Logique et Théorie Axiomatiques

(*mise à jour le 15 mars 2021*)

J.L. Krivine

Ce document est issu d'un polycopié de licence (actuelle L3) de l'université Paris 7 écrit dans les années 1970 par Jean-Louis Krivine, et distribué sous sa forme originale jusqu'à très récemment. Il en reprend la première partie, consacrée à la théorie des ensembles, avec quelques (très peu) de modifications de l'auteur.

Raphaël Giromini en a produit une première version au format LaTeX en août 2004, version complétée et corrigée par Jean-Louis Krivine, Yves Legrandgérard et Paul Rozière.

N'hésitez pas à envoyer les corrections de coquilles ou autres erreurs que vous pourriez repérer à `roziere@irif.fr` ou à l'auteur `krivine@irif.fr`.

Table des matières

Première partie

Éléments de Théorie des Ensembles.

En mathématique, toutes les notions, ou presque, sont définies à partir de la notion d'ensemble. On ne peut donc pas espérer définir ce qu'est un ensemble. Néanmoins, on voudrait pouvoir utiliser les ensembles et faire des démonstrations avec eux sans avoir de doute sur la rigueur. Pour cela, on va se baser sur l'idée intuitive que chacun possède plus ou moins, sur les ensembles, en la précisant quelque peu.

On peut admettre sans difficulté que certaines choses méritent d'être appelées « ensemble » et certaines non : si la notion d'ensemble recouvrait n'importe quoi, ce n'aurait pas été la peine de lui donner un nom. Les ensembles décrivent donc un certain domaine que nous nous proposons d'étudier.

Chapitre 1

Ensembles, relations, fonctions.

1.1 Axiomes de Zermelo

Il y a deux relations fondamentales entre les ensembles : l'égalité et l'appartenance. Nous ne les définissons pas : on considère que chacun sait ce que veut dire « les ensembles a, b sont égaux (ou identiques) » (ce qu'on écrit $a = b$ bien entendu) et « l'ensemble a appartient à l'ensemble b », (ce qu'on écrit $a \in b$, on dit aussi « a est un élément de b »). Toutes les autres relations seront, elles, définies à partir de ces deux-là.

Par exemple, on dira que « l'ensemble a est une partie (ou un sous-ensemble) de l'ensemble b » ou encore que « a est contenu dans b » si chaque élément de a est aussi élément de b. La notation est $a \subset b$.

Nous allons, dans ce qui suit, énoncer certaines propriétés des ensembles, sous forme de règles : comme on ne peut pas définir les ensembles, on se contente de dire ce qu'on peut faire et ce qu'on ne peut pas faire avec eux. Ces règles constituent donc le « mode d'emploi » de la notion d'ensemble. On les appelle « axiomes de la théorie des ensembles » ou « axiomes de Zermelo ». A part le premier, tous ces axiomes ont l'allure générale suivante : certains ensembles étant donnés, il existe un ensemble ayant telle et telle propriété vis-à-vis des ensembles donnés. Traditionnellement, ils portent les noms suivants : *axiome d'extensionnalité*, *axiome de compréhension* (ou *séparation*), *axiome de la paire*, *axiome de la réunion*, *axiome de l'ensemble des parties*, *axiome de l'infini*, *axiome du choix*.

Ils expriment des propriétés plus ou moins évidentes de la notion d'ensemble, ce qui fait qu'à première vue on ne voit pas, au moins pour plusieurs d'entre eux, l'intérêt qu'il y a à les énoncer. L'intérêt existe néanmoins, pour la raison suivante : les axiomes de Zermelo expriment, de façon *exhaustive*, les propriétés des ensembles. Ce qui fait que chaque fois que l'on aura un doute sur la validité de la construction de tel ou tel ensemble, c'est aux axiomes qu'il faudra se référer pour voir s'ils permettent de faire cette construction. Or, en mathématiques, en toute rigueur, il faudrait avoir un doute à chaque pas en avant qu'on se propose de faire ...

AXIOME 1 (AXIOME D'EXTENSIONNALITÉ)
Pour que l'ensemble a soit égal à l'ensemble b, il faut et il suffit que tout élément de a soit élément de b et inversement. Autrement dit,

$$a = b \Leftrightarrow (a \subset b \text{ et } b \subset a)$$

AXIOME 2 (AXIOME DE SÉPARATION (OU DE COMPRÉHENSION))
Étant donné un ensemble a et une propriété $P(x)$ (portant sur un ensemble variable x) il existe un ensemble b dont les éléments sont ceux, parmi les éléments de a, qui ont la propriété $P(x)$.

Notons que d'après l'axiome d'extensionnalité, un tel ensemble b est déterminé de façon unique.

Remarque. On a pensé à énoncer l'axiome suivant : étant donné une propriété $P(x)$, il existe un ensemble b dont les éléments sont les ensembles qui ont la propriété $P(x)$.
Mais cela mène à une contradiction quand on prend comme propriété $P(x)$: $x \notin x$ (autrement dit la

propriété pour un ensemble de ne pas s'appartenir à lui-même). En effet, l'énoncer précédent donne alors un ensemble b tel que pour tout ensemble x on ait $x \in b \Leftrightarrow x \notin x$. En particulier, pour $x = b$ on obtient $b \in b \Leftrightarrow b \notin b$ ce qui est évidemment faux.
Cette remarque a été faite par B. Russell (d'où son nom : le paradoxe de Russell) et a imposé l'axiome de compréhension tel que nous l'avons énoncé.

Ensembles et propriétés. L'ensemble des éléments de l'ensemble a qui ont la propriété $P(x)$ est noté : $\{x \in a\ ;\ P(x)\}$

Il existe un ensemble et un seul qui n'a aucun élément, on le note $\emptyset$ et on l'appelle « ensemble vide ». Pour montrer son existence, on prend n'importe quel ensemble a et on considère $\{x \in a\ ;\ x \neq x\}$; cet ensemble n'a aucun élément. L'unicité est due à l'axiome d'extensionnalité.

Il n'existe aucun ensemble qui ait tous les ensembles comme éléments : en effet, si a est un tel ensemble, on pose $b = \{x \in a\ ;\ x \notin x\}$. Alors pour tout ensemble x, on a $x \in b \Leftrightarrow x \notin x$, d'où une contradiction comme pour le paradoxe de Russell.

AXIOME 3 (AXIOME DE LA PAIRE)
a, b étant des ensembles, il existe un ensemble qui a comme éléments a, b et eux seulement.

D'après l'axiome d'extensionnalité, il existe un seul ensemble ayant cette propriété, on le note $\{a, b\}$. En particulier, lorsque $a = b$, on voit qu'il existe un ensemble dont a est le seul élément. On le note $\{a\}$.

AXIOME 4 (AXIOME DE LA RÉUNION)
Étant donné un ensemble a, il existe un ensemble b dont les éléments sont les ensembles qui appartiennent à un élément de a.

Cet ensemble b (unique d'après l'axiome d'extensionnalité) est appelé « réunion des éléments de a » et noté $\bigcup_{x \in a} x$.

Étant donné deux ensembles a, b, on appelle réunion de a et b (et on note $a \cup b$) la réunion des éléments de l'ensemble $\{a, b\}$. Pour tout ensemble x, on a donc

$$x \in a \cup b \Leftrightarrow (x \in a \text{ ou } x \in b)\ .$$

A l'aide de l'axiome d'extensionnalité, on voit aisément que

$$a \cup b = b \cup a\ ;\ a \cup (b \cup c) = (a \cup b) \cup c\ .$$

Ce dernier ensemble est noté $a \cup b \cup c$ et est appelé réunion des ensemble a, b, c. On définit de même la réunion de quatre ensemble a, b, c, d, etc.

Étant donné trois ensemble a, b, c, il existe un ensemble qui a comme éléments a, b, c et eux seulement : c'est $\{a\} \cup \{b\} \cup \{c\}$. On le note $\{a, b, c\}$. On définit de même l'ensemble $\{a, b, c, d\}$, etc.

Étant donné deux ensembles a, b, on appelle intersection de a et b (et on note $a \cap b$) l'ensemble $\{x \in a\ ;\ x \in b\}$ (défini grâce à l'axiome de compréhension). On a donc pour tout ensemble x,

$$x \in a \cap b \Leftrightarrow x \in a \text{ et } x \in b\ .$$

On voit immédiatement, à l'aide de l'axiome d'extensionnalité que

$$a \cap b = b \cap a\ ;\ a \cap (b \cap c) = (a \cap b) \cap c\ .$$

Ce dernier ensemble est noté $a \cap b \cap c$ et appelé intersection des ensemble a, b, c. On défini de même l'intersection de quatre ensemble a, b, c, d, etc. Toujours à l'aide de l'axiome d'extensionnalité, on voit que

$$a \cup (b \cap c) = (a \cup b) \cap (a \cup c) \text{ et } a \cap (b \cup c) = (a \cap b) \cup (a \cap c)\ .$$

Étant donné un ensemble A et une partie X de A, l'ensemble $\{x \in A\ ;\ x \notin X\}$ (défini grâce à l'axiome de compréhension) est appelé complémentaire de X par rapport à A et noté $\complement_A X$ (ou encore $A - X$). On voit aisément, à l'aide de l'axiome d'extensionnalité que si $X, Y \subset A$, on a

$$\complement_A(X \cup Y) = \complement_A X \cap \complement_A Y \text{ et } \complement_A(X \cap Y) = \complement_A X \cup \complement_A Y\ .$$

AXIOME 5 (AXIOME DE L'ENSEMBLE DES PARTIES)
Pour tout ensemble a, il existe un ensemble b dont les éléments sont les sous-ensembles de a.

Cet ensemble b (unique d'après l'axiome d'extensionnalité) est appelé ensemble des parties de a et noté $\mathscr{P}(A)$.

Nous énoncerons plus tard les deux derniers axiomes de la théorie des ensembles : l'axiome de l'infini et l'axiome du choix.

1.2 Quelques notions élémentaires.

1.2.1 Couple ordonné.

Étant donné deux ensembles a, b, on appelle « couple ordonné dont le premier élément est a et le second b », l'ensemble $\{\{a\},\{a,b\}\}$. On le note (a,b).

THÉORÈME 1.2.1
Si $(a,b)=(a',b')$, alors $a=a'$ et $b=b'$.

On a en effet $\{\{a\},\{a,b\}\}=\{\{a'\},\{a',b'\}\}$ et il y a donc deux possibilités :

1. $\{a\}=\{a'\}$ et $\{a,b\}=\{a',b'\}$; d'où $a=a'$ et $b=b'$.
2. $\{a\}=\{a',b'\}$ et $a'=\{a,b\}$; d'où $a'=b'=a$ et $a=b=b'$, donc $a=a'=b=b'$.

C.Q.F.D.

Étant donné trois ensembles a,b,c, on appelle « triplet ordonné dont le premier élément est a, le second b et le troisième c » l'ensemble $(a,(b,c))$. On le note (a,b,c).

THÉORÈME 1.2.2
Si $(a,b,c)=(a',b',c')$ alors $a=a'$, $b=b'$ et $c=c'$.

En effet on a $(a,(b,c))=(a',(b',c))$ donc $a=a'$ et $(b,c)=(b',c')$ d'après le théorème précédent. D'où $b=b'$ et $c=c'$. C.Q.F.D.

On définit de même le quadruplet ordonné (a,b,c,d) en posant $(a,b,c,d)=(a,(b,c,d))$. Donc si $(a,b,c,d)=(a',b',c',d')$ alors $a=a'$, $b=b'$, $c=c'$ et $d=d'$. Et ainsi de suite.

1.2.2 Ensemble produit.

Étant donné deux ensembles A,B, il existe un ensemble P donc les éléments sont les couples ordonnés (x,y) avec $x\in A$ et $y\in B$.

En effet, si $x\in A$, $y\in B$, alors $\{x\}$, $\{x,y\}$ appartiennent à $\mathscr{P}(A\cup B)$. Donc $\{\{x\},\{x,y\}\}\subset\mathscr{P}(A\cup B)$ et donc $\{\{x\},\{x,y\}\}\in\mathscr{P}(\mathscr{P}(A\cup B))$. On définit alors P à l'aide de l'axiome de compréhension, en posant

$$P=\{z\in\mathscr{P}(\mathscr{P}(A\cup B))\}\ ;\ z \text{ est un couple } (x,y) \text{ avec } x\in A, y\in B\}$$

et il est clair que P est l'ensemble cherché. Cet ensemble P est appelé produit de A et B, et noté $A\times B$.

Étant donné trois ensemble A, B, C, l'ensemble des triplets (x,y,z) avec $x\in A$, $y\in B$ et $z\in C$ est l'ensemble $A\times(B\times C)$. On le note $A\times B\times C$ et on l'appelle produit des ensembles A, B, C.

1.2.3 Relations binaires.

Une relation binaire R sur un ensemble E est, par définition, un sous-ensemble de E^2, c'est-à-dire un ensemble de couples (x,y) avec $x,y\in E$.

— R est dite réflexive si $(x,x)\in R$ pour tout $x\in R$.
— R est dite symétrique si $(x,y)\in R\Rightarrow(y,x)\in R$
— R est dite antisymétrique si $\big[(x,y)\in R \text{ et } (y,x)\in R\big]\Rightarrow x=y$.

— R est dite transitive si $\big[(x,y)\in R \text{ et } (y,z)\in R\big] \Rightarrow (x,z)\in R$.
— R est dite totale si $x,y\in E \Rightarrow \big[(x,y)\in R \text{ ou } (y,x)\in R\big]$.

Si R est une relation binaire sur E, réflexive, antisymétrique et transitive, on dit que R est une relation d'ordre sur E. Si, de plus, R est totale, on dit que R est une relation d'ordre total sur E.
Si R est une relation binaire sur E, réflexive, symétrique et transitive, on dit que R est une relation d'équivalence sur E.
Si $a\in E$, l'ensemble $\{x\in E\ ;\ (x,a)\in R\}$ est appelé classe d'équivalence de $a\pmod R$.
Notons $\bar{a}$ la classe d'équivalence de a. Il existe un ensemble E' qui a comme éléments les classes d'équivalence des éléments de E.
En effet, si $a\in E$, alors $\bar{a}\subset E$, donc $\bar{a}\in\mathscr{P}(E)$. Donc, si on pose

$$E' = \{X\in\mathscr{P}(E)\ ;\ \text{il existe } a\in E \text{ tel que } X=\bar{a}\}$$

(E' est défini grâce à l'axiome de compréhension) E' est l'ensemble cherché. On l'appelle ensemble quotient de E par la relation d'équivalence R, et on le note E/R.
On appelle « partition de E » un sous-ensemble P de $\mathscr{P}(E)$ tel que :
— $X\in P \Rightarrow X\neq\emptyset$
— $X,Y\in P$ et $X\neq Y \Rightarrow X\cap Y=\emptyset$
— $\bigcup_{X\in P} X = E$

Alors, E/R est une partition de E, comme on le voit immédiatement. Inversement, si P est une partition de E, on lui associe une relation d'équivalence R sur E définie par :

$$(x,y)\in R \Leftrightarrow \text{il existe un élément } X \text{ de } P \text{ tel que } x,y\in X$$

Les relations d'équivalence sur l'ensemble E correspondent donc canoniquement aux partitions de E.

1.2.4 Fonctions.

Une application de l'ensemble A dans l'ensemble B (ou encore une fonction définie sur l'ensemble A à valeurs dans B), est par définition, un sous ensemble f de $A\times B$ qui a la propriété suivante : pour tout élément $x\in A$, il existe un élément $y\in B$ et un seul tel que $(x,y)\in f$. On écrit alors $y=f(x)$ au lieu de $(x,y)\in f$. On écrit $f:A\to B$ pour « f est une application de A dans B ».

Il existe un ensemble C dont les éléments sont les applications de A dans B. En effet, si f est une application de A dans B, alors $f\subset A\times B$, donc $f\in\mathscr{P}(A\times B)$. On peut donc (au moyen de l'axiome de compréhension) définir l'ensemble

$$C = \{f\in\mathscr{P}(A\times B)\ ;\ f \text{ est une application de } A \text{ dans } B\}$$

qui est l'ensemble cherché. L'ensemble des applications de A dans B est noté B^A.

Par exemple, si $A=\emptyset$, $B^A=\{\emptyset\}$ ($\emptyset$ est une fonction de domaine $\emptyset$ et c'est la seule). Si $B=\emptyset$ et $A\neq\emptyset$ on a $B^A=\emptyset$ (il n'y a aucune fonction de domaine $A\neq\emptyset$ à valeur dans $\emptyset$).

Une application $f:A\to B$ est dite :
— injective si $x,x'\in A$, $x\neq x' \Rightarrow f(x)\neq f(x')$
— surjective si pour tout $y\in B$, il existe $x\in A$ tel que $y=f(x)$
— bijective (ou biunivoque de A sur B) si elle est à la fois injective et surjective.

Si f est une application biunivoque de A sur B, l'ensemble des couples (y,x) avec $x\in A$, $y\in B$ et $(x,y)\in f$ est alors une application de B sur A qu'on note f^{-1} et qu'on appelle application inverse (ou réciproque) de f.

1.2.5 Applications composées.

Soit $f:A\to B$ et $g:B\to C$. Désignons par φ l'ensemble des couples (x,z) avec $x\in A$, $z\in C$, tels qu'il existe $y\in B$ avec $(x,y)\in f$ et $(y,z)\in g$. Il est facile de montrer que φ est alors une application de A dans C. On l'appelle « application composée de f et de g » et on la note $g\circ f$. Pour tout $x\in A$, on a donc $g\circ f(x)=g(f(x))$.

Soient $f : A \to B$, $g : B \to C$, $h : C \to D$ on a alors

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

En effet, on a :

$(x,t) \in h \circ (g \circ f) \Leftrightarrow$ il existe $z \in C$ tel que $(x,z) \in g \circ f$ et $(z,t) \in h$.
$\Leftrightarrow$ il existe $z \in C$ et $y \in B$ tels que $(x,y) \in f$, $(y,z) \in g$ et $(z,t) \in h$.
$\Leftrightarrow$ il existe $y \in B$ tel que $(x,y) \in f$ et $(y,t) \in h \circ g$.
$\Leftrightarrow (x,t) \in (h \circ g) \circ f$. C.Q.F.D.

Soit f une application de A dans B ; on lui associe deux applications : $\dot{f} : \mathscr{P}(A) \to \mathscr{P}(B)$ et $f^{-1} : \mathscr{P}(B) \to \mathscr{P}(A)$ (à ne pas confondre, malgré la notation identique, avec l'application réciproque de f, qui n'est définie que lorsque f est biunivoque). Elles sont définies de la façon suivante :

— si $X \in \mathscr{P}(A)$, $\dot{f}(X) = \{y \in B$ il existe $x \in X$ tel que $y = f(x)\}$
— si $Y \in \mathscr{P}(B)$, $f^{-1}(Y) = \{x \in A \ ; \ f(x) \in Y\}$

$\dot{f}(X)$ est appelé l'image de X par la fonction f et $f^{-1}(Y)$ l'image réciproque de Y par la fonction f. En particulier, $\dot{f}(A)$ est appelé l'image de la fonction f.

Soient $X, X' \subset A$ et $Y, Y' \subset B$, on a les propriétés suivantes :

$$\begin{array}{rcl|rcl} \dot{f}(X \cup X') & = & \dot{f}(X) \cup \dot{f}(X') & f^{-1}(Y \cup Y') & = & f^{-1}(Y) \cup f^{-1}(Y') \\ \dot{f}(X \cap X') & \subset & \dot{f}(X) \cap \dot{f}(X') & f^{-1}(Y \cap Y') & = & f^{-1}(Y) \cap f^{-1}(Y') \\ & & & f^{-1}(\complement_B Y) & = & \complement_A f^{-1}(Y) \end{array}$$

1.2.6 Familles d'ensembles.

Étant donné un ensemble I, une fonction de domaine I (à valeur dans un ensemble quelconque A) est aussi appelée *famille d'ensembles indexée par I*. On la note alors $(a_i)_{i \in I}$; I est appelé l'ensemble d'indices de la famille I.
On appelle réunion de la famille $(a_i)_{i \in I}$ (et on la note $\bigcup_{i \in I} a_i$) la réunion d'éléments de l'image de la fonction f (on utilise ici l'axiome de la réunion).
En particulier, si $I = \emptyset$, on a $f = \emptyset$. L'image de f est donc $\emptyset$ et donc aussi la réunion des éléments de l'image de f. Donc il n'y a qu'une seule famille dont l'ensemble d'indices est vide (on l'appelle la famille vide). Sa réunion est $\emptyset$.
Supposons maintenant $I \neq \emptyset$. Alors il existe un ensemble C dont les éléments sont les ensembles qui appartiennent à tous les éléments de l'image de f. En effet, puisque $I \neq \emptyset$, on prend $i_0 \in I$. Un ensemble qui appartient à tous les a_i appartient en particulier à a_{i_0}.

On peut donc définir C en posant

$$C = \{x \in a_{i_0} \ ; \ x \text{ appartient à tous les éléménts de l'image de } f\}$$

(on utilise ici l'axiome de compréhension).
Cet ensemble C est appelé « intersection de la famille $(a_i)_{i \in I}$ » et noté $\bigcap_{i \in I} a_i$. Notons que cette intersection n'est définie que pour une famille non vide.
Étant donné une famille d'ensemble $(a_i)_{i \in I}$, il existe un ensemble C dont les éléments sont les fonctions φ de domaine I telle que $\varphi(i) \in a_i$ pour tout $i \in I$.
En effet, une telle fonction φ est une application de I dans $\bigcup_{i \in I} a_i$ donc un éléments de $(\bigcup_{i \in I} a_i)^I$. On peut donc poser, en utilisant l'axiome de compréhension :

$$C = \{\varphi \in (\bigcup_{i \in I} a_i)^I \ ; \ \varphi(i) \in a_i \text{ pour tout } i \in I\} \ .$$

Cet ensemble est appelé « produit de la famille $(a_i)_{i \in I}$ » et noté $\prod_{i \in I} a_i$.

Chapitre 2

Entiers Naturels.

2.1 Définition des entiers naturels.

On commence par définir chacun des entiers naturels 0, 1, 2, ... L'idée de la définition est la suivante : l'entier 5, par exemple, doit être un ensemble qui a cinq éléments. Si on a déjà défini 0, 1, 2, 3, 4, il est alors naturel de poser $5 = \{0,1,2,3,4\}$.

On définit donc successivement :

$$0 = \emptyset \ ; \ 1 = \{0\} \ ; \ 2 = \{0,1\} \ ; \ 3 = \{0,1,2\} \ ; \ \ldots$$

On a donc :

$$1 = \{\emptyset\} \ ; \ 2 = \{\emptyset, \{\emptyset\}\} \ ; \ 3 = \{\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\} \ ; \ \ldots$$

L'opération qui permet de passer d'un entier au suivant est une opération très simple sur les ensembles : celle qui, à l'ensemble x, associe l'ensemble $x \cup \{x\}$ (c'est-à-dire l'ensemble dont les éléments sont x et les éléments de x). En effet on a, par exemple, $12 = \{0,1,2,\ldots,11\}$ et $13 = \{0,1,2,\ldots,11,12\}$ donc $13 = 12 \cup \{12\}$.

Dans la suite de ce chapitre on utilisera la notation x^+ pour désigner l'ensemble $x \cup \{x\}$.

On se propose de définir l'ensemble des entiers. Cet ensemble A doit avoir les propriétés suivantes :

$$\begin{cases} \emptyset \in A \\ \text{si } x \in A \text{ alors } x^+ \in A \, . \end{cases} \qquad (*)$$

On ne peut pas déduire des axiomes déjà énoncés l'existence d'un ensemble A ayant les propriétés $(*)$. On énonce donc un nouvel axiome :

AXIOME 6 (AXIOME DE L'INFINI)
Il existe un ensemble A tel que $\emptyset \in A$ et si $x \in A$ alors $x \cup \{x\} \in A$.

On montre alors le théorème suivant :

THÉORÈME 2.1.1
Il existe un ensemble et un seul qui a les propriétés $(*)$ et qui est contenu dans tout ensemble A qui a les propriétés $(*)$.

On considère un ensemble A qui a les propriétés $(*)$, il en existe un d'après l'axiome de l'infini. Soit B l'intersection de tous les sous-ensembles de A qui ont les propriétés $(*)$. Il est immédiat que B a encore les propriétés $(*)$.

Soit C un ensemble quelconque ayant les propriétés $(*)$; alors $C \cap A$ a encore cette propriété et c'est un sous-ensemble de A ; donc $B \subset C \cap A$, par définition de B. Par suite $B \subset C$, ce qui montre que B est l'ensemble cherché.

Si B' a la propriété $(*)$ et est inclus dans tout ensemble ayant la propriété $(*)$, alors $B' \subset B$ et $B \subset B'$ donc $B = B'$. C.Q.F.D.

L'ensemble défini par le théorème précédent est appelé ensemble des entiers naturels et désigné par $\mathbb{N}$. Par définition un entier naturel est donc un élément de $\mathbb{N}$, autrement dit un ensemble qui appartient à tout ensemble ayant la propriété $(*)$.

2.2 Relation d'ordre sur les entiers.

Il s'agit maintenant de démontrer pour l'ensemble $\mathbb{N}$ ainsi défini, les propriétés que l'on sait intuitivement être vraies pour l'ensemble des entiers naturels. Pour cela on va définir sur $\mathbb{N}$ une relation d'ordre et montrer qu'elle a les propriétés suivantes :

1. $\mathbb{N}$ est totalement ordonné et a pour plus petit élément 0.
2. Tout élément n de $\mathbb{N}$ a un successeur $n^+ = n \cup \{n\}$ (autrement dit l'ensemble des majorants stricts de n a un plus petit élément qui est n^+).
3. Tout élément $n \neq 0$ de $\mathbb{N}$ a un prédécesseur (c'est-à-dire un entier m tel que $m^+ = n$).
4. Si une propriété $P(x)$ est vraie pour 0, et si $P(n) \Rightarrow P(n^+)$ pour tout $n \in \mathbb{N}$, alors $P(n)$ est vraie pour tout $n \in \mathbb{N}$.

On montre d'abord la propriété (4) : Soit A l'ensemble des entiers qui ont la propriété $P(x)$ ($A = \{x \in \mathbb{N} : P(x)\}$ est défini à l'aide de l'axiome de compréhension). Alors $0 \in A$ et si $x \in A$ alors $x^+ \in A$. Donc A a les propriétés $(*)$ et (par définition de $\mathbb{N}$) on a $\mathbb{N} \subset A$. Tout entier a donc la propriété $P(x)$. La propriété (4) s'appelle le *principe d'induction*; on va l'utiliser constamment dans toutes les démonstrations sur $\mathbb{N}$.

Montrons maintenant la propriété (3). Soient $B = \{n \in \mathbb{N}\ ;\ $ il existe $m \in \mathbb{N}$ tel que $m^+ = n\}$ et $A = B \cup \{0\}$. Alors $0 \in A$, si $x \in A$ on a évidement $x^+ \in B$, donc $x^+ \in A$. Donc A a les propriétés $(*)$ et par suite $\mathbb{N} \subset A$; cela veut dire que tout entier non nul a un prédécesseur.

LEMME 2.2.1
Si $n \in \mathbb{N}$ et $m \in n$ alors $m \in \mathbb{N}$ (tous les éléments d'un entier sont des entiers).

On le montre par induction sur n. La propriété est vraie si $n = 0$ (n n'a pas d'élément). Supposons-la vraie pour n. Si $m \in n^+$ comme $n^+ = n \cup \{n\}$ on a ou bien $m \in n$ donc m est entier (hypothèse d'induction), ou bien $m = n$ et m est encore entier. C.Q.F.D.

LEMME 2.2.2
Si n est entier et si $m \in n$, alors $m \subset n$.

Par induction sur n. C'est évident si $n = 0$. On suppose que c'est vraie pour n et soit $m \in n \cup \{n\}$. Alors ou bien $m \in n$ donc $m \subset n$ (hypothèse d'induction) donc $m \subset n \cup \{n\}$; ou bien $m = n$, donc $m \subset n \cup \{n\}$. C.Q.F.D.

LEMME 2.2.3
Si n est entier, $n \notin n$.

C'est évident si $n = 0$ (n n'a pas d'élément). Supposons que $n \notin n$ et que $n \cup \{n\} \in n \cup \{n\}$. On a alors ou bien $n \cup \{n\} = n$ ou bien $n \cup \{n\} \in n$. Dans le premier cas on a $n \in n$ (car $n \in n \cup \{n\}$) ce qui contredit l'hypothèse. Dans le second, on a $n \cup \{n\} \subset n$ (lemme 2.2.2). Or $n \in n \cup \{n\}$, donc $n \in n$ contrairement à l'hypothèse. C.Q.F.D.

LEMME 2.2.4
Si m, n sont entiers et $m \subset n$ alors $m = n$ ou bien $m \in n$.

On le montre par induction sur n : la propriété $P(n)$ est alors « pour tout entier m, si $m \subset n$ alors $m = n$ ou $m \in n$ ».

C'est évident si $n = 0$ (car $m \subset n \Rightarrow m = 0$). Supposons que l'on ait $P(n)$ et soit $m \subset n \cup \{n\}$; si $n \notin m$ alors $m \subset n$ et donc (hypothèse d'induction) $m = n$ ou $m \in n$; dans ce cas $m \in n \cup \{n\}$. Si $n \in m$ on a $n \subset m$ (lemme 2.2.2) et $\{n\} \subset m$ (car cela équivaut à $n \in m$). Donc $n \cup \{n\} \subset m$ et comme par hypothèse on a l'inclusion inverse, $m = n \cup \{n\}$. C.Q.F.D.

LEMME 2.2.5
Si m est un entier non nul, alors $0 \in m$.

On montre par induction sur m que $m = 0$ ou $0 \in m$. C'est évident si $m = 0$, si c'est vrai pour m on a nécessairement $0 \in m \cup \{m\}$: car ou bien $m = 0$ et on sait que $m \in m \cup \{m\}$ ou bien $m \neq 0$ donc $0 \in m$ (hypothèse d'induction) et $0 \in m \cup \{m\}$. C.Q.F.D.

LEMME 2.2.6
Si m, n sont entiers, alors un et un seul de trois cas suivants est réalisé : $n \in m$, $n = m$ ou $m \in n$.

Montrons d'abord l'unicité : si $n \in m$ et $n = m$ on a $m \in m$ ce qui est impossible (lemme 2.2.3) ; si $n \in m$ et $m \in n$ on a $n \subset m$ (lemme 2.2.2) et comme $m \in n$, on a $m \in m$ ce qui est impossible.

On montre alors, par induction sur n, la propriété $P(n)$: « pour tout entier m, $m \in n$, $m = n$ ou $n \in m$. »

Si $n = 0$ c'est vrai d'après le lemme 2.2.4.

Supposons $P(n)$ et considérons $n \cup \{n\}$ et un entier quelconque m. Par hypothèse on a donc $m \in n$, $m = n$ ou $n \in m$.

— Si $m \in n$ ou $m = n$ alors $m \in n \cup \{n\}$.
— Si $n \in m$ on a $n \subset m$ (lemme 2.2.2) et $\{n\} \subset m$ donc $n \cup \{n\} \subset m$. Par suite (lemme 2.2.4) $n \cup \{n\} = m$ ou $n \cup \{n\} \in m$.

C.Q.F.D.

On définit alors une relation d'ordre sur $\mathbb{N}$ en posant $m \leqslant n$ si et seulement si $m \subset n$. D'après le lemme 2.2.4, on a $m < n$ si et seulement si $m \in n$. D'après le lemme 2.2.6 cette relation d'ordre est totale. D'après le lemme 2.2.5, elle a 0 pour plus petit élément (propriété (1)).

Si m, n sont entiers on a $m < n \Leftrightarrow m \in n$ et $m \leqslant n \Leftrightarrow m \subset n$, donc $m < n \Leftrightarrow (m \in n$ et $m \subset n)$, c'est-à-dire $m < n \Leftrightarrow m \cup \{m\} \subset n$, et donc $m < n \Leftrightarrow m \cup \{m\} \leqslant n$. L'ensemble des majorants stricts de m a donc un plus petit élément qui est $m \cup \{m\}$. On a ainsi terminé la démonstration des propriétés (1), (2), (3) et (4). C.Q.F.D.

Dans toute la suite, nous ne nous servirons plus explicitement de la définition de $\mathbb{N}$, mais seulement du fait que $\mathbb{N}$ satisfait les propriétés (1), (2), (3) et (4).

THÉORÈME 2.2.7
Si $m, n \in \mathbb{N}$ et $m^+ = n^+$, alors $m = n$.

En effet, si $m \neq n$, on a par exemple $m < n$. Donc $m^+ \leqslant n$ (par définition du successeur). Comme $n < n^+$, on a $m^+ < n^+$ ce qui contredit l'hypothèse. C.Q.F.D.

THÉORÈME 2.2.8
Tout ensemble d'entier qui est non vide a un plus petit élément.

Soit X un sous-ensemble de $\mathbb{N}$ qui n'a pas de plus petit élément. On considère la propriété $P(n)$: « n est un entier et aucun entier $m \leqslant n$ n'est élément de X ».

Comme X n'a pas de plus petit élément, en particulier $0 \notin X$ et donc $P(0)$. De plus $P(n) \Rightarrow P(n^+)$ pour tout entier n : car n^+ n'étant pas le plus petit élément de X, si aucun entier inférieur ou égal à n n'est élément de X, aucun entier inférieur ou égal à n^+ ne peut être élément de X.

D'après le principe d'induction, $P(n)$ est donc vrai pour tout entier n mais cela implique que X est vide. On en déduit le résultat par contraposée. C.Q.F.D.

2.3 Fonction sur les entiers.

On considère un ensemble quelconque $E \neq \emptyset$, un élément a de E et une application $H : \mathbb{N} \times E \to E$.

THÉORÈME 2.3.1
Il existe une application f de $\mathbb{N}$ dans E, et une seule, telle que $f(0) = a$ et $f(n^+) = H(n, f(n))$ pour tout entier n.

Unicité. Considérons deux fonctions f, g ayant cette propriété. Si $f \neq g$, l'ensemble $\{n \in \mathbb{N}\ ;\ f(n) \neq g(n)\}$ est non vide, donc a un plus petit élément m; $m \neq 0$ car $f(0) = g(0) = a$. Donc m a un prédécesseur p; on a $f(p) = g(p)$, donc $H(p, f(p)) = H(p, g(p))$ soit $f(p^+) = g(p^+)$ c'est-à-dire $f(m) = g(m)$, ce qui contredit la définition de m.

Existence. On considère les sous-ensembles M de $\mathbb{N} \times E$ qui ont les propriétés suivantes :

$$\begin{cases} (0, a) \in M \\ \text{si } (n, y) \in M \text{ alors } (n^+, H(n, y)) \in M \end{cases}$$

Il est clair que l'intersection M_0 de tous ces ensembles M de $\mathbb{N} \times E$ a encore ces propriétés. C'est donc le plus petit sous-ensemble de $\mathbb{N} \times E$ qui a ces propriétés. On va en déduire que c'est le graphe d'une application de $\mathbb{N}$ dans E.
Pour tout entier n, il existe $y \in E$ tel que $(n, y) \in M_0$: c'est vrai pour $n = 0$, puisque $(0, a) \in M_0$; si c'est vrai pour n, c'est vrai pour n^+ d'après la deuxième propriété satisfaite par M_0.
Pour tout entier n, si $(n, y) \in M_0$ et $(n, z) \in M_0$ alors $y = z$. On raisonne par l'absurde et on considère le premier entier m tel qu'il existe $y, z \in E$, $y \neq z$, $(m, y) \in M_0$, $(m, z) \in M_0$.
Si $m = 0$, on a par exemple $y \neq a$; soit M_0' l'ensemble obtenu en ôtant $(0, y)$ de M_0 ($M_0' = M_0' - \{(0, y)\}$) ; Alors M_0' a les deux propriétés ci-dessus. et est strictement inclus dans M_0, ce qui contredit la définition de M_0.
On a donc $m \neq 0$ et par suite $m = p^+$. D'après la définition de m, il existe un élément t et un seul de E tel que $(p, t) \in M_0$; alors $(p^+, H(p, t)) \in M_0$ et on a, par exemple $y \neq H(p, t)$. On pose $M_0' = M_0 - \{(m, y)\} = M_0 - \{(p^+, y)\}$. Alors M_0' a les deux propriétés ci-dessus : car $(0, a) \in M_0$ et $(0, a) \neq (m, y)$, donc $(0, a) \in M_0'$. Si $(n, u \in M_0'$ alors $(n^+, H(n, u)) \in M_0$ et $(n^+, H(n, u)) \neq (m, y)$: c'est évident si $n^+ \neq m$ et si $n^+ = m$ alors $n = p$, donc $u = t$ et $y \neq H(p, t)$. Donc $(n^+, H(n, u)) \in M_0'$. Comme M_0' est strictement inclus dans M_0 , on a encore contredit la définition de M_0.
M_0 est donc le graphe d'une application f de $\mathbb{N}$ dans E et on a bien $f(0) = a$, $f(n^+) = H(n, f(n))$ pour chaque entier n.

C.Q.F.D.

Quand on utilise ce théorème pour définir une fonction f, on dit que f est *définie par induction* sur les entiers.

2.3.1 L'addition des entiers.

Elle est définie par induction. Étant donné un entier k, on défini $k + n$ par induction sur n par les conditions :

$$\begin{cases} k + 0 &= k \\ k + n^+ &= (k + n)^+ \end{cases}$$

D'après cette définition on a $k+1 = k^+$, et nous utiliserons la notation $k+1$ pour le successeur de l'entier k.

Associativité de l'addition. $k + (n + p) = (k + n) + p$, ce qu'on montre par induction sur p. C'est évident si $p = 0$, et on a

$$k + (n + p^+) = k + (n + p)^+ = [k + (n + p)]^+$$

D'après l'hypothèse d'induction, $k + (n + p) = (k + n) + p$ et donc

$$k + (n + p^+) = [(k + n) + p]^+ = (k + n) + p^+ .$$

Commutativité de l'addition. On montre d'abord, par induction sur k, que $0+k=k+0=k$. C'est évident si $k=0$, et on a

$$0+k^+=(0+k)^+=k^+=k^++0\,.$$

On montre ensuite $1+k=k+1$; c'est évident si $k=0$, et on a

$$1+k^+=(1+k)^+=(k+1)^+=k^{++}=k^++1\,.$$

On montre alors, par induction sur k que $k+n=n+k$: c'est déjà fait si $k=0$; et on a

$$\begin{aligned}k^++n &= (k+1)+n=(1+k)+n=1+(k+n)=1+(n+k)\\ &= (1+n)+k=(n+1)+k=n+(1+k)=n+k^+\,.\end{aligned}$$

2.3.2 Le produit de deux entiers.

Il est aussi défini par induction. Étant donné un entier k, on définit $n.k$ par induction sur k, par les conditions suivantes :

$$\begin{cases} n\cdot 0 &= 0\\ n(k+1) &= n\cdot k+n\end{cases}$$

Distributivité du produit par rapport à l'addition. On montre par induction sur k que $n(m+k)=n\cdot m+n\cdot k$. C'est évident si $k=0$, et on a :

$$\begin{aligned}n(m+k^+) &= n[(m+k)+1]\\ &= n(m+k)+n\\ &= nm+nk+n \qquad \text{(hypothèse d'induction)}\\ &= nm+nk^+\,.\end{aligned}$$

associativité du produit. On montre par induction sur k que $n(mk)=(nm)k$. C'est évident si $k=0$, et on a

$$\begin{aligned}(nm)k^+ &= (nm)k+nm\\ &= n(mk)+nm \qquad \text{(hypothèse d'induction)}\\ &= n(mk+m) \qquad\qquad \text{(distributivité)}\\ &= n(mk^+)\,.\end{aligned}$$

Commutativité du produit. On montre d'abord par induction sur k que $0\cdot k=k\cdot 0=0$. C'est évident si $k=0$, et :

$$\begin{aligned}0\cdot k^+ &= 0\cdot k+0\cdot 1 \qquad \text{(distributivité)}\\ &= 0\,.\end{aligned}$$

On montre ensuite, par induction sur k, que $(n+1)k=nk+k$. C'est évident si $k=0$; et on a :

$$\begin{aligned}(n+1)k^+ &= (n+1)k+(n+1) \qquad\qquad \text{(distributivité)}\\ &= nk+k+n+1 \qquad \text{(hypothèse d'induction)}\\ &= (nk+n)+(k+1)\\ &= n(k+1)+(k+1)\,.\end{aligned}$$

On montre enfin que $kn=nk$ par induction sur k. C'est évident si $k=0$ et on a :

$$\begin{aligned}nk^+ &= nk+n\\ &= kn+n \qquad\qquad \text{(hypothèse d'induction)}\\ &= (k+1)n \quad \text{(d'après ce qu'on vient de montrer).}\end{aligned}$$

2.3.3 Exponentiation.

Étant donné un entier k, on définit k^n par induction sur n par les conditions :

$$\begin{cases} k^0 & = & 1 \\ k^{n+1} & = & k^n \cdot k\,. \end{cases}$$

On montre aisément par induction sur p les propriétés :

$$k^{n+p} = k^n \cdot k^p\ ;\ \ (k^n)^p = k^{n\cdot p}\,.$$

Chapitre 3

Ensembles finis et dénombrables.

3.1 Ensembles finis.

Un ensemble a est dit *fini* s'il existe une bijection de a sur un entier.

THÉORÈME 3.1.1
Si a est un ensemble fini, il existe un et un seul entier qui puisse être mis en bijection avec a. Cet entier est appelé le cardinal de a ou encore le nombre d'éléments de a. Il est noté $\overline{\overline{a}}$

On montre d'abord le lemme suivant :

LEMME 3.1.2
Soient a un ensemble non vide, $x_0 \in a$, f une bijection de a sur un entier n. Alors $n \neq 0$ et il existe une bijection de $a - \{x_0\}$ sur $n-1$.

Il est clair que $n \neq 0$, donc $n = p+1 = p \cup \{p\}$. Si $f(x_0) = p$, la restriction de f à $a - \{x_0\}$ est une bijection de cet ensemble sur $p = n-1$. Si $f(x_0) = m < p$, soit y_0 l'élément tel que $f(y_0) = p$; on définit $g : a \to n$ en posant $g(x_0) = p$, $g(y_0) = m$ et $g(x) = f(x)$ pour tout élément $x \in a$, $x \neq x_0, y_0$. Alors g est une bijection de a sur n et $g(x_0) = p$, on est ramené au cas précédent. C.Q.F.D.

On montre alors le théorème par l'absurde.

Pour cela, on choisit le plus petit entier n tel qu'il existe un ensemble a, en bijection avec n et avec un entier $n' \neq n$. On a alors $n < n'$ donc $n' \neq 0$ et donc $a \neq \emptyset$ (car a est en bijection avec n') ; donc $n \neq 0$. On pose $n = p+1$, $n' = p'+1$, donc $p < p'$. On prend $x_0 \in a$, d'après le lemme précédent $a - \{x_0\}$ est en bijection avec p et avec p'. Comme $p < n$, on a une contradiction avec la définition de n. C.Q.F.D.

THÉORÈME 3.1.3
Si a est un ensemble fini et b une partie de a, alors b est fini et $\overline{\overline{b}} \leqslant \overline{\overline{a}}$. De plus, si $b \neq a$, alors $\overline{\overline{b}} < \overline{\overline{a}}$.

On le montre par induction sur sur le cardinal n de a. Si $n = 0$, $a = \emptyset$ et le théorème est évident. Supposons-le vrai pour un entier n et soient a un ensemble fini de cardinal $n+1$, $b \subset a$. Si $b = a$, le résultat est évident. Si $b \neq a$, on choisit $x_0 \in b - a$. Alors $b \subset a - \{x_0\}$ et $a - \{x_0\}$ a pour cardinal n, d'après le lemme précédent. Donc $\overline{\overline{b}} \leqslant n$ (hypothèse d'induction) et par suite $\overline{\overline{b}} < \overline{\overline{a}}$. C.Q.F.D.

THÉORÈME 3.1.4
si a, b sont deux ensembles finis disjoints, $\overline{\overline{a \cup b}} = \overline{\overline{a}} + \overline{\overline{b}}$. Si a, b sont deux ensembles finis quelconques, $\overline{\overline{a \times b}} = \overline{\overline{a}}.\overline{\overline{b}}$ et $\overline{\overline{\mathscr{P}(a)}} = 2^{\overline{\overline{a}}}$.

Démonstration par induction sur $n = \overline{\overline{a}}$.

3.2 Équipotence

Deux ensembles quelconques a, b seront dit *équipotents* s'il existe une bijection de a sur b. On dit aussi que a et b ont le même cardinal ou encore la même puissance. Mais on ne définira pas, comme dans le cas des ensembles finis, le cardinal d'un ensemble quelconque a. On utilisera provisoirement la notation $a \sim b$ pour « a est équipotent à b ». On a évidemment les propriétés suivantes :

$$a \sim a$$
$$a \sim b \Leftrightarrow b \sim a$$
$$(a \sim b \text{ et } b \sim c) \Rightarrow a \sim c$$

Un ensemble est donc fini si et seulement s'il est équipotent à un entier. Dans le cas contraire, il est dit *infini*.

3.3 Ensembles dénombrables.

Un ensemble est dit *dénombrable* s'il est équipotent à $\mathbb{N}$. Notons qu'un ensemble dénombrable est nécessairement infini (sinon $\mathbb{N}$ lui-même serait un ensemble fini, donc équipotent à un entier n. Or $n+1 \subset \mathbb{N}$, donc le cardinal de $\mathbb{N}$ est supérieur ou égal à $n+1$ d'où $n \geqslant n+1$; contradiction.

THÉORÈME 3.3.1
Tout sous ensemble d'un ensemble dénombrable est fini ou dénombrable

On peut supposer que l'ensemble dénombrable considéré est $\mathbb{N}$ lui-même. On a donc un ensemble $A \subset \mathbb{N}$: supposons A infini. On définit par induction une bijection $f : \mathbb{N} \to A$. Pour cela on pose :

$f(0)$ = le plus petit élément de A
$f(n+1)$ = le plus petit élément de A qui est strictement supérieur à $f(n)$
(il existe un tel élément dans A sinon $A \subset \{x \in \mathbb{N} \,;\, x \leqslant f(n)\}$ donc A est fini).

On a donc $f(n+1) > f(n)$ pour tout $n \in \mathbb{N}$, donc f est une injection de $\mathbb{N}$ dans A.

La fonction f est surjective. En effet soit x_0 le plus petit élément de A non atteint par f, s'il en existe ; l'ensemble $\{x \in A \,;\, x < x_0\}$ a un plus grand élément $x_1 = f(n)$ (il est atteint par f). Mais x_0 est le plus petit élément de A qui est strictement supérieur à $f(n)$ donc $x_0 = f(n+1)$. C.Q.F.D.

THÉORÈME 3.3.2
S'il existe une injection de A dans $\mathbb{N}$ ou bien une surjection de $\mathbb{N}$ dans A alors A est fini ou dénombrable.

Si f est une injection de A dans $\mathbb{N}$, l'image de f est un sous ensemble de $\mathbb{N}$ qui est équipotent à A. Donc A est fini ou dénombrable d'après le théorème précédent.
Soit g une surjection de $\mathbb{N}$ sur A. On définit une injection $h : A \to \mathbb{N}$ en posant

$$h(x) = \text{ le premier entier } n \text{ tel que } \left[g(n) = x\right] .$$

C.Q.F.D.

THÉORÈME 3.3.3
$\mathbb{N} \times \mathbb{N}$ *est dénombrable.*

Comme $\mathbb{N} \times \mathbb{N}$ n'est pas fini, il suffit, d'après le théorème précédent de trouver une injection $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. On peut poser, par exemple $f(x, y) = 2^x.3^y$ (unicité de la décomposition d'un nombre en facteurs premiers). C.Q.F.D.

COROLLAIRE 3.3.4
Le produit de deux ensembles dénombrables est dénombrable.

COROLLAIRE 3.3.5
$\mathbb{N}^p$ est dénombrable, pour tout entier $p \geqslant 1$.

Démonstration par induction sur p.

COROLLAIRE 3.3.6
$\mathbb{Z}$ et $\mathbb{Q}$ sont des ensembles dénombrables.

On définit une surjection $f : \mathbb{N}^3 \to \mathbb{Q}$ en posant

$$f(x,y,z) = \epsilon \cdot \frac{y}{z} \text{ si } z \neq 0 \text{ avec } \epsilon = 1 \text{ si } x \text{ est pair } \epsilon = -1 \text{ sinon ; } f(x,y,0) = 0 \,.$$

Donc $\mathbb{Q}$ est dénombrable, ainsi que $\mathbb{Z}$ puisque $\mathbb{Z} \subset \mathbb{Q}$. C.Q.F.D.

Définition. Une famille d'ensembles $(a_n)_{n\in\mathbb{N}}$ indexée par $\mathbb{N}$ (autrement dit une fonction de domaine $\mathbb{N}$) est aussi appelée une « suite d'ensembles ». Au lieu d'écrire « la suite d'ensembles $(a_n)_{n\in\mathbb{N}}$ » on écrit souvent « la suite d'ensembles $a_0, a_1, \ldots, a_n, \ldots$ ». On a donc trois noms (et trois notations) différents pour la même notion :

— une fonction de domaine $\mathbb{N}$, notée $f : \mathbb{N} \to E$
— une famille d'ensembles indexée par $\mathbb{N}$, notée $(a_n)_{n\in\mathbb{N}}$
— une suite d'ensembles, notée $a_0, a_1, \ldots, a_n, \ldots$

Lorsqu'on a un ensemble dénombrable E, on choisit souvent une bijection $f = \mathbb{N} \to E$ qui est donc aussi une suite $a_0, a_1, \ldots, a_n, \ldots$. On dit que cette suite énumère E et on écrit :

$$E = \{a_0, a_1, \ldots, a_n, \ldots\} \,.$$

Chapitre 4

Comparaison des ensembles infinis

4.1 Axiome du choix.

À l'aide des axiomes de la théorie des ensembles énoncés jusqu'ici, on ne parvient pas à rendre compte correctement des propriétés intuitives des ensembles infinis. Par exemple, on n'arrive pas à prouver que tout ensemble infini contient un sous-ensemble dénombrable. Nous énonçons donc maintenant le dernier axiome de la théorie des ensembles.

AXIOME 7 (AXIOME DU CHOIX)
Pour toute famille $(A_i)_{i\in I}$ d'ensembles non vides, il existe une fonction f de domaine I telle que $f(i) \in A_i$ pour tout $i \in I$.

Autrement dit : le produit d'une famille d'ensemble non vides est non vide.

Étant donné un ensemble E, on appelle fonction de choix sur E une application f dont le domaine est l'ensemble des parties non vides de E (c'est à dire $\mathscr{P}(E) - \{\emptyset\}$) à valeur dans E, telle que $f(X) \in X$ pour toute partie non vide X de E

COROLLAIRE 4.1.1
Sur tout ensemble E, il existe une fonction de choix.

En effet, il suffit d'appliquer l'axiome du choix à la famille des parties non vide de E (c'est à dire l'application identique dont le domaine est l'ensemble des parties non vides de E). C.Q.F.D.

THÉORÈME 4.1.2
Tout ensemble infini possède un sous-ensemble dénombrable.

Soit E un ensemble infini. Il s'agit de trouver une application injective $\varphi : \mathbb{N} \to E$. soit $f : \mathscr{P}(E) \to E$ une fonction telle que $f(X) \in X$ pour toute partie X non vide E (il en existe d'après ce qui précède).

On défini $\Phi : \mathbb{N} \to \mathscr{P}(E)$ par induction :

$$\begin{array}{lcl} \Phi(0) & = & \{f(E)\} \\ \Phi(n+1) & = & \Phi(n) \cup \{f(E - \Phi(n))\} \end{array}$$

$\Phi(n+1)$ est donc une partie de E, obtenue en ajoutant à $\Phi(n)$ un élément de E. Il en résulte que $\Phi(n)$ est fini pour tout $n \in \mathbb{N}$. Comme E n'est pas fini, $E - \Phi(n) \neq \emptyset$, donc (par définition de f) $f(E - \Phi(n)) \notin \Phi(n)$ et, par suite, $\Phi(n+1) - \Phi(n)$ possède un élément et un seul. On peut alors définir $\varphi : \mathbb{N} \to E$ en posant

$$\begin{array}{lcl} \varphi(0) & = & f(E) \\ \varphi(n+1) & = & \text{le seul élément de } \Phi(n+1) - \Phi(n) \end{array}$$

Il est clair que φ est injective. C.Q.F.D.

Le corollaire suivant donne une caractérisation des ensembles infinis.

COROLLAIRE 4.1.3
Un ensemble est infini si et seulement s'il est équipotent à une partie propre.

On a vu que : si E est un ensemble fini et F une partie propre de E ($F \subset E$, $F \neq E$) alors le cardinal de F est strictement inférieur à celui de E, donc E n'est pas équipotent à F.

Soit maintenant E un ensemble infini ; on prend une partie D dénombrable de E, $D = \{a_0, a_1, \ldots, a_n, \ldots\}$. On définit $f : E \to E$ en posant :

$$\begin{aligned} f(x) &= x \text{ si } x \in E - D \\ f(a_i) &= a_{i+1} \text{ pour } i \in \mathbb{N} \end{aligned}$$

Il est clair que f est une bijection de E sur $E - \{a_0\}$; donc une bijection de E sur une partie propre de E. C.Q.F.D.

THÉORÈME 4.1.4
La réunion d'une suite d'ensemble dénombrables ou finis est un ensemble dénombrable ou fini.

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'ensemble dénombrables ou finis. Il suffit de trouver une application surjective de $\mathbb{N}$ sur $\bigcup_{n\in\mathbb{N}} A_n$.

Pour chaque $n \in \mathbb{N}$, soit S_n l'ensemble des applications surjectives de $\mathbb{N}$ sur A_n. Par hypothèse, $S_n \neq \emptyset$ pour chaque $n \in \mathbb{N}$. D'après l'axiome du choix, il existe une fonction φ, de domaine $\mathbb{N}$, telle que $\varphi(n) \in S_n$ pour tout $n \in \mathbb{N}$. Autrement dit, $\varphi(n)$ est, pour chaque entier n, une surjection de $\mathbb{N}$ sur A_n.

On définit $\Phi : \mathbb{N}\times\mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ en posant $\Phi(n,p) = \varphi(n)(p)$. Alors Φ est surjective. car si $x \in \bigcup_{n\in\mathbb{N}} A_n$, on a $x \in A_m$ donc $x = \varphi(m)(a)$ pour un certain entier q, puisque $\varphi(m)$ est surjective.

Comme $\mathbb{N}\times\mathbb{N}$ est dénombrable, on a le résultat cherché. C.Q.F.D.

On en déduit immédiatement que l'ensemble des suites finies d'entiers — autrement dit l'ensemble $\bigcup_{p\in\mathbb{N}} \mathbb{N}^p$) — est dénombrable.

L'ensemble P des polynômes à une variable, à coefficient dans $\mathbb{Z}$ est dénombrable : si P_k est l'ensemble des polynômes de degré $\leqslant k$, P_k est équipotent à $\mathbb{Z}^{k+1}$ (un polynôme de degré $\leqslant k$ est une suite de $k+1$ entiers relatifs) donc est dénombrable. L'ensemble considéré est $\mathrm{P} = \bigcup_{k\in\mathbb{N}} \mathrm{P}_k$ donc est dénombrable.

L'ensemble des nombres algébriques réels (nombres réels qui est racine d'un polynôme à une variable à coefficient dans $\mathbb{Z}$) est dénombrable. En effet, à chaque polynôme $u \in \mathrm{P}$, associons l'ensemble fini R_u de ses racines réelles. L'ensemble étudié est $\bigcup_{u\in\mathrm{P}} R_u$. Comme P est dénombrable c'est la réunion d'une suite d'ensemble finis, donc un ensemble dénombrable d'après le théorème précédent (ce n'est évidemment pas un ensemble fini).

4.2 Ensembles non dénombrables.

Tous les ensembles infinis rencontrés jusqu'ici se sont révélés dénombrables. Les deux théorèmes suivants, dus à Cantor, montrent qu'il existe des ensembles infinis non dénombrables.

THÉORÈME 4.2.1
$\mathscr{P}(\mathbb{N})$ *n'est pas dénombrable.*

Supposons qu'il existe une application surjective $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$. On définit $X \in \mathscr{P}(\mathbb{N})$ en posant $X = \{n \in \mathbb{N} \ ; \ n \notin f(n)\}$. Puisque f est surjective, il existe $n_0 \in \mathbb{N}$ tel que $f(n_0) = X$. Par définition de X, on a $n_0 \in X \Leftrightarrow n_0 \notin f(n_0)$. Autrement dit $n_0 \in X \Leftrightarrow n_0 \notin X$, ce qui est une contradiction. C.Q.F.D.

THÉORÈME 4.2.2
L'intervalle $[0,1[$ *de* $\mathbb{R}$ *n'est pas dénombrable.*

Soit f une application surjective de $\mathbb{N}$ sur $[0,1[$. Le réel $f(n)$ possède une représentation décimale $0, a_n^0 a_n^1 \cdots a_n^k \cdots$ et une seule (où $a_n^k \in \{0,1,\ldots,9\}$), la suite $(a_n^k)_{k\in\mathbb{N}}$ n'étant pas formée de 9 à partir d'un certain rang. On définit alors une suite $(b_k)_{k\in\mathbb{N}}$ d'entiers ($0 \leqslant b_k \leqslant 9$), en posant $b_k = 0$ si $a_k^k \neq 0$ et $b_k = 1$

si $a_k^k = 0$. La suite b_k n'est pas formée de 9 à partir d'un certain rang (elle ne contient en fait pas de 9). Le réel $0, b_1 b_2 \cdots b_k \cdots$ est donc un élément de $[0,1[$, et comme f est surjective, il existe un entier n_0 tel que $f(n_0) = 0, b_1 b_2 \cdots b_k \cdots$. On a donc :

$$0, a_{n_0}^1 a_{n_0}^2 \cdots a_{n_0}^k \cdots = 0, b_1 b_2 \cdots b_k \cdots .$$

En particulier $b_{n_0} = a_{n_0}^{n_0}$, ce qui contredit la définition de b_{n_0}. C.Q.F.D.

Bien entendu, il en résulte que $\mathbb{R}$ lui-même n'est pas dénombrable puisque $\mathbb{R} \supset [0,1[$.

Comme on a démontré plus haut que l'ensemble des nombres algébriques réels est dénombrable on a ainsi prouvé :

il existe un nombre réel transcendant (c'est-à-dire non algébrique),
et même : l'ensemble des nombres réels transcendants est non dénombrable (car la réunion de deux ensembles dénombrables est dénombrable).

THÉORÈME 4.2.3
Si E est infini et F est dénombrable ou fini, alors $E \cup F$ est équipotent à E.

En effet on a un ensemble dénombrable $D \subset E$. Donc $E \cup F = (E - D) \cup (D \cup F)$. Mais $D \cup F$ est équipotent à D (tous deux sont dénombrables) donc $E \cup F \sim (E - D) \cup E = E$. C.Q.F.D.

THÉORÈME 4.2.4
Si E est infini non dénombrable, on obtient un ensemble équipotent en lui retranchant une partie dénombrable ou finie.

En effet, si $A \subset E$ est dénombrable ou fini, $E' = E - A$ est infini donc d'après le théorème précédent, $E' \sim E' \cup A = E$. C.Q.F.D.

On en déduit que les divers ensembles non dénombrables rencontrés jusqu'à présent sont tous équipotents :

THÉORÈME 4.2.5
$\mathbb{R}$, $\mathscr{P}(\mathbb{N})$, les intervalles $[a,b]$, $[a,b[$, $]a,b[$ de $\mathbb{R}$ $(a < b)$ sont des ensembles équipotents.

Il est clair que $[a,b]$, $[a,b[$, $]a,b[$ sont équipotents (d'après le théorème 4.2.4). Or $]-1,1[$ est équipotent à $]a,b[$ (considérer la fonction $y = a + \frac{1}{2}(b-a)(x+1)$ et $]-1,1[$ est équipotent à $\mathbb{R}$ (considérer la fonction $y = \frac{x}{1-x^2}$).

Il suffit donc de montrer que $\mathscr{P}(\mathbb{N})$ est équipotent à un intervalle $[a,b[$, par exemple à $[0,1[$. On définit une application $\varphi : [0,1[\to \mathscr{P}(\mathbb{N})$ de la façon suivante : si $r \in [0,1[$, r possède une développement binaire et un seul, soit $0, \varepsilon_0 \varepsilon_1 \cdots \varepsilon_n \cdots$ qui ne doit pas formé exclusivement de 1 à partir d'un certain rang.

On pose $\varphi(r) = \{n \in \mathbb{N} \ ; \ \varepsilon_n = 0\}$. On montre aisément que φ est injective et que l'image de φ est l'ensemble $\mathscr{P}_\infty(\mathbb{N})$ des parties infinies de $\mathbb{N}$. Il en résulte que $[0,1[$ est équipotent à $\mathscr{P}_\infty(\mathbb{N})$.

Or $\mathscr{P}_\infty(\mathbb{N})$ est obtenu en retranchant de $\mathscr{P}(\mathbb{N})$ un ensemble dénombrable (l'ensemble des parties finies de $\mathbb{N}$) donc (théorème 4.2.4) $\mathscr{P}_\infty(\mathbb{N})$ est équipotent à $\mathscr{P}(\mathbb{N})$. C.Q.F.D.

THÉORÈME 4.2.6
$\mathbb{R}^n$, $\mathscr{P}(\mathbb{N})^n$ sont équipotents à $\mathbb{R}$ pour tout $n \in \mathbb{N}$.

On montre d'abord que $\mathscr{P}(\mathbb{N})^2$ est équipotent à $\mathscr{P}(\mathbb{N})$: soient A, B deux ensembles dénombrables disjoints. Il est clair que $\mathscr{P}(\mathbb{N})$ est équipotent à $\mathscr{P}(A)$, $\mathscr{P}(B)$, $\mathscr{P}(A \cup B)$ (car $A \cup B$ est aussi dénombrable). Or, se donner une partie de $A \cup B$ revient à se donner une partie de A et une partie de B, donc $\mathscr{P}(A) \times \mathscr{P}(B) \sim \mathscr{P}(A \cup B)$. Il en résulte que $\mathscr{P}(\mathbb{N})^2 \sim \mathscr{P}(\mathbb{N})$.

Il en résulte que $\mathbb{R}^2$ est équipotent à $\mathbb{R}$. On montre alors immédiatement par induction sur n, que $\mathbb{R}^n \sim \mathbb{R}$ pour tout $n \in \mathbb{N}$ C.Q.F.D.

Définition. On dit qu'un ensemble a la *puissance du continu* s'il est équipotent à $\mathbb{R}$.

Les ensembles infinis rencontrés jusqu'ici se rangent en deux classes : ceux qui sont dénombrables, ceux qui ont la puissance du continu. Nous allons voir qu'il existe des ensembles infinis qui ne se rangent dans aucune de ces classes. Pour pouvoir comparer entre-eux les divers ensembles infinis on introduit la définition suivante :

Définition. On dit que le cardinal (ou la puissance) de l'ensemble E est inférieur ou égal à celui de F s'il existe une injection de E dans F. On écrit alors $\overline{\overline{E}} \leqslant \overline{\overline{F}}$.

On a immédiatement les propriétés suivantes :

— Si $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{F}} \leqslant \overline{\overline{G}}$, alors $\overline{\overline{E}} \leqslant \overline{\overline{G}}$.

— Si $E \sim F$ alors $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{F}} \leqslant \overline{\overline{E}}$. La réciproque est vraie (mais non évidente) et sera démontrée un peu plus loin (théorème 4.2.8 de Cantor-Bernstein).

— Si $\overline{\overline{E}} \leqslant \overline{\overline{E'}}$ et $\overline{\overline{F}} \leqslant \overline{\overline{F'}}$, alors $\overline{\overline{E \times E'}} \leqslant \overline{\overline{F \times F'}}$

— Si E et E' sont des ensembles disjoints, ainsi que F et F', et si $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{E'}} \leqslant \overline{\overline{F'}}$, alors $\overline{\overline{E \cup E'}} \leqslant \overline{\overline{F \cup F'}}$.

THÉORÈME 4.2.7
Soient E et F deux ensembles tel que $E \neq \emptyset$. Alors $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ si et seulement s'il existe une surjection de F sur E.

Si $\overline{\overline{E}} \leqslant \overline{\overline{F}}$, on a une injection $f : E \to F$. Comme $E \neq \emptyset$, on prend $a \in E$, on définit $g : F \to E$ en posant $g(y) =$ le seul élément x de E tel que $y = f(x)$ s'il existe ; $g(y) = a$ sinon. Alors g est une surjection de F sur E.

Inversement, soit $g : F \to E$, surjective. D'après l'axiome du choix, il existe une application $\varphi : \mathscr{P}(E) \to E$ telle que $\varphi(X) \in X$ pour toute partie X non vide de E. On définit alors $f : E \to F$ en posant $f(x) = \varphi(g^{-1}(\{x\}))$ ($g^{-1}(\{x\}) = \{y \in F \ ; \ g(y) = x\}$ est non vide puisque g est surjective). Alors f est une injection de E dans F. C.Q.F.D.

THÉORÈME 4.2.8 (CANTOR-BERNSTEIN)
Si $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{F}} \leqslant \overline{\overline{E}}$, alors E est équipotent à F.

Ce théorème justifie qu'on emploie à partir de maintenant la notation $\overline{\overline{E}} = \overline{\overline{F}}$ au lieu de $E \sim F$. Le symbole $\leqslant$ se comporte alors comme une relation d'ordre.

Pour montrer le théorème, il suffit de montrer le lemme suivant :

LEMME 4.2.9
Si $A \subset E$, et s'il existe $\varphi : E \to A$ injective, alors $A \sim E$.

(Théorème 4.2.8.) En effet, en admettant le lemme, considérons deux ensembles E, F tels que $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{F}} \leqslant \overline{\overline{E}}$. On a donc deux injections $f : E \to F$ et $g : F \to E$. Si $A \subset E$ est l'image de g, F est équipotent à A. Or $g \circ f$ est une injection de E dans A. D'après le lemme 4.2.9, E est équipotent à A, donc à F C.Q.F.D.

(Lemme 4.2.9.) Pour chaque $x \in E$, on définit par induction $\varphi^n(x)$ pour tout $n \in \mathbb{N}$:

$$\begin{array}{lcl} \varphi^0(x) & = & 0 \\ \varphi^{n+1}(x) & = & \varphi(\varphi^n(x)) \end{array}$$

On désigne par B l'ensemble des éléments de E de la forme $\varphi^n(x)$, n décrivant $\mathbb{N}$ et x décrivant $E - A$ (on obtient cet ensemble à l'aide de l'axiome de compréhension : $B = \{y \in E \ ; \ \text{il existe } x \in E - A \text{ et } n \in \mathbb{N} \text{ tels que } y = \varphi^n(x)\}$).

Évidemment, $B \supset E - A$ (puisque $\varphi^0(x) = x$. De plus φ envoie B dans B et comme φ est une injection de E dans A, on voit que la restriction φ_B de φ à B est une injection de B dans $B \cap A$.

En fait, tout élément de $B \cap A$ est atteint par φ_B : si $u \in B \cap A$, on a $u = \varphi^n(x)$, $x \in E - A$; donc $n \neq 0$ (sinon $u \notin A$) soit $n = p+1$ et $u = \varphi(v)$, $v = \varphi^p(x)$ donc $v \in B$. On a ainsi montré que φ_B est une bijection de B sur $B \cap A$.

Donc B est équipotent à $B \cap A$. Il en résulte que $(E-B) \cup B$ est équipotent à $(E-B) \cup (B \cap A)$ (car $E-B$ et B sont disjoints ainsi que $E-B$ et $B \cap A$).

Comme $(E-B) \cup B = E$ on aura le résultat cherché si on montre que $(E-B) \cup (B \cap A) = A$. Or $B \supset E-A$, donc $E-B \subset A$ et comme $B \cap A \subset A$, on a $(E-B) \cup (B \cap A) \subset A$. Inversement, si $x \in A$, ou bien $x \in B$ d'où $x \in A \cap B$, ou bien $x \notin B$ et $x \in E-B$. Donc $(E-B) \cup (B \cap A) \supset A$. C.Q.F.D.

A titre d'application du théorème de Cantor-Bernstein, notons le théorème suivant :

THÉORÈME 4.2.10
L'ensemble $\mathbb{N}^{\mathbb{N}}$ des fonctions définies sur les entiers à valeur entières a la puissance du continu.

Si, à chaque partie de $\mathbb{N}$, on associe sa fonction caractéristique, on obtient une injection de $\mathscr{P}(\mathbb{N})$ dans $\mathbb{N}^{\mathbb{N}}$.

D'autre part, une fonction $f : \mathbb{N} \to \mathbb{N}$ est un ensemble de couples d'entiers donc $\mathbb{N}^{\mathbb{N}} \subset \mathscr{P}(\mathbb{N}^2)$. Comme $\mathscr{P}(\mathbb{N}^2)$ est équipotent à $\mathscr{P}(\mathbb{N})$ (puisque $\mathbb{N}^2$ est dénombrable) on en déduit une injection de $\mathbb{N}^{\mathbb{N}}$ dans $\mathscr{P}(\mathbb{N})$.

D'après le théorème de Cantor-Bernstein, $\mathbb{N}^{\mathbb{N}}$ est équipotent à $\mathscr{P}(\mathbb{N})$. C.Q.F.D.

On dira que l'ensemble E a une puissance strictement inférieure à celle de F (ce qu'on écrit $\overline{\overline{E}} < \overline{\overline{F}}$) si $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{E}} \neq \overline{\overline{F}}$. Autrement dit, s'il existe une injection de E dans F mais pas d'injection de F dans E. Par exemple, $\mathbb{R}$ a une puissance strictement supérieure à celle de $\mathbb{N}$.

Le théorème suivant montre que , pour tout ensemble, il en existe un de puissance strictement supérieure.

THÉORÈME 4.2.11 (CANTOR)
Pour tout ensemble E, $\overline{\overline{\mathscr{P}(E)}} > \overline{\overline{E}}$.

L'application $x \mapsto \{x\}$ de E dans $\mathscr{P}(E)$ est injective. Donc $\overline{\overline{E}} \leqslant \overline{\overline{\mathscr{P}(E)}}$.
Si f est une bijection de E sur $\mathscr{P}(E)$, on définit $X \in \mathscr{P}(E)$ en posant $X = \{x \in E \,;\, x \notin f(x)\}$. Alors il existe $x_0 \in E$ tel que $f(x_0) = X$. Par définition de X on a $x_0 \in X \Leftrightarrow x_0 \notin f(x_0)$ soit $x_0 \in X \Leftrightarrow x_0 \notin X$ ce qui est une contradiction. C.Q.F.D.

Chapitre 5

Le théorème de Zorn.

Grâce au théorème de Cantor-Bernstein, on a montré que les notations $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ et $\overline{\overline{E}} = \overline{\overline{F}}$ (qui, rappelons-le, signifient respectivement « il existe une injection de E dans F » et « il existe une bijection de E sur F ») se comportent comme une relation d'ordre, c'est-à-dire :

$$\begin{array}{l}\overline{\overline{E}} \leqslant \overline{\overline{E}}\,;\\ \left(\overline{\overline{E}} \leqslant \overline{\overline{F}} \text{ et } \overline{\overline{F}} \leqslant \overline{\overline{E}}\right) \Leftrightarrow \overline{\overline{E}} = \overline{\overline{F}}\,;\\ \left(\overline{\overline{E}} \leqslant \overline{\overline{F}} \text{ et } \overline{\overline{F}} \leqslant \overline{\overline{G}}\right) \Rightarrow \overline{\overline{E}} \leqslant \overline{\overline{G}}\,.\end{array}$$

Il serait maintenant intéressant de montrer qu'on peut comparer les puissances de deux ensembles quelconques. Autrement dit : étant donné deux ensembles E et F, ou bien $\overline{\overline{E}} \leqslant \overline{\overline{F}}$ ou bien $\overline{\overline{F}} \leqslant \overline{\overline{E}}$. Ce résultat sera obtenu comme application d'un théorème très important de théorie des ensembles attribué à Zorn.

5.1 Théorème de Zorn.

THÉORÈME 5.1.1 (ZORN)
Soit E un ensemble ordonné qui a la propriété suivante : tout sous-ensemble de E qui est totalement ordonné a un majorant. Alors E a un élément maximal.

Rappelons que, si E est un ensemble ordonné et X une partie de E, un majorant de X est un élément m de E tel que $m \geqslant x$ pour tout $x \in X$. Si, de plus, $m \notin X$, on dit que m est un majorant strict de X.

Un élément maximal de E est un élément a de E qui n'a pas de majorant strict (pour aucun x de E, on n'a $x > a$).

Avant de prouver le théorème de Zorn, donnons l'application indiquée à la comparaison des puissances de deux ensembles.

THÉORÈME 5.1.2
Quels que soient les ensembles E, F, il existe ou bien une injection de E dans F, ou bien une injection de F dans E.

(Théorème 5.1.2.) Soit $\mathscr{E}$ l'ensemble des applications f injectives donc le domaine est une partie A_f de E et l'image une partie B_f de F. On met sur $\mathscr{E}$ une relation d'ordre en posant :

$$f \leqslant g \Leftrightarrow \left(A_f \subset A_g \text{ et } f \text{ est la restriction de } g \text{ à } A_f\right)$$

autrement dit, $f \leqslant g$ si et seulement si g est un prolongement de f.

L'hypothèse du théorème de Zorn est alors satisfaite : en effet, soit $\mathscr{X}$ une partie de $\mathscr{E}$ totalement ordonnée. On pose $A = \bigcup_{f \in \mathscr{X}} A_f$ et on définit une fonction φ de domaine A en posant pour chaque $x \in A$:

$$\varphi(x) = f(x) \text{ pour n'importe quel } f \in \mathscr{X} \text{ tel que } x \in A_f\,.$$

En effet, si $x \in A_f \bigcap A_{f'}$, avec $f,\ f' \in \mathscr{X}$, comme $\mathscr{X}$ est totalement ordonné on a par exemple $f \leqslant f'$, donc $f(x) = f'(x)$) ; φ est le prolongement commun de tous les éléments de $\mathscr{X}$, donc est un majorant de $\mathscr{X}$.

Soit $f_0 : A_0 \to B_0$ un élément maximal de $\mathscr{E}$; il en existe d'après le théorème de Zorn. Si $A_0 = E$, f_0 est une injection de E dans F. Si $B_0 = F$, f_0^{-1} est une injection de F dans E.

Supposons alors que $A_0 \neq E$, $B_0 \neq F$. On prend $x_0 \in E - A_0$, $y \in F - B_0$ et on définir $g_0 : A_0 \cup \{x_0\} \to B_0 \cup \{y_0\}$ comme le prolongement de f qui donne à x_0 la valeur y_0. On a alors $g_0 \in \mathscr{E}$ et $g_0 > f_0$ ce qui contredit la maximalité de f_0. C.Q.F.D.

Démonstration du théorème de Zorn.

Soient E un ensemble ordonné, X une partie de E. Si l'ensemble des majorants de X a un plus petit élément, on l'appelle *borne supérieure de* X et on le désigne par $\sup(X)$. De même, si l'ensemble des minorants de X a un plus grand élément, on l'appelle *borne inférieure de* X et on le désigne par $\inf(X)$.

Un *élément minimal* de X est par définition, un élément de X qui n'a pas de minorant strict dans X.

Une partie de E qui est totalement ordonnée est appelée *une chaîne de* E. Deux éléments x, y de E sont dits *comparables* si on a $x \leqslant y$ ou $y \leqslant x$. Une chaîne de E est donc une partie de E dont les éléments sont deux à deux comparables.

On montre d'abord ce théorème.

THÉORÈME 5.1.3
Soit E un ensemble ordonné ayant les propriétés suivantes :

— *toute chaîne de E a une borne supérieure ;*

— *si $x \in E$ a un majorant strict, l'ensemble des majorants stricts de x a un élément minimal.*

Alors E a un élément maximal.

On raisonne par l'absurde, en supposant que tout élément de E a un majorant strict. En utilisant l'axiome du choix, on considère une fonction $f : \mathscr{P}(E) \to E$ telle que $f(X) \in X$ pour toute partie non vide X de E.

Pour chaque $x \in E$, soit M_x l'ensemble des majorants stricts minimaux de x. Par hypothèse $M_x \neq \emptyset$ pour tout $x \in E$.

On définit l'application $\mu : E \to E$ en posant $\mu(x) = f(M_x)$. il en résulte que $\mu(x)$ est un majorant strict minimal de x ; autrement dit, $x < \mu(x)$ et, si $y \in E$ est tel que $x \leqslant y \leqslant \mu(x)$, on a $x = y$ ou $y = \mu(x)$.

Soit $\mathscr{X}$ l'ensemble de toutes les parties X de E ayant les propriétés suivantes :

— si $Y \subset X$ est une chaîne, alors $\sup(Y) \in X$;

— si $x \in X$, $\mu(x) \in X$.

On désigne par X_0 l'intersection de tous les $X \in \mathscr{X}$, qui est non vide car $E \in \mathscr{X}$. On vérifie immédiatement que $X_0 \in \mathscr{X}$; X_0 est donc le plus petit élément de $\mathscr{X}$ (c'est-à-dire qu'il est inclus dans tout élément de $\mathscr{X}$).

Si on montre que X_0 est une chaîne, on aura la contradiction cherchée : en effet, d'après les propriétés définissant les éléments de $\mathscr{X}$, on aura alors $\sup(X_0) = a \in X_0$ d'où $\mu(a) \in X_0$. Or $\mu(a)$ est un majorant strict de $\sup(X_0)$, donc de X_0 et ne peut être élément de X_0.

Soit $x \in X_0$, x étant comparable à tout élément de X_0 ; alors pour tout $y \in X_0$ on a $y \leqslant x$ ou $y \geqslant \mu(x)$.

Soit $X = \{y \in X_0\ ;\ y \leqslant x$ ou $y \geqslant \mu(x)\}$. On montre que $X \in \mathscr{X}$ et donc $X \supset X_0$, ce qui est le résultat cherché.

En effet, soit $Y \subset X$, Y étant une chaîne. Si tout élément de Y est inférieur ou égal à x on a $\sup(Y) \leqslant x$, donc $\sup(Y) \in X$; si l'un des éléments de Y est supérieur ou égal à $\mu(x)$, on a $\sup(Y) \geqslant \mu(x)$, donc $\sup(Y) \in X$.

Soit $y \in X$.

— Si $y \geqslant \mu(x)$, on a $\mu(y) > y \geqslant \mu(x)$, donc $\mu(y) \in X$.

— Si $y \leqslant x$, on remarque que $\mu(y) \in X_0$ (car $y \in X_0$) donc $\mu(y)$ est comparable à x ; si $\mu(y) \leqslant x$ on a $\mu(y) \in X$; si $\mu(y) \geqslant x$, on a $y \leqslant x \leqslant \mu(y)$ et donc $y = x$ ou $x = \mu(y)$. Donc $\mu(y) = \mu(x)$ ou $\mu(y) = x$. Dans les deux cas $\mu(y) \in X$.

On a ainsi montré que :

si $x \in X_0$ est comparable à tout élément de X_0, il en est de même de $\mu(x)$.

Soit alors $Z = \{x \in X_0 \ ; \ x$ est comparable à tout élément de $X_0\}$. On a donc $x \in Z \Rightarrow \mu(x) \in Z$. D'autre part, soit Y une chaîne incluse dans Z et $u \in X_0$. Tout élément de Y est comparable à u. Si $y \leqslant u$ pour tout $y \in Y$ on a $\sup(Y) \leqslant u$. Si l'un des $y \in Y$ vérifie $y \geqslant u$ on a $\sup(Y) \geqslant u$. Dans les deux cas, $\sup(Y)$ est comparable à u. Comme u est un élément quelconque de X_0, on a $\sup(Y) \in Z$.

Finalement, on a montré que $Z \in \mathscr{X}$ donc $Z \supset X_0$. Cela prouve que tous les éléments de X_0 sont comparables, autrement dit que X_0 est une chaîne. C.Q.F.D.

THÉORÈME 5.1.4 (HAUSDORFF)
Soit C un ensemble ordonné. Alors l'ensemble des chaînes de C (ordonné par inclusion) a un élément maximal.

Soit E l'ensemble des chaînes de C. Il suffit de prouver que E satisfait aux hypothèses du théorème précédent.

Soit $X \subset E$ une chaîne de E, donc si $A, B \in X$ A et B sont des chaînes de C et on a $A \subset B$ ou $B \subset A$. il est alors immédiat de voir que $\bigcup_{A \in X} A$ est une chaîne de C qui est la borne supérieure de X.

Soit $A \in E$, A ayant un majorant strict B. On a donc $A \subset B$, $A \neq B$, A,B étant des chaînes de C. On choisit $a \in B - A$; alors $A \cup \{a\}$ est un majorant strict minimal de A. C.Q.F.D.

On peut alors montrer le théorème de Zorn (théorème 5.1.1) :

Soit E un ensemble ordonné, dont toute chaîne est majorée. Soit $A \subset E$ une chaîne maximale de E (obtenue à l'aide du théorème précédent). A a un majorant $a \in E$. Si a n'est pas maximal, on prend $b \in E$, $b > a$. Alors $b \notin A$ et $A \cup \{b\}$ est une chaîne qui contient strictement A, ce qui est impossible. Donc a est maximal. C.Q.F.D.

5.2 Applications du théorème de Zorn.

Les théorèmes suivants, qui sont des applications du théorème de Zorn, permettent d'évaluer la puissance de la réunion et du produit de deux ensembles infinis.

THÉORÈME 5.2.1
Soient E_1, E_2 deux ensembles équipotents à un ensemble infini E. Alors

$$\overline{\overline{E_1 \cup E_2}} = \overline{\overline{E}}$$

Supposons d'abord E_1, E_2 disjoints. On prend des bijections $f_1 : E \to E_1$, $f_2 : E \to E_2$. Pour chaque partie X de E, on désignera par X_1 (resp. X_2) l'image de X par f_1 (resp. f_2).

Soit $\mathscr{E}$ l'ensemble des bijections $\varphi : X \to X_1 \cup X_2$, X décrivant $\mathscr{P}(E)$. Si $\varphi, \psi \in \mathscr{E}$ on pose $\varphi \leqslant \psi$ si ψ est un prolongement de φ.

L' ensemble $\mathscr{E}$ est alors ordonné, non vide; car E est infini, donc il y a une partie dénombrable X, alors X_1 et X_2 sont dénombrables aussi, donc aussi $X_1 \cup X_2$ et il existe une bijection de X sur $X_1 \cup X_2$.

L'ensemble $\mathscr{E}$ satisfait l'hypothèse du théorème de Zorn : en effet si $\mathscr{X}$ est une partie totalement ordonnée de $\mathscr{E}$, on voit aisément que les $\varphi \in \mathscr{X}$ ont un prolongement commun qui est un majorant (et même la borne supérieure) de $\mathscr{X}$.

Soit alors $\mu : A \to A_1 \cup A_2$ un élément maximal de $\mathscr{X}$. Comme μ est bijective, on a $\overline{\overline{A}} = \overline{\overline{A_1 \cup A_2}}$. On va montrer que $E - A$ est un ensemble fini, donc aussi $E_1 - A_1$, $E_2 - A_2$. Il en résulte, d'après le théorème 4.2.4, que $\overline{\overline{E}} = \overline{\overline{A}}$, $\overline{\overline{E_1 \cup E_2}} = \overline{\overline{A_1 \cup A_2}}$ d'où le résultat cherché.

Supposons que $E - A$ soit infini et soit D une partie dénombrable de $E - A$. On pose $B = A \cup D$; on a $B_1 = A_1 \cup D_1$, $B_2 = A_2 \cup D_2$. Comme D, D_1 et D_2 sont dénombrables, il existe une bijection $\nu : D \to D_1 \cup D_2$. La fonction $\mu' : B \to B_1 \cup B_2$, égale à μ sur A et à ν sur D est bijective, donc $\mu \in \mathscr{E}$ et μ' est un majorant strict de μ ce qui est une contradiction.

Le cas où E_1 et E_2 ne sont pas disjoints est conséquence du corollaire suivant. C.Q.F.D.

COROLLAIRE 5.2.2
Soient $E_1, \ldots, E_n$ des ensembles tels que $\overline{\overline{E_1}} \geqslant \overline{\overline{E_2}}, \overline{\overline{E_1}} \geqslant \overline{\overline{E_3}}, \ldots, \overline{\overline{E_1}} \geqslant \overline{\overline{E_n}}$, E_1 étant infini. Alors :

$$\overline{\overline{E_1 \cup E_2 \cup \cdots \cup E_n}} = \overline{\overline{E_1}} .$$

On le montre d'abord lorsque $n = 2$. On choisit un ensemble E_2' équipotent à E_1 et disjoint de E_1. D'après le théorème précédent $E_1 \cup E_2'$ est équipotent à E_1.

Comme $\overline{\overline{E_2}} \leqslant \overline{\overline{E_2'}} = \overline{\overline{E_1}}$, il existe une surjection de E_2' sur E_2, donc une surjection de $E_1 \cup E_2'$ sur $E_1 \cup E_2$. Donc $\overline{\overline{E_1 \cup E_2}} \leqslant \overline{\overline{E_1 \cup E_2'}} = \overline{\overline{E_1}}$.

Comme $\overline{\overline{E_1}} \leqslant \overline{\overline{E_1 \cup E_2}}$ de façon évidente, on a bien $\overline{\overline{E_1 \cup E_2}} = \overline{\overline{E_1}}$.

On montre alors aisément le résultat par induction sur n. C.Q.F.D.

THÉORÈME 5.2.3
Soient E un ensemble infini, E_1, E_2 deux ensembles équipotents à E. Alors $\overline{\overline{E}} = \overline{\overline{E_1 \times E_2}}$.

On choisit deux bijections $f_1 : E \to E_1$, $f_2 : E \to E_2$; pour chaque partie X de E, on désignera par X_1 (resp. X_2) l'image de X par f_1 (resp. f_2).

Soit $\mathscr{E}$ l'ensemble des bijections $\varphi : X \to X_1 \times X_2$ (X décrit l'ensemble des parties de $\mathscr{E}$). Si φ, $\psi \in \mathscr{E}$ on pose $\varphi \leqslant \psi$ si ψ est un prolongement de φ.

L'ensemble $\mathscr{E}$ est alors ordonné non vide, car $\mathscr{E}$ étant infini, il a une partie dénombrable X, alors X_1, X_2 sont dénombrables aussi, donc aussi $X_1 \times X_2$ et il existe une bijection de X sur $X_1 \times X_2$.

L'ensemble $\mathscr{E}$ satisfait l'hypothèse du théorème de Zorn : en effet si $\mathscr{X}$ est une partie totalement ordonnée de $\mathscr{E}$, on voit aisément que les $\varphi \in \mathscr{X}$ ont un prolongement commun qui est un majorant de $\mathscr{X}$.

Soit alors $\mu : A \to A_1 \times A_2$ un élément maximal de $\mathscr{X}$. Comme μ est bijective, on a $\overline{\overline{A}} = \overline{\overline{A_1 \times A_2}}$. On va montrer que $\overline{\overline{E - A}} < \overline{\overline{A}}$. Il en résultera que $\overline{\overline{E}} = \overline{\overline{(E - A) \cup A}} = \overline{\overline{A}}$ (théorème précédent). Donc $\overline{\overline{A}} = \overline{\overline{A_1}} = \overline{\overline{A_2}} = \overline{\overline{E_1}} = \overline{\overline{E_2}} = \overline{\overline{E}}$. Comme $\overline{\overline{A}} = \overline{\overline{A_1 \times A_2}}$, on aura bien montré que $\overline{\overline{E}} = \overline{\overline{E_1 \times E_2}}$.

Supposons alors que $\overline{\overline{E - A}} \geqslant \overline{\overline{A}}$. Il en résulte qu'il existe $B \subset E - A$, $\overline{\overline{B}} = \overline{\overline{A}}$. On a alors :

$$B_1 \subset E_1 - A_1,\ B_2 \subset E_2 - A_2 ;$$

donc si on pose $C = A \cup B$, on a :

$$C_1 \times C_2 = (A_1 \times A_2) \cup (A_1 \times B_2) \cup (A_2 \times B_1) \cup (B_1 \times B_2) .$$

Les ensembles A, A_1, A_2, $A_1 \times A_2$, B_1, B_2 sont équipotents. Donc ils sont aussi équipotents à $A_1 \times B_2$, $A_2 \times B_1$, $B_1 \times B_2$.

D'après le théorème précédent, ils sont aussi équipotents à $(A_1 \times B_2) \cup (A_2 \times B_1) \cup (B_1 \times B_2)$. Soit alors ν une bijection de B sur ce dernier ensemble. La fonction μ' de domaine C qui est égale à μ sur A, et à ν sur B est donc une bijection de C sur $C_1 \times C_2$. Donc $\mu' \in \mathscr{E}$ et μ' est un majorant strict de μ ce qui est une contradiction. C.Q.F.D.

COROLLAIRE 5.2.4
Soient $E_1, \ldots, E_n$ des ensembles non vides tels que $\overline{\overline{E_1}} \geqslant \overline{\overline{E_2}}, \overline{\overline{E_1}} \geqslant \overline{\overline{E_3}}, \ldots, \overline{\overline{E_1}} \geqslant \overline{\overline{E_n}}$, E_1 étant infini. Alors

$$\overline{\overline{E_1 \times E_2 \times \cdots \times E_n}} = \overline{\overline{E_1}} .$$

On le montre par induction sur n : en admettant le résultat pour $n-1$, on a $\overline{\overline{E_1 \times E_2 \times \cdots \times E_{n-1}}} = \overline{\overline{E_1}}$. Donc

$$\overline{\overline{E_1 \times \cdots \times E_n}} = \overline{\overline{E_1 \times E_n}} \leqslant \overline{\overline{E_1 \times E_1}} \qquad (\text{puisque } \overline{\overline{E_n}} \leqslant \overline{\overline{E_1}}) .$$

Donc $\overline{\overline{E_1 \times \cdots \times E_n}} \leqslant \overline{\overline{E_1}}$. L'inégalité inverse est évidente (parce que $E_2, \ldots, E_n$ sont non vides). C.Q.F.D.