordonnée. C'est un segment initial de C, ainsi que de D, comme réunion de segments initiaux.

Si $\bigcup \mathscr{S} = C$ ou $\bigcup \mathscr{S} = D$ on a la conclusion du lemme. Raisonnons par l'absurde et supposons que $\bigcup \mathscr{S} \neq C$ et $\bigcup \mathscr{S} \neq D$. Comme $\bigcup \mathscr{S}$ est un segment initial strict de ces deux ensembles bien ordonnés, il existe un élément x de C, et un élément x' de D tel que $\bigcup \mathscr{S} = \{y \in C \;/\; y < x\} = \{y \in D \;/\; y < x'\}$. Mais comme C et D sont des g-chaînes, $x = x' = g(\bigcup \mathscr{S}\}$. La chaîne bien ordonnée $\bigcup \mathscr{S} \cup \{g(\bigcup \mathscr{S})\}$ est alors un segment initial de C, ainsi que de D. Ceci contredit la définition de $\mathscr{S}$. ■

Démonstration (lemme de Zorn). Le théorème 2 (lemme de Zorn pour les chaînes bien ordonnées) se déduit des trois lemmes précédents.

Soit $(E, \leqslant)$ un ensemble ordonné dont toutes les chaînes bien ordonnées possèdent un majorant. D'après le lemme 2 (axiome du choix), il existe une fonction g qui a toute chaîne bien ordonnée strictement majorée dans $(E, \leqslant)$ associe un majorant strict. l'ensemble des g-chaînes satisfait les hypothèses du lemme 1 d'après le lemme 3. La réunion G de toutes les g-chaînes de $(E, \leqslant)$ est donc encore une chaîne bien ordonnée (lemme 1 (ii)). Soit $x \in G$, alors $x \in C$ pour une g-chaîne C, et par définition $x = g(\{y \in C \;/\; y < x\})$. Or C est un segment initial de G (lemme 1 (i)), $\{y \in C \;/\; y < x\}) = \{y \in G \;/\; y < x\}$, d'où $x = g(\{y \in G \;/\; y < x\})$; G est donc une g-chaîne.

La g-chaîne bien ordonnée G ne possède pas de majorant strict, car sinon on pourrait la prolonger en une nouvelle g-chaîne par $g(G)$. Or, par hypothèse, G possède un majorant, qui est donc un élément maximal. ■