

Fondements des mathématiques

2. Premiers développements

Note de vocabulaire : les mots “proposition”, “théorème”, “lemme”, “corollaire” sont synonymes, et désignent un énoncé qu’on prouve vrai, aux nuances près qu’un théorème est plus important qu’une proposition, qu’un lemme sert à démontrer un théorème, et un corollaire en résulte. On appelle “axiome” un énoncé qu’on choisit (ou qu’on envisage) de supposer toujours vrai.

2.1. Quelques propriétés des quantificateurs

Soit E un ensemble. Sous-entendant le domaine E pour tous les quantificateurs, les propriétés suivantes utilisant suivant les cas, des relations unaires A, B sur E , et une variable propositionnelle C sont toujours vraies. Nous ne les justifierons que brièvement, laissant le lecteur compléter les explications et méthodes pour les déduire les unes des autres. (Il n’y a ici aucune notation de couple, les ponctuations articulent les quantificateurs avec leurs énoncés...).

Quelques évidences:

$$\begin{aligned} & \forall x, (\text{vrai}) \\ & \forall x, (A(x) \Rightarrow \exists y, A(y)) \\ & \forall x, ((\forall y, A(y)) \Rightarrow A(x)) \end{aligned}$$

Que C soit vrai ou faux, on a toujours

$$\begin{aligned} & (\exists x, C) \Rightarrow C \\ & C \Rightarrow \forall x, C \\ & (\exists x, C) \Leftrightarrow (C \text{ et } \exists x, \text{vrai}) \end{aligned}$$

Exemple: s’il existe une chaise telle que la Terre est ronde, alors la Terre est ronde. Puis, si la Terre est ronde, alors, quelle que soit la chaise que l’on considérerait, la Terre serait toujours ronde. Mais la réciproque est fautive: la phrase “si la Terre est ronde, alors il existe une chaise telle que la Terre est ronde”, est équivalente à “il existe une chaise, ou la Terre n’est pas ronde”.

Sur l’énoncé $(A(x), B(x))(y)$, de variables $x \in E, y \in \mathcal{V}$, on peut commuter les quantificateurs de même espèce (de $\forall x \forall y$ à $\forall y \forall x$, et de même pour les \exists):

$$\begin{aligned} & ((\exists x, A(x)) \text{ ou } (\exists x, B(x))) \Leftrightarrow (\exists x, A(x) \text{ ou } B(x)) \\ & ((\forall x, A(x)) \text{ et } (\forall x, B(x))) \Leftrightarrow (\forall x, A(x) \text{ et } B(x)) \end{aligned}$$

Autres en vrac:

$$\begin{aligned} & (\exists x, C \text{ ou } A(x)) \Rightarrow (C \text{ ou } \exists x, A(x)) \\ & (C \text{ et } \forall x, A(x)) \Rightarrow (\forall x, C \text{ et } A(x)) \\ & (\exists x, C \text{ et } A(x)) \Leftrightarrow (C \text{ et } \exists x, A(x)) \\ & (\forall x, C \text{ ou } A(x)) \Leftrightarrow (C \text{ ou } \forall x, A(x)) \\ & (\forall x, C \Rightarrow A(x)) \Leftrightarrow (C \Rightarrow \forall x, A(x)) \\ & (\forall x, A(x) \Rightarrow C) \Leftrightarrow ((\exists x, A(x)) \Rightarrow C) \\ & ((\forall x, A(x) \Rightarrow B(x)) \text{ et } (\forall x, A(x))) \Rightarrow (\forall x, B(x)) \\ & ((\forall x, A(x) \Rightarrow B(x)) \text{ et } (\exists x, A(x))) \Rightarrow (\exists x, B(x)) \\ & (\forall x, A(x) \text{ ou } B(x)) \Rightarrow ((\exists x, A(x)) \text{ ou } (\forall x, B(x))) \\ & ((\exists x, A(x)) \text{ et } (\forall x, B(x))) \Rightarrow \exists x, (A(x) \text{ et } B(x)) \\ & (\forall x, A(x)) \Rightarrow (\forall x, (A(x) \text{ ou } B(x))) \\ & (\exists x, A(x)) \Rightarrow (\exists x, (A(x) \text{ ou } B(x))) \\ & (\forall x, A(x) \text{ et } B(x)) \Rightarrow (\forall x, A(x)) \\ & (\exists x, A(x) \text{ et } B(x)) \Rightarrow (\exists x, A(x)) \end{aligned}$$

Un cas particulier de deux de ces formules donne les propriétés de distributivité entre les connecteurs (et) et (ou): pour trois variables propositionnelles A, B, C on a

$$\begin{aligned} ((A \text{ et } B) \text{ ou } C) &\Leftrightarrow ((A \text{ ou } C) \text{ et } (B \text{ ou } C)) \\ ((A \text{ ou } B) \text{ et } C) &\Leftrightarrow ((A \text{ et } C) \text{ ou } (B \text{ et } C)). \end{aligned}$$

Soient maintenant deux ensembles E et F et une relation R entre E et F . On a alors

$$(\exists x \in E, \forall y \in F, R(x, y)) \Rightarrow (\forall y \in F, \exists x \in E, R(x, y)).$$

Changements de domaine

Soient deux ensembles E et F , une application f de domaine E , et une relation unaire ou autre énoncé A de domaine contenant F ainsi que $\text{Im } f$. On a alors

$$\begin{aligned} (\exists y \in F, y \in \text{Im } f \text{ et } A(y)) &\Leftrightarrow \exists y \in F, \exists x \in E, f(x) = y \text{ et } A(y) \\ &\Leftrightarrow \exists x \in E, \exists y \in F, f(x) = y \text{ et } A(f(x)) \\ &\Leftrightarrow \exists x \in E, f(x) \in F \text{ et } A(f(x)) \end{aligned}$$

On en tire comme cas particuliers les conséquences suivantes.

Lemme. Pour tous ensembles $E \subset F$, et toute relation unaire A sur F ,

$$\begin{aligned} (\exists x \in E, A(x)) &\Leftrightarrow (\exists x \in F, (x \in E \text{ et } A(x))) \\ (\forall x \in E, A(x)) &\Leftrightarrow (\forall x \in F, (x \in E \Rightarrow A(x))) \end{aligned}$$

La première s'obtient par $f = \text{Id}_E$, et la deuxième s'en déduit en remplaçant A par sa négation. En procédant de même avec $A = \text{vrai}$, ou en passant par $\exists x \in E \cup F, x \in E$ et $x \in F$, on obtient $(\exists x \in E, x \in F) \Leftrightarrow (\exists x \in F, x \in E)$, d'où par négation

Définition. On dit que deux ensembles E et F sont disjoints ssi $(\forall x \in E, x \notin F)$, ce qui équivaut à $(\forall x \in F, x \notin E)$.

Passons à une troisième conséquence, qui peut d'ailleurs être intuitivement tirée de la définition de l'image d'une application f comme ensemble des valeurs possibles de $f(x)$.

Lemme. Soit une application f d'un ensemble E dans un ensemble F , et soit R une relation unaire sur F . Alors

$$\begin{aligned} (\exists x \in E, R(f(x))) &\Leftrightarrow (\exists y \in \text{Im } f, R(y)) \\ (\forall x \in E, R(f(x))) &\Leftrightarrow (\forall y \in \text{Im } f, R(y)) \end{aligned}$$

La deuxième formule est, là encore, simple reformulation de la première.

Ainsi, tout énoncé comportant une variable x apparaissant uniquement sous forme de $f(x)$ et liée par un quantificateur de même domaine que f , est équivalente à la formule modifiée en remplaçant le domaine du quantificateur par $\text{Im } f$ et en remplaçant $f(x)$ par x .

De là vient le fait que souvent on peut employer de manière semblable la notion d'ensemble et celle de famille: pour une famille u indexée par I , un énoncé avec quantificateur sur $i \in I$, où i n'apparaît que par son image u_i , est équivalent à l'énoncé obtenu en portant ce quantificateur sur $x \in \text{Im } u$, où u_i est remplacé par x . C'est ainsi par exemple que les notions d'unions et d'intersections s'appliqueront aussi bien aux familles d'ensembles qu'aux ensembles d'ensembles.

Quantificateur d'unicité

Pour tout objet x et tout ensemble F on a

$$\begin{aligned} \{x\} \subset F &\Leftrightarrow x \in F \\ F \subset \{x\} &\Leftrightarrow (\forall y \in F, y = x) \\ F = \{x\} &\Leftrightarrow (\{x\} \subset F \text{ et } F \subset \{x\}) \Leftrightarrow (x \in F \text{ et } \forall y \in F, y = x) \end{aligned}$$

L'énoncé $F \neq \emptyset$ que F a un élément, s'écrit $(\exists x \in F, \text{vrai})$.

On notera $\exists!(F)$ l'énoncé qu'il n'en a pas d'autre, i.e. que F un singleton:

$$\exists!(F) \Leftrightarrow \exists x \in F, \forall y \in F, x = y.$$

Le fait que F comporte au moins deux éléments, peut s'écrire $\exists x, y \in F, x \neq y$.

La négation de ce dernier énoncé, affirmation qu'il a au plus un élément (il est vide ou un singleton), sera notée $!(F)$:

$$!(F) \Leftrightarrow \forall x, y \in F, x = y$$

Soient un objet x , un ensemble F , et B une relation unaire ayant un sens sur F et aussi sur l'élément x . Alors:

Si $x \in F$ on a

$$(\forall y \in F, B(y)) \Rightarrow B(x) \Rightarrow (\exists y \in F, B(y))$$

Ces deux implications sont symétriques à travers le remplacement de B par non B . De là vient

$$F \neq \emptyset \Rightarrow ((\forall y \in F, B(y)) \Rightarrow (\exists y \in F, B(y))).$$

De même, si $\forall y \in F, y = x$, on a

$$(\exists y \in F, B(y)) \Rightarrow B(x) \Rightarrow (\forall y \in F, B(y)).$$

La deuxième implication dans le cas où $B(x)$ est l'énoncé $\forall y \in F, y = x$, donne

$$(\forall y \in F, y = x) \Rightarrow !(F)$$

ce qui permet de redémontrer en formules que $\exists!(F) \Leftrightarrow (F \neq \emptyset \text{ et } !(F))$ (" F est un singleton" équivaut à " F a un et un seul élément").

On peut voir directement, ou comme résultant de la dernière double implication (avec $x \in F$ si $F \neq \emptyset$ ou x quelconque sinon), que

$$!(F) \Rightarrow ((\exists y \in F, B(y)) \Rightarrow (\forall y \in F, B(y))).$$

Rassemblant tout cela, on a

$$\begin{aligned} F = \{x\} &\Rightarrow ((\exists y \in F, B(y)) \Leftrightarrow B(x) \Leftrightarrow (\forall y \in F, B(y))) \\ \exists!(F) &\Rightarrow ((\exists y \in F, B(y)) \Leftrightarrow (\forall y \in F, B(y))). \end{aligned}$$

Soit maintenant une relation unaire A sur un ensemble E , et soit $F = \{x \in E | A(x)\}$.

On notera " $!x \in E, A(x)$ " et on dira *il y a unicité de $x \in E$ tel que $A(x)$* , l'énoncé que F a au plus un élément, autrement dit

$$!x \in E, A(x) \Leftrightarrow !\{x \in E | A(x)\} \Leftrightarrow \forall x, y \in E, ((A(x) \text{ et } A(y)) \Rightarrow x = y).$$

De même on notera " $\exists!x \in E, A(x)$ " et on lira *il existe un unique $x \in E$ tel que $A(x)$* , l'énoncé que F est un singleton, ce qui s'écrit

$$\begin{aligned} \exists!x \in E, A(x) &\Leftrightarrow \exists!\{x \in E | A(x)\} \\ &\Leftrightarrow \exists x \in E, (A(x) \text{ et } \forall y \in E, (A(y) \Rightarrow y = x)) \\ &\Leftrightarrow ((\exists x \in E, A(x)) \text{ et } (!x \in E, A(x))). \end{aligned}$$

Puis, soient deux relations unaires A et B sur un ensemble E , et un élément $x \in E$ tels que $A(x)$ et $\forall y \in E, (A(y) \Rightarrow y = x)$. Ainsi $\forall y \in E, A(y) \Leftrightarrow y = x$. Alors on a les équivalences

$$(\exists y \in E, A(y) \text{ et } B(y)) \Leftrightarrow B(x) \Leftrightarrow (\forall y \in E, A(y) \Rightarrow B(y)).$$

De cette manière, si $\exists!x \in E, A(x)$ alors les énoncés B qu'on peut formuler sur l'unique objet x tel que $A(x)$ peuvent également s'exprimer sans le symbole de la variable libre " x ", au moyen d'une variable liée par un quantificateur, et de la relation unaire A .

Par ailleurs, en remplaçant $A(y)$ par son expression équivalente “ $y = x$ ”, on retrouve (dans le cas particulier $x \in E$) deux équivalences dont on avait déjà initialement vu la première

$$(\exists y \in E, y = x \text{ et } B(y)) \Leftrightarrow B(x) \Leftrightarrow (\forall y \in E, y = x \Rightarrow B(y)).$$

2.2. Opérations sur les ensembles; l’axiome des parties

Objets canoniques

On dira qu’un objet x est *canonique par rapport* à des objets désignés par des symboles, disons y, z , si x est le seul objet à satisfaire une certaine formule sans autre symbole de variable libre que y et z . Cette formule constitue donc une définition de x . En pratique, la mention des symboles y, z par rapport auxquels on considère la canonicité, est sous-entendue comme étant donnée par le contexte, à savoir que ce sont les symboles apparaissant dans les termes alors utilisés.

Ceci n’est pas une notion mathématique au sens où elle ne se situe pas au même niveau de théorie que les autres notions que nous exposons autour, mais une notion métamathématique (située au-dessus du langage de ces dernières). Précisément, c’est l’évocation d’une formule de définition implicitement utilisée, et qu’on ne prend pas nécessairement la peine de réexpliquer. Il faudrait réexpliquer cette formule pour traduire cette notion en un travail mathématique proprement dit au même niveau que le reste.

Par exemple, si on considère un singleton, son élément est canonique (par rapport à lui); mais si on considère un ensemble à plus d’un élément (autrement dit ni vide ni singleton), sans préciser d’autre contexte, et en particulier le considérant comme un ensemble d’éléments purs, alors aucun de ses éléments n’est canonique.

Union et intersection d’une famille d’ensembles

Soit une famille d’ensembles $(E_i)_{i \in I}$. Notons son ensemble image $\mathcal{E} = \{E_i | i \in I\}$. Pour tout objet x on a

$$(\exists i \in I, x \in E_i) \Leftrightarrow (\exists F \in \mathcal{E}, x \in F) \Leftrightarrow x \in \bigcup \mathcal{E}.$$

On définit alors l’*union de la famille* $(E_i)_{i \in I}$, notée $\bigcup_{i \in I} E_i$ (notation par laquelle on peut du même coup définir cette famille en remplaçant ici E_i par le terme qui le définit), comme étant $\bigcup \mathcal{E}$, de sorte qu’on a

$$x \in \bigcup_{i \in I} E_i \Leftrightarrow (\exists i \in I, x \in E_i).$$

Si $I \neq \emptyset$, on définit de même l’*intersection de la famille* $(E_i)_{i \in I}$ comme étant l’ensemble défini comme classe par l’énoncé de variable x :

$$x \in \bigcap_{i \in I} E_i \Leftrightarrow \forall i \in I, x \in E_i.$$

Ceci définit bien un ensemble pour la raison suivante: I étant non vide, soit un certain $j \in I$. Alors cet énoncé de définition (partie de droite) implique $x \in E_j$, de sorte que cette intersection peut se voir comme définie par compréhension dans E_j . De même on définit l’*intersection de* \mathcal{E} où \mathcal{E} est un ensemble non vide d’ensembles, comme classe des x tels que $\forall F \in \mathcal{E}, x \in F$, de sorte que si \mathcal{E} est comme ci-dessus l’image de la famille $(E_i)_{i \in I}$, son intersection est égale à celle de cette famille.

Par abus de langage, quand on étudie les intersections d’ensembles (ou de familles) de parties d’un ensemble E précis, on prolongera cette notion au cas de l’intersection de la famille vide (ou de l’ensemble vide) définie comme égale à E lui-même, entendant l’opération comme définie par compréhension dans E .

L’algèbre des parties d’un ensemble

Les parties d’un ensemble E se traduisant en relations unaires sur E , les opérations sur les relations unaires se traduisent en opérations entre parties. Appliquant des relations unaires sur E à une même variable libre $x \in E$, elles deviennent des variables propositionnelles libres entre lesquelles opèrent les connecteurs logiques. Soient A et B des parties de E , et soient $\mathcal{A} \Leftrightarrow x \in A$, $\mathcal{B} \Leftrightarrow x \in B$ leurs traductions en variables propositionnelles. Voici des opérations entre parties traduisant des connecteurs, à commencer par ceux d’arité 0 ou 1.

$$\begin{array}{l} \text{faux} \quad \emptyset \\ \text{non } \mathcal{A} \quad \complement_E A = E \setminus A \quad (\text{complémentaire de } A \text{ dans } E) \end{array}$$

Cette notion de complémentaire nécessite de préciser E afin de ramener la classe ainsi définie à un ensemble par compréhension. Pour éviter ce problème, seules seront nommées les opérations d'arité 2 entre parties issues des connecteurs qui donnent faux quand \mathcal{A} et \mathcal{B} sont faux. (En effet, la classe obtenue étant alors incluse dans $A \cup B$, définit dedans par compréhension un ensemble qui ne dépend que de A et B , et non de E):

\mathcal{A} ou \mathcal{B}	$A \cup B$	(union)
\mathcal{A} et \mathcal{B}	$A \cap B$	(intersection)
\mathcal{A} et (non \mathcal{B})	$A \setminus B$	(différence)
\mathcal{A} XOR \mathcal{B}	$A \Delta B$	(différence symétrique)

Dire que A et B sont disjoints se traduit par $A \cap B = \emptyset$.

On remarque que, tout comme les connecteurs (et) et (ou) sont des cas particuliers d'usage des quantificateurs (\forall) et (\exists), l'union (respectivement l'intersection) de deux ensembles est un cas particulier de celle d'une famille d'ensembles :

$$E \cup F = \bigcup(E, F) = \bigcup\{E, F\}$$

$$E \cap F = \bigcap(E, F) = \bigcap\{E, F\}.$$

La relation d'inclusion entre deux parties F et G d'un ensemble E se traduit par

$$F \subset G \Leftrightarrow (\forall x \in F, x \in G) \Leftrightarrow (\forall x \in E, x \in F \Rightarrow x \in G)$$

Tout comme avec les implications, on notera des chaînes d'inclusions successives entre ensembles :

$$F \subset G \subset H \Leftrightarrow (F \subset G \text{ et } G \subset H) \Rightarrow F \subset H.$$

De même, les opérations d'union et d'intersections sont associatives et distributives l'une sur l'autre:

$$(A \cup B) \cup C = A \cup (B \cup C) = \bigcup(A, B, C)$$

$$(A \cap B) \cap C = A \cap (B \cap C) = \bigcap(A, B, C)$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

ce qui tout comme pour les connecteurs est un cas particulier des formules plus générales

$$\left(\bigcup_{i \in I} A_i\right) \cap C = \bigcup_{i \in I} (A_i \cap C), \quad \left(\bigcap_{i \in I} A_i\right) \cup C = \bigcap_{i \in I} (A_i \cup C).$$

Axiomes des parties et de la puissance

Fixant deux ensembles E et F , considérons la classe des applications f de E vers F , i.e. telles que $\text{Dom } f = E$ et $\text{Im } f \subset F$. C'est un cas particulier de la notion de produit, évoquée précédemment. Or, pas plus que d'ensemble de tous les ensembles, il n'existe d'ensemble de toutes les applications dans lequel celui de toutes celles qui vont de E vers F se définirait par compréhension. Mais les conditions semblent restreindre la classe de manière appréciable, rendant envisageable qu'elle soit un ensemble, i.e. qu'on puisse abstraitement "trouver" tous ses éléments. Alors donc, est-ce un ensemble ? Dans toute la suite nous postulerons que oui, par les axiomes suivants:

Axiome des parties. *Pour tout ensemble E , la classe de toutes les parties de E est un ensemble, noté $\mathcal{P}(E)$. Formellement, pour tout F , on a: $F \in \mathcal{P}(E) \Leftrightarrow (F \text{ ensemble et } F \subset E)$.*

Axiome de la puissance. *Pour tous ensembles E et F , la classe des applications de E dans F est un ensemble noté F^E . Formellement, pour tout f on a: $f \in F^E \Leftrightarrow (f \text{ application, } \text{Dom } f = E \text{ et } \text{Im } f \subset F)$.*

En fait, ce ne sont pas exactement des axiomes, mais chacun est constitué d'un enrichissement du langage de la théorie des ensembles par un symbole d'opérateur (respectivement \mathcal{P} et le symbole invisible de puissance), suivi d'un axiome portant sur ce symbole. (On appelle ici *symbole d'opérateur*

un symbole de la théorie des ensembles au même titre que $=$ ou \in , i.e. quelque chose qui ressemble à un symbole d'opération mais dont les variables n'ont pas pour domaines des ensembles, de sorte qu'il ne s'agit d'opérations objets du même univers).

Cet ajout au langage de la théorie des ensembles, d'un symbole d'opérateur désignant l'ensemble postulé égal à une classe donnée dépendant de paramètres, est ainsi a priori nécessaire pour rendre effectif le postulat suivant lequel cette classe serait un ensemble. Pour expliquer cela, appelons *normal* un énoncé dont tous les quantificateurs ont pour domaines des ensembles. La question est donc de formuler l'égalité entre une classe d'énoncé $P(x)$ et un ensemble K . L'inclusion de K dans cette classe s'écrit $\forall x \in K, P(x)$. Mais, l'autre inclusion aurait la forme anormale $\forall x, P(x) \Rightarrow x \in K$ (avec un quantificateur universel sur l'univers).

Ce problème se résoud dans le cas des concepts d'union, de compréhension, d'image et de produit cartésien, car les quantificateurs de domaines les résultats de ces opérateurs sont remplaçables par des constructions n'utilisant que des quantificateurs rapportés aux ensembles initiaux. En effet: pour tout ensemble E d'ensembles, le quantificateur $\forall x \in \bigcup E$, est traduisible par $(\forall A \in E, \forall x \in A)$; pour toute application f , le quantificateur $\forall x \in \text{Im } f, \dots x$ se traduit par $\forall x \in \text{Dom } f, \dots f(x)$; et on peut en faire autant avec la compréhension et le produit cartésien.

Mais il y a d'autres classes comme celles des parties d'un ensemble ou des applications d'un ensemble dans un autre, telles que même en admettant qu'elles soient par ailleurs des ensembles (si une telle hypothèse pouvait trouver un sens), un quantificateur indéfini portant sur cette classe ne peut pas se traduire en une construction d'énoncé équivalent normal. Il a alors un grand risque qu'un tel ensemble ne soit identifiable par (i.e. ne soit l'unique objet qui satisfasse) aucun énoncé normal. Seul le fait d'ajouter de l'extérieur un symbole spécifique pour désigner les ensembles en question, et d'inclure l'énoncé anormal qui les identifie, comme axiome de la théorie des ensembles, permet de les reconnaître effectivement comme tels, ce que ne permettrait pas le seul axiome de leur existence.

En l'occurrence, l'axiome des parties ne serait au fond que l'énoncé d'une relation entre $\mathcal{P}(E)$ et l'univers, telle que dans chaque univers donné il existe au plus un objet $\mathcal{P}(E)$ qui la satisfasse. Finalement, il définit $\mathcal{P}(E)$ comme dépendant de l'univers. Or, l'intention profonde serait de le concevoir comme objet absolument défini sans considération de l'univers, lequel serait pour cela supposé suffisamment grand pour contenir "vraiment" toutes les parties de E , et par là, le "vrai" $\mathcal{P}(E)$. Donc, lorsqu'on parlera de $\mathcal{P}(E)$, on oubliera qu'il s'agit en fait d'un aspect de l'univers.

Les "axiomes" composites des parties et de la puissance, bien que n'étant ainsi pas vraiment des énoncés, sont néanmoins "équivalents", autrement dit il suffit d'en postuler un des deux pour que l'autre en résulte. Voici ce que cela signifie: l'axiome de la puissance "implique" l'axiome des parties, en ce sens qu'il est possible de désigner l'ensemble $\mathcal{P}(E)$ en termes de l'opération de puissance (\mathcal{V}^E). Réciproquement, on peut désigner l'ensemble puissance F^E en termes de $\mathcal{P}(E \times F)$, au moyen de la représentation des applications par leur graphe (comme précisé plus loin).

Cet "axiome" des parties ou de la puissance, enrichissant le langage de la théorie des ensembles, est introduit parce qu'il est indispensable pour pouvoir s'exprimer raisonnablement en mathématiques, en énonçant les nombreuses définitions dont on a besoin. Sans lui, on ne pourrait plus dire grand-chose. On ne pourrait pas toujours distinguer si un ensemble est fini ou infini: un ensemble ne pourrait être déclaré fini que s'il est construit "à la main", élément par élément, ou du moins s'il satisfait une propriété qui limite formellement son nombre d'éléments. Il n'y aurait plus de \mathbb{N} , ni encore moins de définition de suites par récurrence. On pourrait seulement établir l'existence d'un ensemble puissance F^E lorsqu'on saurait que E est fini.

Heureusement, accepter le langage et les axiomes des parties et de la puissance n'entraîne pas de contradiction (du moins on espère qu'on en trouvera pas, tout en sachant que la non-contradiction d'une théorie des ensembles comme celle-ci ou d'autres est de toute manière indémontrable).

Mais, lorsqu'on explore les fondements des mathématiques, on découvre que la question posée par ces axiomes, de la désignation ou de l'existence des ensembles des parties ou des ensembles puissances, est un problème central, sur lequel repose l'essentiel de l'incomplétude des mathématiques.

Précisément, contrairement au cas d'un ensemble démontrablement fini évoqué ci-dessus, si E est un ensemble infini et F a plus d'un élément, il s'avère impossible de formaliser totalement l'affirmation qu'un ensemble donné d'applications de E dans F , même noté F^E , contient réellement toutes les applications de E dans F : bien qu'effectivement d'après l'axiome il contient toutes celles qui sont dans notre univers, on ne peut exclure qu'il puisse exister ailleurs, dans un autre univers plus grand, d'autres applications de E dans F qui n'appartiennent pas à notre F^E . L'axiome de la puissance peut être aussi vrai dans cet autre univers, mais avec une autre interprétation des ensembles de

parties et de puissance. C'est par ces jeux de Grandes Illusions se démontrent nombre de résultats d'indécidabilités des mathématiques (un énoncé est dit indécidable si ni lui ni sa négation n'est démontrable).

Alors donc, nous poserons les axiomes (équivalents) des parties et de la puissance parce que nous en aurons besoin au départ des mathématiques. En aurons-nous vraiment besoin ? Si on examine attentivement les mathématiques courantes, et même celles qui constituent tout l'essentiel du cycle fondateur des mathématiques, il s'avère que, en gros, on n'utilise guère que les ensembles finis, l'ensemble \mathbb{N} des entiers naturels et l'ensemble $\mathcal{P}(\mathbb{N})$ de ses parties (qui permet de construire l'ensemble \mathbb{R} des nombres réels); qu'il est rare d'aller plus loin, et que même si on se permet parfois de parler de "parties quelconques de \mathbb{R} ", ce n'est souvent en pratique que pour en dire des choses qu'on pourrait avec plus ou moins de difficultés réécrire en termes de \mathbb{R} seul, sans lui appliquer l'axiome des parties. Mais ici, comme d'ailleurs suivant la tradition ZF, nous accepterons l'axiome des parties, lequel permet à partir de \mathbb{N} d'invoquer non seulement $\mathcal{P}(\mathbb{N})$ mais aussi $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$ et ainsi de suite, donc bien plus que ce qui est nécessaire en pratique.

Pourrait-on alors se contenter d'un axiome plus faible ? Bien sûr, sauf que pour cela il faudrait écrire par exemple "on désigne par $\mathcal{P}(\mathbb{N})$ l'ensemble de toutes les parties de \mathbb{N} " (ce qui est d'ailleurs un pléonasme puisqu'on utilise un énoncé avec quantificateur sur $\mathcal{P}(\mathbb{N})$ pour formuler que \mathbb{N} désigne bien l'ensemble des entiers naturels), ou encore "il existe un ensemble infini dont on a l'ensemble des parties" (avec une remarque du même style). Mais de telles subtilités n'ont guère d'intérêt dans une première approche des mathématiques, et, étant donnée la relative complexité de l'introduction de \mathbb{N} ou de la notion d'ensemble infini, compliqueraient encore l'exposé, par les précautions prises à marcher sur des oeufs. Or nous voulions partir des fondements les plus simples possibles. Par mesure de simplicité donc, nous accepterons l'axiome des parties dans son intégralité.

Nous n'avons pas encore ici complété la formulation d'une théorie des ensembles suffisante pour fonder les mathématiques. Pour cela, il restera à poser l'axiome de l'infini: "Il existe un ensemble infini" (une fois définie la notion d'ensemble infini), permettant de construire \mathbb{N} et bien d'autres ensembles intéressants. Car bien sûr, sans ensemble infini, toute la force de l'axiome des parties que nous venons de commenter resterait quasiment sans objet.

On abrègera l'expression $\forall A \in \mathcal{P}(E)$ en $\forall A \subset E$, et de même pour \exists .

Produit

Montrons que l'axiome de la puissance est aussi "équivalent" à la donnée des ensembles produit.

On définit le produit de toute famille d'ensembles $(E_i)_{i \in I}$, grâce aux axiomes de la réunion et de la puissance, par

$$\prod_{i \in I} E_i = \{f \in (\bigcup_{i \in I} E_i)^I \mid \forall i \in I, f(i) \in E_i\}$$

$$f \in \prod_{i \in I} E_i \Leftrightarrow (f \text{ application, } \text{Dom } f = I \text{ et } \forall i \in I, f(i) \in E_i)$$

Réciproquement, l'ensemble puissance F^E est égal à au produit de la famille constante $\prod_{i \in I} F$.

Pour tout $i \in I$ on appelle *i-ième projection canonique*, l'application π_i de $\prod_{i \in I} E_i$ dans E_i qui à toute famille x associe l'image de i par x : $\pi_i(x) = x_i$.

On peut voir cela comme une sorte de renversement des rôles entre l'application (famille) et sa variable (son indice), c'est-à-dire, si on voit l'expression x_i comme une opération entre les deux objets x et i , le fait de fixer i et laisser x variable de domaine le produit, au lieu de fixer x et de laisser i variable comme on le conçoit à la base.

Somme ou union disjointe

Etant donnée une famille $(E_i)_{i \in I}$ d'ensembles, on définit leur *somme* ou *union disjointe* (même si les E_i ne sont pas disjoints) comme étant l'union de copies des E_i deux à deux disjointes. Comment construit-on de telles copies ? Un élément d'une telle copie comporte deux informations: l'indice i et l'élément x de E_i qu'il représente. C'est donc le couple (i, x) :

$$\coprod_{i \in I} E_i = \{(i, x) \mid i \in I \text{ et } x \in E_i\} \subset I \times \bigcup_{i \in I} E_i$$

$$\coprod_{i \in I} E = I \times E.$$

Opérations et inclusions

$$\begin{aligned}
 F \subset F' &\Rightarrow F^E \subset F'^E \\
 E \subset E', F \subset F' &\Rightarrow E \times F \subset E' \times F' \\
 (\forall i \in I, E_i \subset E'_i) &\Rightarrow \prod_{i \in I} E_i \subset \prod_{i \in I} E'_i \quad \text{et} \quad \prod_{i \in I} E_i \subset \prod_{i \in I} E'_i.
 \end{aligned}$$

2.3. Etude des applications

Graphe d'une application

Pour toute relation R entre E et F , on avait déjà défini son graphe comme étant

$$\text{Gr}(R) = \{(x, y) \in E \times F \mid R(x, y)\}.$$

Soit f une application de E dans F . On appelle *graphe de f* l'ensemble (indépendant de f)

$$\text{Gr } f = \text{Im}(\text{Id}_E \times f) = \{(x, f(x)) \mid x \in E\} \subset E \times F$$

$$\begin{aligned}
 \forall x \in E, \forall y \in F, \quad (x, y) \in \text{Gr } f &\Leftrightarrow \exists x' \in E, (x, y) = (x', f(x')) \\
 &\Leftrightarrow \exists x' \in E, x = x' \text{ et } y = f(x'). \\
 &\Leftrightarrow y = f(x)
 \end{aligned}$$

Pour tous $f, g \in F^E$ et toutes relations R, S entre E et F on a

$$\begin{aligned}
 \text{Gr}(R) \subset \text{Gr}(S) &\Leftrightarrow \forall x \in E, \forall y \in F, (R(x, y) \Rightarrow S(x, y)) \\
 \text{Gr}(R) = \text{Gr}(S) &\Leftrightarrow \forall x \in E, \forall y \in F, (R(x, y) \Leftrightarrow S(x, y)) \\
 \text{Gr}(f) \subset \text{Gr}(R) &\Leftrightarrow \forall x \in E, R(x, f(x)) \\
 \text{Gr}(R) \subset \text{Gr}(f) &\Leftrightarrow \forall x \in E, \forall y \in F, (R(x, y) \Rightarrow y = f(x)) \\
 \text{Gr}(f) = \text{Gr}(R) &\Leftrightarrow \forall x \in E, \forall y \in F, (R(x, y) \Leftrightarrow y = f(x)) \\
 \text{Gr}(f) \subset \text{Gr}(g) &\Leftrightarrow f = g.
 \end{aligned}$$

La dernière ligne se vérifie en écrivant

$$\text{Gr}(f) \subset \text{Gr}(g) \Leftrightarrow (\forall x \in E, (x, f(x)) \in \text{Gr}(g)) \Leftrightarrow (\forall x \in E, f(x) = g(x)) \Leftrightarrow f = g.$$

De plus, deux applications ayant même graphe sont égales. En effet, elles ont même domaine $\text{Dom } f = \pi_1[\text{Gr } f]$, puis le reste vient d'être vu.

Comme toute partie $P \subset E \times F$ est le graphe d'une unique relation $(x, y) \mapsto ((x, y) \in P)$ entre E et F , la donnée d'une application f de E dans F se traduit donc, à travers son graphe, sous forme de la relation $(y = f(x))$ entre E et F .

Soit R une relation entre deux ensembles E et F . Alors on a

$$(\forall x \in E, \exists! y \in F, R(x, y)) \Rightarrow \exists! f \in F^E, \text{Gr}(f) \subset \text{Gr}(R)$$

En effet, si la partie de gauche est vraie alors $\forall f, g \in F^E, \text{Gr}(f) \subset \text{Gr}(R)$ et $\text{Gr}(g) \subset \text{Gr}(R) \Rightarrow (\forall x \in E, R(x, f(x)) \text{ et } R(x, g(x))) \Rightarrow \forall x \in E, f(x) = g(x) \Rightarrow f = g$.

Proposition. Soit R une relation entre deux ensembles E et F . Il y a équivalence entre

1. $\forall x \in E, \exists! y \in F, R(x, y)$
2. $\exists f \in F^E, \text{Gr}(f) = \text{Gr}(R)$
3. $\exists! f \in F^E, \text{Gr}(f) = \text{Gr}(R)$
4. $\exists! f \in F^E, \text{Gr}(f) \subset \text{Gr}(R)$.

Preuves:

1. \Rightarrow 2. peut être vue comme un axiome: l'expression du fait que lorsqu'on trouve une relation qui a manifestement la propriété qui en fait un graphe d'application, alors, l'application qu'elle définit, sera effectivement reconnue comme application qui existe formellement dans F^E .

2. \Rightarrow 1. est immédiat: $\forall x \in E, \exists! y \in F, y = f(x)$.

2. \Leftrightarrow 3. est évident.

(1. et 2.) \Rightarrow 4. : 2. donne l'existence du f ; 1. donne son unicité par le résultat plus haut.

4. \Rightarrow 2. : soit f tel que $\text{Gr}(f) \subset \text{Gr}(R)$. Alors $\forall(x, y) \in \text{Gr}(R), \text{Gr}(x' \mapsto (y, f(x'))(x = x')) \subset \text{Gr}(R)$, donc $f = (x' \mapsto (y, f(x'))(x = x'))$, donc $y = f(x)$. Finalement $\text{Gr}(f) = \text{Gr}(R)$.

Les cours traditionnels présentent la notion d'application comme n'étant pas première, mais comme ramenée à la notion d'ensemble au moyen du graphe. Or cela pose deux problèmes: d'une part, pour traiter de même la notion de relation, il faudrait ajouter à la donnée de son graphe celle des domaines de ses variables, ce qui en fait une notion composite; d'autre part cela utilise la notion de couple. Nous n'avons pas procédé ainsi à cause de l'intérêt de définir les couples comme cas particuliers d'applications, et des spécificités de la notion d'application avec ses usages et notations, auxquels le formalisme des relations n'est guère adapté.

Injections, surjections, bijections

Théorème et définition. *Pour toute application f d'un ensemble E dans un ensemble F , les énoncés suivants sont équivalents, et seront désignés en disant que f est injective (ou : une injection):*

$$\begin{aligned} \forall x, x' \in E, f(x) = f(x') \Rightarrow x = x' \\ \forall y \in F, \exists! x \in E, f(x) = y \end{aligned}$$

Preuve: les équivalences s'enchaînent ainsi (utilisant que $f(x) \in F$):

$$\begin{aligned} \forall y \in F, \exists! x \in E, f(x) = y &\Leftrightarrow \forall y \in F, \forall x, x' \in E, (f(x) = y \text{ et } f(x') = y) \Rightarrow x = x' \\ &\Leftrightarrow \forall x, x' \in E, \forall y \in F, f(x) = y \Rightarrow (f(x') = y \Rightarrow x = x') \\ &\Leftrightarrow \forall x, x' \in E, f(x) = f(x') \Rightarrow x = x' \end{aligned}$$

Remarque: la première formule peut aussi s'écrire sous la forme plus intuitive mais moins utilisée en pratique dans les démonstrations:

$$\forall x, x' \in E, x \neq x' \Rightarrow f(x) \neq f(x').$$

Définition. *On dit qu'une application de E dans F est surjective (ou une surjection) lorsque $\text{Im } f = F$, autrement dit lorsque $\forall y \in F, \exists x \in E, f(x) = y$.*

C'est donc en quelque sorte non une propriété de f en elle-même mais une relation entre f et l'ensemble d'arrivée mentionné. On dit aussi, pour insister, une surjection de E sur F .

Une *application bijective* (ou *bijection*) f de E sur F est une application injective et surjective de E sur F , autrement dit telle que $\forall y \in F, \exists! x \in E, f(x) = y$.

Une bijection d'un ensemble sur lui-même s'appelle une *permutation* (on dit aussi une *transformation* pour un espace géométrique).

Identité, composition et restriction

Pour toutes applications $f \in F^E$ et $g \in G^F$, on définit leur *composée* $g \circ f \in G^E$ par :

$$g \circ f = (E \ni x \mapsto g(f(x))).$$

De même pour la composée de toute chaîne d'applications entre ensembles successifs :

$$h \circ g \circ f = (h \circ g) \circ f = h \circ (g \circ f) = (\text{Dom } f \ni x \mapsto h(g(f(x)))).$$

Pour tout ensemble E on appelle *identité sur E* l'application

$$\text{Id}_E = (E \ni x \mapsto x) \in E^E.$$

qu'on appellera aussi l'*injection canonique* de E dans tout ensemble dans lequel E est inclus.

Pour toute application $f \in F^E$ et $A \subset E$, on appelle *restriction de f à A* l'application notée $f|_A \in F^{E'}$, définie par

$$f|_A = (A \ni x \mapsto f(x)) = f \circ \text{Id}_A \in F^A.$$

ce qui définit une surjection de F^E sur F^A si $F \neq \emptyset$ (voir plus loin un résultat plus général).

Image directe, image réciproque

Soient $f \in F^E$ et $B \subset F$. On appelle *image réciproque* de B par f et on note $f^*(B)$ l'ensemble

$$f^*(B) = \{x \in E \mid f(x) \in B\}.$$

Cette notation f^* ne peut être vue comme désignant une application qu'à condition de choisir un ensemble d'arrivée F de f , dont l'ensemble des parties servira de domaine.

Si $A \subset B \subset F$ alors $f^*(A) \subset f^*(B)$ et $f^*(\mathbb{C}_F A) = \mathbb{C}_F f^*(A)$.

Pour toute famille d'ensembles $(A_i)_{i \in I}$,

$$\begin{aligned} f^*\left(\bigcup_{i \in I} A_i\right) &= \bigcup_{i \in I} f^*(A_i) \\ f^*\left(\bigcap_{i \in I} A_i\right) &= \bigcap_{i \in I} f^*(A_i). \end{aligned}$$

Soit maintenant un ensemble $A \subset E$. On appelle *image directe* de A par f et on note $f[A]$ l'ensemble

$$f[A] = \text{Im}(f|_A) = \{f(x) \mid x \in A\} = \{y \in F \mid \exists x \in A, y = f(x)\} \subset \text{Im } f \subset F.$$

(Cette notation avec crochets, évitant les ambiguïtés de l'usage classique des parenthèses, est celle du Wikipedia anglophone.) C'est une surjection $f|_{\mathbb{P}}$ de $\mathcal{P}(E)$ sur $\mathcal{P}(\text{Im } f)$ car $\forall B \subset \text{Im } f, f[f^*(B)] = B$.

Pour tous $A \subset B \subset E$ on a $f[A] \subset f[B]$. Pour toute famille $(A_i)_{i \in I}$ de parties de E ,

$$\begin{aligned} f\left[\bigcup_{i \in I} A_i\right] &= \bigcup_{i \in I} f[A_i] \\ f\left[\bigcap_{i \in I} A_i\right] &\subset \bigcap_{i \in I} f[A_i] \end{aligned}$$

où l'inclusion devient une égalité notamment si (f est injective et $I \neq \emptyset$).

Proposition. Soient deux applications $f \in F^E$ et $g \in G^F$. On a:

- 1) Si f et g sont injectives alors $g \circ f$ est injective.
- 2) $\text{Im}(g \circ f) = g[\text{Im } f] \subset \text{Im } g$
- 3) Si f est surjective (i.e. $\text{Im } f = F$) alors $\text{Im}(g \circ f) = \text{Im } g$.
- 4) Si f et g sont surjectives alors $g \circ f$ est surjective (i.e. $\text{Im}(g \circ f) = G$).
- 5) Si $g \circ f$ est surjective alors g est surjective.
- 6) Si $g \circ f$ est injective alors f est injective.
- 7) Si f et g sont bijectives alors $g \circ f$ est bijective.

Preuves:

- 1) Si f et g sont injectives, $\forall x, y \in E, g(f(x)) = g(f(y)) \Rightarrow f(x) = f(y) \Rightarrow x = y$.
- 2) $\forall z \in G, z \in \text{Im}(g \circ f) \Leftrightarrow (\exists x \in E, g(f(x)) = z) \Leftrightarrow (\exists y \in \text{Im } f, g(y) = z) \Leftrightarrow z \in g[\text{Im } f]$.
- 3) résulte de 2)
- 4) résulte de 3)
- 5) résulte de 2)
- 6) Si $g \circ f$ est injective alors $\forall x, y \in E, f(x) = f(y) \Rightarrow (g \circ f)(x) = (g \circ f)(y) \Rightarrow x = y$.
- 7) vient de 1) et 4). □

Transposition

Une transposition est une permutation qui échange deux éléments et laisse fixe les autres.

En particulier, dans une paire on s'intéressera à l'unique transposition, qu'on notera σ . C'est l'unique cas d'une permutation d'un ensemble sans structure, qui soit canonique et différente de l'identité. Soit maintenant une opération f à deux variables de domaines E et F . Voyant f comme application de domaine $E \times F$, on appellera *transposée de f* l'opération ${}^t f$ entre F et E obtenue par transposition des positions des variables dans l'écriture de f :

$$\forall y \in F, \forall x \in E, {}^t f(y, x) = f(x, y) = f((y, x) \circ \sigma).$$

Ceci s'applique en particulier au cas d'une relation entre deux ensembles, et, suivant la correspondance entre ces relations et les ensembles de couples, on parlera aussi de transpositions sur les ensembles C de couples:

$${}^tC = \{(y, x) | (x, y) \in C\}.$$

Inversion d'applications et propriétés de la composition

Soient E et F deux ensembles, $f \in F^E$ et $g \in E^F$. Alors on a équivalence entre

- 1) $g \circ f = \text{Id}_E$
- 2) $\forall x \in E, \forall y \in F, f(x) = y \Rightarrow g(y) = x$
- 3) $\text{Gr } f \subset {}^t\text{Gr } g$.
- 4) $\forall y \in F, f^*(\{y\}) \subset \{g(y)\}$
- 5) $\forall x \in E, f(x) \in g^*(\{x\})$

Vérification facile, dans l'ordre suivant: $5 \Leftrightarrow 1 \Leftrightarrow 2 \Leftrightarrow 3 \Leftrightarrow 4$ (et aussi $2 \Leftrightarrow 4$).

On remarque que ces conditions impliquent que f est injective.

De même en combinant ces énoncés avec ceux où on échange f et g : on a équivalence entre

- 1) $g \circ f = \text{Id}_E$ et $f \circ g = \text{Id}_F$.
- 2) $\forall x \in E, \forall y \in F, f(x) = y \Leftrightarrow g(y) = x$
- 3) $\text{Gr } g = {}^t\text{Gr } f$.
- 4) $\forall y \in F, f^*(\{y\}) = \{g(y)\}$
- 5) $\forall x \in E, \{f(x)\} = g^*(\{x\})$

D'après les propriétés des graphes d'applications,

$$\begin{aligned} \forall f \in F^E, \quad f \text{ bijective} &\Leftrightarrow \exists g \in E^F, \text{Gr } g = {}^t\text{Gr } f \\ &\Leftrightarrow \exists !g \in E^F, \text{Gr } g = {}^t\text{Gr } f \\ &\Leftrightarrow \exists !g \in E^F, \text{Gr } g \subset {}^t\text{Gr } f. \end{aligned}$$

Par ailleurs, on remarque que quelles que soient les applications f et g sans hypothèse sur les domaines et images, l'égalité $\text{Gr } g = {}^t\text{Gr } f$ implique à elle seule que $\text{Dom } f = \text{Im } g$, $\text{Dom } g = \text{Im } f$ et que f et g sont injectives. Toute injection pouvant être regardée comme bijective sur son image, on peut poser:

Définition. Pour toute injection f , on appelle inverse de f et on note f^{-1} l'application définie par $\text{Gr}(f^{-1}) = {}^t\text{Gr } f$; elle est bijective de $\text{Im } f$ sur $\text{Dom } f$.

Comme la condition $\text{Gr } g = {}^t\text{Gr } f$ équivaut à $\text{Gr } f = {}^t\text{Gr } g$, on a $(f^{-1})^{-1} = f$.

Proposition. Soient deux ensembles E et F et trois applications $f, h \in F^E, g \in E^F$ telles que $g \circ f = \text{Id}_E$ et $h \circ g = \text{Id}_F$. Alors $f = h$, de sorte que f et g sont l'inverse l'un de l'autre.

Preuve: $\forall x \in E, f(x) = h(g(f(x))) = h(x)$.

Proposition. Soient deux applications $f \in F^E$ et $g \in G^F$ bijectives. Alors $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

On peut écrire la preuve

$$\forall x \in E, \forall y \in G, g \circ f(x) = y \Leftrightarrow f(x) = g^{-1}(y) \Leftrightarrow x = f^{-1} \circ g^{-1}(y)$$

ou encore $(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ \text{Id}_F \circ g^{-1} = \text{Id}_G$, et de même $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \text{Id}_E$.

Théorème. Soient trois ensembles E, F, G , soit $f \in F^E$, et soit $\phi = (G^F \ni g \mapsto g \circ f)$ l'application de composition à droite par f , arrivant dans G^E . On a alors:

- 1) Si f est surjective alors ϕ est injective
- 2) Si f est injective et $G \neq \emptyset$ alors ϕ est surjective
- 3) Si ϕ est injective et $\exists z, z' \in G, z \neq z'$ alors f est surjective.
- 4) Si ϕ est surjective et $\exists z, z' \in G, z \neq z'$ alors f est injective.

Preuves:

- 1) $\forall g, h \in G^F, \phi(g) = \phi(h) \Leftrightarrow (\forall x \in E, g(f(x)) = h(f(x))) \Leftrightarrow \forall y \in F, g(y) = h(y) \Leftrightarrow g = h$.
- 2) Soient $h \in G^E$ et $z \in G$. Alors, f étant injective, $\phi(F \ni y \mapsto (h \circ f^{-1}(y), z)(y \in \text{Im } f)) = h$.
- 3) $\phi(y \mapsto z) = \phi(y \mapsto (z, z')(y \in \text{Im } f)) \Rightarrow (\forall y \in F, (y \in \text{Im } f \text{ ou } z = z')) \Rightarrow \text{Im } f = F$.
- 4) $\forall x \in E, \exists g \in G^F, \forall y \in E, g(f(y)) = (z, z')(y = x)$ donc $f(y) = f(x) \Rightarrow g(f(y)) = g(f(x)) = z \Rightarrow y = x$.

Prenant dans 2) le cas $G = E$ on a en particulier

Corrolaire 1. Soit une injection $f \in F^E$ où $E \neq \emptyset$, alors $\exists g \in E^F, g \circ f = \text{Id}_E$.

Par ailleurs le cas $G = \mathcal{V}$ se traduit par

Corrolaire 2. Soient deux ensembles E, F , soit $f \in F^E$, et considérons f^* comme application de $\mathcal{P}(F)$ dans $\mathcal{P}(E)$. Alors on a (f injective $\Leftrightarrow f^*$ surjective), et (f surjective $\Leftrightarrow f^*$ injective).

Théorème. Soient trois ensembles E, F, G , soit $g \in G^F$, et soit $\psi = (F^E \ni f \mapsto g \circ f)$ l'application de composition à gauche par g , arrivant dans G^E . On a alors:

- 1) Si g est injective alors ψ est injective
- 2) (Si g est surjective alors ψ est surjective) est une expression de l'axiome du choix.
- 3) Si ψ est injective et $E \neq \emptyset$ alors g est injective.
- 4) Si ψ est surjective et $E \neq \emptyset$ alors g est surjective.

Preuves:

- 1) $\forall f, f' \in F^E, \psi(f) = \psi(f') \Leftrightarrow \forall x \in E, g(f(x)) = g(f'(x)) \Rightarrow \forall x \in E, f(x) = f'(x) \Rightarrow f = f'$.
- 2) sera étudié avec l'axiome du choix.
- 3) $\forall y, y' \in F, g(y) = g(y') \Rightarrow \psi(x \mapsto y) = \psi(x \mapsto y') \Rightarrow (\forall x \in E, y = y') \Rightarrow y = y'$ car $E \neq \emptyset$.
- 4) $\forall z \in G, \exists f \in F^E, g \circ f = (x \mapsto z)$ donc $E \neq \emptyset \Rightarrow z \in \text{Im } g$.

Proposition. Soient $f \in F^E, g \in E^F$ tels que $g \circ f = \text{Id}_E$. Alors f est injective, g est surjective, et on a les équivalences : (f surjective) \Leftrightarrow (g injective) $\Leftrightarrow f \circ g = \text{Id}_F$.

Preuve:

Les premiers résultats découlent de l'injectivité et la surjectivité de $\text{Id}_E = g \circ f$.

De $f \circ g = \text{Id}_F$ on tire les résultats analogues en échangeant f et g .

Si f est surjective alors $f \circ g \circ f = f \Rightarrow f \circ g = \text{Id}_F$

Si g est injective alors $g \circ f \circ g = g \Rightarrow f \circ g = \text{Id}_F$. □

En particulier, si f ou g est bijective et $g \circ f = \text{Id}_E$ alors f et g sont l'inverse l'un de l'autre.

Points fixes; applications idempotentes

Définition. Etant donnée une application f d'un ensemble E dans lui-même, on dit qu'un élément $x \in E$ est un point fixe de f ssi $f(x) = x$. L'ensemble des points fixes de f sera noté $\text{Fix } f$.

Définition. Une application f d'un ensemble dans lui-même est dite idempotente ssi $f \circ f = f$.

Pour tous ensembles E et F , tous $f \in F^E$ et $g \in F^F$,

$$\begin{aligned} \text{Fix } g &\subset \text{Im } g \\ g \circ f = f &\Leftrightarrow \text{Im } f \subset \text{Fix } g \\ g \circ g = g &\Leftrightarrow \text{Im } g = \text{Fix } g \end{aligned}$$

2.4. Bijections canoniques remarquables

Les identités remarquables usuelles reliant les opérations entre entiers se retrouveront ici en remplaçant l'égalité par l'existence de bijections canoniques. L'existence d'une bijection canonique entre deux ensembles E et F (donnés sous forme de termes dépendant d'autres noms d'ensembles), sera notée $E \simeq F$.

Tout comme pour la notion d'objet canonique dont ceci est un cas particulier, cette notation $E \simeq F$, est un énoncé métamathématique qui a pour objet de sous-entendre un autre énoncé, à savoir un énoncé de définition d'une bijection f entre E et F . En pratique, cela pourra être un énoncé P de variables $x \in E, y \in F$, de paramètres les noms des ensembles figurant dans les termes définissant E et F , tel que $P(x, y)$ désigne f suivant $P(x, y) \Leftrightarrow y = f(x)$.

Dans les situations auxquelles nous nous intéresserons, lorsque $E \simeq F$ et $F \simeq G$ on aura $E \simeq G$. En effet, en assemblant les énoncés définissant les bijections entre E et F et entre F et G on obtient une définition d'une bijection entre E et G .

Mais dans des situations auxquelles nous ne nous intéresserons pas, cela pourrait être faux, comme par exemple si ce sont des paires d'éléments purs, que $E \cap G = \emptyset$ et que F est formé d'un élément de E et d'un élément de G : dans ce cas, la bijection définie entre E et G dépend de F et n'est donc pas canonique. Mais, dans les situations auxquelles nous nous intéresserons, donc, il sera possible d'éliminer le paramètre F , par exemple en le remplaçant par sa définition à partir de ce qui constitue E et/ou G , pour aboutir à une définition sans paramètre.

Des bijections canoniques entre des ensembles et d'autres ensembles permettent de définir des bijections canoniques entre ensembles construits à partir des premiers et ceux construits de même à partir des seconds, par exemple $(E \simeq E' \text{ et } F \simeq F') \Rightarrow (E^F \simeq E'^{F'} \text{ et } E \times F \simeq E' \times F')$.

La transposition (composition des couples avec σ) donne une formule de commutativité du produit cartésien: $E \times F \simeq F \times E$.

Nous avons précédemment mentionné la bijection canonique

$$G^{E \times F} \simeq (G^F)^E$$

$$f \mapsto (x \mapsto (y \mapsto f(x, y)))$$

d'inverse $g \mapsto ((x, y) \mapsto g(x)(y))$. Il en résulte $(G^F)^E \simeq G^{E \times F} \simeq G^{F \times E} \simeq (G^E)^F$.

Dans le cas $G = \mathcal{V}$, ces identités traduites par $\mathcal{V}^E \simeq \mathcal{P}(E)$ donnent

$$(\mathcal{P}(F))^E \simeq \mathcal{V}^{E \times F} \simeq \mathcal{P}(E \times F) \simeq (\mathcal{P}(E))^F.$$

Pour toute relation R entre deux ensembles E et F , notons $\vec{R} \in \mathcal{P}(F)^E$ et $\overleftarrow{R} \in \mathcal{P}(E)^F$ les images de R par ces bijections canoniques:

$$\forall x \in E, \vec{R}(x) = \{y \in F \mid x R y\}$$

$$\forall y \in F, \overleftarrow{R}(y) = \{x \in E \mid x R y\}$$

$$\forall x \in E, \forall y \in F, x R y \Leftrightarrow x \in \overleftarrow{R}(y) \Leftrightarrow y \in \vec{R}(x).$$

Etant donnée une famille d'ensembles $\prod_{i \in I} F_i$, la formule $(F^I)^E \simeq (F^E)^I$ appliquée à l'union F des F_i donne par restriction

$$\left(\prod_{i \in I} F_i\right)^E \simeq \prod_{i \in I} (F_i^E)$$

$$h \mapsto (f_i)_{i \in I}$$

définie par $\forall i \in I, f_i = \pi_i \circ h$, ou inversement par $\forall x \in E, h(x) = (f_i(x))_{i \in I}$. On dit que h est le *produit* de la famille (f_i) , et on le note $h = \prod_{i \in I} f_i$.

En particulier, étant données deux applications f et g de même domaine E , leur produit est

$$f \times g = (E \ni x \mapsto (f(x), g(x))).$$

De cette manière, $(F \times G)^E \simeq F^E \times G^E$, qui se raffine en la formule du développement

$$\left(\prod_{i \in I} F_i\right)^E \simeq \prod_{h \in I^E} \prod_{x \in E} F_{h(x)}.$$

La bijection $G^{E \times F} \simeq (G^F)^E$ se raffine en

$$F^{\prod_{i \in I} E_i} \simeq \prod_{i \in I} F^{E_i}$$

où une application f de $\prod_{i \in I} E_i$ dans F est liée à une famille d'applications $f_i \in F^{E_i}$ par $\forall i \in I, f_i = f \circ j_i$ où j_i est l'injection canonique $(x \mapsto (i, x))$ de E_i dans $\prod_{i \in I} E_i$. On écrira alors

$$f = \coprod_{i \in I} f_i$$

Cette opération de somme d'une famille d'applications, est ainsi nommée parce que son graphe est essentiellement la somme des graphes respectifs, $(\text{Gr}(f) = \bigcup_{i \in I} \text{Im}(j_i \times f_i))$ et elle s'applique à toute famille $(f_i)_{i \in I}$; elle ne dépend pas de l'ensemble d'arrivée F . On peut aussi l'écrire sous la forme $f(x) = f_{\pi(x)}(j_{\pi(x)}^{-1}(x))$ où π est la première projection de $\prod_{i \in I} E_i$ sur I (avec $j(x) = i \Leftrightarrow x \in \text{Im } j_i$).

Par restriction à des sous-ensembles, cette bijection canonique donne également

$$\prod_{i \in I} \prod_{y \in E_i} F_{(i, y)} \simeq \prod_{x \in \prod_{i \in I} E_i} F_x$$

dont un cas très particulier donne $(E \times F) \times G \simeq E \times F \times G$.

Nous ne détaillerons pas $E^{\{x\}} \simeq E$, $E \times \{x\} \simeq E$, $E \times \emptyset = \emptyset$, $\{x\}^E \simeq \{x\}$, $E^\emptyset \simeq \{x\}$, et pour $E \neq \emptyset$, $\emptyset^E = \emptyset$.

2.5. Notions sur les relations binaires.

On appelle *relation binaire* sur un ensemble E , une relation à deux variables de même domaine E . Par exemple sur tout ensemble la relation d'égalité est une relation binaire.

Dans ce qui suit nous noterons ces deux variables de part et d'autre du symbole de relation (comme $x R y$) au lieu de les noter à droite entre parenthèse (comme $R(x, y)$); bien sûr cela ne change rien au fond.

Une relation binaire R sur un ensemble E est dite:

- *réflexive* ssi $\forall x \in E, x R x$
- *antiréflexive* ssi $\forall x \in E, \text{non}(x R x)$
- *symétrique* ssi $\forall x, y \in E, x R y \Rightarrow y R x$
- *antisymétrique* ssi $\forall x, y \in E, (x R y \text{ et } y R x) \Rightarrow x = y$.
- *transitive* ssi $\forall x, y, z \in E, (x R y \text{ et } y R z) \Rightarrow x R z$

Toute relation binaire transitive et antiréflexive est antisymétrique.

Préordre. On appelle *préordre* toute relation binaire réflexive et transitive. Un ensemble muni d'une relation de préordre est dit un ensemble préordonné. Un préordre antisymétrique est appelé un *ordre* (ou: relation d'ordre). Un ensemble muni d'un ordre est appelé un *ensemble ordonné*.

Relation d'équivalence. On nomme ainsi une relation de préordre symétrique.

Sous-entendons les quantificateurs comme portant sur E dans la proposition suivante.

Proposition. 1) Si R est un préordre alors $x R y \Leftrightarrow \overleftarrow{R}(x) \subset \overleftarrow{R}(y)$, i.e.

$$\forall x, y, x R y \Leftrightarrow \forall z, (z R x \Rightarrow z R y)$$

2) Si de plus R est symétrique (donc, une relation d'équivalence) alors $x R y \Leftrightarrow \overleftarrow{R}(x) = \overleftarrow{R}(y)$, i.e.

$$\forall x, y, x R y \Leftrightarrow \forall z, (z R x \Leftrightarrow z R y)$$

3) Si R est réflexive et $\forall x, y, z, (x R y \text{ et } z R y) \Rightarrow z R x$ alors R est une relation d'équivalence.

Preuves:

1) La transitivité se réécrit $\forall x, y, x R y \Rightarrow \forall z, (z R x \Rightarrow z R y)$.

Puis, R étant réflexive, $\forall x, y, (\forall z, z R x \Rightarrow z R y) \Rightarrow (x R x \Rightarrow x R y) \Rightarrow x R y$.

2) $\forall x, y, x R y \Leftrightarrow (x R y \text{ et } y R x) \Leftrightarrow (\overleftarrow{R}(x) \subset \overleftarrow{R}(y) \text{ et } \overleftarrow{R}(y) \subset \overleftarrow{R}(x)) \Leftrightarrow (\overleftarrow{R}(x) = \overleftarrow{R}(y))$.

3) on vérifie la symétrie: $\forall x, y, (x R y \text{ et } y R y) \Rightarrow y R x$. La transitivité en découle. \square

On voit facilement que les réciproques de 1) et 2) sont vraies aussi. Ainsi, leurs formules étant respectivement équivalentes aux notions de préordre et de relation d'équivalence, peuvent donc leur servir de définitions.

2.6. Etude des relations d'équivalence

Partitions et familles-partitions

Soit E un ensemble.

On appellera *famille-partition de E* une famille $(A_i)_{i \in I}$ de parties de E non vides, deux à deux disjointes et dont l'union est E , autrement dit

$$\begin{aligned} \forall i \in I, A_i \neq \emptyset \\ \forall i, j \in I, i \neq j \Rightarrow A_i \cap A_j = \emptyset \\ \bigcup_{i \in I} A_i = E \end{aligned}$$

Reformulons la deuxième condition:

$$\begin{aligned} (\forall i, j \in I, i \neq j \Rightarrow A_i \cap A_j = \emptyset) &\Leftrightarrow \forall i, j \in I, i \neq j \Rightarrow \forall x \in E, \text{non}(x \in A_i \text{ et } x \in A_j) \\ &\Leftrightarrow \forall i, j \in I, \forall x \in E, i \neq j \Rightarrow \text{non}(x \in A_i \text{ et } x \in A_j) \\ &\Leftrightarrow \forall x \in E, \forall i, j \in I, (x \in A_i \text{ et } x \in A_j) \Rightarrow i = j \\ &\Leftrightarrow \forall x \in E, \exists ! i \in I, x \in A_i \end{aligned}$$

Par conséquent, le système des 3 conditions pour qu'une famille $(A_i)_{i \in I}$ de parties de E soit une famille-partition de E se résume en un système de deux conditions

$$\begin{aligned} \forall i \in I, \exists x \in E, x \in A_i \\ \forall x \in E, \exists! i \in I, x \in A_i. \end{aligned}$$

On appelle *partition de E* un ensemble P d'ensembles non vides, deux à deux disjoints et dont l'union est E . Ceci équivaut à dire que Id_P est une famille-partition de E .

Nous allons examiner les correspondances entre les notions suivantes, concernant un même ensemble E :

- Surjection de domaine E ;
- Famille-partition de E ;
- Partition de E ;
- Relation d'équivalence sur E .

Surjection et famille-partition

Soit une application quelconque f de domaine un ensemble E , et $I = \text{Im } f$. Alors, définissons l'application f^\bullet de I dans $\mathcal{P}(E)$ (autrement dit une famille de parties de E indexée par I), par

$$\forall i \in I, f^\bullet(i) = f^*(\{i\}) = \{x \in E \mid f(x) = i\}.$$

Autrement dit, f^\bullet est définie comme étant l'unique application de I dans $\mathcal{P}(E)$ telle que

$$\forall i \in I, \forall x \in E, x \in f^\bullet(i) \Leftrightarrow f(x) = i.$$

On remarque qu'à travers la bijection canonique $\mathcal{P}(E)^I \simeq \mathcal{P}(E \times I)$, f^\bullet correspond au graphe de f . Or par cette même correspondance, le système de formules définissant la notion de famille-partition de E , se traduit en celui caractérisant les graphes de surjections de E sur I . Ceci définit une bijection canonique entre l'ensemble des surjections de E sur I , et celui des familles-partitions de E indexées par I .

De surjection à relation d'équivalence

Soit une application f de domaine E . On appelle *relation d'équivalence sur E associée à f* la relation binaire $\underset{f}{\sim}$ sur E définie par

$$\forall x, y \in E, x \underset{f}{\sim} y \Leftrightarrow f(x) = f(y).$$

En effet, les propriétés de réflexivité, symétrie et transitivité d'une relation ainsi définie se vérifient immédiatement.

Relation d'équivalence et partition, surjection canonique

Soit R une relation binaire sur E et $P = \text{Im } \overleftarrow{R}$.

Le fait que R soit une relation d'équivalence, se réexprime par les formules équivalentes

$$\begin{aligned} \forall x, y \in E, x R y &\Leftrightarrow \overleftarrow{R}(x) = \overleftarrow{R}(y) \\ \forall x, y \in E, x \in \overleftarrow{R}(y) &\Leftrightarrow \overleftarrow{R}(x) = \overleftarrow{R}(y) \\ \forall x \in E, \forall A \in P, x \in A = \text{Id}_P(A) &\Leftrightarrow \overleftarrow{R}(x) = A \\ \text{Id}_P &= \overleftarrow{R}^\bullet. \end{aligned}$$

L'ensemble des partitions de E , autrement dit des $P \subset \mathcal{P}(E)$ tels que Id_P est une famille-partition de E , donc de la forme $\overleftarrow{R}^\bullet$ pour une certaine relation binaire R sur E finalement unique, est ainsi en bijection canonique avec l'ensemble des relations d'équivalence.

Dans les constructions ci-dessus, lorsque R est une relation d'équivalence, et que donc P est une partition, l'ensemble P est appelé le *quotient de E par R* et noté E/R ; et l'application $\overleftarrow{R} : E \rightarrow P$ est appelée *surjection canonique* de E sur E/R . Pour tout $x \in E$, l'élément $\overleftarrow{R}(x)$, unique élément A de P tel que $x \in A$, est appelé la *classe de x par R* .

De surjection à partition

A toute surjection f de E sur I nous avons associé une relation d'équivalence R sur E par $\forall x, y \in E, x R y \Leftrightarrow f(x) = f(y)$, et montré que toute relation d'équivalence R est égale à celle associée à \overleftarrow{R} . Puis nous avons associé à une telle relation R une partition $P = \text{Im } \overleftarrow{R}$ de E .

Nous allons maintenant voir que $P = \text{Im}(f^\bullet)$, tout comme il était égal à $\text{Im}(\overleftarrow{R}^\bullet)$ où $\overleftarrow{R}^\bullet = \text{Id}_P$. En effet, la définition de R se traduit par

$$\forall x, y \in E, x \in \overleftarrow{R}(y) \Leftrightarrow f(x) = f(y) \Leftrightarrow x \in f^\bullet(f(y))$$

autrement dit $\overleftarrow{R} = f^\bullet \circ f$, d'où $P = \text{Im } \overleftarrow{R} = \text{Im } f^\bullet$ puisque f est surjective.

L'ensemble I muni de f , étant naturellement par f^\bullet en bijection avec E/R , pourra être utilisé comme jouant le rôle de E/R , autrement dit être vu comme un autre quotient (une copie du quotient) de E par R ; le rôle de la surjection canonique est alors joué par f .

Remarque. f^\bullet est injective.

On peut le voir directement, ou en notant que $\underset{f}{\sim} = \underset{f^\bullet \circ f}{\sim}$. Cette injectivité devient fautive lorsqu'on étend f^\bullet à plus d'un élément hors de $\text{Im } f$:

Extension de notation. *Ultérieurement, pour toute application f et tout ensemble d'arrivé F de f donné, on notera encore (abusivement) f^\bullet l'application $F \ni y \mapsto f^\bullet(\{y\}) = \{x \in E \mid f(x) = y\}$, qui prolonge le f^\bullet précédemment défini, par \emptyset hors de $\text{Im } f$.*

Autre résultat

Lemme. *Soient trois ensembles E, F, G , deux applications $f \in F^E, g \in G^E$, soit $H = \text{Im}(f \times g) = \{(f(x), g(x)) \mid x \in E\} \subset F \times G$, et soit R une relation entre F et G . Alors*

$$(\forall x \in E, R(f(x), g(x))) \Leftrightarrow (\forall (y, z) \in H, y R z)$$

En effet, $R(f(x), g(x))$ peut également s'écrire $R((f \times g)(x))$.

Théorème. *Avec les notations ci-dessus, si $\text{Im } f = F$ et $\forall x, x' \in E, f(x) = f(x') \Rightarrow g(x) = g(x')$ (ce qu'on peut abréger en $\underset{f}{\sim} < \underset{g}{\sim}$) alors il existe un unique $h \in G^F$ tel que $g = h \circ f$.*

Preuve: Par le lemme, $g = h \circ f \Leftrightarrow (\forall (y, z) \in H, z = h(y)) \Leftrightarrow H \subset \text{Gr } h$.

Il ne reste plus qu'à vérifier que H est le graphe d'une application de F dans G .

De la surjectivité de f il vient $\forall y \in F, \exists z \in G, (y, z) \in H$ (par $(y = f(x) \Rightarrow (y, g(x)) \in H)$).

Enfin, par le lemme,

$$\begin{aligned} \underset{f}{\sim} < \underset{g}{\sim} &\Leftrightarrow \forall (y, z) \in H, \forall (y', z') \in H, y = y' \Rightarrow z = z' \\ &\Leftrightarrow \forall y \in F, !z \in G, (y, z) \in H. \end{aligned}$$

□

Remarque. Ce théorème a une sorte de réciproque: l'existence d'un h tel que $g = h \circ f$, même sans hypothèse sur son domaine, implique que $\underset{f}{\sim} < \underset{g}{\sim}$. Finalement, pour une surjection f fixée, l'injection $G^F \ni h \mapsto h \circ f$ a pour image l'ensemble des $g \in G^E$ tels que $\underset{f}{\sim} < \underset{g}{\sim}$. Sa restriction à l'ensemble des injections de F dans G a pour image $\{g \in G^E \mid \underset{f}{\sim} = \underset{g}{\sim}\}$.

Notation. *Soit $g \in F^E$ et R une relation d'équivalence sur E telle que $R < \underset{g}{\sim}$. On note alors g/R l'application de domaine E/R définie par $g = (g/R) \circ \overleftarrow{R}$. Si $R = \underset{g}{\sim}$ on l'appellera l'injection canonique de $E/\underset{g}{\sim}$ dans F .*

2.7. Axiome du choix

Théorème et définition. Pour tout ensemble X fixé, les énoncés suivants sont équivalents, et pareillement nommés *axiome du choix de base X* et notés AC_X :

- 1) Tout produit indexé par X d'ensembles non vides est non vide
- 2) Pour tout ensemble E et toute relation R entre X et E ,

$$(\forall x \in X, \exists y \in E, R(x, y)) \Rightarrow (\exists f \in E^X, \forall x \in X, R(x, f(x)))$$

- 3) Pour toute application g d'image X , $\exists f \in (\text{Dom } g)^X, g \circ f = \text{Id}_X$.

1) \Rightarrow 2) est immédiat; 2) \Rightarrow 1) en définissant E comme union de la famille.

On a 2) \Rightarrow 3) en définissant $R(x, y) \Leftrightarrow (x = g(y))$. Autrement dit on a 1) \Rightarrow 3) en prenant la famille g^\bullet d'ensembles non vides. Réciproquement, on montre 3) \Rightarrow 1) en prenant la somme de la famille, ou encore on montre 3) \Rightarrow 2) à l'aide du graphe de R . \square

Axiome du choix (AC). Pour tout ensemble X , AC_X .

Déjà, AC_X est vrai pour tout ensemble fini X , comme on peut facilement le voir à “la main” et on le démontrera dans le texte suivant.

Mais, les logiciens professionnels ont réussi à démontrer que l'axiome du choix est indécidable, c'est-à-dire ni démontrable ni réfutable. Précisément, que s'il existe un univers de la théorie des ensembles dans lequel il est vrai, alors il en existe aussi un dans lequel il est faux (AC_X devient faux pour certains ensembles infinis X), et inversement. Comment est-ce possible ? Nous avons prévenu que les indécidabilités venaient du fait que pour deux ensembles X et Y où X est infini, rien ne peut garantir que ce qui sert d'ensemble puissance Y^X dans un univers donné, soit le vrai. Autrement dit, il pourrait toujours y avoir des applications de X dans Y qui n'appartiennent pas à cet univers, mais seulement à un autre univers plus grand. La véracité d'une formule énoncée en termes d'ensembles puissance pourrait n'être qu'une illusion due à sa restriction à l'univers donné. En fabriquant des petits univers ainsi illusoire, certains énoncés comme l'axiome du choix peuvent y apparaître de valeur vraie ou faux contraire à ce qu'ils seraient dans une supposée “réalité” extérieure.

Ainsi l'axiome du choix peut sembler faux dans un univers U , alors qu'il serait vrai dans un univers plus vaste U' , du fait que les éléments dans U' d'un certain produit d'ensembles non vides, n'apparaissent pas dans U . Et au contraire il peut sembler vrai dans un univers U , alors qu'il est faux dans un univers plus vaste U' , parce qu'une certaine famille d'ensembles non vides dans U' , dont le produit est vide, n'existe simplement pas dans U . Mais les détails de ces constructions sont bien trop complexes pour être abordés ici.

En pratique, comme l'axiome du choix est conforme à l'intuition et plus facile à affirmer qu'à nier (comme il y a plusieurs manières de le nier), la majorité des travaux de mathématiques sur les questions qui en dépendent le supposent vrai. Cependant, bien des questions n'en dépendent pas, ou se satisfont d'une version plus faible (notamment $AC_{\mathbb{N}}$).

Pour terminer, citons d'autres équivalents simples de l'axiome du choix:

Théorème. Les énoncés suivants sont équivalents à l'axiome du choix:

- 4) Pour tous ensembles E, F, G et toute $g \in G^F$ surjective, $\{g \circ f \mid f \in F^E\} = G^E$.
- 5) Pour tout ensemble E et toute relation d'équivalence R sur E , $\exists A \subset E, \forall x \in E, \exists! y \in A, xRy$.
- 6) Pour tout ensemble E d'ensembles, $\emptyset \notin E \Rightarrow (\prod_{A \in E} A) \neq \emptyset$.

Preuves:

$AC_E \Rightarrow$ 4) par $\forall h \in G^E, (\forall x \in E, \exists y \in F, g(y) = h(x)) \Rightarrow (\exists f \in F^E, \forall x \in E, g(f(x)) = h(x))$

$AC_G \Rightarrow$ 4) par $\exists i \in F^G, g \circ i = \text{Id}_G$ et $\forall h \in G^E, i \circ h \in F^E$ et $g \circ i \circ h = h$.

4) \Rightarrow 3) : avec $E = G$, par $\text{Id}_E \in \{g \circ f \mid f \in F^E\}$.

3) \Rightarrow 5) : $\exists g \in E^{E/R}, \overline{R} \circ g = \text{Id}_{E/R}$ de sorte que $A = \text{Im } g$ convient.

5) \Rightarrow 3) : soit $E = \text{Dom } g$, et $A \subset E$ tel que $\forall x \in E, \exists! y \in A, g(x) = g(y)$. Alors $g|_A$ est bijective de A sur X , et son inverse $f \in A^X \subset E^X$ vérifie $g \circ f = g|_A \circ f = \text{Id}_X$.

1) \Rightarrow 6) : il suffit de prendre la famille Id_E .

6) \Rightarrow 1) : soit $(A_i)_{i \in I}$ une famille d'ensembles non vides, et soit E son image $\{A_i \mid i \in I\}$. On a alors $\emptyset \notin E$, donc il existe $f \in \prod_{A \in E} A$. Alors $(f(A_i))_{i \in I} \in \prod_{i \in I} A_i$. \square

Les prochains textes sur la théorie des ensembles s'appuieront sur l'axiome des parties mais non pas l'axiome du choix (sauf cas particuliers explicites). Non que le premier soit plus vrai (on pourrait même estimer le contraire), seulement il sera le seul des deux à y être indispensable.