LA THÉORIE ZFC DES ENSEMBLES

1. Les débuts de la théorie des ensembles

Georg Cantor (1845–1918) est considéré comme l'inventeur de la théorie des ensembles. La théorie de Cantor reposait sur deux principes, que nous allons maintenant examiner l'un après l'autre. Voici le premier :

I. *Deux ensembles sont égaux si et seulement si ils ont les mêmes éléments.*

1.1. **Exemple.** Supposons que A et B sont des ensembles satisfaisant :

- $2 \in A$, $3 \in A$, $4 \in A$, et $2, 3, 4$ sont les seuls éléments de A
- $2 \in B$, $3 \in B$, $4 \in B$, et $2, 3, 4$ sont les seuls éléments de B.

Alors le principe I implique que $A = B$.

Le symbole $\{2, 3, 4\}$ désigne l'unique ensemble dont les éléments sont $2, 3, 4$ (et rien d'autre). En général, si $a_1, \ldots, a_n$ sont des objets quelconques alors $\{a_1, \ldots, a_n\}$ désigne l'unique ensemble dont les éléments sont précisément $a_1, \ldots, a_n$. Remarquez aussi que

$$\{2, 3, 4\} = \{4, 2, 3\} = \{2, 3, 2, 4\}$$

puisque tous ces ensembles ont les mêmes éléments.

1.2. **Exemple.** Montrons que $\{2, 3, 4\} \neq \{2, 3\}$.

Preuve. On a $4 \in \{2, 3, 4\}$ et $4 \notin \{2, 3\}$, donc $\{2, 3, 4\}$ et $\{2, 3\}$ n'ont pas les mêmes éléments, donc le principe I implique $\{2, 3, 4\} \neq \{2, 3\}$. □

II. *Étant donnée une condition P, il existe un ensemble dont les éléments sont précisément tous les objets qui satisfont cette condition.*

On utilise la notation suivante pour désigner cet ensemble :

$$\big\{\, x \mid x \text{ satisfait la condition } P \,\big\}.$$

1.3. **Exemple.** D'après le principe II, les ensembles suivants **existent** :

- $A = \big\{\, x \mid x \text{ est un nombre premier} \,\big\} = \{2, 3, 5, 7, 11, 13, 17, \ldots\}$
  (A est l'ensemble de tous les nombres premiers)
- $B = \big\{\, x \mid x \text{ est un sous-ensemble de } \mathbb{N} \,\big\}$
  (B est l'ensemble de tous les sous-ensembles de $\mathbb{N}$)
- $C = \big\{\, x \mid x \text{ est une fonction de } \mathbb{N} \text{ vers } \mathbb{N} \,\big\}$
  (C est l'ensemble de toutes les fonctions de $\mathbb{N}$ vers $\mathbb{N}$)

- $U = \{\, x \mid x \text{ est un ensemble} \,\}$
  (U est l'ensemble de tous les ensembles, donc en particulier $U \in U$).

Vers le tournant du siècle, plusieurs mathématiciens découvrirent que le Principe II avait des conséquences impossibles. Voici un exemple, découvert par Bertrand Russell.

Paradoxe de Russell (1901). Considérons les objets X qui satisfont la condition P suivante :

$$P : \quad \text{“}X \text{ est un ensemble et } X \notin X\text{”}.$$

Alors d'après le principe II l'ensemble $R = \{\, X \mid X \text{ est un ensemble et } X \notin X \,\}$ existe (R est l'ensemble de tous les ensembles qui ne sont pas éléments d'eux-même). Il y a alors deux possibilités : soit $R \in R$, soit $R \notin R$.

Si $R \in R$ alors R satisfait P (car chaque élément de R satisfait cette condition), donc R est un ensemble qui n'est pas élément de lui-même, donc $R \notin R$, donc

$$R \in R \text{ et } R \notin R.$$

Si $R \notin R$ alors R est un ensemble qui n'est pas élément de lui-même, autrement dit R satisfait P, mais alors $R \in R$ (car tout objet qui satisfait la condition P est élément de R), donc on arrive à la même conclusion :

$$R \in R \text{ et } R \notin R.$$

Donc, dans un cas comme dans l'autre, R satisfait $R \in R$ et $R \notin R$.

En résumé, le principe II implique qu'il existe un ensemble R qui satisfait les deux conditions $R \in R$ et $R \notin R$, ce qui est impossible.

D'autres paradoxes ont été découverts, dont au moins un par Cantor lui-même. Au tournant du siècle, il était devenu clair que la théorie des ensembles se contredisait elle-même et qu'il était nécessaire de la reconstruire sur des bases logiques plus solides. Cette période est connue sous le nom de la crise des fondements (parce que toutes les mathématiques s'appuient sur la théorie des ensembles).

2. Introduction à la théorie ZFC

On a vu que les principes I et II de Cantor donnent lieu à une théorie des ensembles qui se contredit elle-même. Pour résoudre cette difficulté, Zermelo proposa en 1908 une liste d'axiomes destinée à remplacer les deux principes de Cantor ; en 1922 Fraenkel et Skolem apportèrent quelques améliorations aux axiomes de Zermelo et le système d'axiomes qui en résulta est connu sous le nom de la théorie “ZF” des ensembles (pour Zermelo-Fraenkel). Plus tard on ajouta un axiome supplémentaire à la liste : l'axiome du choix. “ZFC” désigne le système d'axiomes ZF auquel on a ajouté l'axiome du choix. C'est le système qui est utilisé aujourd'hui, donc toutes les mathématiques s'appuient sur ZFC.

Avant d'énoncer les axiomes de ZFC, expliquons dans quel contexte nous nous plaçons.

Au commencement nous avons deux choses :

- nous avons des objets
- nous avons une relation "$\in$" entre ces objets.

Ces objets et cette relation forment ce que nous appellerons *l'univers*. Pour l'instant, la seule chose que nous savons à propos de cet univers c'est que si x, y sont des objets quelconques alors on a soit $x \in y$, soit $\neg(x \in y)$. On écrira "$x \notin y$" comme abbréviation de $\neg(x \in y)$. Les axiomes que nous allons bientôt énoncer sont des conditions que les objets et la relation $\in$ doivent satisfaire.

Exemple. Imaginez un univers qui contient exactement deux objets, disons a et b, et tel que la relation $\in$ est définie par :

$$a \in b, \quad a \notin a, \quad b \notin b, \quad b \notin a.$$

Alors cet univers satisfait les trois premiers axiomes ci-dessous mais ne satisfait pas le quatrième. En fait on peut montrer qu'un univers qui satisfait les quatre premiers axiomes contient nécessairement une infinité d'objets.

La théorie des ensembles est l'étude des univers qui satisfont **tous** les axiomes de ZFC. Ces univers sont compliqués et personne n'en a une compréhension complète.

Pour parler de l'univers il est commode d'utiliser des mots connus : les objets seront appelés des *ensembles*, et lorsque $x \in y$ on dira que x *est un élément de* y. Il vaut la peine d'insister : **tous** les objets de l'univers sont des ensembles, donc il n'existe rien d'autre que des ensembles et nous ne rencontrerons jamais une "chose" qui n'est pas un ensemble. En particulier tous les éléments d'un ensemble donné sont eux-même des ensembles. Lorsqu'on écrit "$x \in y$", non seulement y est un ensemble mais x aussi. Une formule $\forall_x(\dots)$ doit toujours être interprétée de la manière suivante :

"tout objet x de l'univers satisfait la condition $(\dots)$"

et puisque "ensemble" est synonyme de "objet", la même formule peut aussi être traduite par

"tout ensemble x satisfait la condition $(\dots)$".

Similairement, $\exists_x(\dots)$ se traduit par "il existe au moins un objet x de l'univers qui satisfait la condition $(\dots)$", ou encore par "il existe au moins un ensemble x qui satisfait la condition $(\dots)$".

On a dit que "$\in$" était la seule relation entre les objets mais évidemment on a aussi la relation d'égalité : deux objets x, y peuvent être égaux, $x = y$, et comme d'habitude ceci signifie que x et y sont en fait le même objet (on n'a pas deux objets mais un seul, qu'on a nommé deux fois).

3. Les trois premiers axiomes de ZFC

Étant donnés des ensembles x et y, la phrase "x et y ont les mêmes éléments" s'écrit

$$\forall_z(z \in x \iff z \in y).$$

Si $x = y$ alors forcément x et y ont les mêmes éléments, puisque nous parlons du même ensemble. Donc dans n'importe quel univers, quel qu'il soit, l'énoncé suivant est vrai :

$$\forall_{x,y}\big(x = y \implies \forall_z(z \in x \iff z \in y)\big).$$

On peut se demander si la réciproque est vraie : si x et y sont des ensembles qui ont les mêmes éléments, s'ensuit-il nécessairement que $x = y$? Dans certains univers, la réponse est "non" (voir Ex. 3.1). Nous voulons une théorie des ensembles dans laquelle la réponse à cette question est "oui", donc nous nous donnons l'axiome suivant (qui est exactement le principe "I" de Cantor) :

Axiome 1 (Axiome d'extensionalité)**.** *Deux ensembles sont égaux si et seulement si ils ont les mêmes éléments.*

$$\forall_{x,y}\big(x = y \iff \forall_z(z \in x \iff z \in y)\big)$$

3.1**. Exemple.** Considérez un univers qui a précisément trois objets distincts, a, b_1, b_2, et tel que $\in$ est définie par :

$$a \in b_1, \quad a \in b_2, \quad a \notin a, \quad \forall_i(b_i \notin a), \quad \forall_{i,j}(b_i \notin b_j).$$

Alors $b_1 \neq b_2$ mais b_1, b_2 ont les mêmes éléments, donc cet univers ne satisfait pas l'Axiome 1.

Le deuxième axiome est une version affaiblie du principe "II" de Cantor :

Axiome 2 (Axiome de spécification)**.** *Étant donnés un ensemble A et une condition $P(x)$ sur x, il existe un ensemble B dont les éléments sont précisément les éléments x de A qui satisfont la condition $P(x)$.*

Ainsi, étant donnés A et $P(x)$, l'axiome affirme l'existence d'un ensemble B qui satisfait

$$\forall_x\big(x \in B \iff [x \in A \ \wedge \ P(x)]\big).$$

Cet ensemble B est unique, en vertu de l'axiome d'extensionalité. On le désigne par la notation suivante :

$$B = \big\{\, x \in A \mid P(x) \,\big\}.$$

La différence entre l'Axiome 2 et le principe II de Cantor est celle-ci : au lieu de former un ensemble avec *tous les objets x de l'univers* qui satisfont la condition $P(x)$, on se restreint maintenant aux éléments d'un ensemble A.

3.2**. Exemple.** L'axiome de spécification implique que si A est un ensemble alors l'ensemble $\big\{\, x \in A \mid x \notin x \,\big\}$ existe, mais l'axiome ne permet pas d'affirmer que $\big\{\, x \mid x \notin x \,\big\}$ existe.

3.3. **Exemple.** Considérez un univers dans lequel il y a exactement trois objets, a, b, c, et tel que $\in$ est définie par

$$a \in c, \quad b \in c, \quad b \in b \quad \text{(et } x \notin y \text{ dans tous les cas qui ne sont pas nommés).}$$

- Remarquez que a n'a aucun élément, b a un seul élément, et c a deux éléments ; donc il n'existe pas deux objets différents qui ont les mêmes éléments, donc cet univers satisfait l'Axiome 1.
- Dans cet univers il n'existe aucun objet x qui satisfait : "$a \in x$ et a est le seul élément de x"; autrement dit, le singleton $\{a\}$ n'existe pas.
- Cet univers ne satisfait pas l'Axiome 2. En effet, si cet axiome était satisfait alors l'ensemble $B = \{\, x \in c \mid x \notin x \,\}$ existerait, et en fait B serait le singleton $\{a\}$ (car il existe exactement un élément x de c qui satisfait $x \notin x$, et cet élément est $x = a$). Puisque $\{a\}$ n'existe pas, l'Axiome 2 n'est pas satisfait.

Pour compléter la présentation de l'Axiome 2 il nous reste à préciser qu'est-ce qu'on entend par "une condition $P(x)$ sur x". Nous dirons que $P(x)$ est une condition sur x si $P(x)$ est une formule dont la seule variable libre est x. Par exemple, "$x \notin x$" est une formule dont la seule variable libre est x, donc est une condition sur x.

Voici un exemple plus compliqué : l'expression

$$\exists_{y,z}(x \in y \ \wedge \ y \in z)$$

est une condition sur x (car c'est une formule dont la seule variable libre est x). On peut donc utiliser cette condition avec l'Axiome 2 pour définir des ensembles : l'axiome affirme que si A est un ensemble alors l'ensemble $\{\, x \in A \mid \exists_{y,z}(x \in y \ \wedge \ y \in z) \,\}$ existe.

Voici un autre exemple. La formule $x \notin y$ a deux variables libres x, y, mais si on remplace la variable y par un ensemble spécifique B on obtient une formule "$x \notin B$" dont la seule variable libre est x. Donc "$x \notin B$" est une condition sur x et peut donc être utilisée avec l'Axiome 2 pour définir des ensembles : si A est un ensemble alors l'ensemble $\{\, x \in A \mid x \notin B \,\}$ existe. Prenons note de ce qu'on vient de démontrer :

3.4. *Si A, B sont des ensembles alors l'ensemble $\{\, x \in A \mid x \notin B \,\}$ existe.*

L'ensemble $\{\, x \in A \mid x \notin B \,\}$ est désigné par $A \setminus B$ (on lit "A moins B"). Remarquez :

$$\forall_x \big[x \in A \setminus B \iff (x \in A \ \wedge \ x \notin B)\big].$$

Similairement, l'Axiome 2 implique :

3.5. *Si A, B sont des ensembles alors l'ensemble $\{\, x \in A \mid x \in B \,\}$ existe.*

L'ensemble $\{\, x \in A \mid x \in B \,\}$ est désigné par $A \cap B$. Il satisfait :

$$\forall_x \big[x \in A \cap B \iff (x \in A \ \wedge \ x \in B)\big].$$

Remarque. Les axiomes qu'on a vus jusqu'ici ne nous permettent **pas** d'affirmer que si A et B sont des ensembles alors l'ensemble $A \cup B$ existe. En effet, considérez un

univers avec trois objets a_0, a_1, a_2 et tel que $a_i \in a_j \iff j = i+1$; cet univers satisfait les axiomes 1 et 2 (et aussi 3 ci-dessous), mais $a_1 \cup a_2$ n'existe pas.

Même si on n'a vu que deux des axiomes, on peut déjà démontrer quelque-chose d'intéressant : que l'ensemble de tous les ensembles n'existe pas.

3.6. **Proposition.** *Il n'existe pas d'ensemble A qui satisfait $\forall_x(x \in A)$.*

Démonstration. (Par contradiction.) Supposons que :

$(*)$ *Il existe un ensemble A tel que $\forall_x(x \in A)$.*

Puisque A est un ensemble et puisque "$x \notin x$" est une condition sur x, l'axiome de spécification implique que l'ensemble $R = \{\, x \in A \mid x \notin x \,\}$ existe. Notons que $R \in A$, puisque $\forall_x(x \in A)$. Notons aussi que R doit satisfaire une des conditions $R \in R$ ou $R \notin R$.

- Si $R \in R$ alors (puisque chaque $x \in R$ satisfait $x \notin x$) on a $R \notin R$, donc on a $R \in R$ et $R \notin R$.
- Si $R \notin R$ alors $R \in A$ et $R \notin R$, donc $R \in \{\, x \in A \mid x \notin x \,\}$, donc $R \in R$, donc $R \in R$ et $R \notin R$.

Donc (par séparation des cas) R satisfait les deux conditions $R \in R$ et $R \notin R$.

Puisque l'hypothèse $(*)$ implique qu'il existe un ensemble R qui satisfait $R \in R$ et $R \notin R$, on conclut que $(*)$ est fausse et la preuve est terminée. □

L'univers vide (aucun objet) satisfait les axiomes 1 et 2. Donc pour s'assurer qu'il existe au moins un ensemble, nous avons besoin d'un nouvel axiome :

Axiome 3 (Axiome d'existence)**.** *Il existe au moins un ensemble.*

$$\exists_x(x = x)$$

3.7. **Proposition.** *Il existe un et un seul ensemble qui n'a aucun élément.* (On l'appelle l'ensemble vide, et on le désigne par le symbole $\varnothing$.)

Démonstration. En vertu de l'axiome d'existence, il existe au moins un ensemble ; disons que A est un ensemble. Alors l'axiome de spécification affirme que l'ensemble suivant existe :

$$B = \{\, t \in A \mid t \neq t \,\}.$$

Les éléments de B sont les éléments t de A qui satisfont $t \neq t$, donc B n'a aucun élément. Ceci montre qu'il existe *au moins un* ensemble qui n'a aucun élément. Si B, B' sont des ensembles qui n'ont aucun élément alors $B = B'$ en vertu de l'axiome d'extensionalité. Donc il existe *exactement un* ensemble qui n'a aucun élément. □

3.8. **Exercice.** Imaginez un univers avec un seul objet a, et tel que $a \in a$. Montrez que les axiomes 1 et 3 sont satisfaits mais pas 2.

3.9. **Exercice.** Considérez un univers avec un seul objet a, et tel que $a \notin a$ (autrement dit, le seul objet de l'univers est l'ensemble vide). Montrez que les trois axiomes 1, 2 et 3 sont satisfaits.

3.10. **Exercice.** Considérez un univers avec deux objets $a \neq b$, et tel que

$$a \in b, \quad a \notin a, \quad b \notin b, \quad b \notin a.$$

Montrez que les trois axiomes 1, 2 et 3 sont satisfaits.

Résumé. Dans un univers qui satisfait les trois premiers axiomes,

- il n'existe aucun ensemble A qui satisfait $\forall_x(x \in A)$
- il existe exactement un ensemble vide.

Remarque. À chaque fois qu'on démontre un théorème (ou une proposition, ou un lemme, etc), ce théorème est valide dans tout univers qui satisfait les axiomes vus avant le théorème. Par exemple, 3.4, 3.5 et 3.6 sont valides dans tout univers qui satisfait les deux premiers axiomes, et 3.7 est valide dans tout univers qui satisfait les trois premiers axiomes.

4. L'intersection

Montrons maintenant que si A est un ensemble non vide, alors l'intersection de tous les éléments de A est un ensemble. En fait les deux premiers axiomes suffisent pour démontrer ceci.

4.1. **Proposition.** *Si A est un ensemble non vide alors il existe un et un seul ensemble V dont les éléments sont précisément tous les ensembles x qui satisfont :*

$$x \text{ est élément de chaque élément de } A.$$

Démonstration. Puisque $A \neq \varnothing$, on peut choisir E tel que $E \in A$. Puisque E est un ensemble et $\forall_a(a \in A \implies x \in a)$ est une condition sur x, l'axiome de spécification implique que

$$V = \{\, x \in E \mid \forall_a(a \in A \implies x \in a) \,\}$$

est un ensemble. Pour un ensemble x quelconque, on a :

$$\begin{aligned} x \in V &\iff x \in E \text{ et } \forall_a(a \in A \implies x \in a) \\ &\iff x \in E \text{ et } x \text{ est élément de chaque élément de } A \\ &\iff x \text{ est élément de chaque élément de } A. \end{aligned}$$

Donc l'ensemble V a la propriété voulue. L'axiome d'extensionalité implique que V est unique. □

4.2. **Notation.** Soient A et V comme dans 4.1. Alors on écrit $V = \bigcap A$, ou encore, $V = \bigcap_{X \in A} X$.

Donc si A est un ensemble non vide alors l'ensemble $\cap A$ existe et satisfait :

$$\textit{Quel que soit l'ensemble } x, \qquad x \in \cap A \iff \forall_{E \in A}\ (x \in E).$$

5. Les paires (non ordonnées)

5.1. **Notation.** Soient $x_1, \ldots, x_n$ des ensembles. S'il existe un ensemble z dont les éléments sont précisément $x_1, \ldots, x_n$ (c'est à dire que $x_1, \ldots, x_n$ sont éléments de z et sont les seuls éléments de z), alors cet ensemble z est unique, en vertu de l'axiome d'extensionalité. Lorsqu'il existe, on le désignera par $z = \{x_1, \ldots, x_n\}$.

5.2. **Exemple.** L'univers de l'Exercice 3.10 satisfait les axiomes 1 à 3. Dans cet univers, il n'existe aucun ensemble z tel que $b \in z$. Donc $\{b\}$ et $\{a, b\}$ n'existent pas.

Axiome 4 (Axiome de la paire)**.** *Étant donnés des ensembles x et y, il existe un ensemble z tel que $x \in z$ et $y \in z$.*

$$\forall_{x,y} \exists_z (x \in z \ \wedge \ y \in z)$$

5.3. **Proposition.**

(1) *Si a est un ensemble alors l'ensemble $\{a\}$ existe.*
(2) *Si a et b sont des ensembles alors l'ensemble $\{a, b\}$ existe.*

Démonstration. Prouvons d'abord l'assertion (b). Soient a, b des ensembles. En vertu de l'axiome de la paire, il existe un ensemble c' qui satisfait $a \in c'$ et $b \in c'$. Alors l'axiome de spécification affirme que l'ensemble suivant existe :

$$c = \big\{\, x \in c' \mid x = a \ \vee \ x = b \,\big\}.$$

Alors a et b sont éléments de c, et sont les seuls éléments de c. Donc $c = \{a, b\}$, donc l'ensemble $\{a, b\}$ existe et (b) est démontré. Pour démontrer (a) il suffit de remarquer que l'assertion (b) est aussi valide dans le cas où $a = b$. □

On sait maintenant que plusieurs ensembles existent :

$$\varnothing \text{ existe} \overset{5.3}{\Longrightarrow} \{\varnothing\} \text{ existe} \overset{5.3}{\Longrightarrow} \{\{\varnothing\}\} \text{ existe} \overset{5.3}{\Longrightarrow} \{\{\{\varnothing\}\}\} \text{ existe} \overset{5.3}{\Longrightarrow} \ldots$$
$$\varnothing \text{ et } \{\varnothing\} \text{ sont des ensembles} \overset{5.3}{\Longrightarrow} \big\{\varnothing, \{\varnothing\}\big\} \text{ existe.}$$

Remarquez que $\varnothing \neq \{\varnothing\}$ puisque ces ensembles n'ont pas les mêmes éléments ($\{\varnothing\}$ a un élément mais $\varnothing$ n'a aucun élément). Puisque $\varnothing \neq \{\varnothing\}$, on voit que l'ensemble $\big\{\varnothing, \{\varnothing\}\big\}$ a précisément deux éléments, donc

$$x = \varnothing, \quad y = \{\varnothing\}, \quad z = \big\{\varnothing, \{\varnothing\}\big\}$$

sont trois ensembles différents (x n'a aucun élément, y en a 1, z en a 2). On peut se demander si l'ensemble $\{x, y, z\}$ existe mais, à ce stade de la théorie, rien ne nous permet de l'affirmer (dans certains univers satisfaisant tous les axiomes qu'on a vus jusqu'ici, on peut trouver des ensembles x, y, z tels que $\{x, y, z\}$ n'existe pas).

5.4. **Exercice.** Montrez que $\{\varnothing\} \neq \{\{\varnothing\}\}$ et que $\{\{\varnothing\}\} \neq \{\{\{\varnothing\}\}\}$.

6. La réunion

Axiome 5 (Axiome de la réunion)**.** *Quel que soit l'ensemble A, il existe au moins un ensemble W qui satisfait : tout élément d'un élément de A est élément de W.*

$$\forall_A \exists_W \forall_x \big(\exists_a (a \in A \ \wedge \ x \in a) \implies x \in W \big)$$

Par exemple, si on a l'ensemble $A = \big\{ \{a, b\}, \{c, d\} \big\}$ alors il existe au moins un ensemble W tel que $a \in W$, $b \in W$, $c \in W$ et $d \in W$ (mais W peut avoir plus d'éléments que ceux-là). Comme on l'a fait pour les axiomes 3 et 4, on peut utiliser l'axiome de spécification pour obtenir une version plus précise de l'Axiome 5 :

6.1**. Proposition.** *Si A est un ensemble alors il existe un et un seul ensemble W dont les éléments sont précisément tous les ensembles x qui satisfont :*

$$x \text{ est élément d'au moins un élément de } A.$$

Démonstration. En vertu de l'axiome de la réunion, il existe un ensemble W' qui satisfait : tout élément d'un élément de A est élément de W'. Autrement dit, pour tout objet x l'implication suivante est vraie :

$$\exists_a (a \in A \ \wedge \ x \in a) \implies x \in W'. \tag{1}$$

Puisque W' est un ensemble et $\exists_a (a \in A \ \wedge \ x \in a)$ est une condition sur x, l'axiome de spécification affirme que l'ensemble suivant existe :

$$W = \big\{ x \in W' \mid \exists_a (a \in A \ \wedge \ x \in a) \big\}.$$

Alors quel que soit l'objet x on a

$$\begin{aligned} x \in W &\iff x \in W' \text{ et } \exists_a (a \in A \wedge x \in a) \\ &\overset{(1)}{\iff} \exists_a (a \in A \wedge x \in a) \\ &\iff x \text{ est élément d'au moins un élément de } A. \end{aligned}$$

Donc l'ensemble W a la propriété voulue. En vertu de l'axiome d'extensionalité, il ne peut y avoir qu'un seul W ayant cette propriété. □

6.2**. Notation.** Si A et W sont comme dans la proposition ci-dessus, on écrit $W = \bigcup A$, ou encore $W = \bigcup_{X \in A} X$.

Par exemple, si on a l'ensemble $A = \big\{ \{a, b\}, \{c, d\} \big\}$ alors $\bigcup A = \{a, b, c, d\}$.

Donc si A est un ensemble quelconque alors l'ensemble $\cup A$ existe et satisfait :

$$\textit{Quel que soit l'ensemble } x, \qquad x \in \cup A \iff \exists_{E \in A} \ (x \in E).$$

6.3**. Corollaire.** *Si $x_1, \ldots, x_n$ sont des ensembles, alors l'ensemble $\{x_1, \ldots, x_n\}$ existe.*

On sait que 6.3 est vrai lorsque $n = 0, 1, 2$, montrons le cas $n = 3$ à titre d'exemple. Supposons que x_1, x_2, x_3 sont des ensembles.

x_1, x_2 sont des ensembles $\overset{5.3}{\Longrightarrow}$ l'ensemble $\{x_1, x_2\}$ existe
x_3 est un ensemble $\overset{5.3}{\Longrightarrow}$ l'ensemble $\{x_3\}$ existe
$\{x_1, x_2\}$ et $\{x_3\}$ sont des ensembles $\overset{5.3}{\Longrightarrow}$ l'ensemble $\big\{\{x_1, x_2\}, \{x_3\}\big\}$ existe
$A = \big\{\{x_1, x_2\}, \{x_3\}\big\}$ est un ensemble $\overset{6.1}{\Longrightarrow}$ l'ensemble $\bigcup A = \{x_1, x_2, x_3\}$ existe

6.4. **Exercice.** Démontrez 6.3 par induction sur n.

6.5. **Définition.** Étant donné un nombre fini d'ensembles $E_1, \ldots, E_n$, on définit les ensembles $E_1 \cup \cdots \cup E_n$ et $E_1 \cap \cdots \cap E_n$ par

$$E_1 \cup \cdots \cup E_n = \bigcup\{E_1, \ldots, E_n\}, \qquad E_1 \cap \cdots \cap E_n = \bigcap\{E_1, \ldots, E_n\}.$$

Remarquez que l'ensemble $\{E_1, \ldots, E_n\}$ existe en vertu de 6.3, donc $\bigcup\{E_1, \ldots, E_n\}$ et $\bigcap\{E_1, \ldots, E_n\}$ existent en vertu de 6.1 et 4.1.

7. Sous-ensembles

7.1. **Définition.** Soit B un ensemble. Un *sous-ensemble* de B est un ensemble A qui satisfait :

chaque élément de A est élément de B.

Pour indiquer que A est un sous-ensemble de B on écrit $A \subseteq B$ (ou parfois $B \supseteq A$). Ainsi,

$$A \subseteq B \quad \text{est une abbréviation de} \quad \forall_x(x \in A \Rightarrow x \in B).$$

On écrira $A \not\subseteq B$ comme abbréviation de $\neg(A \subseteq B)$. Puisque

$$\neg\,\forall_x(x \in A \Rightarrow x \in B) \;\equiv\; \exists_x \neg(x \in A \Rightarrow x \in B) \;\equiv\; \exists_x(x \in A \;\wedge\; x \notin B),$$

on voit que $A \not\subseteq B$ est équivalent à

$$\exists_x(x \in A \;\wedge\; x \notin B).$$

Remarque. On a les synonymes suivants : A est un sous-ensemble de B, A est une partie de B, A est inclus dans B.

7.2. **Proposition.** *Pour tout ensemble A, on a $\varnothing \subseteq A$ et $A \subseteq A$.*

Démonstration. Soit A un ensemble.

Puisque la condition $\forall_x(x \in \varnothing \Rightarrow x \in A)$ est satisfaite, on a $\varnothing \subseteq A$. Puisque la condition $\forall_x(x \in A \Rightarrow x \in A)$ est satisfaite, on a $A \subseteq A$. □

Remarque. On peut aussi démontrer $\varnothing \subseteq A$ par contradiction : supposons que $\varnothing \not\subseteq A$, alors $\exists_x(x \in \varnothing \;\wedge\; x \notin A)$, donc $\exists_x(x \in \varnothing)$, contradiction.

Remarque. Si A est un ensemble et $P(x)$ une condition sur x alors l'ensemble $\big\{\, x \in A \mid P(x) \,\big\}$ est un sous-ensemble de A.

7.3. **Théorème.** *Étant donnés des ensembles A et B,*

$$A = B \quad \Longleftrightarrow \quad (A \subseteq B \;\wedge\; B \subseteq A).$$

Démonstration. Exercice. □

7.4. **Définition.** Soit z un ensemble. S'il existe un ensemble p satisfaisant

$$\forall_x(x \in p \iff x \subseteq z),$$

alors p est appelé *l'ensemble des parties de* z et est désigné par le symbole $\wp(z)$, ou simplement $\wp\, z$.

7.5. **Proposition.** *Si z est un ensemble fini alors $\wp\, z$ existe.*

Démonstration. Puisque z a un nombre fini d'éléments, il s'ensuit qu'il existe un nombre fini d'ensembles x qui satisfont $x \subseteq z$. Donc on peut considérer la liste complète

$$x_1, \ldots, x_n$$

de toutes les parties x_i de z. En vertu de 6.3 l'ensemble $\{x_1, \ldots, x_n\}$ existe, donc il existe un ensemble dont les éléments sont les parties de z. □

7.6. **Exemple.** Si $z = \{a, b, c\}$ alors

$$\wp\, z = \big\{\, \varnothing,\, \{a\},\, \{b\},\, \{c\},\, \{a, b\},\, \{a, c\},\, \{b, c\},\, \{a, b, c\} \,\big\}.$$

Nous voulons une théorie des ensembles dans laquelle $\wp\, z$ existe même lorsque z est un ensemble infini. Pour cela, nous avons besoin d'un nouvel axiome :

Axiome 6 (Axiome de l'ensemble des parties)**.** *Quel que soit l'ensemble z, il existe au moins un ensemble p ayant la propriété suivante : tout sous-ensemble de z est élément de p.*

$$\forall_z \exists_p \forall_x (x \subseteq z \implies x \in p)$$

7.7. **Proposition.** *Quel que soit l'ensemble z, $\wp\, z$ existe.*

Démonstration. Soit z un ensemble.

L'Axiome 6 implique qu'il existe au moins un ensemble p' qui satisfait : tout sous-ensemble de z est élément de p'. On choisit un tel ensemble p', alors on a :

$$\forall_x (x \subseteq z \implies x \in p'). \tag{2}$$

Puisque p' est un ensemble et "$x \subseteq z$" est une condition sur x, l'axiome de spécification implique que l'ensemble suivant existe :

$$p = \big\{\, x \in p' \mid x \subseteq z \,\big\}.$$

Pour un ensemble x quelconque,

$$x \in p \iff (x \in p' \text{ et } x \subseteq z) \overset{(2)}{\iff} x \subseteq z$$

donc l'ensemble p satisfait la condition $\forall_x(x \in p \iff x \subseteq z)$, donc $p = \wp\, z$. □

Remarque. Dans la démonstration ci-dessus on affirme que "$x \subseteq z$" est une condition sur x. Voici pourquoi. Remarquons d'abord que "$x \subseteq z$" est une abbréviation de

$$\forall_t(t \in x \Rightarrow t \in z). \tag{3}$$

Si on considère que x, z sont des variables, alors (3) est une formule avec deux variables libres (x et z). Cependant, z n'est pas une variable, mais plutôt un ensemble spécifique (parce qu'on a commencé la preuve en disant "Soit z un ensemble"). Donc x est la seule variable libre de (3) et c'est pourquoi cette formule est une condition sur x.

7.8. **Définition.** Si les ensembles A et B satisfont $A \subseteq B$ et $A \neq B$, on dit que A est un *sous-ensemble propre* de B et on écrit $A \subset B$ (ou parfois $B \supset A$).

8. Digression : le langage de la théorie des ensembles

Le langage de la théorie des ensembles, que nous désignerons par le symbole $\mathcal{L}$, est constitué de toutes les formules logiques qu'on peut écrire en utilisant seulement les symboles suivants :

- une infinité de variables (les lettres, qu'on peut aussi indicer ; par exemple $a, \ldots, z, A, \ldots, Z$, $a_0, a_1, a_2, \ldots$, etc.)
- les sept symboles logiques : $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, \forall, \exists$
- les parenthèses "(" et ")"
- les deux symboles $\in$ et $=$.

Les symboles ci-dessus constituent notre *alphabet* pour écrire les formules ; nous l'appellerons l'alphabet $\mathcal{A}$. Par exemple, les trois formules

$$x \in y \ \wedge \ \neg(x \in y) \qquad \exists_y(x \in y \ \vee \ x = y) \qquad \forall_x \exists_y(x \in y \ \vee \ x = y)$$

appartiennent à $\mathcal{L}$, mais $\forall_x \exists_y(x \subseteq y)$ n'appartient pas à $\mathcal{L}$ car le symbole $\subseteq$ n'est pas dans $\mathcal{A}$.

Il est permis d'élargir le langage $\mathcal{L}$ en ajoutant des nouveaux symboles à l'alphabet, à condition qu'il soit possible de définir chaque nouveau symbole en utilisant seulement des symboles de $\mathcal{A}$. Par exemple, on définit les nouveaux symboles $\notin$, $\neq$, $\subseteq$, $\not\subseteq$ et $\subset$ par :

$x \notin y$ signifie $\neg(x \in y)$

$x \neq y$ signifie $\neg(x = y)$

$x \subseteq y$ signifie $\forall_z(z \in x \implies z \in y)$

$x \not\subseteq y$ signifie $\neg(x \subseteq y)$, qui signifie $\neg \forall_z(z \in x \implies z \in y)$

$x \subset y$ signifie $x \subseteq y \ \wedge \ x \neq y$, qui signifie $\forall_z(z \in x \implies z \in y) \ \wedge \ \neg(x = y)$.

On obtient alors un langage élargi (appelons-le $\mathcal{L}'$ pour l'instant) qui contient plus de symboles que $\mathcal{L}$. Par exemple, $\forall_y \forall_z \big((y \subseteq z \ \wedge \ z \subseteq y) \ \Rightarrow \ y = z\big)$ est une formule de $\mathcal{L}'$ mais pas de $\mathcal{L}$. Remarquez que chaque formule de $\mathcal{L}'$ est une abbréviation d'une

formule (plus compliquée) de $\mathcal{L}$. Par exemple $\forall_y\forall_z\big((y \subseteq z \ \wedge \ z \subseteq y) \ \Rightarrow \ y = z\big)$ est une abbréviation de

$$\forall_y\forall_z\big(\big[\forall_x(x \in y \Rightarrow x \in z) \ \wedge \ \forall_x(x \in z \Rightarrow x \in y)\big] \ \Rightarrow \ y = z\big).$$

Ainsi, même si $\mathcal{L}'$ contient plus de symboles que $\mathcal{L}$, il ne permet pas d'exprimer plus d'idées. C'est seulement pour des raisons de commodité (pour avoir des formules moins compliquées) qu'on élargit le langage.

Dans les sections précédentes, plusieurs autres symboles ont été ajoutés à l'alphabet : $\varnothing$, $\cup$, $\cap$, $\setminus$, $\wp$, ainsi que des expressions telles que $\{x, y\}$. N'importe quelle formule qui utilise ces nouveaux symboles peut être remplacée par une formule de $\mathcal{L}$. Par exemple, $\varnothing \in z$ est une abbréviation de $\exists_v[v \in z \wedge \forall_t\neg(t \in v)]$. Voici un exemple plus compliqué.

8.1. **Exemple.** Soit φ' la formule $\exists_y(\cup x \in y)$ du langage élargi. On sait que φ' est une abbréviation d'une formule φ de $\mathcal{L}$. Cherchons une telle formule φ. Remarquons d'abord que "$\cup x \in y$" est une abbréviation de $\exists_w[w \in y \ \wedge \ w = \cup x]$. Ensuite :

$$\exists_y(\cup x \in y)$$
$$\exists_y\exists_w[w \in y \ \wedge \ w = \cup x]$$
$$\exists_y\exists_w[w \in y \ \wedge \ \forall_t(t \in w \iff t \in \cup x)]$$
$$\exists_y\exists_w\big[w \in y \ \wedge \ \forall_t(t \in w \iff \exists_{x_0}[t \in x_0 \ \wedge \ x_0 \in x])\big] \tag{4}$$

Désignons par φ la formule (4). Alors $\exists_y(\cup x \in y)$ est une abbréviation de φ et on peut constater que φ est une formule de $\mathcal{L}$.

Les remarques ci-dessus ont leur importance relativement à l'utilisation de l'axiome de spécification (voir Axiome 2). Pour comprendre cet axiome, il faut comprendre ce qu'est une "condition sur x". On a donné une définition provisoire juste après 3.3, on peut maintenant être un peu plus précis :

8.2. **Définition.**

(i) Si φ est une formule de $\mathcal{L}$ dont la seule variable libre est x, alors φ est une condition sur x.

(ii) Si φ' est une formule du langage élargi qui satisfait
φ' est une abbréviation d'une formule φ de $\mathcal{L}$
et la seule variable libre de φ est x,
alors φ' est une condition sur x.

8.3. **Exemple.** Supposons que A est un ensemble. L'axiome de spécification affirme que si $\exists_y(\cup x \in y)$ est une condition sur x alors l'ensemble

$$B = \big\{\, x \in A \mid \exists_y(\cup x \in y) \,\big\}$$

existe. Donc on doit s'assurer que $\exists_y(\cup x \in y)$ est une condition sur x. On a vu en 8.1 que $\exists_y(\cup x \in y)$ est une abbréviation d'une formule φ de $\mathcal{L}$ (φ est (4)). On constate que x est la seule variable libre de φ donc, en vertu de la partie (ii) de la Définition 8.2, $\exists_y(\cup x \in y)$ est une condition sur x. Donc B existe.

Remarquez que dans l'Exemple 8.1 φ et φ' ont les mêmes variables libres (la seule variable libre est x, pour les deux formules). Le principe suivant affirme que c'est toujours le cas :

Si la formule φ' (du langage élargi) est une abbréviation de la formule φ (de $\mathfrak{L}$), alors φ et φ' ont les mêmes variables libres.

En combinant l'affirmation ci-dessus avec la Définition 8.2, on obtient la très utile

8.4. **Proposition.** *Supposons que φ' est une formule du langage élargi et que x est la seule variable libre de φ'. Alors φ' est une condidition sur x.*

8.5. **Exemple.** Soit A un ensemble. L'axiome de spécification affirme que si $\exists_y(\wp(y) \in \cup x)$ est une condition sur x alors l'ensemble

$$B = \{\, x \in A \mid \exists_y(\wp(y) \in \cup x) \,\}$$

existe. Il est clair que $\exists_y(\wp(y) \in \cup x)$ est une formule (du langage élargi) dont la seule variable libre est x. Grâce à 8.4, on peut tout de suite conclure que $\exists_y(\wp(y) \in \cup x)$ est une condition sur x, sans qu'on ait besoin de trouver une formule φ de $\mathfrak{L}$ dont $\exists_y(\wp(y) \in \cup x)$ est une abbréviation.

9. Les paires ordonnées

On sait que si a, b sont des ensembles alors les ensembles $\{a\}$ et $\{a, b\}$ existent, donc l'ensemble $\{\{a\}, \{a, b\}\}$ existe aussi.

9.1. **Définition.** Étant donnés des ensembles a et b, l'ensemble $\{\{a\}, \{a, b\}\}$ sera désigné par le symbole (a, b). L'ensemble (a, b) est appelé une paire ordonnée. Plus précisément, on dit qu'un ensemble x *est une paire ordonnée* s'il existe des ensembles a et b tels que $x = (a, b)$.

9.2. **Exemple.** Si $a = \varnothing$ et $b = \{\varnothing\}$ alors

$$(a, b) = \{\{a\}, \{a, b\}\} = \big\{\{\varnothing\}, \{\varnothing, \{\varnothing\}\}\big\}$$
$$(b, a) = \{\{b\}, \{b, a\}\} = \big\{\{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\big\}$$

9.3. **Exercice.** L'ensemble $\{\varnothing, \{\varnothing\}\}$ est-il une paire ordonnée ? Et $\{\{\varnothing\}\}$?

9.4. Montrons que la phrase *"x est une paire ordonnée"* est une condition sur x. Cette phrase s'écrit

$$\exists_{a,b}\,[\, x = (a, b)\,]$$

et ceci est une formule du langage élargi dont la seule variable libre est x. En vertu de la Proposition 8.4, c'est une condition sur x.

9.5. **Théorème** (Propriété fondamentale des paires ordonnées)**.**
Quels que soient les ensembles a, b, a' et b', $\quad (a, b) = (a', b') \iff (a = a' \ \wedge \ b = b')$.

Démonstration. L'implication "$\Leftarrow$" est évidente.
Preuve de "$\Rightarrow$". Soient a, b, a' et b' des ensembles qui satisfont $(a,b) = (a',b')$. Écrivons $E = (a,b) = (a',b')$. Puisque

$$\cap E = \cap(a,b) = \cap\{\{a\},\{a,b\}\} = \{a\} \cap \{a,b\} = \{a\}$$

et

$$\cap E = \cap(a',b') = \cap\{\{a'\},\{a',b'\}\} = \{a'\} \cap \{a',b'\} = \{a'\},$$

on obtient $\{a\} = \{a'\}$. Puisque $a \in \{a\}$ et $\{a\} = \{a'\}$ on a $a \in \{a'\}$, donc $a = a'$. Similairement, on a

$$\cup E = \cup(a,b) = \cup\{\{a\},\{a,b\}\} = \{a\} \cup \{a,b\} = \{a,b\}$$

et

$$\cup E = \cup(a',b') = \cup\{\{a'\},\{a',b'\}\} = \{a'\} \cup \{a',b'\} = \{a',b'\},$$

donc $\{a,b\} = \{a',b'\}$. Puisque $\{a,b\} = \{a',b'\}$ et $\{a\} = \{a'\}$, on a aussi

$$\{a,b\} \setminus \{a\} = \{a',b'\} \setminus \{a'\}. \tag{5}$$

Montrons que $b = b'$ par séparation des cas :

- Si $a = b$ alors $\{a,b\} = \{b\}$, donc $b' \in \{a',b'\} = \{a,b\} = \{b\}$, donc $b' = b$.
- Si $a \neq b$ alors $\{a,b\} \setminus \{a\} = \{b\}$, donc (5) implique que $\{a',b'\} \setminus \{a'\} = \{b\}$. Puisque $b \in \{b\}$ on a donc $b \in \{a',b'\} \setminus \{a'\}$, donc $b \in \{a',b'\}$ et $b \neq a'$, donc $b = b'$.

Donc $b = b'$. On a montré que $a = a'$ et $b = b'$, donc la preuve est terminée. □

9.6. **Théorème.** *Étant donnés des ensembles A et B, il existe un ensemble P dont les éléments sont toutes les paires (a,b) telles que $a \in A$ et $b \in B$.*

Remarque. L'ensemble P est unique, en vertu de l'Axiome 1. On appelle P le *produit cartésien* de A et B, et on écrit $P = A \times B$. Donc le théorème affirme que, étant donnés des ensembles A, B quelconques, $A \times B$ existe.

Démonstration de 9.6. Soient A et B des ensembles. Commençons par prouver les deux affirmations suivantes:

$$\text{il existe un ensemble } E \text{ tel que } \forall_{a\in A}\, \forall_{b\in B}\, [(a,b) \in E], \tag{6}$$

$$\text{l'expression } \exists_{a\in A}\, \exists_{b\in B}\, [x = (a,b)] \text{ est une condition sur } x. \tag{7}$$

Montrons (6). Puisque A et B sont des ensembles, les ensembles $A \cup B$, $\wp(A \cup B)$ et $\wp\,\wp(A \cup B)$ existent. Si $a \in A$ et $b \in B$ alors $\{a\}$ et $\{a,b\}$ sont deux sous-ensembles de $A \cup B$, donc $\{a\} \in \wp(A \cup B)$ et $\{a,b\} \in \wp(A \cup B)$, donc $\{\{a\},\{a,b\}\} \subseteq \wp(A \cup B)$, donc $(a,b) = \{\{a\},\{a,b\}\} \in \wp\,\wp(A \cup B)$. On vient de montrer que

$$\forall_{a\in A}\, \forall_{b\in B}\, \big[(a,b) \in \wp\,\wp(A \cup B)\big],$$

donc (6) est démontré. Pour montrer (7) on commence par récrire $\exists_{a\in A}\,\exists_{b\in B}\,[x=(a,b)]$ sous la forme

$$\exists_a \exists_b [a \in A \ \wedge \ b \in B \ \wedge \ x = (a,b)]. \tag{8}$$

Remarquez que dans la formule (8), A et B ne sont pas des variables mais des ensembles spécifiques car on a commencé la preuve par "Soient A et B des ensembles". Donc x est la seule variable libre de (8) et, en vertu de 8.4, (7) est démontré.

Puisque les affirmations (6) et (7) sont vraies, l'axiome de spécification affirme que l'ensemble $P = \{\, x \in E \mid \exists_{a\in A}\,\exists_{b\in B}\,[x=(a,b)] \,\}$ existe. Alors on a pour tout ensemble x

$$\begin{aligned} x \in P &\iff x \in E \ \wedge \ \exists_{a\in A}\,\exists_{b\in B}\,[x=(a,b)] \\ &\iff \exists_{a\in A}\,\exists_{b\in B}\,[x=(a,b)] \\ &\iff x \text{ est une paire } (a,b) \text{ telle que } a\in A \text{ et } b\in B, \end{aligned}$$

autrement dit P satisfait la condition demandée. □

9.7. **Théorème.** *Soit R un ensemble de paires ordonnées (chaque élément de R est une paire ordonnée). Alors il existe des ensembles R_1 et R_2 tels que :*

(i) *les éléments de R_1 sont les ensembles x qui satisfont $\exists_t[(x,t)\in R]$*
(ii) *les éléments de R_2 sont les ensembles x qui satisfont $\exists_t[(t,x)\in R]$.*

Remarque. Dans le théorème, R_1 est l'ensemble des premières coordonnées des éléments de R, et R_2 est l'ensemble des deuxièmes coordonnées. On dit que R_1 est la *première projection* de R, et que R_2 est la *deuxième projection.* Remarquez que les ensembles R_1 et R_2 satisfaisant (i) et (ii) sont uniques, en vertu de l'Axiome 1. Remarquez aussi (exercice!) que (i) et (ii) impliquent que $R \subseteq R_1 \times R_2$. (Par exemple, si $R = \{(a,a),(a,b),(a,c)\}$ alors $R_1 = \{a\}$ et $R_2 = \{a,b,c\}$.)

Démonstration de 9.7. Soit $(a,b)\in R$. Puisque $\{a,b\}\in(a,b)\in R$, on voit que $\{a,b\}$ est élément d'au moins un élément de R, donc $\{a,b\}\in\cup R$. Puisque $a\in\{a,b\}\in\cup R$, a est élément d'au moins un élément de $\cup R$, donc $a\in\cup\cup R$. Puisque $b\in\{a,b\}\in\cup R$, on a aussi $b\in\cup\cup R$. Ceci montre que :

$$\textit{Si } (a,b)\in R \textit{ alors } a\in\cup\cup R \textit{ et } b\in\cup\cup R. \tag{9}$$

Puisque $\cup\cup R$ est un ensemble et puisque (en vertu de 8.4) chacune des formules

$$\exists_t[(x,t)\in R], \qquad \exists_t[(t,x)\in R]$$

est une condition sur x, l'Axiome 2 affirme que les ensembles suivants existent :

$$R_1 = \{\, x\in\cup\cup R \mid \exists_t[(x,t)\in R]\,\}, \quad R_2 = \{\, x\in\cup\cup R \mid \exists_t[(t,x)\in R]\,\}.$$

Montrons que R_1 satisfait (i). Pour un ensemble x quelconque,

$$\begin{aligned} x\in R_1 &\iff x\in\cup\cup R \ \wedge \ \exists_t[(x,t)\in R] \\ &\overset{(9)}{\iff} \exists_t[(x,t)\in R], \end{aligned}$$

donc effectivement R_1 satisfait (i). (Expliquons "$\Leftarrow$" dans la dernière biconditionnelle : si x satisfait $\exists_t[(x,t) \in R]$ alors (9) implique que $x \in \cup\cup R$, donc x satisfait $x \in \cup\cup R \ \wedge \ \exists_t[(x,t) \in R]$.)

La preuve que R_2 satisfait (ii) est similaire, et nous l'omettons. □

Les sous-ensembles de $A \times B$

Supposons que A et B sont des ensembles. Alors l'ensemble

$$U = \big\{\, (x,y) \in A \times B \mid x \in y \,\big\}$$

existe-t-il ? On a envie de répondre que oui, en vertu de l'axiome de spécification, mais il y a un problème technique : la formule logique "$x \in y$" a deux variables libres, et pour utiliser l'axiome on a besoin d'une formule ayant une seule variable libre. On peut résoudre cette difficulté en remarquant que le même ensemble U de ci-dessus peut aussi s'écrire

$$U = \big\{\, z \in A \times B \mid \exists_{x,y}\,(z = (x,y) \wedge x \in y) \,\big\}$$

et que z est la seule variable libre de $\exists_{x,y}\,(z = (x,y) \wedge x \in y)$. Donc, en vertu de l'axiome de spécification, l'ensemble U existe.

L'exemple ci-dessus est un cas particulier du résultat suivant :

9.8. **Lemme.** *Soient A et B des ensembles, et soit $P(x,y)$ une formule (du langage élargi) dont toutes les variables libres sont parmi x, y. Alors l'ensemble*

$$\big\{\, (x,y) \in A \times B \mid P(x,y) \,\big\}$$

existe.

Dans le lemme ci-dessus, dire que les variables libres de $P(x,y)$ sont parmi x, y signifie que $P(x,y)$ peut avoir deux variables libres (x et y), mais elle peut aussi en avoir une seule (seulement x, ou seulement y), et elle peut même n'en avoir aucune.

Démonstration de 9.8. La notation $\big\{\, (x,y) \in A \times B \mid P(x,y) \,\big\}$ n'est qu'une abbréviation de

$$\big\{\, z \in A \times B \mid \exists_{x,y}[z = (x,y) \ \wedge \ P(x,y)] \,\big\}$$

et ce dernier ensemble existe en vertu de l'Axiome 2. □

Les n-uplets ordonnés

9.9. **Définition.** Étant donnés des ensembles $a_1, \ldots, a_n$, on définit le n-uplet ordonné $(a_1, \ldots, a_n)$ récursivement :

- $(a_1, a_2) = \{\{a_1\}, \{a_1, a_2\}\}$
- si $(a_1, \ldots, a_{n-1})$ a déjà été défini, on pose $(a_1, \ldots, a_n) = ((a_1, \ldots, a_{n-1}), a_n)$.

Par exemple dans le cas $n = 3$ la définition donne

$$(a,b,c) = \big((a,b),c\big) = \Big\{\{(a,b)\},\{(a,b),c\}\Big\}$$
$$= \Big\{\big\{\{\{a\},\{a,b\}\}\big\},\ \big\{\{\{a\},\{a,b\}\},c\big\}\Big\}$$

et dans le cas $n = 4$, $(a,b,c,d) = ((a,b,c),d) = (((a,b),c),d) =$ etc. Notez bien que tout n-uplet ordonné $(a_1,\ldots,a_n)$ est un ensemble.

9.10. **Théorème.**

(a) *Quels que soient les ensembles* $a_1,\ldots,a_n$ *et* $a'_1,\ldots,a'_n$,

$$(a_1,\ldots,a_n) = (a'_1,\ldots,a'_n) \iff (a_1 = a'_1 \ \wedge\ \cdots\ \wedge\ a_n = a'_n).$$

(b) *Étant donnés des ensembles* $A_1,\ldots,A_n$, *il existe un ensemble* P *dont les éléments sont tous les* n*-uplets ordonnés* $(a_1,\ldots,a_n)$ *tels que* $a_1 \in A_1,\ \ldots,\ a_n \in A_n$.

Remarques. (1) Le Théorème 9.7 peut également être généralisé au cas des n-uplets, mais nous ne le ferons pas.

(2) Dans l'assertion (b) du Théorème 9.10, l'ensemble P est unique, et est appelé le *produit cartésien* de $A_1,\ldots,A_n$. On écrit $P = A_1\times\cdots\times A_n$. Donc l'assertion (b) affirme que $A_1 \times \cdots \times A_n$ existe, quels que soient les ensembles $A_1,\ldots,A_n$.

Démonstration de 9.10(a). "$\Longleftarrow$" est évident. Prouvons "$\Longrightarrow$" par induction sur n. Le cas $n = 2$ est déjà démontré (9.5). Supposons que $n \geq 3$ et que l'assertion est vraie pour les $(n-1)$-uplets.

Supposons que $a_1,\ldots,a_n$ et $a'_1,\ldots,a'_n$ sont des ensembles tels que $(a_1,\ldots,a_n) = (a'_1,\ldots,a'_n)$. Alors

$$((a_1,\ldots,a_{n-1}),a_n) = (a_1,\ldots,a_n) = (a'_1,\ldots,a'_n) = ((a'_1,\ldots a'_{n-1}),a'_n).$$

Puisque $((a_1,\ldots,a_{n-1}),a_n) = ((a'_1,\ldots a'_{n-1}),a'_n)$, 9.5 implique que

$$(a_1,\ldots,a_{n-1}) = (a'_1,\ldots a'_{n-1}) \ \text{ et } \ a_n = a'_n.$$

Puisque $(a_1,\ldots,a_{n-1}) = (a'_1,\ldots a'_{n-1})$, l'hypothèse d'induction implique que

$$a_1 = a'_1 \ \wedge\ \cdots\ \wedge\ a_{n-1} = a'_{n-1}.$$

Donc $a_1 = a'_1 \ \wedge\ \cdots\ \wedge\ a_n = a'_n$. □

La partie (b) se démontre de façon similaire, par induction sur n. Nous omettons cette preuve, mais remarquez tout de même qu'on a les égalités suivantes :

$$A_1 \times A_2 \times A_3 = (A_1 \times A_2) \times A_3,$$
$$A_1 \times A_2 \times A_3 \times A_4 = (A_1 \times A_2 \times A_3) \times A_4 = ((A_1 \times A_2) \times A_3) \times A_4, \ \text{ etc.}$$

10. Les familles

10.1. **Définition.** Une *famille* est un ensemble F qui satisfait les conditions suivantes :

- chaque élément de F est une paire ordonnée
- $\forall_{x,y_1,y_2}\ \big([\,(x,y_1)\in F\ \wedge\ (x,y_2)\in F\,]\implies y_1=y_2\big)$

10.2. **Exemple.** Définissons les ensembles $0,1,2$ par

$$0=\varnothing,\quad 1=\{\varnothing\},\quad 2=\big\{\varnothing,\{\varnothing\}\big\}$$

et remarquons que $0\neq 1$, $0\neq 2$ et $1\neq 2$. Soient u,v des ensembles quelconques. Alors l'ensemble

$$F=\{(0,v),(1,u),(2,v)\}$$

est une famille. Notez que les projections de F sont $F_1=\{0,1,2\}$ et $F_2=\{u,v\}$. Remarquez aussi que si $u\neq v$ alors l'ensemble $F'=\{(0,u),(0,v),(1,u),(2,v)\}$ n'est pas une famille.

10.3. **Vocabulaire et notations.** Supposons que F est une famille. Notons I la première projection de F (donc $I=F_1$, voir 9.7). Alors on dit que I est l'ensemble des indices et que F est une famille indicée par I. La définition de famille implique que, pour chaque $i\in I$, il existe un et un seul ensemble y qui satisfait $(i,y)\in F$; si cet ensemble y est noté A_i on écrit alors $F=\big(A_i\big)_{i\in I}$.

Attention : N'oubliez pas qu'une famille est toujours un ensemble de paires ordonnées (voir déf. 10.1) même lorsqu'on utilise la notation $\big(A_i\big)_{i\in I}$ pour la représenter.

10.4. **Exemple.** Soient $I=\{0,1,2\}$, $A_0=v$, $A_1=u$, $A_2=v$. Alors $\big(A_i\big)_{i\in I}$ est la famille de l'exemple 10.2, c'est à dire que

$$\big(A_i\big)_{i\in I}=\{(0,v),(1,u),(2,v)\}.$$

Donc $F=\big(A_i\big)_{i\in I}$ et $F=\{(0,v),(1,u),(2,v)\}$ sont deux notations qui désignent le même ensemble F (la même famille F).

10.5. **Notation.** Si $F=\big(A_i\big)_{i\in I}$ est une famille quelconque alors F est un ensemble de paires ordonnées, donc la seconde projection de F existe. On définit l'ensemble $\big\{\,A_i\mid i\in I\,\big\}$ de la manière suivante :

$$\big\{\,A_i\mid i\in I\,\big\}=\text{ la seconde projection de }F.$$

On définit ensuite les ensembles $\displaystyle\bigcup_{i\in I}A_i$ et $\displaystyle\bigcap_{i\in I}A_i$ par :

$$\bigcup_{i\in I}A_i=\cup\big\{\,A_i\mid i\in I\,\big\},\qquad \bigcap_{i\in I}A_i=\cap\big\{\,A_i\mid i\in I\,\big\}$$

(dans le cas de l'intersection, on a besoin de supposer que l'ensemble $\big\{\,A_i\mid i\in I\,\big\}$ soit non vide, ce qui revient à supposer que $F\neq\varnothing$, ou encore que $I\neq\varnothing$). On a donc

pour tout x :

$$x \in \bigcup_{i\in I} A_i \iff \exists_{i\in I}\,(x \in A_i), \qquad x \in \bigcap_{i\in I} A_i \iff \forall_{i\in I}\,(x \in A_i).$$

10.6. **Exemple.** Soit $\big(A_i\big)_{i\in I}$ la famille de l'exemple 10.4. Alors

$$\big\{\, A_i \mid i \in I \,\big\} = \{A_0, A_1, A_2\} = \{v, u, v\} = \{u, v\}$$

donc $\bigcup_{i\in I} A_i = \cup\{u, v\} = u \cup v$ et $\bigcap_{i\in I} A_i = \cap\{u, v\} = u \cap v$.

11. Les fonctions

11.1. **Définition.** Voici deux définitions équivalentes du concept de fonction :

(1) Une *fonction* est un triplet ordonné (A, B, F) satisfaisant :

$$F \subseteq A \times B \quad \wedge \quad \forall_{x\in A}\,\exists!_{y\in B}\,[(x, y) \in F].$$

(2) Une *fonction* est un triplet ordonné (A, B, F) satisfaisant :

$$F \text{ est une famille, } F_1 = A \text{ et } F_2 \subseteq B,$$

où F_i est la i-ième projection de F (voir 9.7).

11.2. **Définition.** Supposons que $f = (A, B, F)$ est une fonction.

(1) A est appelé le *domaine* de f ($\mathrm{dom}(f) = A$).
(2) B est appelé le *codomaine* de f ($\mathrm{codom}(f) = B$).
(3) On dit que f est une *fonction de A vers B* ; on écrit $f : A \to B$ ou $A \xrightarrow{f} B$.
(4) Si $(a, b) \in F$ alors on écrit $f(a) = b$. Autrement dit, si $a \in A$ alors $f(a)$ désigne l'unique élément de B qui satisfait $(a, f(a)) \in F$.
(5) Puisque F est un ensemble de paires ordonnées, on peut considérer les projections F_1 et F_2 de F. La deuxième projection satisfait $F_2 \subseteq B$, mais n'est pas nécessairement égale à B. Le sous-ensemble F_2 de B est appelé *l'image* de f. Autrement dit,

$$\mathrm{im}(f) = F_2 = \big\{\, y \in B \mid \exists_{x\in A}[(x, y) \in F] \,\big\} = \big\{\, y \in B \mid \exists_{x\in A}[f(x) = y] \,\big\}.$$

L'image de f est aussi désignée par $f(A)$.

11.3. **Exercice.** Considérons le singleton $A = \{\varnothing\}$ et soit $f : A \to A$ l'unique fonction de A vers A (donc $f(\varnothing) = \varnothing$). D'après la définition de fonction ci-dessus, on a $f = (A, A, F)$ où $F = \{(\varnothing, \varnothing)\}$. Montrez que $f = \{\{\{\{\varnothing\}\}\}\}$.

11.4. Montrons que si E est un ensemble quelconque, la fonction identité $i_E : E \to E$ existe.

En vertu de 9.8, l'ensemble $D = \big\{\, (x, y) \in E \times E \mid x = y \,\big\}$ existe, donc le triplet (E, E, D) existe aussi. Il est clair que $D \subseteq E \times E$ et vous pouvez vérifier que $\forall_{x\in E}\,\exists!_{y\in E}\,[(x, y) \in D]$, donc le triplet (E, E, D) est une fonction de E vers E. On définit $i_E = (E, E, D)$. Alors pour chaque $x \in E$ on a $(x, x) \in D$, donc $i_E(x) = x$.

En résumé, pour un ensemble E quelconque on a défini une fonction $i_E : E \to E$ qui satisfait $\forall_{x\in E}\ i_E(x) = x$. On l'appelle la *fonction identité* de E.

11.5. Soient $A \xrightarrow{f} B \xrightarrow{g} C$ des fonctions, disons $f = (A, B, F)$ et $g = (B, C, G)$. Montrons que la fonction composée $g \circ f : A \to C$ existe.

En vertu de 9.8, l'ensemble $H = \{\, (x, z) \in A \times C \mid \exists_{y\in B}[(x, y) \in F \ \wedge\ (y, z) \in G] \,\}$ existe, donc le triplet (A, C, H) existe. Il est clair que $H \subseteq A \times C$ et vous pouvez vérifier que $\forall_{x\in A}\, \exists!_{z\in C}\, [(x, z) \in H]$, donc le triplet (A, C, H) est une fonction de A vers C. On définit $g \circ f = (A, C, H)$. Vérifiez que cette fonction $g \circ f : A \to C$ satisfait

$$\forall_{x\in A}\ (g \circ f)(x) = g(f(x)).$$

11.6. **Exercice.** Soient $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ des fonctions. Montrez que

$$(h \circ g) \circ f = h \circ (g \circ f).$$

11.7. **Proposition.** *Étant donnés des ensembles A et B, il existe un ensemble dont les éléments sont précisément toutes les fonctions de A vers B.*

Démonstration. Soient A et B des ensembles. On doit démontrer les deux affirmations suivantes :

(10) il existe un ensemble E satisfaisant : toute fonction de A vers B est un élément de E

(11) la phrase *"f est une fonction de A vers B"* est une condition sur f.

En effet, si (10) et (11) sont vrais alors l'Axiome 2 implique que l'ensemble

$$\{\, f \in E \mid f \text{ est une fonction de } A \text{ vers } B \,\}$$

existe, et il est clair que cet ensemble satisfait la condition demandée. Donc il suffit de prouver (10) et (11).

Vous pouvez vérifier que chaque fonction de A vers B est un élément de

$$\{A\} \times \{B\} \times \wp(A \times B).$$

Donc (10) est vrai. La phrase *"f est une fonction de A vers B"* s'écrit

$$\exists_F \left(f = (A, B, F) \quad \wedge \quad F \subseteq A \times B \quad \wedge \quad \forall_{x\in A}\, \exists!_{y\in B}\, [(x, y) \in F] \right)$$

Puisque A, B ne sont pas des variables (ce sont des ensembles donnés d'avance), f est la seule variable libre de la formule ci-dessus. En vertu de 8.4, cette formule est une condition sur f. Donc (11) est démontré. □

11.8. **Notation.** L'ensemble de toutes les fonctions de A vers B est désigné par B^A. La Proposition 11.7 montre que si A, B sont des ensembles alors B^A existe.

11.9. **Exercice.** Soit B un ensemble quelconque. Montrez que le triplet $f = (\varnothing, B, \varnothing)$ est une fonction de $\varnothing$ vers B, et est la seule fonction de $\varnothing$ vers B. Donc $B^{\varnothing} = \{f\}$ est un singleton. L'unique fonction f de $\varnothing$ vers B est appelée la "fonction vide", même si $f \neq \varnothing$. L'ensemble f possède combien d'éléments ?

Quelques propriétés des fonctions

11.10. **Définition.** Soient A, B des ensembles et $f : A \to B$ une fonction.

(1) Dans le cas particulier où $\mathrm{im}(f) = B$, on dit que f est une fonction *surjective.* Autrement dit, f est surjective si $\forall_{y\in B}\, \exists_{x\in A}\, [f(x) = y]$.
(2) f est *injective* si $\forall_{x_1,x_2\in A}\, (f(x_1) = f(x_2) \implies x_1 = x_2)$.
(3) f est *bijective* si elle est injective et surjective.

11.11. **Exercice.** Montrez que $i_E : E \to E$ est bijective (quel que soit E).

11.12. **Exercice.** Soit $f : A \to B$ une fonction. Montrez que $f \circ i_A = f$ et $i_B \circ f = f$.

Avant de lire la prochaine définition, remarquez que si $A \underset{g}{\overset{f}{\rightleftarrows}} B$ sont des fonctions alors on a $A \xrightarrow{g\circ f} A$ et $B \xrightarrow{f\circ g} B$.

11.13. **Définition.** Une fonction $f : A \to B$ est *inversible* si il existe au moins une fonction $g : B \to A$ satisfaisant

$$g \circ f = i_A \quad \wedge \quad f \circ g = i_B.$$

11.14. **Lemme.** *Si $f : A \to B$ est une fonction inversible, alors il existe exactement une fonction $g : B \to A$ satisfaisant $g \circ f = i_A$ et $f \circ g = i_B$.* (Définition : g est appelée *l'inverse* de f. On écrit parfois $g = f^{-1}$.)

11.15. **Théorème.** *Une fonction est inversible si et seulement si elle est bijective.*

Fonctions et familles

11.16. Si I et B sont des ensembles et $f : I \to B$ est une fonction alors f est un triplet, disons $f = (I, B, F)$, et la composante F de ce triplet est une famille. Ainsi, toute fonction f détermine une famille F. Si on représente F en utilisant la notation des familles, on a $F = \big(f(i)\big)_{i\in I}$.

11.17. Soit $F = \big(A_i\big)_{i\in I}$ une famille quelconque (voir 10.1, 10.3). Si B est n'importe quel ensemble tel que $F_2 \subseteq B$, alors le triplet $f = (I, B, F)$ est une fonction de I vers B. Cette fonction satisfait

$$\forall_{i\in I} \quad f(i) = A_i.$$

Notez que la famille F ne détermine pas une fonction unique : pour chaque choix d'un B tel que $F_2 \subseteq B$, on obtient une fonction de I vers B. Toutes les fonctions ainsi obtenues à partir de F ont le même domaine I et sont définies par la même "règle" (la règle est la famille F).

12. L'axiome de l'infini

En vertu des cinq premiers axiomes, si y est un ensemble quelconque alors l'ensemble $y \cup \{y\}$ existe et est unique. Écrivons $y^+ = y \cup \{y\}$ comme abbréviation. On dit que y^+ est le *successeur* de y. Par exemple, on a

$$\varnothing^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\} \quad \text{et} \quad \{\varnothing\}^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}.$$

12.1. **Exercice.** Montrez que pour tout ensemble y on a $y \in y^+$ et $y \subseteq y^+$.

12.2. **Définition.** On dit qu'un ensemble A est *inductif* si il satisfait les deux conditions

$$\varnothing \in A \quad \text{et} \quad \forall_y (y \in A \Rightarrow y^+ \in A).$$

12.3. **Exercice.** Montrez que si A et B sont des ensembles inductifs alors $A \cap B$ est un ensemble inductif.

12.4. **Lemme.** *Supposons que E est un ensemble non vide qui satisfait :*

Chaque élément de E est un ensemble inductif.

Alors $\cap E$ est un ensemble inductif.

12.5. **Exercice.** Démontrez le Lemme 12.4.

12.6. **Lemme.** *La phrase "x est un ensemble inductif" est une condition sur x.*

Démonstration. La phrase "x est un ensemble inductif" peut s'écrire

$$\varnothing \in x \ \wedge \ \forall_y (y \in x \Rightarrow y^+ \in x). \tag{12}$$

L'expression (12) est une formule (du langage élargi) dont la seule variable libre est x, donc en vertu de 8.4 on conclut que (12) est une condition sur x. □

Axiome 7 (Axiome de l'infini)**.** *Il existe au moins un ensemble inductif.*

12.7. **Théorème.** *Il existe exactement un ensemble ω qui satisfait les deux conditions suivantes :*

- *ω est un ensemble inductif*
- *ω est inclus ($\subseteq$) dans tout ensemble inductif.*

Démonstration. En vertu de l'Axiome 7, il existe un ensemble inductif A. Considérons l'ensemble

$$E = \big\{\, x \in \wp A \mid x \text{ est un ensemble inductif} \,\big\}.$$

Cet ensemble existe, parce que $\wp A$ est un ensemble et "x est un ensemble inductif" est une condition sur x (voir 12.6). Notons que $E \neq \varnothing$ (puisque $A \in E$), et que chaque élément de E est un ensemble inductif ; donc le Lemme 12.4 implique que $\cap E$ est un ensemble inductif. On définit

$$\omega = \cap E.$$

Il reste à montrer que ω est inclus dans tout ensemble inductif. Soit B un ensemble inductif quelconque. Alors $A \cap B$ est un ensemble inductif (Exercice 12.3) et $A \cap B \in$

$\wp A$, donc $A \cap B \in E$. Puisque $\cap E \subseteq C$ pour tout $C \in E$, on a $\cap E \subseteq A \cap B \subseteq B$, donc $\omega \subseteq B$.

On a montré l'existence d'au moins un ensemble inductif ω qui est inclus dans tout ensemble inductif. Un tel ensemble est forcément unique. En effet, supposons que ω' est aussi un ensemble inductif inclus dans tout ensemble inductif. Alors on a $\omega \subseteq \omega'$ et $\omega' \subseteq \omega$, donc $\omega = \omega'$. □

12.8. **Définition.** Les éléments de ω sont appelés les *nombres naturels.*

En particulier, on définit les nombres naturels $0, 1, 2, 3, 4, 5$ de la manière suivante :

$$0 = \varnothing$$
$$1 = 0^+ = 0 \cup \{0\} = \{0\} = \{\varnothing\}$$
$$2 = 1^+ = 1 \cup \{1\} = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$$
$$3 = 2^+ = 2 \cup \{2\} = \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$
$$4 = 3^+,\ 5 = 4^+.$$

12.9. **Exercice.** Prouvez que $2 \neq 3$.

Le théorème suivant résume les propriétés les plus importantes des nombres naturels.

12.10. **Théorème** (Axiomes de Peano). *L'ensemble ω, l'élément 0 et l'opération "$^+$" ont les propriétés suivantes.*

(I) $0 \in \omega$
(II) *si* $n \in \omega$ *alors* $n^+ \in \omega$
(III) *il n'existe aucun* $n \in \omega$ *qui satisfait* $n^+ = 0$
(IV) *si* $m, n \in \omega$ *satisfont* $m^+ = n^+$ *alors* $m = n$
(V) *si* S *est une partie de* ω *satisfaisant* $0 \in S$ *et* $\forall_n (n \in S \Rightarrow n^+ \in S)$, *alors* $S = \omega$.

Démonstration de (I), (II), (III) *et* (V). Puisque ω est un ensemble inductif, (I) et (II) sont vrais. S'il existe un ensemble n tel que $n^+ = 0$ alors $n \cup \{n\} = 0 = \varnothing$, donc $n \in \varnothing$, contradiction (donc (III) est vrai). Puisque ω est inclus dans tout ensemble inductif, (V) est vrai (car le S dans (V) est un ensemble inductif). □

La Proposition 12.16 ci-dessous fournira la preuve de l'assertion (IV), et complètera donc la démonstration des Axiomes de Peano. L'assertion (V) est appelée le principe d'induction.

12.11. **Définition.** On dit qu'un ensemble A est *transitif* si il satisfait la condition

$$\forall_{x,y} (x \in y \in A \Rightarrow x \in A).$$

12.12. **Exercice.** Vérifiez que la phrase "x est un ensemble transitif" est une condition sur x.

12.13. **Exercice.** Montrez qu'un ensemble A est transitif si et seulement si il satisfait la condition $\forall_y(y \in A \Rightarrow y \subseteq A)$.

12.14. **Exercice.** Montrez que si s est un ensemble transitif alors s^+ est un ensemble transitif.

12.15. **Proposition.** *Chaque élément de ω est un ensemble transitif.*

Démonstration. Puisque ω est un ensemble et (voir Ex. 12.12) la phrase "x est un ensemble transitif" est une condition sur x, l'axiome de spécification implique que l'ensemble

$$S = \big\{\, x \in \omega \mid x \text{ est un ensemble transitif} \,\big\}$$

existe. Il est clair que $S \subseteq \omega$, et pour démontrer la proposition on doit prouver que $\omega \subseteq S$. Donc il suffit de montrer que S est un ensemble inductif (car ω est inclus dans tout ensemble inductif).

- Puisque $\varnothing \in \omega$ et $\varnothing$ est un ensemble transitif, on a $\varnothing \in S$.
- Soit $s \in S$. Puisque $s^+ \in \omega$ (car $s \in \omega$ et ω est un ensemble inductif) et puisque s^+ est transitif (voir Ex. 12.14), on a $s^+ \in S$. Ainsi, $\forall_s(s \in S \Rightarrow s^+ \in S)$.

Les deux points ci-dessus montrent que S est un ensemble inductif, ce qui termine la démonstration. □

12.16. **Proposition.** *Si $x, y \in \omega$ satisfont $x^+ = y^+$, alors $x = y$.*

Démonstration. Soient $x, y \in \omega$.

Supposons que $x^+ \subseteq y^+$. Alors $x \in x^+ = y^+ = y \cup \{y\}$, donc $x \in y$ ou $x = y$. On déduit que $x \subseteq y$ par séparation des cas : si $x \in y$ alors $x \subseteq y$ puisque y est transitif (cf. 12.15, 12.13) ; et si $x = y$ alors $x \subseteq y$. Donc $x \subseteq y$.

On a montré que $x^+ \subseteq y^+$ implique $x \subseteq y$. Donc

$$x^+ = y^+ \implies (x^+ \subseteq y^+ \ \wedge\ y^+ \subseteq x^+) \implies (x \subseteq y \ \wedge\ y \subseteq x) \implies x = y.$$

□

La démonstration du Théorème 12.10 est maintenant complète.

Pour conclure, on remarque que l'opération successeur $(x \mapsto x^+)$ est en fait une fonction de ω vers ω. En effet, l'ensemble $S = \big\{\, (x, y) \in \omega \times \omega \mid y = x^+ \,\big\}$ existe, donc l'ensemble $s = (\omega, \omega, S)$ existe ; ce triplet est une fonction $s : \omega \to \omega$ et $s(x) = x^+$ pour tout $x \in \omega$. Remarquez aussi que s est injective, en vertu de 12.16 (ou de 12.10(IV)). Donc :

12.17. **Proposition.** *La fonction $s : \omega \to \omega$, $s(x) = x^+$, existe et est injective.*

13. La relation d'ordre dans ω

Nous allons démontrer plusieurs propriétés de l'ensemble ω des nombres naturels.

13.1. **Lemme.** *Quel que soit $n \in \omega$, il n'existe aucun x tel que $n \subseteq x \in n$.*

Démonstration. Il faut montrer que l'ensemble $S = \{\, n \in \omega \mid \nexists_x (n \subseteq x \in n) \,\}$ est égal à ω. En vertu du principe d'induction (la partie (V) de 12.10) il suffit de montrer que

(1) $0 \in S$ (rappel : $0 = \varnothing$)
(2) $\forall_n (n \in S \Rightarrow n^+ \in S)$

L'assertion (1) est claire (si $0 \notin S$ alors il existe x tel que $0 \subseteq x \in 0 = \varnothing$, ce qui est faux). Montrons (2) par contradiction : supposons que $n \in \omega$ satisfait

$$n \in S \ \wedge \ n^+ \notin S.$$

Puisque $n^+ \notin S$, il existe x tel que

$$n^+ \subseteq x \in n^+.$$

On a $x \in n^+ = n \cup \{n\}$, donc $x \in n$ ou $x = n$.

Si $x \in n$ alors $n \subseteq n^+ \subseteq x \in n$ contredit $n \in S$.

Si $x = n$ alors $n \in n^+ \subseteq x = n$ implique $n \in n$, donc $n \subseteq n \in n$, et ceci contredit $n \in S$.

Donc dans les deux cas on a une contradiction. Ceci démontre (2) et complète la démonstration du lemme. □

13.2. **Corollaire.** *Chaque $n \in \omega$ satisfait $n \notin n$.*

Démonstration. Si $n \in n$ alors avec $x = n$ on a $n \subseteq x \in n$, ce qui contredit 13.1. □

13.3. **Définition.** On définit un sous-ensemble $<$ de $\omega \times \omega$ par

$$< \,= \{\, (m, n) \in \omega \times \omega \mid m \in n \,\}.$$

Autrement dit, on définit la relation binaire $<$ sur l'ensemble ω en stipulant que

$$\forall_{m,n \in \omega} \quad m < n \Leftrightarrow m \in n.$$

13.4. **Lemme.** *La relation $<$ est transitive et satisfait $\nexists_{x,y \in \omega} (x < y \ \wedge \ y < x)$.*

Démonstration. Supposons que $x, y, z \in \omega$ satisfont $x < y$ et $y < z$. Alors $x \in y \in z$ et (voir 12.15) z est un ensemble transitif, donc $x \in z$, donc $x < z$. Ainsi, $<$ est transitive.

Supposons qu'il existe $x, y \in \omega$ satisfaisant $x < y$ et $y < x$. Alors $x < x$ par transitivité, donc $x \in x$, ce qui contredit 13.2. Cette contradiction montre que $\nexists_{x,y \in \omega} (x < y \ \wedge \ y < x)$. □

13.5. **Définition.** On définit la relation binaire $\leq$ sur ω par

$$x \leq y \quad \Leftrightarrow \quad (x < y \ \vee \ x = y).$$

Le lemme 13.4 implique que $\leq$ est un ordre partiel sur ω :

- $\forall_{x \in \omega} \ (x \leq x)$
- $\forall_{x,y,z \in \omega} \big((x \leq y \ \wedge \ y \leq z) \Rightarrow x \leq z\big)$
- $\forall_{x,y \in \omega} \big((x \leq y \ \wedge \ y \leq x) \Rightarrow x = y\big)$

Nous allons montrer que $(\omega, \leq)$ est bien ordonné ; cependant la première étape est de montrer qu'il est totalement ordonné.

On définit les symboles $>$ et $\geq$ de la manière habituelle : $x > y$ signifie $y < x$ et $x \geq y$ signifie $y \leq x$.

13.6. **Lemme.** *Chaque $n \in \omega$ satisfait $n < n^+$ et $n \geq 0$.*

Proof. Quel que soit $n \in \omega$, on a $n \in n^+$, donc $n < n^+$ par définition de $<$.

Soit $S = \{ n \in \omega \mid n \geq 0 \}$. Alors $0 \in S$ est évident. De plus, si $n \in S$ alors $0 \leq n < n^+$ donc $n^+ \geq 0$, donc $n^+ \in S$. En vertu de la partie (V) de 12.10, $S = \omega$. □

Remarque. Le Lemme 13.6 implique que 0 est élément minimum de $(\omega, \leq)$.

13.7. **Lemme.** $\forall_{m,n\in\omega}\ (m < n^+ \iff m \leq n)$.

Démonstration. Soient $m, n \in \omega$. Alors

$$m < n^+ \iff m \in n^+ = n \cup \{n\} \iff (m \in n \text{ ou } m = n) \iff (m < n \text{ ou } m = n).$$

□

13.8. **Proposition.** $(\omega, \leq)$ *est totalement ordonné.*

Proof. Pour chaque $n \in \omega$, on définit l'ensemble

$$S(n) = \{ x \in \omega \mid n < x \ \vee \ n = x \ \vee \ n > x \}.$$

Pour démontrer la proposition, on doit montrer que $\forall_{n\in\omega}\,(S(n) = \omega)$. Autrement dit il faut prouver que $\mathcal{S} = \omega$, où on définit $\mathcal{S} = \{ n \in \omega \mid S(n) = \omega \}$. En vertu du principe d'induction (12.10.V), pour prouver que $\mathcal{S} = \omega$ il suffit de prouver les deux assertions suivantes :

(1) $0 \in \mathcal{S}$
(2) $\forall_n\,(n \in \mathcal{S} \Rightarrow\ n^+ \in \mathcal{S})$

Preuve de (1). Soit $m \in \omega$. Alors 13.6 implique que $m \geq 0$, donc

$$0 < m \ \vee \ 0 = m \ \vee \ 0 > m,$$

donc $m \in S(0)$. Ceci montre que $S(0) = \omega$, donc $0 \in \mathcal{S}$.

Preuve de (2). Supposons que $n \in \mathcal{S}$ (donc $n \in \omega$ est tel que $S(n) = \omega$). On doit montrer que $n^+ \in \mathcal{S}$; autrement dit, on doit prouver que $S(n^+) = \omega$. Pour montrer que $S(n^+) = \omega$, il suffit de vérifier que

(2.1) $0 \in S(n^+)$
(2.2) $\forall_m\,(m \in S(n^+)\ \Rightarrow\ m^+ \in S(n^+))$.

En résumé, la preuve de la Proposition sera terminée aussitôt qu'on aura prouvé (2.1) et (2.2).

Le lemme 13.6 implique que $n^+ \geq 0$, donc (2.1) est clair. Pour prouver (2.2) on considère $m \in S(n^+)$ et on doit montrer que $m^+ \in S(n^+)$. Puisque $m \in S(n^+)$,

$$n^+ \leq m \ \vee \ n^+ > m$$

et puisque $m^+ \in \omega = S(n)$,

$$n < m^+ \ \vee \ n \geq m^+.$$

On a donc

$$\text{(i) } n^+ \leq m \quad \text{ou} \quad \text{(ii) } n \geq m^+ \quad \text{ou} \quad \text{(iii) } (n^+ > m \ \wedge \ n < m^+).$$

- Si (i) est satisfaite alors $n^+ \leq m < m^+$, donc $n^+ < m^+$, donc $m^+ \in S(n^+)$.
- Si (ii) est satisfaite alors $n^+ > n \geq m^+$, donc $n^+ > m^+$, donc $m^+ \in S(n^+)$.
- Si (iii) est satisfaite alors $m < n^+$, donc (voir 13.7) $m \leq n$; et $n < m^+$, donc $n \leq m$. Ainsi, $m = n$ et conséquemment $m^+ = n^+$, donc $m^+ \in S(n^+)$.

Donc $m^+ \in S(n^+)$, par séparation des cas. Ceci prouve (2.2) et complète la démonstration de la proposition. □

13.9. **Théorème.** $(\omega, \leq)$ *est bien ordonné.*

Démonstration. Pour chaque $n \in \omega$, on définit $(-\infty, n] = \{\, x \in \omega \mid x \leq n \,\}$. Nous dirons (seulement dans cette preuve) que le nombre naturel n est "bon" si :

toute partie A de ω qui satisfait $(-\infty, n] \cap A \neq \varnothing$ possède un élément minimum.

Pour démontrer le théorème, il suffit de montrer que chaque élément de ω est bon. En effet, si c'est le cas et si $\varnothing \neq A \subseteq \omega$ alors on peut choisir $n \in A$, alors n est bon et $(-\infty, n] \cap A \neq \varnothing$, donc A possède un élément minimum.

Montrons que chaque élément de ω est bon. Par induction (12.10.V), il suffit de montrer que 0 est bon et que $\forall_{n \in \omega}\, (n \text{ bon} \Rightarrow n^+ \text{ bon})$.

Le Lemme 13.6 implique que 0 est élément minimum de $(\omega, \leq)$. Il s'ensuit que 0 est bon. En effet, $(-\infty, 0] = \{0\}$, donc $(-\infty, 0] \cap A \neq \varnothing$ implique que $0 \in A$, qui implique que 0 est élément minimum de A.

Supposons que $n \in \omega$ est bon, et montrons que n^+ est bon. Soit A une partie de ω telle que $(-\infty, n^+] \cap A \neq \varnothing$. Notons que

$$(-\infty, n^+] \cap A = \{n^+\} \quad \vee \quad (-\infty, n^+] \cap A \neq \{n^+\}.$$

Montrons que, dans les deux cas, A possède un élément minimum.

- Si $(-\infty, n^+] \cap A = \{n^+\}$ alors n^+ est l'élément minimum de A. En effet, soit $x \in A$. Puisque $(\omega, \leq)$ est totalement ordonné,

  $$n^+ \geq x \quad \vee \quad n^+ < x.$$

  Si $n^+ \geq x$ alors $x \in (-\infty, n^+] \cap A = \{n^+\}$, donc $x = n^+$, donc $n^+ \leq x$; si $n^+ < x$ alors $n^+ \leq x$. Donc $\forall_{x \in A}\, (n^+ \leq x)$.
- Si $(-\infty, n^+] \cap A \neq \{n^+\}$ alors (puisque $(-\infty, n^+] \cap A \neq \varnothing$) il existe $m \in (-\infty, n^+] \cap A$ tel que $m \neq n^+$. Puisque $m \in (-\infty, n^+]$ et $m \neq n^+$, on a $m < n^+$, donc (13.7) $m \leq n$, donc $m \in (-\infty, n] \cap A$. Puisque $(-\infty, n] \cap A \neq \varnothing$ et n est bon, A possède un élément minimum.

Ainsi, A possède un élément minimum. Ceci montre que n^+ est bon, ce qui termine la preuve (par induction) que chaque élément de ω est bon. Donc $(\omega, \leq)$ est bien ordonné. □

13.10. **Corollaire.** *Quels que soient $m, n \in \omega$, on a $m < n \Leftrightarrow m \subset n$.*

Démonstration. Rappelons que si $n \in \omega$ alors (12.15) n est transitif, donc (12.13) tout élément de n est un sous-ensemble de n. Ceci implique que si $m, n \in \omega$ satisfont $m < n$ alors $m < n \Rightarrow m \in n \Rightarrow m \subseteq n$. En fait on a $m \subset n$, car on a $m < n$ (donc $m \neq n$) par hypothèse. Donc $m < n \implies m \subset n$.

Réciproquement, supposons que $m, n \in \omega$ satisfont $m \subset n$. En vertu de 13.8 on a

$$m < n \ \vee \ m = n \ \vee \ m > n.$$

Or, $m = n$ est impossible puisque $m \subset n$, et $m > n$ est impossible puisque le premier paragraphe montre que cette condition implique $m \supset n$, qui est incompatible avec l'hypothèse $m \subset n$. La condition $m < n$ doit donc être satisfaite. □

14. Un peu d'arithmétique

Supposons que A est un ensemble, $\alpha : A \to A$ est une fonction et $a_0 \in A$. Alors on aimerait définir une fonction $f : \omega \to A$ qui satisfait :

$$f(0) = a_0, \quad f(1) = \alpha(a_0), \quad f(2) = \alpha(\alpha(a_0)), \quad f(3) = \alpha(\alpha(\alpha(a_0))), \quad \text{etc.}$$

Cette façon de définir une fonction s'appelle une **definition récursive**, parce que $f(n^+)$ est défini en termes de $f(n)$:

$$f(0) = a_0, \quad f(1) = \alpha(f(0)), \quad f(2) = \alpha(f(1)), \quad f(3) = \alpha(f(2)), \text{ etc.}$$

Le théorème suivant affirme que cette fonction f existe et est unique, donc il sera dorénavant permis de définir une fonction de manière récursive. Ce résultat porte le nom de Théorème de récursion. Il existe aussi un théorème de récursion transfinie, plus général que celui-ci, mais nous n'en parlerons pas.

14.1. **Théorème.** *Supposons que A est un ensemble, $\alpha : A \to A$ est une fonction et $a_0 \in A$. Alors il existe exactement une fonction $f : \omega \to A$ satisfaisant*

$$f(0) = a_0 \quad \textit{et} \quad \forall_{n \in \omega} \, f(n^+) = \alpha(f(n)).$$

Pour la démonstration, voir p. 48 de *Naive Set Theory*, de P. Halmos.

Nous allons maintenant appliquer ce théorème au cas où $A = \omega$ et $\alpha : \omega \to \omega$ est la fonction successeur, c'est à dire que $\alpha(n) = n^+$ pour tout $n \in \omega$ (on a vu en 12.17 que $n \mapsto n^+$ est effectivement une fonction de ω vers ω).

Soit $m \in \omega$. En vertu du Théorème de récursion 14.1, il existe une et une seule fonction $s_m : \omega \to \omega$ satisfaisant

$$s_m(0) = m \quad \text{et} \quad \forall_{n \in \omega} \, s_m(n^+) = (s_m(n))^+.$$

Nous venons de définir une infinité de fonctions : nous avons $s_m : \omega \to \omega$ pour chaque $m \in \omega$. Ces fonctions satisfont

$$\forall_{x\in\omega}\ s_x(0) = x \tag{13}$$

$$\forall_{x,y\in\omega}\ s_x(y^+) = s_x(y)^+. \tag{14}$$

14.2. **Définition.** Étant donnés $m, n \in \omega$, on définit $m + n = s_m(n)$.

Pour le lemme suivant, rappelons que la définition de 1 est $1 = 0^+$.

14.3. **Lemme.** *Pour chaque $n \in \omega$, on a $n + 0 = n$ et $n + 1 = n^+$.*

Démonstration. Soit $n \in \omega$, alors $n + 0 \overset{14.2}{=} s_n(0) \overset{(13)}{=} n$ et $n + 1 \overset{14.2}{=} s_n(1) = s_n(0^+) \overset{(14)}{=} s_n(0)^+ \overset{(13)}{=} n^+$. □

14.4. **Lemme.** $\forall_{k,m,n\in\omega}\ \ (k + m) + n = k + (m + n)$

Démonstration. Soit $S = \{\, n \in \omega \mid \forall_{k,m\in\omega}\ \ (k + m) + n = k + (m + n) \,\}$. Alors il faut montrer que $S = \omega$, donc il suffit de montrer que $0 \in S$ et que si $n \in S$ alors $n^+ \in S$.

Quels que soient $k, m \in \omega$, on a

$$(k + m) + 0 \overset{14.2}{=} s_{k+m}(0) \overset{(13)}{=} k + m \overset{(13)}{=} k + s_m(0) \overset{14.2}{=} k + (m + 0),$$

donc $0 \in S$. Supposons que $n \in S$, donc n satisfait :

$$\forall_{k,m\in\omega}\ \ (k + m) + n = k + (m + n). \tag{15}$$

Soient $k, m \in \omega$, alors

$$(k + m) + n^+ \overset{14.2}{=} s_{k+m}(n^+) \overset{(14)}{=} \big(s_{k+m}(n)\big)^+ \overset{14.2}{=} \big((k + m) + n\big)^+ \overset{(15)}{=} \big(k + (m + n)\big)^+$$
$$\overset{14.2}{=} s_k(m + n)^+ \overset{(14)}{=} s_k((m + n)^+) \overset{14.2}{=} s_k(s_m(n)^+) \overset{(14)}{=} s_k(s_m(n^+)) \overset{14.2}{=} k + (m + n^+)$$

et on a montré que $\forall_{k,m\in\omega}\ \ (k + m) + n^+ = k + (m + n^+)$. Donc $n^+ \in S$ et la preuve est terminée. □

14.5. **Lemme** (Loi de la simplification). $\forall_{x_1,x_2,y\in\omega}\,[x_1 + y = x_2 + y \ \Rightarrow\ x_1 = x_2]$

Démonstration. On doit montrer que $S = \omega$, où on définit :

$$S = \{\, y \in \omega \mid \forall_{x_1,x_2\in\omega}\,[x_1 + y = x_2 + y \ \Rightarrow\ x_1 = x_2] \,\}.$$

Donc il suffit de montrer que (i) $0 \in S$ et (ii) $\forall_y\,(y \in S \Rightarrow y^+ \in S)$.

(i) Il faut montrer que $\forall_{x_1,x_2\in\omega}\,[x_1 + 0 = x_2 + 0 \ \Rightarrow\ x_1 = x_2]$. Soient $x_1, x_2 \in \omega$ tels que $x_1 + 0 = x_2 + 0$; en vertu de 14.3, $x_1 = x_1 + 0 = x_2 + 0 = x_2$. Donc (i) est clair.

(ii) Supposons que $y \in S$, donc : $\forall_{x_1,x_2\in\omega}\,[x_1 + y = x_2 + y \ \Rightarrow\ x_1 = x_2]$. On doit montrer que $\forall_{x_1,x_2\in\omega}\,[x_1 + y^+ = x_2 + y^+ \ \Rightarrow\ x_1 = x_2]$. Soient $x_1, x_2 \in \omega$ tels que $x_1 + y^+ = x_2 + y^+$, alors

$$x_1 + y^+ = s_{x_1}(y^+) = s_{x_1}(y)^+ = (x_1 + y)^+$$

et similairement $x_2 + y^+ = (x_2 + y)^+$. Donc $(x_1 + y)^+ = (x_2 + y)^+$ donc (en vertu de 12.16 ou de 12.10(IV)) $x_1 + y = x_2 + y$. Puisque $y \in S$ on conclut que $x_1 = x_2$. Ceci démontre (ii), et complète la démonstration de la loi de la simplification. □

Si on voulait continuer la théorie des nombres naturels, il faudrait maintenant :

- montrer que l'addition est commutative
- montrer que $\forall_{x_1,x_2,y\in\omega}\,[x_1 < x_2 \;\Rightarrow\; x_1 + y < x_2 + y]$
- définir la multiplication et démontrer ses propriétés
- etc.

15. Bref aperçu de la construction des ensembles de nombres $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$

À partir de maintenant on écrira $\mathbb{N} = \omega$ et on supposera qu'on a déjà démontré toutes les propriétés élémentaires de $\mathbb{N}$ (les propriétés de l'addition, de la multiplication et de la relation d'ordre). Nous allons indiquer (très brièvement) comment on peut construire $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$. Tout ceci peut se faire en se basant sur les Axiomes 1 à 7.

De $\mathbb{N}$ à $\mathbb{Z}$

On définit la relation $\sim$ sur l'ensemble $\mathbb{N} \times \mathbb{N}$ par

$$(a, b) \sim (a', b') \quad \Longleftrightarrow \quad (a + b' = b + a' \text{ dans } \mathbb{N}).$$

15.1. **Exercice.** Montrez que $\sim$ est une relation d'équivalence sur l'ensemble $\mathbb{N} \times \mathbb{N}$ (pour la transitivité, on a besoin d'utiliser la loi de la simplification 14.5).

On définit

$$\mathbb{Z} \;=\; \text{l'ensemble des classes d'équivalence} \;\; (\text{pour la relation } \sim \text{ sur } \mathbb{N} \times \mathbb{N}).$$

Disons que $[a, b]$ désigne la classe d'équivalence de $(a, b) \in \mathbb{N} \times \mathbb{N}$.

15.2. **Lemme.** *Supposons que $X, Y \in \mathbb{Z}$. Si $(a, b), (a', b') \in X$ et $(u, v), (u', v') \in Y$ alors alors $(a + u, b + v) \sim (a' + u', b' + v')$.*

On définit l'addition dans $\mathbb{Z}$ en se servant de l'addition dans $\mathbb{N}$. Soient $X, Y \in \mathbb{Z}$. Pour additionner X et Y,

- choisir un élément quelconque (a, b) de X
- choisir un élément quelconque (u, v) de Y
- définir $X + Y = [a + u, b + v]$.

En vertu de 15.2, $X + Y$ ne dépend pas des choix de (a, b) et (u, v). Voici une manière plus compacte de décrire l'addition dans $\mathbb{Z}$:

$$\boxed{[a, b] + [u, v] \;=\; [a + u, b + v]}$$

15.3. **Lemme.** (1) *L'addition dans $\mathbb{Z}$ est associative et commutative.*
(2) $[0, 0]$ *est élément neutre de l'addition dans $\mathbb{Z}$.*

(3) *Pour tout* $[a,b] \in \mathbb{Z}$, *on a* $[a,b]+[b,a]=[0,0]$.
Autrement dit, $(\mathbb{Z},+)$ *est un groupe abélien.*

On définit une relation $\leq$ dans $\mathbb{Z}$ par

$$[a,b] \leq [u,v] \quad \Longleftrightarrow \quad (a+v \leq b+u \text{ dans } \mathbb{N}),$$

et une multiplication dans $\mathbb{Z}$ par

$$[a,b]\cdot[u,v] = [au+bv, av+bu].$$

Ici il faudrait vérifier que ces deux définitions ont un sens, c'est à dire qu'il faudrait démontrer deux lemmes semblables à 15.2. Il faudrait aussi démontrer les propriétés de la multiplication et de $\leq$, par exemple le fait que $\leq$ est un ordre total sur $\mathbb{Z}$, et le fait que si $x,y \in \mathbb{Z}$ satisfont $xy=0$ alors $x=0$ ou $y=0$.

On aimerait que $\mathbb{N} \subseteq \mathbb{Z}$, mais ce n'est malheureusement pas le cas. Pour résoudre cette difficulté on considère la fonction injective $f : \mathbb{N} \to \mathbb{Z}$ définie par $f(n)=[n,0]$ (vérifiez que f est injective). Si l'image de f est notée $\mathbb{N}'$ alors $\mathbb{N}'$ est une copie de $\mathbb{N}$ qui satisfait $\mathbb{N}' \subset \mathbb{Z}$.

$$\mathbb{N}' = \{[0,0],[1,0],[2,0],[3,0],\dots\}.$$

La restriction à $\mathbb{N}'$ de l'addition de $\mathbb{Z}$ est une addition dans $\mathbb{N}'$. En fait, on peut montrer que la fonction $n \mapsto [n,0]$ est un isomorphisme de $\mathbb{N}$ vers $\mathbb{N}'$, et il s'ensuit que $\mathbb{N}'$ a exactement les mêmes propriétés que $\mathbb{N}$. À partir de maintenant on considérera que $\mathbb{N}'$ est l'ensemble des nombres naturels. Autrement dit on redéfinit le symbole $\mathbb{N}$ en posant $\mathbb{N} = \{[0,0],[1,0],[2,0],[3,0],\dots\}$. On a alors $\mathbb{N} \subset \mathbb{Z}$.

Autrement dit, le nombre "5" est en fait $[5,0] \in \mathbb{Z}$, et le nombre "-5" est $[0,5] \in \mathbb{Z}$.

De $\mathbb{Z}$ à $\mathbb{Q}$

À partir de maintenant on suppose qu'on a déjà démontré toutes les propriétés élémentaires de l'addition, de la multiplication et de la relation d'ordre de $\mathbb{Z}$. Soit $E = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. On définit une relation $\sim$ sur l'ensemble E par

$$(a,b) \sim (u,v) \quad \Longleftrightarrow \quad (av = bu \text{ dans } \mathbb{Z}).$$

15.4. **Exercice.** Montrez que $\sim$ est une relation d'équivalence sur E. De quelles propriétés de $\mathbb{Z}$ avez-vous besoin pour prouver la transitivité ?

On définit

$$\mathbb{Q} \;=\; \text{l'ensemble des classes d'équivalence (pour la relation } \sim \text{ sur } E).$$

Disons que $\frac{a}{b}$ désigne la classe d'équivalence de $(a,b) \in E$. Par exemple, on a $(4,6) \sim (2,3)$, donc $\frac{4}{6} = \frac{2}{3}$.

L'addition, la multiplication et la relation d'ordre dans $\mathbb{Q}$ sont définies par :

$$\frac{a}{b}+\frac{u}{v}=\frac{av+bu}{bv}$$
$$\frac{a}{b}\cdot\frac{u}{v}=\frac{au}{bv}$$
$$\frac{a}{b}\leq\frac{u}{v}\quad\Longleftrightarrow\quad\begin{cases}av\leq bu & \text{si } bv>0\\ av\geq bu & \text{si } bv<0\end{cases}$$

Comme toujours, il faudrait vérifier que les définitions ont un sens, et démontrer que ces opérations et la relation d'ordre ont les propriétés qu'on s'attend qu'elles aient.

De $\mathbb{Q}$ à $\mathbb{R}$

On suppose que les propriétés élémentaires de $\mathbb{Q}$ (concernant l'addition, la multiplication et la relation d'ordre) ont été démontrées. En particulier, on sait que $(\mathbb{Q},\leq)$ est un ensemble totalement ordonné. On peut maintenant construire les nombres réels à partir de $\mathbb{Q}$.

15.5**. Définition.** Une *coupure de Dedekind* est une paire ordonnée (A,B) satisfaisant :

(1) A et B sont des sous-ensembles de $\mathbb{Q}$
(2) $A\neq\varnothing$ et $B\neq\varnothing$
(3) $A\cap B=\varnothing$
(4) $A\cup B=\mathbb{Q}$
(5) $\forall_{a\in A}\forall_{b\in B}\ \ a<b$
(6) A n'a pas délément maximum (relativement à la relation d'ordre de $\mathbb{Q}$).

15.6**. Exemple.** Si on définit

$$A_0=\left\{\,x\in\mathbb{Q}\mid x^2<2\ \vee\ x<0\,\right\}\quad\text{et}\quad B_0=\left\{\,x\in\mathbb{Q}\mid x^2>2\ \wedge\ x>0\,\right\}$$

alors la paire ordonnée (A_0,B_0) est une coupure de Dedekind.

Remarquons que l'ensemble $\wp\mathbb{Q}\times\wp\mathbb{Q}$ existe (puisque $\mathbb{Q}$ existe) et que chaque coupure de Dedekind est un élément de $\wp\mathbb{Q}\times\wp\mathbb{Q}$. Remarquons aussi que la phrase "x est une coupure de Dedekind" est une condition sur x. Donc l'Axiome de spécification implique que l'ensemble suivant existe :

$$\mathbb{R}\stackrel{\text{def}}{=}\left\{\,x\in\wp\mathbb{Q}\times\wp\mathbb{Q}\mid x\text{ est une coupure de Dedekind}\,\right\}.$$

Donc, par définition, $\mathbb{R}$ est l'ensemble de toutes les coupures de Dedekind. Toujours par définition, un nombre réel est une coupure de Dedekind.

- On définit une relation $\leq$ sur $\mathbb{R}$ de la manière suivante :

  étant donnés $(A,B),(A',B')\in\mathbb{R}$, $\quad(A,B)\leq(A',B')\ \Longleftrightarrow\ A\subseteq A'$.

  On peut alors montrer que $(\mathbb{R},\leq)$ est totalement ordonné.

- Pour chaque $q \in \mathbb{Q}$ on définit
$$f(q) = \big(\{\, x \in \mathbb{Q} \mid x < q \,\}, \{\, x \in \mathbb{Q} \mid x \geq q \,\}\big) \in \mathbb{R}.$$
Alors $f : \mathbb{Q} \to \mathbb{R}$ est une fonction injective. Le nombre réel $f(q)$ est identifié au nombre rationnel q ; autrement dit, on considère que le sous-ensemble $f(\mathbb{Q})$ de $\mathbb{R}$ est l'ensemble des nombres rationnels. Par exemple, soit $r = (A_0, B_0) \in \mathbb{R}$ la coupure de Dedekind définie dans 15.6. Alors $r \notin f(\mathbb{Q})$, donc r est un nombre irrationnel (en fait, r est le nombre $\sqrt{2}$).
- On définit l'addition dans $\mathbb{R}$ de la manière suivante : si $(A, B), (A', B') \in \mathbb{R}$ alors
$$(A, B) + (A', B') = (C, D)$$
où on définit $C = \{\, a + a' \mid a \in A \text{ et } a' \in A' \,\}$ et $D = \mathbb{Q} \setminus C$. On peut vérifier que (C, D) est effectivement une coupure de Dedekind, donc un nombre réel. On peut montrer que cette addition est commutative et associative, etc. Le nombre réel $f(0)$ est l'élément neutre de l'addition.
- Nous omettons la définition de la multiplication dans $\mathbb{R}$ (c'est un peu compliqué).

Il faudrait ensuite démontrer que $(\mathbb{R}, +, \cdot)$ est un corps, que $f(\mathbb{Q})$ est dense dans $\mathbb{R}$, que $\mathbb{R}$ est complet, et plusieurs autres propriétés.

À partir d'ici, on suppose que les ensembles $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ ont été définis et que leurs propriétés de base ont été établies. Remarquez que tout cela peut être fait en n'utilisant rien de plus que les 7 premiers axiomes de la théorie ZF des ensembles.

16. Les deux derniers axiomes de ZF

Axiome 8 (Axiome de fondation)**.** *Pour tout ensemble x non vide, il existe au moins un élément $y \in x$ qui satisfait $y \cap x = \varnothing$.*
$$\forall_x \big[x \neq \varnothing \Rightarrow \exists_{y \in x}(y \cap x = \varnothing)\big]$$

16.1**. Proposition.** *Aucun ensemble n'est élément de lui-même.*

Démonstration. On procède par contradiction : supposons qu'il existe un ensemble a qui satisfait $a \in a$. On sait que l'ensemble $\{a\}$ existe. Puisque $\{a\} \neq \varnothing$, l'axiome de fondation implique qu'il existe $y \in \{a\}$ tel que $y \cap \{a\} = \varnothing$. On a forcément $y = a$, donc $a \cap \{a\} = \varnothing$. Cependant, $a \in a \cap \{a\}$, contradiction. □

16.2**. Proposition.** *Il n'existe pas d'ensembles $a_1, \ldots, a_n$ tels que*
$$a_1 \in a_2 \in \cdots \in a_n \in a_1.$$

Démonstration. Supposons que de tels ensembles existent. On sait que l'ensemble $x = \{a_1, \ldots, a_n\}$ existe. Puisque $x \neq \varnothing$, l'axiome de fondation implique qu'il existe $y \in x$ tel que $y \cap x = \varnothing$. Puisque $y \in x$, il existe i tel que $y = a_i$. Remarquons que $i > 1$ ou $i = 1$. Par séparation des cas, on montre que $y \cap x \neq \varnothing$:

- Si $i > 1$ alors $a_{i-1} \in a_i = y$, donc $a_{i-1} \in y \cap x$, donc $y \cap x \neq \varnothing$.
- Si $i = 1$ alors $a_n \in a_1 = y$, donc $a_n \in y \cap x$, donc $y \cap x \neq \varnothing$.

Donc $y \cap x \neq \varnothing$, contradiction. □

Remarquez en particulier que 16.2 implique qu'il n'existe pas d'ensembles a, b satisfaisant $a \in b$ et $b \in a$.

16.3. **Proposition.** *Si x, y sont des ensembles tels que $x^+ = y^+$, alors $x = y$.*

Remarque. Le cas particulier où $x, y \in \omega$ a déjà été démontré, voir 12.16.

Démonstration de 16.3. Par contradiction : supposons que $x^+ = y^+$ et $x \neq y$.
$x \in x^+ = y^+ = y \cup \{y\} \Rightarrow (x \in y \ \vee \ x = y) \Rightarrow x \in y$
$y \in y^+ = x^+ = x \cup \{x\} \Rightarrow (y \in x \ \vee \ x = y) \Rightarrow y \in x$
donc $x \in y \ \wedge \ y \in x$, contradiction. □

Axiome 9 (Axiome de remplacement)**.** *Étant donnés un ensemble A et une formule $Q(x, y)$ avec variables libres x, y, si l'énoncé*

$$\forall_{a \in A} \exists!_b Q(a, b)$$

est vrai alors il existe un ensemble B tel que

$$\forall_{a \in A} \exists!_{b \in B} Q(a, b).$$

17. L'Axiome du Choix

Les 9 axiomes qu'on a vus ci-dessus constituent la théorie "ZF" des ensembles (l'axiomatization de Zermelo et Fraenkel). La théorie ZFC des ensembles est obtenue en ajoutant aux axiomes de ZF un axiome supplémentaire, appelé l'Axiome du Choix.

Observez la démonstration du lemme suivant.

17.1. **Lemme.** *Supposons que E est un sous-ensemble de $\mathbb{R}$ satisfaisant*

$$\forall_{a \in (0,\infty)} \ \ (0, a) \cap E \neq \varnothing.$$

Alors il existe une fonction $f : (0, \infty) \to E$ telle que $\forall_{a \in (0,\infty)} \ \ 0 < f(a) < a$.

Démonstration. Pour chaque $a \in (0, \infty)$ on peut choisir un élément y_a de $(0, a) \cap E$, puisque cet ensemble est non vide. On peut donc définir une fonction

$$f : (0, \infty) \to E$$

par la condition $f(a) = y_a$. Il est clair que cette fonction satisfait la condition $\forall_{a \in (0,\infty)} \ \ 0 < f(a) < a$. □

Dans la démonstration ci-dessus on prétend définir une fonction en effectuant une infinité de choix (un choix pour chaque nombre réel positif). Or, aucun des axiomes de la théorie ZF des ensembles n'implique qu'une telle fonction existe. Nous ajoutons donc un nouvel axiome à notre liste :

17.2. **Axiome du choix** (Version #1).
Soit $(A_i)_{i\in I}$ *une famille telle que* $\forall_{i\in I}\ (A_i \neq \varnothing)$. *Alors il existe au moins une famille* $(a_i)_{i\in I}$ *(indicée par le même ensemble* I*) satisfaisant* $\forall_{i\in I}\ (a_i \in A_i)$.

17.3. **Axiome du choix** (Version #2).
Soit $(A_i)_{i\in I}$ *une famille telle que* $\forall_{i\in I}\ (A_i \neq \varnothing)$. *Alors il existe au moins une fonction* $c : I \to \bigcup_{i\in I} A_i$ *telle que* $\forall_{i\in I}\ (c(i) \in A_i)$.

Les deux axiomes ci-dessus sont équivalents, dans le sens suivant : à partir des 9 axiomes de ZF et de 17.2, on peut déduire 17.3 ; et à partir de ZF et de 17.3, on peut déduire 17.2. En effet, les paragraphes 11.16 et 11.17 nous permettent de voir que l'existence de la famille $(a_i)_{i\in I}$ (dans 17.2) est équivalente à l'existence de la fonction c (dans 17.3). Nous insistons sur le fait qu'il est impossible de démontrer l'axiome du choix à partir des axiomes 1 à 9 de la théorie ZF des ensembles ; il s'agit donc d'un nouvel axiome, indépendant des autres.

En vertu de l'axiome du choix il est permis de définir une fonction en effectuant une infinité de choix, donc la preuve de 17.1 qu'on a donnée ci-dessus est valide. Il n'est pas nécessaire de corriger cette preuve, mais nous allons quand même le faire pour bien montrer où et comment on utilise l'axiome.

Démonstration "corrigée" de 17.1. Pour chaque $a \in (0,\infty)$ on définit l'ensemble $E_a = (0,a) \cap E$. On considère la famille d'ensembles $(E_a)_{a\in(0,\infty)}$ et on note que $E_a \neq \varnothing$ pour tout $a \in (0,\infty)$. En vertu de l'axiome du choix, il existe une fonction

$$c : (0,\infty) \to \textstyle\bigcup_{a\in(0,\infty)} E_a$$

telle que $\forall_{a\in(0,\infty)}\ \ c(a) \in E_a$. La composition

$$(0,\infty) \ \xrightarrow{c}\ \textstyle\bigcup_{a\in(0,\infty)} E_a \ \hookrightarrow\ E$$

est une fonction $f : (0,\infty) \to E$ qui satisfait la condition $\forall_{a\in(0,\infty)}\ \ 0 < f(a) < a$. □

Remarque. Pour que la démonstration ci-dessus soit vraiment complète, il faudrait aussi nous assurer que la famille $(E_a)_{a\in(0,\infty)}$ existe.

17.4. **Axiome du choix** (Version #3).
Supposons que U *et* V *sont des ensembles et que* $P(x,y)$ *est une formule dont les variables libres sont comprises parmi* x, y. *Si la condition*

$$\forall_{u\in U}\exists_{v\in V}\ P(u,v) \tag{16}$$

est satisfaite alors il existe au moins une fonction $c : U \to V$ *qui satisfait*

$$\forall_{u\in U}\ P(u,c(u)).$$

Preuve que $(ZF \wedge 17.3) \Rightarrow 17.4$. Pour chaque $u \in U$ on définit $E_u = \{\, v \in V \mid P(u,v) \,\}$. Alors la famille $(E_u)_{u\in U}$ existe ; puisqu'on suppose que (16) est satisfaite, cette famille satisfait $\forall_{u\in U}\,(E_u \neq \varnothing)$. En vertu de 17.3, il existe une fonction $c' : U \to \bigcup_{u\in U} E_u$ telle que $\forall_{u\in U}\,(c'(u) \in E_u)$. Alors la composition $U \xrightarrow{c'} \bigcup_{u\in U} E_u \hookrightarrow V$ est une fonction $c : U \to V$ satisfaisant la condition $\forall_{u\in U}\ P(u,c(u))$. □

Preuve que $(ZF \wedge 17.4) \Rightarrow 17.3$. Soit $F = \big(A_i\big)_{i\in I}$ une famille telle que $\forall_{i\in I}\ (A_i \neq \varnothing)$. Considérons la formule

$$P(x,y) : \quad \exists_t \big((x,t) \in F \ \wedge\ y \in t\big),$$

dont les variables libres sont x et y. Si on écrit $V = \bigcup_{i\in I} A_i$, alors la condition

$$\forall_{i\in I} \exists_{v\in V}\ P(i,v) \tag{17}$$

est satisfaite. En effet, (17) équivaut à

$$\forall_{i\in I}\, \exists_{v\in V}\, \exists_t\ \big((i,t) \in F \ \wedge\ v \in t\big) \tag{18}$$

et (18) est vraie puisque $\forall_{i\in I}\ (A_i \neq \varnothing)$. Alors 17.4 implique l'existence d'une fonction $c : I \to V = \bigcup_{i\in I} A_i$ satisfaisant $\forall_{i\in I}\ P(i, c(i))$, donc satisfaisant $\forall_{i\in I}\ (c(i) \in A_i)$. □

Donnons une nouvelle démonstration de 17.1 mais cette fois utilisons 17.4 plutôt que 17.3.

Démonstration de 17.1. Puisque la condition $\forall_{a\in(0,\infty)}\ \ (0,a) \cap E \neq \varnothing$ est satisfaite, il s'ensuit que la condition

$$\forall_{a\in(0,\infty)}\, \exists_{b\in E}\ (0 < b < a)$$

est satisfaite. Donc 17.4 implique qu'il existe une fonction $f : (0,\infty) \to E$ telle que $\forall_{a\in(0,\infty)}\ \ 0 < f(a) < a$. (On applique 17.4 à la formule $P(x,y) =$ " $0 < y < x$ ".) □

Voici un deuxième exemple de démonstration qui utilise l'axiome du choix.

17.5. **Proposition.** *Soient A et B des ensembles et $f : A \to B$ une fonction surjective. Alors il existe au moins une une fonction $g : B \to A$ telle que $f \circ g = 1_B$.*

Démonstration de 17.5 *utilisant* 17.3. Pour chaque $b \in B$, on définit l'ensemble $E_b = \big\{\, a \in A \mid f(a) = b \,\big\}$. On considère la famille d'ensembles $\big(E_b\big)_{b\in B}$ et on note que $E_b \neq \varnothing$ pour tout $b \in B$ (parce que f est surjective). En vertu de 17.3, il existe une fonction $g : B \to \bigcup_{b\in B} E_b$ telle que $\forall_{b\in B}\ g(b) \in E_b$. Remarquons que $\bigcup_{b\in B} E_b = A$, donc g est en fait une fonction de B vers A ($g : B \to A$).

Soit $b \in B$. Puisque $g(b) \in E_b$, on a $f(g(b)) = b$. Ainsi, $f \circ g = 1_B$.

On a montré qu'il existe au moins une fonction $g : B \to A$ telle que $f \circ g = 1_B$. □

Démonstration de 17.5 *utilisant* 17.4. Puisque $f : A \to B$ est surjective, la condition

$$\forall_{b\in B} \exists_{a\in A}\ f(a) = b$$

est satisfaite. Alors 17.4 implique qu'il existe une fonction $g : B \to A$ telle que $\forall_{b\in B}\ f(g(b)) = b$. (On utilise 17.4 avec $U = B$, $V = A$ et $P(x,y) =$ "$f(y) = x$".) □

18. Le lemme de Zorn

18.1. **Définition.** Soit $(X, \leq)$ un ensemble partiellement ordonné. Si C est un sous-ensemble de X qui satisfait $\forall_{x_1,x_2 \in C}$ ($x_1 \leq x_2$ ou $x_2 \leq x_1$), on dit que C est une *chaîne* de $(X, \leq)$.

Autrement dit, une chaîne de $(X, \leq)$ est une partie de X totalement ordonnée.

18.2. **Lemme de Zorn.** *Supposons que $(X, \leq)$ est un ensemble partiellement ordonné satisfaisant les deux conditions suivantes :*

(1) $X \neq \varnothing$
(2) *Pour toute chaîne $C \subseteq X$ telle que $C \neq \varnothing$, il existe au moins un $y \in X$ satisfaisant $\forall_{x \in C}\ x \leq y$.*

Alors $(X, \leq)$ a au moins un élément maximal.

Démonstration. Le Lemme de Zorn est une conséquence de l'Axiome du Choix. Nous omettons la démonstration, qui est un peu compliquée (référence : Halmos, *Naive set theory*). □

Remarque. Un élément $y \in X$ satisfaisant $\forall_{x \in C}\ x \leq y$ est appelé une *borne supérieure* de C. Ainsi, la condition (2) demande que toute chaîne non vide de X admette une borne supérieure dans X (notez qu'on ne demande pas que y appartienne à C).

18.3. **Exemple.** $C = [1, 5)$ est une chaîne de $(\mathbb{R}, \leq)$ et 8 est une borne supérieure de C (n'importe quel nombre réel y satisfaisant $y \geq 5$ est une borne supérieure de C). Notez qu'aucun élément de C n'est une borne supérieure de C.

D'autre part, $C' = \mathbb{Z}$ et $C'' = \mathbb{R}$ sont des chaînes de $(\mathbb{R}, \leq)$ qui n'admettent aucune borne supérieure dans $\mathbb{R}$.

18.4. **Exercice.** Soit Ω l'ensemble des sous-ensembles propres de $\mathbb{R}$,

$$\Omega = \big\{\, A \in \wp\,\mathbb{R} \mid A \neq \mathbb{R} \,\big\}.$$

Considérez l'ensemble partiellement ordonné $(\Omega, \subseteq)$.
Pour chaque $a \in \mathbb{R}$, soit $I_a = (-\infty, a) \in \Omega$.

(1) Soit $C = \big\{\, I_a \mid a < 0 \,\big\}$. Montrez que C est une chaîne de $(\Omega, \subseteq)$. Montrez aussi qu'il existe des bornes supérieures $Y \in \Omega$ de C.
(2) Soit $C = \big\{\, I_a \mid a \in \mathbb{R} \,\big\}$. Montrez que C est une chaîne de $(\Omega, \subseteq)$ qui n'a aucune borne supérieure dans Ω. (Donc on ne peut pas utiliser le Lemme de Zorn pour prouver que $(\Omega, \subseteq)$ possède un élément maximal.)

Le Lemme de Zorn est très utile pour démontrer des théorèmes dans plusieurs branches des mathématiques. Voici un exemple provenant de l'algèbre linéaire.

18.5. **Proposition.** *Soit v_0 un vecteur non nul dans un espace vectoriel V. Alors il existe au moins un sous-espace U de V qui satisfait :*

(1) $v_0 \notin U$

(2) *si W est un sous-espace de V tel que $U \subset W$, alors $v_0 \in W$.*

Remarque. Comme d'habitude, le symbole "$\subset$" désigne l'inclusion stricte.

Remarque. Lorsque $\dim(V) < \infty$ on n'a pas besoin du lemme de Zorn pour démontrer la proposition. La démonstration que nous donnons ci-dessous est valide dans tous les cas, que $\dim(V)$ soit finie ou infinie.

Démonstration. On définit

$$\Omega = \big\{\, W \subseteq V \mid W \text{ est un sous-espace de } V \text{ et } v_0 \notin W \,\big\}.$$

Alors $(\Omega, \subseteq)$ est un ensemble partiellement ordonné. Pour démontrer la proposition, il suffit de montrer que $(\Omega, \subseteq)$ possède au moins un élément maximal. En effet, si $U \in \Omega$ est un élément maximal de $(\Omega, \subseteq)$ alors U est un sous-espace de V satisfaisant les conditions requises par la proposition.

Vérifions que $(\Omega, \subseteq)$ satisfait les deux hypothèses du lemme de Zorn.

(1) Le sous-espace nul $\{0\}$ de V est un élément de Ω, donc $\Omega \neq \varnothing$.

(2) Soit $C \subseteq \Omega$ une chaîne de $(\Omega, \subseteq)$ telle que $C \neq \varnothing$. On doit montrer qu'il existe $Y \in \Omega$ tel que $\forall_{W \in C}\, W \subseteq Y$. Soit Y l'union de tous les éléments de C,

$$Y = \cup C = \bigcup_{W \in C} W.$$

L'exercice 18.6 implique que Y est un sous-espace de V. Montrons que $v_0 \notin Y$. Par contradiction : supposons que $v_0 \in Y$. Alors il existe un $W \in C$ tel que $v_0 \in W$, ce qui contredit le fait que $W \in \Omega$ ($W \in C \subseteq \Omega$). Cette contradiction montre que $v_0 \notin Y$, et puisque Y est un sous-espace de V on conclut que

$$Y \in \Omega.$$

Il est clair que $\forall_{W \in C} W \subseteq Y$, donc Y est une borne supérieure de C, ce qui montre que toute chaîne de $(\Omega, \subseteq)$ admet une borne supérieure dans Ω.

Puisque les hypothèses du Lemme de Zorn sont satisfaites, il s'ensuit que $(\Omega, \subseteq)$ possède au moins un élément maximal. □

18.6. **Exercice.** Soit V un espace vectoriel. Si W_1 et W_2 sont des sous-espaces de V, $W_1 \cup W_2$ n'est pas nécessairement un sous-espace de V. Cependant, si $W_1 \subseteq W_2$ ou $W_1 \supseteq W_2$ alors $W_1 \cup W_2$ est un sous-espace de V. Plus généralement, montrez que si C est un ensemble satisfaisant :

- $C \neq \varnothing$
- chaque élément de C est un sous-espace de V
- $\forall_{W_1, W_2 \in C}(W_1 \subseteq W_2 \text{ ou } W_1 \supseteq W_2)$

alors la réunion de tous les éléments de C est un sous-espace de V.

Application du Lemme de Zorn : existence des bases

Une des applications importantes du Lemme de Zorn est la preuve que tout espace vectoriel possède une base. Revoyons d'abord les définitions.

Soit V un espace vectoriel sur un corps K.

18.7. **Définition.** Un sous-ensemble S de V est *linéairement dépendant* s'il existe un sous-ensemble fini et non-vide $\{v_1, \dots, v_n\}$ de S (où $v_1, \dots, v_n$ sont distincts) et des éléments $a_1, \dots, a_n \in K$ non tous nuls tels que $a_1 v_1 + \cdots + a_n v_n = 0$.

Un sous-ensemble S de V est *linéairement indépendant* s'il n'est pas linéairement dépendant.

Dans le cas particulier où S est fini, la définition ci-dessus est équivalente à celle que vous avez vue en algèbre linéaire.

Notez que l'ensemble vide $\varnothing \subset V$ est linéairement indépendant. (En effet, si $\varnothing$ est linéairement dépendant alors il existe un **sous-ensemble non-vide** $\{v_1, \dots, v_n\}$ **de** $\varnothing$ et des éléments $a_1, \dots, a_n \in K$ non tous nuls tels que $a_1 v_1 + \cdots + a_n v_n = 0$. Or, un tel ensemble $\{v_1, \dots, v_n\}$ n'existe pas.)

18.8. **Exercice.** Soit S un sous-ensemble de V.

(1) Si S est linéairement indépendant alors tout sous-ensemble de S l'est aussi.
(2) Si tout sous-ensemble fini de S est linéairement indépendant alors S est linéairement indépendant.

18.9. **Définition.** Étant donné un sous-ensemble S de V, on définit un sous-espace $\langle S \rangle$ de V de la manière suivante :

- Si $S = \varnothing$, on définit $\langle S \rangle = \{0\}$ (le sous-espace nul de V)
- Si $S \neq \varnothing$, on définit $\langle S \rangle =$ l'ensemble de tous les vecteurs de la forme

$$a_1 v_1 + \cdots + a_n v_n,$$

pour tous les choix possibles de $n > 0$, $v_1, \dots, v_n \in S$ et $a_1, \dots, a_n \in K$.

On appelle $\langle S \rangle$ le *sous-espace de V engendré par S*.
Dans le cas où $\langle S \rangle = V$, on dit que S *engendre* V.

18.10. **Définition.** Une *base* de V est un sous-ensemble $S \subseteq V$ qui satisfait les deux conditions suivantes :

- S est linéairement indépendant
- S engendre V.

Remarque. Supposons que $V = \{0\}$ est l'espace vectoriel nul. Alors $\varnothing$ est une base de V. En effet, la définition 18.7 implique que $\varnothing$ est un ensemble linéairement indépendant ; et la définition 18.9 implique que $\langle \varnothing \rangle = \{0\} = V$, donc $\varnothing$ engendre V.

18.11. **Exercice.** Soit $\mathbb{R}[t]$ l'ensemble des polynômes en une variable t et à coefficients réels (par exemple, $2 - (\sqrt{3})t + \frac{1}{2}t^2$ est un élément de $\mathbb{R}[t]$). Alors $\mathbb{R}[t]$ est un espace vectoriel sur $\mathbb{R}$. Démontrez que $\{1, t, t^2, t^3, \dots\}$ est une base de $\mathbb{R}[t]$.

18.12. **Notation.** Soit V un espace vectoriel sur un corps K et soit

$$\mathcal{L} = \{\, S \subseteq V \mid S \text{ est linéairement indépendant} \,\}.$$

On considère l'ensemble partiellement ordonné $(\mathcal{L}, \subseteq)$. L'espace V et l'ensemble $(\mathcal{L}, \subseteq)$ sont fixés pour tout le paragraphe 18.12 (incluant 18.12.1 et 18.12.2).

18.12.1. **Proposition.** *Tout élément maximal de $(\mathcal{L}, \subseteq)$ est une base de V.*

Démonstration. Soit S un élément maximal de $(\mathcal{L}, \subseteq)$. Alors $S \in \mathcal{L}$, donc $S \subseteq V$ et S est linéairement indépendant. Le fait que S soit maximal signifie :

(∗) il n'existe aucun ensemble $S' \subseteq V$ linéairement indépendant et strictement plus grand que S ($S \subset S'$).

On va montrer que S est une base de V. Il suffit de montrer que S engendre V. On considère un élément quelconque v de V et on doit montrer que $v \in \langle S \rangle$. Ceci est clair lorsque $v \in S$, donc on peut supposer que $v \notin S$. Soit $S' = S \cup \{v\}$, alors $S \subset S' \subseteq V$. Puisque S satisfait (∗), S' est linéairement dépendant. Donc il existe une partie finie et non vide $\{v_1, \ldots, v_n\}$ de S' et des $a_1, \ldots, a_n \in K$ non tous nuls tels que $a_1 v_1 + \cdots + a_n v_n = 0$ (où $v_1, \ldots, v_n$ sont distincts).

Si $v \notin \{v_1, \ldots, v_n\}$ alors $\{v_1, \ldots, v_n\} \subseteq S$ donc S est linéairement dépendant, ce qui est faux. Donc $v \in \{v_1, \ldots, v_n\}$. Quitte à renuméroter les v_i, on peut bien supposer que $v = v_1$. Notons que $\{v_2, \ldots, v_n\} \subseteq S$.

Si $a_1 = 0$ alors $\{v_2, \ldots, v_n\} \subseteq S$, $a_2 v_2 + \cdots + a_n v_n = 0$ et $a_2, \ldots, a_n$ sont non tous nuls. Ceci est impossible car S est linéairement indépendant. Donc $a_1 \neq 0$.

En utilisant $a_1 \neq 0$ et $a_1 v_1 + \cdots + a_n v_n = 0$, on peut écrire $v = v_1 = \alpha_2 v_2 + \cdots + \alpha_n v_n$ ($\alpha_i \in K$), donc $v \in \langle S \rangle$ car $\{v_2, \ldots, v_n\} \subseteq S$. Ceci montre que S engendre V, donc S est une base de V. □

Remarque. Il est également vrai que toute base de V est un élément maximal de $(\mathcal{L}, \subseteq)$, mais nous n'aurons pas besoin de ce fait dans la suite.

18.12.2. **Proposition.** $(\mathcal{L}, \subseteq)$ *possède au moins un élément maximal.*

Démonstration. Il suffit de vérifier que $(\mathcal{L}, \subseteq)$ satisfait les deux hypothèses du Lemme de Zorn.

(1) On sait que le sous-ensemble vide $\varnothing \subset V$ est linéairement indépendant, donc $\varnothing \in \mathcal{L}$, donc $\mathcal{L} \neq \varnothing$.

(2) Soit $C \subseteq \mathcal{L}$ une chaîne telle que $C \neq \varnothing$. Soit $T = \bigcup_{S \in C} S$. Montrons que T est linéairement indépendant. En vertu de 18.8, il suffit de montrer que tout sous-ensemble fini de T est linéairement indépendant. Soit $\{v_1, \ldots, v_n\}$ un sous-ensemble fini de T. Alors pour chaque i on a $v_i \in T = \bigcup_{S \in C} S$, donc il existe $S_1, \ldots, S_n \in C$ tels que $v_1 \in S_1$, …, $v_n \in S_n$. Puisque C est une chaîne, il existe $j \in \{1, \ldots, n\}$ tel que $S_1 \cup \cdots \cup S_n = S_j$, donc $\{v_1, \ldots, v_n\} \subseteq S_j$. Puisque $S_j \in C \subseteq \mathcal{L}$, on a $S_j \in \mathcal{L}$, donc S_j est linéairement indépendant. Alors tout sous-ensemble de S_j est linéairement

indépendant, donc en particulier $\{v_1, \ldots, v_n\}$ est linéairement indépendant. Ceci montre que T est linéairement indépendant, donc $T \in \mathcal{L}$. Ainsi, T est une borne supérieure de C dans $\mathcal{L}$ (car il est clair que $\forall_{S \in C}\, S \subseteq T$).

On a montré que $(\mathcal{L}, \subseteq)$ satisfait les hypothèses du Lemme de Zorn et il s'ensuit que $(\mathcal{L}, \subseteq)$ admet un élément maximal S. □

En combinant 18.12.1 et 18.12.2, on obtient :

Théorème. *Tout espace vectoriel admet une base.*

Une petite modification dans la démonstration ci-dessus nous permet de prouver que tout ensemble linéairement indépendant peut être étendu à une base. On commence par démontrer (en utilisant le Lemme de Zorn) une version "forte" du Lemme de Zorn :

18.13. **Lemme.** *Supposons que $(X, \leq)$ est un ensemble partiellement ordonné satisfaisant les deux hypothèses du Lemme de Zorn. Alors, quel que soit $x_0 \in X$, il existe un élément maximal m de $(X, \leq)$ qui satisfait $x_0 \leq m$.*

Démonstration. Soit $X_0 = \{\, x \in X \mid x_0 \leq x \,\}$. On peut vérifier (exercice) que $(X_0, \leq)$ satisfait les deux hypothèses du Lemme de Zorn. Donc, en vertu du Lemme de Zorn, $(X_0, \leq)$ admet un élément maximal m. Alors (exercice) m est un élément maximal de $(X, \leq)$ qui satisfait $x_0 \leq m$. □

18.14. **Proposition.** *Soit V un espace vectoriel sur un corps K et soit S_0 un sous-ensemble de V linéairement indépendant. Alors il existe une base S de V telle que $S_0 \subseteq S$.*

Démonstration. Comme auparavant, on considère l'ensemble partiellement ordonné $(\mathcal{L}, \subseteq)$, où $\mathcal{L} = \{\, S \subseteq V \mid S \text{ est linéairement indépendant} \,\}$. Dans la démonstration de 18.12.2, on a montré que $(\mathcal{L}, \subseteq)$ satisfait les hypothèses du Lemme de Zorn. Puisque $S_0 \in \mathcal{L}$, 18.13 implique qu'il existe un élément maximal S de $(\mathcal{L}, \subseteq)$ qui satisfait $S_0 \subseteq S$. Alors 18.12.1 implique que S est une base de V. □