
Ordinaux et cardinaux

Le Barbuki 2

1^{ère} édition

Florian Langlois

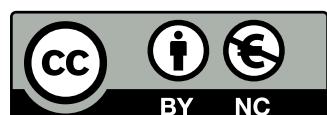
1^{ère} édition rédigée entre mars 2024 et novembre 2024

Collection

Bienvenue dans ce livre ! C'est le deuxième d'une collection qui tente d'exposer et démontrer les mathématiques de niveau licence et master. Le nom BARBUKI est une référence au célèbre groupe BOURBAKI, dont la démarche de cette collection est inspirée.

- 1 – Théorie élémentaire des ensembles
- 2 – Ordinaux et cardinaux

Ce document est mis à disposition selon les termes de la licence Creative Commons “Attribution – Pas d'utilisation commerciale 4.0 International”.



Avant-propos

« *Vers l'infini et au-delà* ». Voilà qui résume bien la théorie des nombres ordinaux dont il est question dans ce livre : s'amuser à compter au-delà même de l'infini, et ce de manière démesurée et vertigineuse. Si cette étude du gigantisme est belle en soi, elle est aussi féconde car elle débouche notamment sur la notion d'entiers naturels, c'est-à-dire les nombres de la vie de tous les jours avec lesquels on a l'habitude de compter. Compter est d'ailleurs un usage important des nombres ordinaux : on apprendra dans le dernier chapitre à le faire rigoureusement à travers les nombres cardinaux. Le célèbre lemme de Zorn pointera aussi le bout de son nez, les ordinaux représentant un cadre idéal pour le démontrer.

Il s'agit ici d'exposer l'une des plus belles théories mathématiques que je connaisse. Je me souviens encore avoir ressenti un vertige métaphysique en découvrant ce dont elle parle. C'était en écoutant la vidéo de la chaîne YouTube VSauce intitulée "*How to count past infinity*". Je vous recommande au passage vivement d'aller visionner sa vidéo.

Cet ouvrage est là pour me permettre de coucher sur le papier les différentes mathématiques que j'ai apprises durant mes études supérieures : je le rédige principalement pour moi-même et il n'a pas pour but d'être pédagogique. Il va me permettre de conserver sur le long terme une trace de ces connaissances, mais aussi d'organiser celles-ci pour en avoir une vue d'ensemble.

Bien que ce livre reste assez personnel, il est possible qu'il vous soit utile. Afin de comprendre pleinement son contenu, il est nécessaire d'être au courant de ce dont parle le premier ouvrage, c'est-à-dire des bases de la théorie des ensembles, notamment à travers les différents axiomes de ZFC.

Il vous faut aussi savoir mener un raisonnement, ou tout du moins en suivre un, puisque c'est l'un des objets principaux de ce livre. Il est à noter que la construction de cette collection se fait sous la manière d'un escalier à gravir : nous n'utiliserons pas des résultats postérieurs pour démontrer des résultats antérieurs, les seules exceptions étant les exemples donnés pour illustrer, puisque ceux-ci ne sont là que pour aider à la lecture, et non permettre une quelconque démonstration, mais aussi certaines digressions abordant d'autres démonstrations que celles proposées.

Remerciements

Merci à Lyra, Chæriss, Shika, Alyssio, GrothenDitQue, et Cassis pour leur pinaillage et leurs précieux éclairages. Sans eux cet ouvrage ne pourrait pas exister.

Les biographiques de mathématiciens sont en parties inspirées de Wikipédia ainsi que de l'excellent ouvrage *Des mathématiciens de A à Z* de Daniel Suratteau et Bertrand Hauchecorne, paru en 2008 aux éditions Ellipses.

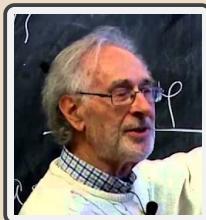
Le contenu à proprement parlé est très fortement inspiré de l'incroyable ouvrage *The Foundations of Mathematics* de Kenneth Kunen dans son édition de 2007, ainsi que de *Théorie des ensembles* de Jean-Louis Krivine, paru aux éditions Cassini en 2007.

Pour la petite histoire



Kenneth Kunen (2 août 1943 – 14 août 2020) est un mathématicien américain, professeur émérite de mathématiques à l'université du Wisconsin à Madison qui travaillait en théorie des ensembles et à ses applications en topologie et en théorie de la mesure.

Pour la petite histoire



Jean-Louis Krivine (1939 –) est professeur à l'Université Paris 7, spécialiste de géométrie algébrique réelle, d'analyse fonctionnelle, de logique et d'informatique théorique. Il a créé, en 1982, l'équipe de logique mathématique, qui est l'un des plus importants laboratoires au monde dans ce domaine. Lauréat de l'Académie des Sciences en 1994 et Prix du Rayonnement français en 2004, il est l'auteur de plusieurs ouvrages de référence en logique.

Table des matières

1 Ordinaux	1
1 Classes et assertions fonctionnelles	3
1.1 Assertions à paramètres	3
1.2 Classes	3
1.3 Assertions fonctionnelles	5
2 Bons ordres	8
3 Ordinaux	15
4 Successeurs, limites et entiers naturels	33
5 Isomorphisme avec les ordinaux	48
6 Récurrence : induction et récursion	64
6.1 Induction	64
6.2 Récursion	67
6.3 Suites	76
2 Opérations sur les ordinaux	89
1 Généralités	90
2 Addition d'ordinaux	94
2.1 Définition et propriétés	94
2.2 Interprétation graphique : la concaténation	111
3 Multiplication d'ordinaux	126
3.1 Définition et propriétés	126
3.2 Interprétation graphique : le produit cartésien	141
4 Exponentiation d'ordinaux	148
4.1 Définition et propriétés	148
4.2 Applications à support fini	159
5 Forme normale de Cantor et ε_0	174
5.1 Logarithme ordinal et forme normale de Cantor	174
5.2 L'ordinal ε_0 et la classe des points fixes	180
3 Cardinaux	199
1 Équipotence et subpotence	200
1.1 Équipotence et subpotence	200
1.2 Théorème de Cantor	209
1.3 Équipotence et opérations	215
2 Nombres cardinaux	230
2.1 Les cardinaux	230
2.2 Le cardinal d'un ensemble	236
3 Les grands théorèmes	242

3.1	Choix, Zorn et Zermelo	242
3.2	Théorème et cardinal de Hartogs	256
4	Opérations sur les cardinaux	265
5	Ensembles finis et ensembles dénombrables	284
5.1	Ensembles finis	284
5.2	Ensembles dénombrables	302
Conclusion		315
Bibliographie		317
Mathématiciens		319

Chapitre 1

Ordinaux

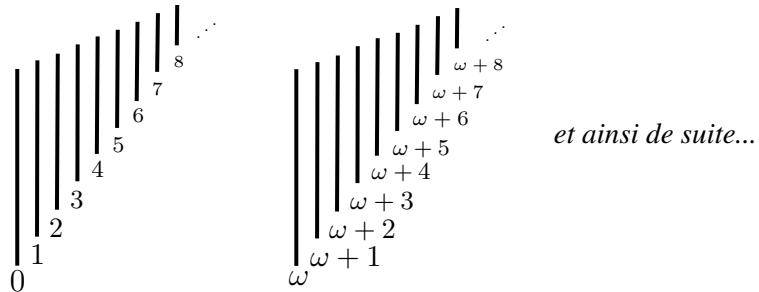
Sommaire

1	Classes et assertions fonctionnelles	3
1.1	Assertions à paramètres	3
1.2	Classes	3
1.3	Assertions fonctionnelles	5
2	Bons ordres	8
3	Ordinaux	15
4	Successseurs, limites et entiers naturels	33
5	Isomorphisme avec les ordinaux	48
6	Récurrence : induction et récursion	64
6.1	Induction	64
6.2	Récursion	67
6.3	Suites	76

Imaginez une course à laquelle vous concourez et à laquelle participe une infinité de coureurs. À la fin de la course, chaque participant se voit attribuer un nombre en fonction de l'ordre dans lequel il est arrivé : le premier arrivé reçoit le nombre 0, le deuxième le nombre 1, le troisième le nombre 2, et ainsi de suite pour chaque entier naturel. Et vous ? Vous arrivez après tous les coureurs ayant reçu un nombre entier naturel ! Quel nombre vous correspond-il ? Certainement pas un entier naturel, puisque ceux-ci ont déjà tous été attribués. Il faut donc introduire un nouveau nombre : on le note généralement ω .



Quel nombre attribuer alors à votre ami arrivé juste après vous ? Le nombre $\omega + 1$ naturellement ! Et $\omega + 2$ pour la personne juste après-lui, puis $\omega + 3$ et ainsi de suite pour les suivants !



Ces nouveaux nombres que nous venons d'introduire font partie de ce que l'on appelle les **nombres ordinaux**, catégorie dans laquelle se trouvent aussi les entiers naturels. L'objet de ce chapitre est justement de définir et développer les nombres ordinaux. Cela s'inscrit dans un contexte plus général qui est celui des **ensembles bien ordonnés**, pour lesquels chaque partie non vide admet un minimum, permettant de répondre notamment à la question « *quel élément vient juste après celui-ci ?* ». S'intéresser aux ordinaux présente différentes vertus :

- ▶ comme les nombres entiers naturels en font partie, nous aurons enfin l'occasion de les définir proprement.
- ▶ même si l'exemple de la course est un peu fantaisiste, des situations où l'on souhaite ordonner des choses avec l'une d'entre elle après une infinité d'autres peuvent se présenter à nous et les ordinaux représentent un outil de choix pour cela.
- ▶ enfin, les ordinaux constituent le cadre idéal pour compter le nombre d'éléments des ensembles, ce qui sera l'objet du chapitre 3.

1 Classes et assertions fonctionnelles

1.1 Assertions à paramètres

Dans le livre précédent, nous avons commencé par expliquer que les objets que nous manipulons sont tous considérés comme des ensembles, au point que la notion d'ensemble est en fait primitive. On ne donne pas de définition a priori de ce qu'est un ensemble, on impose juste des axiomes afin de mimer l'intuition d'ensemble.

Nous avons ensuite indiqué que ce ne sont pas les seuls choses manipulables : il y a aussi les assertions, qui sont des affirmations pouvant être vraies ou fausses. Il s'agit au fond d'une façon de structurer le discours à propos des ensembles. L'auteur de ce livre considère que le lecteur est au clair sur ces choses-là. Cependant, il estime aussi devoir préciser un certain nombre de nuances concernant les assertions. Certaines de ces définitions sont déjà évoquées dans le livre précédent, mais un rappel ne fait jamais de mal.

Définition 1 (Assertion à paramètres)

Une **assertion à paramètres** est une assertion qui nécessite un ou plusieurs paramètres pour être énoncée, et donc la vérité peut varier en fonction de ces paramètres éventuels. Un paramètre est toujours un ensemble.

Exemple :

1. L'assertion $P(n)$ définie par « *n est un entier pair* » dépend de qui est n . C'est en cela que l'on précise entre parenthèses la dépendance de P par rapport à n , pour insister sur ce point.
2. En revanche, l'assertion $Q(x)$ définie par « *x = x* » est toujours vraie, quand bien même elle nécessite le paramètre x pour être énoncée.

1.2 Classes

La notion d'ensemble est née de l'idée de vouloir réunir et regrouper plusieurs objets différents : typiquement \mathbb{Z} est l'ensemble qui contient tous les entiers relatifs. C'est justement le but du premier livre d'expliquer les règles que nous avons choisies ici pour régir les ensembles. Cependant afin d'éviter certains paradoxes et contradictions, nous avons dû restreindre la portée des ensembles : il n'est par exemple pas possible de définir l'ensemble de tous les ensembles, et si on le permettait on aboutirait au paradoxe de Russell. Nous verrons aussi plus tard, après avoir défini la notion d'ordinaux, qu'il est impossible d'avoir un ensemble contenant tous les ordinaux.

Cependant, nous aimerais bien pouvoir simplifier nos discours concernant "*tous les ensembles*" ou "*tous les ordinaux*", c'est-à-dire réunir différents objets sans pour autant craindre de former un ensemble paradoxal, ou même sans être freiné par les axiomes ensemblistes. C'est là qu'interviennent les **classes**. Heureusement, cela ne va pas nécessiter d'introduire autre chose que les ensembles ou que les assertions. En effet, nous allons définir la notion de classe comme étant la même que celle d'assertion à paramètres, le nouveau nom étant simplement associé à un

nouvel usage. Il s'agit d'une approche similaire à celle que nous avons faite dans le premier livre concernant les familles : il n'y a à strictement parler pas de différence entre les familles et les applications, simplement un usage différent et des notations différentes.

L'intérêt des classes est comme nous l'avons dit de pouvoir regrouper différents objets, et donc beaucoup des notions associées aux classes sont inspirées de celles associées aux ensembles, notamment l'appartenance. Il n'est donc pas étonnant qu'on retrouve par exemple le symbole \in .

Définition 2 (Classe)

Soit C une assertion à paramètres.

Si C nécessite un seul paramètre pour être énoncée, on dit parfois que C est une **classe**.

Pour un ensemble x donné, on dit que x **appartient** à C si et seulement si $C(x)$ est vraie, auquel cas on note alors $x \in C$. On dit aussi que x est un **élément** de C .

Dans le cas contraire, c'est-à-dire si $C(x)$ est fausse, on dit que x n'appartient pas à C , ou que x n'est pas un élément de C , et on note $x \notin C$.

Ainsi, la notion de classe généralise celle d'ensemble. En effet, étant donné un ensemble E , on peut lui associer l'assertion à paramètres « $x \in E$ », qui est donc une classe. La définition qui suit précise cela.

Définition 3 (Classe propre)

Soient E un ensemble et C une classe.

1. On appelle classe **issue** de E la classe C_E définie pour tout ensemble x par

$$C_E(x) : \ll x \in E \gg.$$

Autrement dit, pour tout ensemble x on a l'équivalence $x \in C_E \iff x \in E$.

2. On dit que C est une classe **propre** si et seulement si C n'est pas issue d'un ensemble.

Remarque :

Si une classe C est issue d'un ensemble E , alors cet ensemble est unique. En effet, cela vient du fait que l'appartenance caractérise entièrement un ensemble. On commettra souvent l'abus de confondre une classe et l'ensemble dont elle est issue, si celui-ci existe.

Exemple :

Pour avoir des exemples de classes issues d'un ensemble, il suffit simplement de prendre un ensemble de son choix et de former sa classe associée. Voici en revanche quelques exemples de classes propres :

1. La classe U de tous les ensembles. Comme tout paramètre x est nécessairement un

ensemble, on a nécessairement $x \in U$ et donc l'assertion $U(x)$ est toujours vraie. On peut par exemple définir $U(x)$ en posant simplement « $x = x$ ». D'après le paradoxe de Russell, une telle classe est nécessairement propre.

2. La classe ON des ordinaux, que nous aurons l'occasion d'aborder plus tard.
Nous verrons via le paradoxe de Burali-Forti que cette classe est propre.



Notation

Soient C et D deux classes, E un ensemble et C_E la classe issue de E .

1. On note $C \subseteq D$ si et seulement si $\forall x, (x \in C \Rightarrow x \in D)$.
En particulier on note $E \subseteq D$ si et seulement si $C_E \subseteq D$.
Autrement dit, $E \subseteq D$ si et seulement si $\forall x, (x \in E \Rightarrow x \in D)$.
2. On note $C \cap D$ la classe définie pour tout ensemble x par

$$x \in C \cap D \iff (x \in C \text{ et } x \in D)$$

En particulier on note $E \cap D$ la classe $C_E \cap D$.

D'après l'axiome de compréhension que $E \cap D$ est une classe issue d'un ensemble.
En effet, on a

$$E \cap D = \{x \in E \mid x \in D\} = \{x \in E \mid D(x)\}$$

Comme indiqué précédemment on confondra souvent les deux.

3. On note $C \cup D$ la classe définie pour tout ensemble x par

$$x \in C \cup D \iff (x \in C \text{ ou } x \in D)$$

En particulier on note $E \cup D$ la classe $C_E \cup D$.

4. On note $D \setminus C$ la définie pour tout ensemble x par

$$x \in D \setminus C \iff (x \in D \text{ et } x \notin C)$$

En particulier on note $D \setminus E$ la classe $D \setminus C_E$.

5. On note $C \in D$ si et seulement si C est **issue d'un ensemble** F tel que $F \in D$.

1.3 Assertions fonctionnelles

Dans le livre précédent, nous nous sommes intéressés à des assertions à paramètres particulières : les assertions fonctionnelles. Comme le qualificatif *fonctionnelle* le laisse entendre, il s'agit d'une généralisation de la notion de fonction. Redonnons-en la définition.

Définition 4 (Assertion fonctionnelle)

Soit P une assertion à paramètres.

On dit que P est **fondationnelle** si et seulement si elle nécessite deux paramètres pour être énoncée et pour tout ensembles x, y et y' , on a l'implication

$$(P(x, y) \text{ et } P(x, y')) \implies y = y'$$

Ainsi pour un ensemble x donné, il y a au plus un ensemble y qui lui est associé par le biais de P . On dit alors que y est **l'image** de x par P et on note alors $P(x) := y$.

Exemple :

1. L'assertion P définie pour deux ensembles a et b par

$$P(a, b) \iff « a \text{ est un entier naturel et } b = 2a »$$

est une assertion fondationnelle. Pour tout a entier naturel, on a alors $P(a) = 2a$.

2. Étant donnée une application f , on peut naturellement lui associer une assertion fondationnelle P_f en posant pour tout ensembles x et y

$$P_f(x, y) \iff « x \in \text{dom}(f) \text{ et } y = f(x) »$$

Pour tout $x \in \text{dom}(f)$, on a alors $P_f(x) = f(x)$.

Comme l'indiquent ces exemples, la notion d'assertion fondationnelle et la notion de fonctions sont très liées, du fait pour un ensemble x de n'associer qu'au plus un autre ensemble. On retrouve donc naturellement la notion d'image, et les notations $P(x)$ et $f(x)$ qui s'y réfèrent sont identiques. Il est important au passage pour une assertion fondationnelle de ne pas confondre la notation $P(x, y)$ qui se réfère à l'assertion en elle-même et qui est donc soit vraie soit fausse en fonction des paramètres x et y , et la notation $P(x)$ qui désigne l'unique paramètre y tel que $P(x, y)$ soit vraie, à condition bien sûr que celui-ci existe.

Dans le cas d'une fonction f , on peut parler de son domaine $\text{dom}(f)$ comme d'un ensemble, c'est-à-dire l'ensemble de tout ensemble qui admet une image par f . Il n'est pas toujours possible de faire de même pour une assertion fondationnelle : par exemple l'assertion fondationnelle « $x = y$ » aurait pour domaine l'ensemble de tous les ensembles, que nous savons n'existe pas. C'est là qu'interviennent les classes que nous avons introduites plus tôt : la classe de tous les ensembles existe bel et bien !

Définition 5 (Domaine et image d'une assertion fondationnelle)

Soit P une assertion fondationnelle.

1. On appelle **domaine** de P la classe notée $\text{dom}(P)$ définie pour tout ensemble x par

$$x \in \text{dom}(P) \iff \exists y, P(x, y)$$

2. On appelle **image** de P la classe notée $\text{im}(P)$ définie pour tout ensemble y par

$$y \in \text{im}(P) \iff \exists x, P(x, y)$$

Exemple :

1. L'assertion fonctionnelle P définie pour tout ensembles x et y par

$$P(x, y) \iff « x = y »$$

a pour domaine U , la classe de tous les ensembles.

2. L'assertion fonctionnelle Q définie pour tout ensembles x et y par

$$Q(x, y) \iff « x = y \text{ et } x \neq y »$$

a pour domaine la classe issue de \emptyset . Comme dit précédemment, on commettra souvent l'abus de confondre un ensemble et la classe issue de celui-ci, si bien qu'on écrira $\text{dom}(Q) = \emptyset$.

3. Soient f une application et F_f l'assertion fonctionnelle issue de f , c'est-à-dire

$$F_f(x, y) \iff « x \in \text{dom}(f) \text{ et } y = f(x) »$$

On peut voir que $\text{dom}(F_f)$ est tout simplement la classe issue de $\text{dom}(f)$. Comme dit précédemment, on commettra souvent l'abus de confondre un ensemble et la classe issue de celui-ci, si bien qu'on écrira $\text{dom}(F_f) = \text{dom}(f)$.

Remarque :

1. A la manière des images directes et réciproques d'un ensemble par une fonction, on peut se donner une classe C et considérer son **image directe** par P , à savoir la classe notée $P^\rightarrow(C)$ définie pour tout ensemble y par

$$y \in P^\rightarrow(C) \iff \exists x \in C, P(x, y)$$

Nous avons vu dans le livre 1 via l'axiome de remplacement que si E est un ensemble tel que $E \subseteq \text{dom}(P)$ alors $P^\rightarrow(E)$ est un ensemble que l'on a noté $\{P(x) \mid x \in E\}$. De même, on peut considérer l'**image réciproque** de la classe C par P , à savoir la classe notée $P^\leftarrow(C)$ définie pour tout ensemble x par

$$x \in P^\leftarrow(C) \iff \exists y \in C, P(x, y)$$

2. Une façon intuitive de construire une assertion fonctionnelle est de se munir d'une **formule**. Autrement dit, étant donné un ensemble x , on construit $P(x)$ explicitement. Par exemple, on peut définir P en posant que pour tout ensembles x et y , on a

$$P(x, y) \iff y = \bigcup x$$

et dans ce cas-là on a naturellement $P(x) = \bigcup x$. C'est d'ailleurs par ce biais là des formules que l'on s'est déjà donné le moyen de construire des applications dans le précédent livre.

3. Soient P une assertion fonctionnelle et E un ensemble tel que $E \subseteq \text{dom}(P)$. On a montré dans le précédent livre qu'il existe alors une unique application $f : E \rightarrow ?$

telle que $\forall x \in E, f(x) = P(x)$, que l'on note généralement $\begin{pmatrix} E & \rightarrow & ? \\ x & \mapsto & P(x) \end{pmatrix}$.

En cela, on dira que f est la **restriction** de P à E , et on notera $P|_E := f$. Ainsi, même si P est une assertion fonctionnelle sans être une application, toute restriction de celle-ci à un ensemble est nécessairement une application.

4. Pour P une assertion fonctionnelle, C et D deux classes, on notera $P : C \rightarrow ?$ pour signifier $\text{dom}(P) = C$ et on notera $P : C \rightarrow D$ pour signifier $\text{dom}(P) = C$ et $\text{im}(P) \subseteq D$.

2 Bons ordres

Bien souvent en mathématique, nous aimerais étant donné un élément x pouvoir donner du sens à la question « *quel est l'élément qui vient juste après x ?* ». Par exemple chez les entiers, $n + 1$ est l'élément qui vient juste après n . On peut remarquer que c'est le minimum des entiers strictement plus grands que n . C'est là qu'intervient la notion de **bon ordre** : toute partie non vide de l'ensemble va admettre un élément minimum. De fait, l'élément qui suit directement x sera simplement le minimum des éléments strictement plus grands que x .

Concentrons-nous quelques instants sur la notion d'élément minimal. Rappelons qu'un élément a de l'ensemble ordonné (E, \preccurlyeq) est minimal si et seulement si pour tout $x \in E$ on a

$$x \preccurlyeq a \implies x = a$$

c'est-à-dire qu'il n'y a que a pour être plus petit ou égal à a .

Proposition 1 (Élément minimal et ordre strict)

Soient (E, \preccurlyeq) un ensemble ordonné non vide, \prec l'ordre strict associé à \preccurlyeq et $a \in E$.

Alors a est minimal pour (E, \preccurlyeq) si et seulement si $\forall x \in E$, non($x \prec a$).



Démonstration

Raisonnons par double implications.



Supposons que a est minimal pour (E, \preccurlyeq) .

Soit $x \in E$.

Supposons par l'absurde que $x \prec a$.

On a donc $x \preccurlyeq a$ et $x \neq a$.

Comme $x \preccurlyeq a$ et a est minimal pour (E, \preccurlyeq) , on a $x = a$.

Ainsi on a à la fois $x \neq a$ et $x = a$: c'est absurde.

Par l'absurde, on a donc montré que non($x \prec a$).

Donc $\forall x \in E$, non($x \prec a$).

Donc si a est minimal pour (E, \preccurlyeq) alors $\forall x \in E$, non($x \prec a$).



Supposons que $\forall x \in E$, non($x \prec a$).

Soit $x \in E$.

Supposons que $x \preccurlyeq a$.

On a donc $x \prec a$ ou $x = a$.

Or on a non($x \prec a$) par hypothèse donc nécessairement $x = a$.

Donc si $x \preccurlyeq a$ alors $x = a$.
 Donc $\forall x \in E, (x \preccurlyeq a \implies x = a)$.
 Donc a est minimal pour (E, \preccurlyeq) .
 Donc si $\boxed{\forall x \in E, \text{non}(x \prec a) \text{ alors } a \text{ est minimal pour } (E, \preccurlyeq)}$.
CQFD.

Nous l'avons dit dans l'introduction, nous allons dire qu'un ensemble est muni d'un bon ordre lorsque chacune de ses parties (non vides) admet un minimum. Une version plus faible de la notion de bon ordre est la notion d'ordre **bien fondé**, où l'on demande à chaque partie (non vide) d'admettre un élément minimal, mais qui n'est pas nécessairement le minimum de la partie.

Définition 6 (Ordre bien fondé et bon ordre)

Soit E un ensemble ordonné.

1. On dit que E est **bien fondé** si et seulement si toute partie non vide de E admet au moins un élément minimal.
2. On dit que E est **bien ordonné** si et seulement si toute partie non vide de E admet un minimum. On dit aussi que l'ordre sur E est un **bon ordre**.

Remarque :

Dans la suite de cet ouvrage, on va étendre les définitions qui concernent les ordres (larges) aux ordres stricts. Par exemple, considérons (E, \preccurlyeq) un ensemble ordonné, \prec l'ordre strict associé à \preccurlyeq et $a \in E$.

- On dit que a est **minimal** pour (E, \prec) si et seulement si a est minimal pour (E, \preccurlyeq) .
- On dit que a est **le minimum** de (E, \prec) si et seulement si a est le minimum de (E, \preccurlyeq) .
- On dit que \prec est un **ordre strict bien fondé** sur E si et seulement si toute partie non vide de E admet un élément minimal pour \prec . Comme \prec et \preccurlyeq partagent les mêmes éléments minimaux d'après ce qui précède, \prec est un ordre strict bien fondé sur E si et seulement si \preccurlyeq est un ordre (large) bien fondé.
- On dit que \prec est un **bon ordre strict** sur E , ou que (E, \prec) est **strictement bien ordonné**, si et seulement si toute partie non vide de E admet un minimum pour \prec . Comme \prec et \preccurlyeq partagent le même minimum éventuel, \prec est un bon ordre strict sur E si et seulement si \preccurlyeq est un bon ordre (large) sur E .

On pourrait aussi parler d'éléments maximaux et de maximum, mais dans ce livre ce sont avant tout les minimaux et minimum qui vont nous intéresser, bien qu'à quelques endroits les maximaux et maximums reviendront nous voir.

Au premier abord la notion d'élément minimal et la notion d'élément minimum semblent être la même chose. Ce n'est pas vrai, puisqu'un ensemble peut avoir plusieurs éléments minimaux. Pensons par exemple à $\mathbb{N} \setminus \{1\}$ muni de la relation de divisibilité : tous les nombres premiers sont des éléments minimaux sans qu'aucun ne soit un minimum. En réalité, pour qu'un élément minimal soit un minimum, il faut et il suffit qu'il soit comparable à tous les éléments de l'ensemble, ce qui explique pourquoi un bon ordre est nécessairement total.

Proposition 2 (Caractérisation des bons ordres)

Soit E un ensemble ordonné.

Les assertions suivantes sont équivalentes :

1. E est bien ordonné.
2. E est bien fondé et totalement ordonné.



Démonstration

Notons \preccurlyeq l'ordre sur E .

Raisonnons par double implications.

$1 \Rightarrow 2$

Supposons que E est bien ordonné.

Alors toute partie non vide de E admet un minimum.

Or un minimum est un élément minimal (c'est alors le seul).

Donc toute partie non vide de E admet un élément minimal.

Donc E est bien fondé.

Montrons que E est totalement ordonné.

Soient x et y dans E .

Alors $\{x, y\}$ est une partie non vide de E .

Elle admet donc un minimum m .

Si $m = x$ alors on a $x = m \preccurlyeq y$.

Si $m = y$ alors on a $y = m \preccurlyeq x$.

Dans les deux cas on a $x \preccurlyeq y$ ou $y \preccurlyeq x$.

Donc tous les éléments de E sont comparables : E est totalement ordonné.

$1 \Leftarrow 2$

Supposons que E est bien fondé et totalement ordonné.

Soit A une partie non vide de E .

Comme E est bien fondé, A admet au moins un élément minimal m .

Montrons que m est le minimum de A .

Supposons par l'absurde que m n'est pas le minimum de A .

Il existe donc $a \in A$ tel que non($m \preccurlyeq a$).

Or E est totalement ordonné par hypothèse.

On a donc nécessairement $a \preccurlyeq m$.

Or m est un élément minimal de A donc on a $a = m$.

En particulier $m \preccurlyeq a$ par réflexivité de \preccurlyeq .

C'est absurde par définition de a .

Par l'absurde, on vient donc de montrer que m est le minimum de A .

Donc toute partie non vide de E admet un minimum.

Donc $\boxed{E \text{ est bien ordonné}}$.

CQFD.

Le fait pour un ensemble d'être muni d'un ordre bien fondé ou d'un bon ordre se transmet aux parties de cet ensemble, en considérant bien entendu que l'on conserve la même relation d'ordre au passage.

Proposition 3 (Partie d'un ensemble bien ordonné)

Soient E un ensemble ordonné et $A \subseteq E$.

On munit A de la même relation d'ordre que celle de E .

1. Si E est bien fondé alors A est bien fondé.
2. Si E est bien ordonné alors A est bien ordonné.



Démonstration

1. Supposons que E est bien fondé.

Soit B une partie non vide de A .

Comme $A \subseteq E$, B est aussi une partie non vide de E .

Or E est bien fondé par hypothèse.

Donc B admet au moins un élément minimal.

Donc toute partie non vide de A admet au moins un élément minimal.

Donc $\boxed{A \text{ est bien fondé}}$.

2. Supposons que E est bien ordonné.

Soit B une partie non vide de A .

Comme $A \subseteq E$, B est aussi une partie non vide de E .

Or E est bien ordonné par hypothèse.

Donc B admet un minimum.

Donc toute partie non vide de A admet un minimum.

Donc $\boxed{A \text{ est bien ordonné}}$.

CQFD.

Rappelons qu'étant donnés deux ensembles ordonnés (E, \preccurlyeq) et (F, \sqsubseteq) , on peut munir $E \times F$ de l'ordre **lexicographique** associé, c'est-à-dire que pour x et y dans E et s et t dans F , on a

$$(x, s) \preceq (y, t) \iff [x \prec y \text{ ou } (x = y \text{ et } s \sqsubseteq t)]$$

où \prec désigne l'ordre strict associé à \preccurlyeq . L'ordre lexicographique tire son nom du fait que les dictionnaires fonctionnent sur ce principe (par rapport à l'ordre alphabétique) : on compare

d'abord les premières lettres de chaque mot, et éventuellement si ce sont les mêmes on compare les deuxième lettres et ainsi de suite. Ici il s'agit simplement de comparer des mots ayant chacun deux lettres.

Proposition 4 (Bons ordres et ordre lexicographique)

Soient (E, \preccurlyeq) et (F, \sqsubseteq) deux ensembles ordonnés.

Soit \trianglelefteq l'ordre lexicographique associé sur $E \times F$.

Si \preccurlyeq et \sqsubseteq sont des bons ordres alors \trianglelefteq est un bon ordre.

Démonstration

Notons \prec l'ordre strict associé à \preccurlyeq .

Supposons que \preccurlyeq et \sqsubseteq sont des bons ordres.

Soit G une partie non vide de $E \times F$.

Considérons $A := \{x \in E \mid \exists y \in F, (x, y) \in G\}$.

Comme G est non vide, A est une partie non vide de E .

Or E est bien ordonné donc A admet un minimum a_0 .

Considérons alors $B := \{y \in F \mid (a_0, y) \in G\}$.

Par définition on a $a_0 \in A$ donc il existe $y \in F$ tel que $(a_0, y) \in G$ et donc $y \in B$.

Donc B est une partie non vide de F .

Or F est bien ordonné donc B admet un minimum b_0 .

Considérons alors $g_0 := (a_0, b_0)$ et montrons que g_0 est le minimum de G .

Soit $z = (x, y) \in G$.

Par définition de A on a $x \in A$.

Or a_0 est le minimum de A donc $a_0 \preccurlyeq x$.

On a donc $a_0 \prec x$ ou $a_0 = x$.

Si $a_0 \prec x$ alors par définition de \trianglelefteq on a $(a_0, b_0) \trianglelefteq (x, y)$.

Supposons à présent que $a_0 = x$.

On a donc $(a_0, y) = (x, y) \in G$ donc $y \in B$ par définition de B .

Or b_0 est le minimum de B donc $b_0 \sqsubseteq y$.

On a donc $a_0 = x$ et $b_0 \sqsubseteq y$ donc $(a_0, b_0) \trianglelefteq (x, y)$ par définition de \trianglelefteq .

Dans les deux cas on a bien $g_0 \trianglelefteq z$.

Donc pour tout $z \in G$, on a $g_0 \trianglelefteq z$.

Donc g_0 est le minimum de G .

Donc toute partie non vide de $E \times F$ admet un minimum.

Donc $E \times F$ est bien ordonné.

CQFD.

Introduisons à présent la notion de **segment initial**. Une partie d'un ensemble ordonné est un segment initial si et seulement si pour chacun de ses éléments, elle contient aussi tous les éléments qui lui sont inférieurs.

Définition 7 (Segment initial)

Soient (E, \preccurlyeq) un ensemble ordonné et A une partie de E .

On dit que A est un **segment initial** de E si et seulement si pour tout x et y dans E , on a

$$(y \preccurlyeq x \in A) \implies y \in A$$

Exemple :

1. Dans \mathbb{R} muni de l'ordre usuel, $]-\infty, 2[$ est un segment initial. En revanche $]1; 3]$ n'en est pas un car $3 \in]1; 3]$ et $0 \leq 3$ alors que $0 \notin]1; 3]$.
2. Dans \mathbb{N} muni de la relation de divisibilité, $\{1, 2, 4, 8\}$ est un segment initial. En revanche $\{1, 2, 6\}$ n'en est pas un car $6 \in \{1, 2, 6\}$ et $3|6$ alors que $3 \notin \{1, 2, 6\}$.

Notation :

Soient (E, \preccurlyeq) un ensemble ordonné, \prec l'ordre strict associé et $x \in E$.

On pose $E_{\prec x} := \{y \in E \mid y \prec x\}$.

Dans le cas des ensembles bien ordonnés, on a une caractérisation simple des segments initiaux propres. On rappelle au passage qu'une partie A d'un ensemble E est dite propre si et seulement si $A \neq E$.

Proposition 5 (Segments initiaux d'un ensemble bien ordonné)

Soient (E, \preccurlyeq) un ensemble **bien ordonné** et A une partie de E .

Soit \prec l'ordre strict associé à \preccurlyeq .

Les assertions suivantes sont équivalentes :

1. A est un segment initial **propre** de E .
2. Il existe $x \in E$ tel que $A = E_{\prec x}$.

 *Démonstration*

$1 \Rightarrow 2$

Supposons que A est un segment initial propre de E .

Comme A est une partie propre de E , on a $A \subsetneq E$ donc $E \setminus A \neq \emptyset$.

Or E est bien ordonné par définition donc $E \setminus A$ possède un minimum x .

Montrons que $A = E_{\prec x}$.



Soit $a \in A$.

Comme E est bien ordonné, E est totalement ordonné d'après la prop. 2 p. 10.

On a donc $x \preccurlyeq a$ ou $a \prec x$.

Supposons par l'absurde que $x \preccurlyeq a$.

On a donc $x \preccurlyeq a \in A$ et A est un segment initial de E par hypothèse.

Donc $x \in A$, ce qui est absurde car $x \in E \setminus A$.

Donc par l'absurde on a $a \prec x$, c'est-à-dire $a \in E_{\prec x}$.

On a donc $A \subseteq E_{\prec x}$.



Soit $y \in E_{\prec x}$.

On a alors $y \in E$ et $y \prec x$.

Or par définition x est le minimum de $E \setminus A$.

On a donc $y \notin E \setminus A$, donc comme $y \in E$ on a $y \in A$.

Donc $A \supseteq E_{\prec x}$ et donc $\boxed{A = E_{\prec x}}$.

$1 \Leftarrow 2$

Supposons qu'il existe $x \in E$ tel que $A = E_{\prec x}$.

Soient y et z dans E .

Supposons que $z \preccurlyeq y \in A$.

Par hypothèse on a $A = E_{\prec x}$ donc $y \in E_{\prec x}$ et donc $y \prec x$.

Comme $z \preccurlyeq y$ on a donc $z \prec x$ par transitivité et donc $z \in E_{\prec x} = A$.

Donc si $z \preccurlyeq y \in A$ alors $z \in A$.

Donc $\boxed{A \text{ est un segment initial de } E}$.

De plus, on n'a pas $x \prec x$ par antiréflexivité donc $x \notin E_{\prec x} = A$.

Ainsi $x \in E$ et $x \notin A$, donc $E \neq A$ et donc $\boxed{A \text{ est une partie propre de } E}$.

CQFD.

3 Ordinaux

Lors du précédent livre, nous avons vu la notion d'**isomorphisme** entre deux ensembles ordonnés. C'est une façon de dire que ces deux ensembles ordonnés "*se comportent de la même manière*", pour peu que l'on ne s'intéresse qu'à leur structure d'ensembles ordonnés. Nous pouvons donc d'une certaine manière "*identifier*" deux ensembles ordonnés dès lors qu'il existe un isomorphisme entre les deux, et donc dire en ce sens-là qu'ils sont équivalents. Qui dit équivalence dit classe d'équivalence, c'est-à-dire rassembler en un seul endroit tous ces ensembles ordonnés qui sont isomorphes entre eux. Notons au passage que la notion de classe d'équivalence ici n'a pas besoin d'être un ensemble : nous avons justement introduit plus tôt le concept de classe (tout court) pour palier ce problème.

Se pose alors la question suivante : pour chacune de ces classes d'équivalences, peut-on se donner un représentant canonique, c'est-à-dire un ensemble ordonné qui représenterait toute la classe d'équivalence ? Si nous n'allons pas donner de réponse à cette question en toute généralité, nous allons le faire dans le cas particulier où les ensembles sont munis d'un bon ordre : c'est l'objectif derrière la construction des **ordinaux**, car nous verrons après les avoir définis qu'il en existera systématiquement un et un seul dans chacune des classes d'équivalence des ensembles bien ordonnés.

Pour choisir l'ensemble ordonné en question, il faut choisir en particulier sa relation d'ordre. Tout choix de relation pourrait sembler arbitraire, mais il en existe deux qui sortent naturellement du lot : \in et \subseteq , car ce sont les relations les plus fondamentales qui existent chez les ensembles. Nous n'allons d'ailleurs pas avoir besoin de choisir entre les deux : nous allons faire en sorte que \subseteq soit l'ordre (large) et \in l'ordre strict associé.

Définition 8 (Ensemble transitif)

Soit E un ensemble.

On dit que E est **transitif** si et seulement si $\forall x \in E, x \subseteq E$.

Remarque :

Remarquons la chose suivante :

$$\begin{aligned} E \text{ est transitif} &\iff \forall y \in E, y \subseteq E \\ &\iff \forall y, (y \in E \implies y \subseteq E) \\ &\iff \forall x, \forall y, (x \in y \in E \implies x \in E) \end{aligned}$$

Ainsi, la transitivité de E signifie une certaine transitivité de \in .

Cette définition répond aussi au fait que nous allons faire de \in un ordre strict sur E : en particulier \in sera transitif, c'est-à-dire que pour x, y et z dans E , si $x \in y \in z$ alors $x \in z$. Le fait pour E d'être transitif va donc étendre légèrement cette propriété en se permettant en plus de remplacer z par E lui-même : si $x \in y \in E$ alors $x \in E$.

Définition 9 (Ordinaux)

Soit E un ensemble.

On dit que E est un **ordinal** si et seulement si

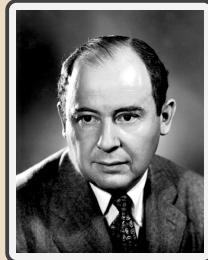
1. \in est un bon ordre strict sur E .
2. E est transitif.

Remarque :

Ainsi, un ordinal est un ensemble E tel que :

1. \in est antiréflexive sur E , c'est-à-dire $\forall x \in E, x \notin x$.
2. \in est transitive sur E , c'est-à-dire $\forall x \in E, \forall y \in E, \forall z \in E, (x \in y \in z \Rightarrow x \in z)$.
3. Pour toute partie non vide F de E , F admet un minimum pour \in , c'est-à-dire qu'il existe $a \in F$ tel que $\forall x \in F, (x \neq a \Rightarrow a \in x)$.
4. E est transitif, c'est-à-dire que $\forall x, \forall y, (x \in y \in E \Rightarrow x \in E)$.

Pour la petite histoire



John von Neumann (28 décembre 1903 – 8 février 1957) est un mathématicien et physicien américano-hongrois. Il a apporté d'importantes contributions en mécanique quantique, en analyse fonctionnelle, en logique mathématique, en informatique théorique, en sciences économiques et dans beaucoup d'autres domaines des mathématiques et de la physique. Il a de plus participé aux programmes militaires américains comme le célèbre projet Manhattan.

C'est à lui que l'on doit cette définition d'ordinaux.

Exemple :

1. \emptyset est un ordinal. En effet, par vérité creuse, on a les quatre points suivants :
 - (a) On a $\forall x \in \emptyset, x \notin x$ donc \in est antiréflexive sur \emptyset .
 - (b) On a $\forall x \in \emptyset, \forall y \in \emptyset, \forall z \in \emptyset, (x \in y \in z \Rightarrow x \in z)$.
Ainsi \in est transitive sur \emptyset .
 - (c) Comme aucune partie de \emptyset n'est non vide, on a bien que toutes les parties non vides de \emptyset admettent un minimum pour \in .
 - (d) On a $\forall x \in \emptyset, x \subseteq \emptyset$ donc \emptyset est transitif.

Les points (a) et (b) font de \in un ordre strict sur \emptyset .

Combinés au point (c), on en conclut que \in est un bon ordre strict sur \emptyset .

Enfin, combinés au point (d) on en conclut que \emptyset est un ordinal.

2. Nous verrons plus tard que tout entier naturel, et même \mathbb{N} l'ensemble des entiers naturels lui-même, est un ordinal.

Remarque :

Il est d'usage de désigner un ordinal par une lettre grecque minuscule.

Par exemple \mathbb{N} sera aussi désigné par la lettre ω , qui lui sera alors réservée.

Notation :

On notera ON la classe de tous les ordinaux, c'est-à-dire que pour un ensemble x , on a l'équivalence $x \in ON \iff x$ est un ordinal.

Tentons de justifier le choix de la notion d'ordinal pour représenter une classe d'équivalence des ensembles bien ordonnés. Nous avons déjà justifié l'usage de \in comme relation de bon ordre strict pour son côté naturel. Il reste donc simplement à justifier la transitivité de l'ensemble lui-même, c'est-à-dire le point 2 de la définition d'ordinal.

Pour cela, intéressons-nous au cas simple d'ensembles à deux éléments, pour la relation d'ordre strict \in . Comme on veut que \in soit un bon ordre strict, on veut en particulier que tous les éléments distincts soient comparables pour l'appartenance, et donc que sur les deux éléments l'un appartienne à l'autre, ce qui impose au représentant α d'être de la forme $\alpha = \{x, E\}$ avec $x \in E$. Pour rendre le choix de α le plus naturel possible, on aimerait épurer au maximum le choix de x et de E : en particulier il semble naturel de demander $x = \emptyset$ pour ne pas s'encombrer avec d'éventuels éléments de x qui seraient nécessairement arbitraires. Pour la même raison, on aimerait que E ne contienne rien d'autre que x , ce qui impose naturellement $E = \{x\}$ et donc $\alpha = \{\emptyset, \{\emptyset\}\}$.

La transitivité va permettre de retirer les éventuels éléments encombrants : si x est un élément de l'ordinal α , alors par transitivité de α on a $x \subseteq \alpha$, c'est-à-dire que tous les éléments de x font aussi partie de α . Ainsi dans l'exemple $\alpha = \{x, E\}$, x ne peut rien contenir car tout élément éventuel de x se retrouverait en plus dans les éléments de α , et pour la même raison E ne peut rien contenir de plus que x .

Proposition 6 (Les éléments d'un ordinal sont des ordinaux)

Soient α un ordinal et x un ensemble.

Si $x \in \alpha$ alors x est un ordinal.



Démonstration

Supposons que $x \in \alpha$.

- Par définition α est un ordinal donc α est transitif et (α, \in) est strictement bien ordonné.

Comme $x \in \alpha$, on a donc $x \subseteq \alpha$ par définition de la transitivité.

Or on vient de dire que (α, \in) est strictement bien ordonné.

Donc (x, \in) est strictement bien ordonné d'après la proposition 3 page 11.

- Il reste donc à montrer que x est transitif.

Soit $y \in x$.

On a vu que $x \subseteq \alpha$ donc $y \in \alpha$ par définition de l'inclusion.

On a donc $y \subseteq \alpha$ car α est transitif.

Montrons que $y \subseteq x$.

Soit $z \in y$.

Comme $y \subseteq \alpha$, on a en particulier $z \in \alpha$ par définition de l'inclusion.

On a donc $z \in y \in x$, et tous les trois sont des éléments de α .

Or (α, \in) est strictement bien ordonné donc \in est transitif sur α .

On a donc $z \in x$ par transitivité de \in .

Donc $\forall z \in y, z \in x$ et donc $y \subseteq x$ par définition de l'inclusion.

Donc $\forall y \in x, y \subseteq x$.

Ainsi x est transitif.

Finalement, (x, \in) est strictement bien ordonné et x est transitif.

Donc x est un ordinal.

CQFD.

Proposition 7 (L'intersection de deux ordinaux est un ordinal)

Soient α et β deux ordinaux.

Alors $\alpha \cap \beta$ est un ordinal.

Démonstration

Comme α est un ordinal, (α, \in) est strictement bien ordonné.

Or $\alpha \cap \beta \subseteq \alpha$ donc $(\alpha \cap \beta, \in)$ est strictement bien ordonné d'après la prop. 3 p. 11.

Il reste à montrer que $\alpha \cap \beta$ est transitif.

Soit $x \in \alpha \cap \beta$.

On a donc $x \in \alpha$ et $x \in \beta$.

Or α et β sont des ordinaux donc sont transitifs donc $x \subseteq \alpha$ et $x \subseteq \beta$.

On a donc $x \subseteq \alpha \cap \beta$.

Donc $\forall x \in \alpha \cap \beta, x \subseteq \alpha \cap \beta$.

Donc $\alpha \cap \beta$ est transitif.

Finalement $(\alpha \cap \beta, \in)$ est strictement bien ordonné et $\alpha \cap \beta$ est transitif.

On a donc $\boxed{\alpha \cap \beta \text{ est un ordinal}}$.

CQFD.

Nous l'avons annoncé quand nous avons introduit la notion d'ordinal : étant donné un ordinal, nous voulons faire de \subseteq l'ordre (large) et de \in l'ordre strict. Par définition d'un ordinal, \in est le bon ordre strict concerné. La proposition suivante va nous montrer que \subseteq est quant à lui l'ordre (large) associé à \in .

Proposition 8 (Ordre large sur les ordinaux)

Soient α et β deux ordinaux.

On a l'équivalence

$$\alpha \subseteq \beta \iff (\alpha \in \beta \text{ ou } \alpha = \beta)$$



Démonstration

Raisonnons par double implications.



Supposons que $\alpha \subseteq \beta$ et $\alpha \neq \beta$.

Montrons que $\alpha \in \beta$.

Posons $X := \beta \setminus \alpha$: par hypothèse on a $X \neq \emptyset$.

Or β est un ordinal donc (β, \in) est strictement bien ordonné.

Donc comme $X \subseteq \beta$ et $X \neq \emptyset$, on en déduit que (X, \in) admet un minimum ξ .

Comme $\xi \in X$ et $X \subseteq \beta$, on a $\xi \in \beta$ par définition de l'inclusion.

On peut donc montrer $\xi = \alpha$ pour conclure.



Soit $\mu \in \xi$.

On a alors $\mu \in \xi \in \beta$ et β est transitif (car ordinal) donc $\mu \in \beta$.

Comme $\mu \in \xi$ et que ξ est le minimum de (X, \in) , on a $\mu \notin X$.

On a donc $\mu \in \beta$ et $\mu \notin X$ donc $\mu \in \beta \setminus X = \alpha$.

Donc $\xi \subseteq \alpha$.



Supposons par l'absurde que $\xi \neq \alpha$, c'est-à-dire $\xi \subsetneq \alpha$ d'après ce qui précède.

On a donc $\alpha \setminus \xi \neq \emptyset$ donc il existe $\mu \in \alpha \setminus \xi$.

En particulier on a $\mu \in \alpha$.

Comme $\alpha \subseteq \beta$ par hypothèse, on a $\mu \in \beta$ par définition de l'inclusion.

Ainsi on a $\xi \in \beta$ et $\mu \in \beta$.

Or β est un ordinal donc (β, \in) est strictement bien ordonné.

Donc \in est un ordre strict total sur β d'après la proposition 2 page 10.

On a donc $\mu \in \xi$ ou $\xi \in \mu$ ou $\mu = \xi$.

► $\mu \in \xi$ est impossible.

En effet par définition on a $\mu \in \alpha \setminus \xi$ donc $\mu \notin \xi$.

► $\xi \in \mu$ est impossible.

En effet on aurait $\xi \in \mu \in \alpha$ donc $\xi \in \alpha$ car α est transitif car ordinal.

Or on a $\xi \in X = \beta \setminus \alpha$ donc $\xi \notin \alpha$.

► $\mu = \xi$ est impossible.

En effet on a $\xi \in X = \beta \setminus \alpha$ donc $\xi \notin \alpha$ alors que $\mu \in \alpha \setminus \xi$ donc $\mu \in \alpha$.

On a donc $\xi \notin \alpha$ et $\mu \in \alpha$ donc on ne peut pas avoir $\mu = \xi$.

On aboutit donc à une contradiction.

Par l'absurde, on a prouvé que $\xi = \alpha$.

Comme $\xi \in \beta$, on a donc $\alpha \in \beta$.

Donc $(\alpha \subseteq \beta \text{ et } \alpha \neq \beta) \implies \alpha \in \beta$.

Donc $\boxed{\alpha \subseteq \beta \implies (\alpha \in \beta \text{ ou } \alpha = \beta)}$.



Supposons que $\alpha \in \beta$ ou $\alpha = \beta$.

Si $\alpha \in \beta$ alors $\alpha \subseteq \beta$ car β est transitif car ordinal.

Si $\alpha = \beta$ alors en particulier $\alpha \subseteq \beta$ par réflexivité de l'inclusion.

Dans tous les cas on a $\alpha \subseteq \beta$.

Donc si $\boxed{\alpha \in \beta \text{ ou } \alpha = \beta \text{ alors } \alpha \subseteq \beta}$.

CQFD.

Remarque :

Désormais, on utilisera régulièrement le fait qu'étant donné un ordinal, il est naturellement muni de \subseteq en tant que relation de bon ordre et que \in est le bon ordre strict associé. En particulier pour deux ordinaux α et β , on a l'équivalence $\alpha \subseteq \beta \iff \alpha \in \beta$.

Le fait d'avoir prouvé ces quelques propriétés générales sur les ordinaux nous permet d'entrevoir le magnifique théorème qui va suivre : celui-ci affirme qu'en fait c'est toute la classe ON qui se comporte comme un ordinal.

Théorème 1 (Bon ordre strict sur les ordinaux)

Soient α , β et γ trois ordinaux.

1. Si $\alpha \in \beta \in \gamma$ alors $\alpha \in \gamma$.

Ainsi \in est **transitif** sur ON .

2. On a $\alpha \notin \alpha$.

Ainsi \in est **antiréflexive** sur ON .

Ainsi par 1 et 2, \in peut être vu comme un **ordre strict** sur ON .

3. On a $\alpha \in \beta$ ou $\beta \in \alpha$ ou $\alpha = \beta$.

Autrement dit \in est un ordre strict **total** sur ON .

4. Soit E un ensemble non vide dont les éléments sont tous des ordinaux.

Alors (E, \in) possède un minimum.

Ainsi \in est un **bon** ordre strict sur ON .

Ainsi, \in est un **bon ordre strict** sur ON .



Démonstration

1. Supposons que $\alpha \in \beta \in \gamma$.

On a alors $\boxed{\alpha \in \gamma}$ car \in est transitif car ordinal.

2.

Supposons par l'absurde que $\alpha \in \alpha$.

En prenant $x := \alpha$, on a l'existence d'un $x \in \alpha$ tel que $x \in x$.

Or α est un ordinal donc (α, \in) est strictement bien ordonné donc \in est antiréflexive sur α . En particulier $\forall x \in \alpha, x \notin x$, d'où l'absurdité.

Par l'absurde, on a donc $\boxed{\alpha \notin \alpha}$.

3. Considérons $\delta := \alpha \cap \beta$.

Alors δ est un ordinal d'après la proposition 7 page 18.

Or on a $\delta \subseteq \alpha$ donc $(\delta \in \alpha \text{ ou } \delta = \alpha)$ d'après la proposition 8 page 19.

De même on a $\delta \subseteq \beta$ donc $(\delta \in \beta \text{ ou } \delta = \beta)$ d'après la proposition 8 page 19.

► Si $\delta = \alpha$ alors comme on a $(\delta \in \beta \text{ ou } \delta = \beta)$ on a $\boxed{\alpha \in \beta \text{ ou } \alpha = \beta}$.

► Si $\delta = \beta$ alors comme on a $(\delta \in \alpha \text{ ou } \delta = \alpha)$ on a $\boxed{\beta \in \alpha \text{ ou } \beta = \alpha}$.

► Sinon si $\delta \neq \alpha$ et $\delta \neq \beta$ alors d'après ce qui précède on a $\delta \in \alpha$ et $\delta \in \beta$.

On a donc $\delta \in \alpha \cap \beta$ par définition de l'intersection.

Mais on a aussi $\delta = \alpha \cap \beta$ par définition de δ , donc $\delta \in \delta$, ce qui contredit 1.

Finalement, on a bien $\boxed{\alpha \in \beta \text{ ou } \beta \in \alpha \text{ ou } \alpha = \beta}$.

4. Comme E est non vide, il existe $\varepsilon \in E$.

Si ε est le minimum de (E, \in) c'est bon.

Supposons donc que ε n'est pas le minimum de (E, \in) .

Il existe donc $\mu \in E$ tel que l'on n'a ni $\varepsilon \in \mu$ ni $\varepsilon = \mu$.

Or tous les éléments de E sont des ordinaux donc $\mu \in \varepsilon$ d'après 3.

Ainsi $\mu \in E$ et $\mu \in \varepsilon$ donc $\mu \in \varepsilon \cap E$ et donc $\varepsilon \cap E \neq \emptyset$.

Donc $\varepsilon \cap E$ est une partie non vide de ε .

Or (ε, \in) est strictement bien ordonné car ε est un ordinal.

Donc $\varepsilon \cap E$ possède un minimum ξ .

Montrons que ξ est le minimum de E .

Soit $\nu \in E$.

Comme tous les éléments de E sont des ordinaux, on a $\nu \in \varepsilon$ ou $\varepsilon \in \nu$ ou $\nu = \varepsilon$ d'après 3.

- Si $\nu \in \varepsilon$ alors $\nu \in \varepsilon \cap E$ donc $\xi \in \nu$ car ξ est le minimum de $\varepsilon \cap E$.
- Si $\varepsilon \in \nu$, comme $\xi \in \varepsilon \cap E$ on a $\xi \in \varepsilon$ donc $\xi \in \varepsilon \in \nu$ et donc $\xi \in \nu$ d'après 1.
- Si $\nu = \varepsilon$, comme $\xi \in \varepsilon \cap E$ on a $\xi \in \varepsilon$ donc $\xi \in \nu$.

Dans tous les cas on a $\xi \in \nu$.

Donc ξ est le minimum de E .

Dans tous les cas, E admet un minimum.

CQFD.

Remarque :

1. Ainsi on dira simplement que (ON, \in) est une classe strictement bien ordonnée, et grâce à la proposition 8 page 19, nous savons que l'ordre associé est \subseteq , donc nous dirons aussi que (ON, \subseteq) est une classe bien ordonnée. Ces affirmations doivent être comprises comme étant un résumé du théorème qui précède.
2. Désormais pour α et β deux ordinaux, il va arriver fréquemment que nous notions $\alpha < \beta$ à la place de $\alpha \in \beta$ et $\alpha \leq \beta$ à la place de $\alpha \subseteq \beta$. **Ce ne sera pas toujours le cas**, mais quand nous le ferons ce sera pour insister sur le fait que c'est en tant que relation d'ordre strict et relation d'ordre (large) sur ON que nous employons ces objets mathématiques. Dans le cas où c'est véritablement l'idée d'appartenance et d'inclusion ensembliste qui nous intéressera, là nous resterons bel et bien avec les symboles \in et \subseteq . À ce titre, nous aurons parfois l'occasion de jongler avec les deux types de symboles.
3. \emptyset est le plus petit des ordinaux. En effet, on a déjà montré dans un exemple précédent que \emptyset est un ordinal, et on sait déjà que pour tout ensemble E on a $\emptyset \subseteq E$. En particulier pour tout ordinal α on a $\emptyset \subseteq \alpha$ et donc $\emptyset \leq \alpha$.

Nous avons expliqué avant le théorème que la classe des ordinaux ON se comporte elle-même comme un ordinal, mais nous n'avons pas montré de propriété qui s'apparente à la transitivité d'un ordinal. En réalité si, c'est l'objet de la proposition 6 page 17 qui affirme que tout élément d'un ordinal est aussi un ordinal. Autrement dit, pour tout $\alpha \in ON$, tous les éléments de α sont des ordinaux et donc $\alpha \subseteq ON$.

Nous avons affirmé pour justifier de l'intérêt des classes qu'il n'existe pas d'ensemble de tous les ordinaux, si bien que ON est une classe qui n'est pas issue d'un ensemble (elle est donc une classe **propre**). Montrons-le enfin : c'est le fameux **paradoxe de Burali-Forti**.

Théorème 2 (Paradoxe de Burali-Forti)

Il n'existe pas d'ensemble contenant tous les ordinaux.

Démonstration

Supposons par l'absurde qu'il existe un ensemble E tel que tout ordinal en est un élément. Potentiellement il y a d'autres éléments dans E qui ne sont pas des ordinaux, mais on est cependant sûr que tout ordinal est un élément de E .

Considérons alors $X := \{x \in E \mid x \text{ est un ordinal}\}$.

X est donc par définition l'ensemble de tous les ordinaux.

Pour arriver à la contradiction, nous allons montrer qu'alors X est en fait lui-même un ordinal, si bien que $X \in X$, ce qui est impossible chez les ordinaux.

- Montrons X est transitif.

Soit $\alpha \in X$.

Par définition α est un ordinal.

Soit $\beta \in \alpha$.

Alors β est un ordinal d'après la proposition 6 page 17.

On a donc $\beta \in E$ par définition de E et donc $\beta \in X$ par définition de X .

Donc $\forall \beta \in \alpha, \beta \in X$ donc $\alpha \subseteq X$ par définition de l'inclusion.

Donc $\forall \alpha \in X, \alpha \subseteq X$ donc X est transitif.

- D'après le théorème 1 page 21 (X, \in) est strictement bien ordonné car tous ses éléments sont des ordinaux.

Ainsi X est transitif et (X, \in) est strictement bien ordonné.

Donc X est un ordinal donc $X \in E$ par définition de E donc $X \in X$ par définition de X .

C'est en contradiction avec l'antiréflexivité de \in chez les ordinaux.

CQFD.

Pour la petite histoire



Cesare Burali-Forti (13 août 1861 – 21 janvier 1931) est un mathématicien italien.

Cesare Burali-Forti est assistant de Giuseppe Peano à Turin de 1894 à 1896. Il a travaillé sur les fondements de la géométrie, sur la géométrie différentielle et le calcul vectoriel. Il a aussi étudié la validité de la théorie de la relativité.

Bertrand Russell a nommé paradoxe de Burali-Forti, le paradoxe du plus grand ordinal en théorie des ensembles, en référence à un article de 1897 où le mathématicien italien, croyant démontrer que deux ordinaux ne sont pas toujours comparables, fait le raisonnement qui conduit au paradoxe décrit par Russell.

Ainsi, il n'existe pas d'ensemble contenant tous les ordinaux : ON n'est pas issue d'un ensemble, c'est donc une **classe propre**. De fait parmi toutes les sous-classes de ON , certaines sont propres (elle-même par exemple). Nous avons vu lors du théorème 1 page 21 que tout ensemble $X \subseteq ON$ possède un minimum. En fait, ce résultat reste vrai si on remplace X par une classe quelconque, pas forcément issue d'un ensemble.

Proposition 9 (Les sous-classes de ON possèdent un minimum)

Soit $C \subseteq ON$ une classe non vide.

Alors (C, \leq) possède un ordinal minimum, c'est-à-dire $\exists \xi \in C, \forall \alpha \in C, \xi \leq \alpha$.

Démonstration

- Supposons que C est issue d'un ensemble X non vide.

En particulier $X \subseteq ON$ par définition de C .

D'après le théorème 1 page 21, X possède un ordinal minimum.

Donc C possède un ordinal minimum.

- Supposons que C est propre.

En particulier elle n'est pas vide car pas issue de l'ensemble vide.

Il existe donc au moins un ordinal $\alpha \in C$.

Posons alors $X := \alpha \cap C = \{\beta \in \alpha \mid \beta \in C\} = \{\beta \in \alpha \mid C(\beta)\}$.

D'après l'axiome de compréhension, X est un ensemble.

► Supposons que X est vide.

Montrons que α est le minimum de C .

Soit $\beta \in C$.

Par définition de C on a $C \subseteq ON$.

Donc α et β sont des ordinaux.

On a donc $\alpha \leq \beta$ ou $\beta < \alpha$ d'après le théorème 1 page 21.

Supposons par l'absurde que $\beta < \alpha$.

On a donc $\beta \in \alpha$ par définition de $<$.

On a donc $\beta \in \alpha \cap C = X$.

Donc X est non vide.

C'est absurde puisqu'on a justement supposé que X est vide.

Par l'absurde on vient donc de montrer que $\alpha \leq \beta$.

Donc $\forall \beta \in C, \alpha \leq \beta$.

Donc α est le minimum de C .

► Supposons que X n'est pas vide.

Comme $X = \alpha \cap C$, on a $X \subseteq \alpha$.

Or α est un ordinal donc tous ses éléments sont des ordinaux.

Donc X est un ensemble non vide d'ordinaux.

Donc X possède un ordinal minimum ξ d'après le théorème 1 page 21.

Montrons que ξ est le minimum de C .

Comme $\xi \in X = \alpha \cap C$, on a déjà $\xi \in C$.

Soit $\beta \in C$.

Par définition de C on a $C \subseteq ON$.

Donc α et β sont des ordinaux.

On a donc $\alpha \leq \beta$ ou $\beta < \alpha$ d'après le théorème 1 page 21.

Supposons que $\alpha \leq \beta$, c'est-à-dire $\alpha \subseteq \beta$.

On a $\xi \in X = \alpha \cap C$ donc $\xi \in \alpha$.

On a donc $\xi \in \beta$ par définition de l'inclusion.

Donc $\xi \subseteq \beta$ car β est transitif car ordinal.

Supposons que $\beta < \alpha$, c'est-à-dire $\beta \in \alpha$.

Comme $\beta \in C$, on a $\beta \in \alpha \cap C = X$.

Or ξ est le minimum de X donc $\xi \subseteq \beta$.

Donc $\forall \xi \in C, \xi \subseteq \beta$, et donc $\forall \xi \in C, \xi \leq \beta$.

Donc ξ est le minimum de C .

Dans tous les cas, C possède un ordinal minimum.

CQFD.

Nous venons de voir que ON est une classe **propre**, et qu'elle se comporte *comme un ordinal*, c'est-à-dire :

1. ON est transitive, au sens où si $\beta \in \alpha \in ON$ alors $\beta \in ON$, puisque les éléments d'un ordinal sont eux aussi des ordinaux.
2. (ON, \in) est strictement bien ordonné, d'après le théorème 1 page 21.

Il s'avère qu'en fait, il s'agit de la seule classe propre à vérifier ces deux propriétés, comme le montre la proposition suivante. En cela, ON est en quelque sorte l'ordinal ultime.

Proposition 10 (ON est la seule classe propre ordinaire)

Soit C une classe propre.

Supposons que :

1. C est transitive, c'est-à-dire $\forall \alpha \in C, \alpha \subseteq C$.
2. (C, \in) est strictement bien ordonné, c'est-à-dire :
 - \in est antiréflexif sur C : $\forall \alpha \in C, \alpha \notin \alpha$
 - \in est transitif sur C : $\forall \alpha \in C, \forall \beta \in C, \forall \gamma \in C, (\alpha \in \beta \in \gamma \implies \alpha \in \gamma)$.
 - Tout ensemble non vide $X \subseteq C$ possède un minimum pour \in .

Alors $C = ON$.



Démonstration

- Commençons par montrer que $C \subseteq ON$.

Autrement dit, montrons que tous les éléments de C sont des ordinaux.

Soit $\alpha \in C$.

Montrons que α est un ordinal.

► Montrons que α est transitif.

Par hypothèse C est transitive donc $\alpha \subseteq C$.

Soit $\beta \in \alpha$.

Comme $\alpha \subseteq C$, on a $\beta \in C$ par définition de l'inclusion.

Donc $\beta \subseteq C$ par transitivité de C .

Soit $\gamma \in \beta$.

Comme $\beta \subseteq C$, on a $\gamma \in C$ par définition de l'inclusion.

Ainsi on a $\gamma \in \beta \in \alpha$ et tous trois sont éléments de C .

Or \in est transitive sur C par hypothèse donc $\gamma \in \alpha$.

Donc $\forall \gamma \in \beta, \gamma \in \alpha$ donc $\beta \subseteq \alpha$ par définition de l'inclusion.

Donc $\forall \beta \in \alpha, \beta \subseteq \alpha$ donc α est transitif.

► Montrons que (α, \in) est strictement bien ordonné.

Soient β, γ et δ dans α .

Supposons que $\beta \in \gamma \in \delta$.

On a dit que $\alpha \subseteq C$ donc β, γ et δ sont dans C .

Or \in est transitive sur C donc $\beta \in \delta$.

Donc si $\beta \in \gamma \in \delta$ alors $\beta \in \delta$.

Donc \in est transitive sur α .

Soit $\beta \in \alpha$.

On a dit que $\alpha \subseteq C$ donc $\beta \in C$ par définition de l'inclusion.

Donc $\beta \notin \beta$ par antiréflexivité de \in sur C .

Donc $\forall \beta \in \alpha, \beta \notin \beta$.

Donc \in est antiréflexive sur α .

Ainsi (α, \in) est strictement ordonné.

Soit X une partie non vide de α .

On a dit que $\alpha \subseteq C$ donc $X \subseteq C$ par transitivité de l'inclusion.

Ainsi X est un ensemble non vide inclus dans C .

Donc X possède un minimum pour \in par hypothèse.

Donc toutes les parties non vides de α possède un minimum pour \in .

On en conclut que (α, \in) est strictement bien ordonné.

Ainsi α est transitif et (α, \in) est strictement bien ordonné.

Donc α est un ordinal.

Donc tout élément de C est un ordinal, et donc $\boxed{C \subseteq ON}$.

• Montrons que $C = ON$.

Supposons par l'absurde que $C \neq ON$.

On a donc $C \subsetneq ON$ par ce qui précède.

Considérons alors $D := ON \setminus C$, qui est donc une classe non vide.

Alors (D, \in) possède un ordinal minimum δ d'après la proposition 9 page 24.

Montrons que $\delta = C$.

$\boxed{\subseteq}$

Soit $\alpha \in \delta$.

Comme δ est le minimum de (D, \in) , on a $\alpha \notin D$.

Or $\delta \in D = ON \setminus C$ donc $\delta \in ON$.

Donc α est un ordinal comme élément de l'ordinal δ .

Ainsi on a $\alpha \in ON$ et $\alpha \notin D$ donc $\alpha \in ON \setminus D = C$.

Donc $\forall \alpha \in \delta, \alpha \in C$ donc $\delta \subseteq C$ par définition de l'inclusion.



Soit $\alpha \in C$.

D'après ce qui précède, $C \subseteq ON$ donc $\alpha \in ON$.

Or on a aussi $\delta \in ON$.

On a donc $\alpha \in \delta$ ou $\alpha = \delta$ ou $\delta \in \alpha$ car (ON, \in) est strictement bien ordonné.

► Plaçons-nous dans le cas où $\alpha = \delta$.

Comme $\alpha \in C$, on a alors $\delta \in C$.

► Plaçons-nous dans le cas où $\delta \in \alpha$.

On a donc $\delta \in \alpha \in C$ et C est transitive par hypothèse.

On a donc $\alpha \subseteq C$ et donc $\delta \in C$ par définition de l'inclusion.

Donc si $\alpha = \delta$ ou $\delta \in \alpha$ alors $\delta \in C$.

C'est absurde puisque $\delta \in D = ON \setminus C$ donc $\delta \notin C$.

On en déduit que $\alpha \in \delta$.

Ainsi $\forall \alpha \in C, \alpha \in \delta$ donc $C \subseteq \delta$ par définition de l'inclusion.

On en déduit donc que $C = \delta$ par ce qui précède.

Ainsi C est un ordinal : en particulier C est un ensemble.

C'est absurde puisque par définition C est une classe propre.

Par l'absurde, on a donc montré que nécessairement $C = ON$.

CQFD.

Cette propriété que nous venons de voir fait donc de ON en quelque sorte l'unique classe propre à pouvoir prétendre généraliser la notion d'ordinaux. Nous aurons l'occasion de la revoir quand nous aurons besoin d'étendre une définition qui initialement ne porte que sur les ordinaux : il semblera légitime de ne l'étendre qu'à ON .

Nous avons vu lors de la proposition 7 page 18 que l'intersection de deux ordinaux est aussi un ordinal. Il en va en fait de même pour l'union de deux ordinaux, et plus généralement pour l'intersection et la réunion d'ensembles d'ordinaux. Cela nous fournit au passage une expression explicite de la borne supérieure et du minimum d'un ensemble d'ordinaux.

Proposition 11 (Union et intersection d'ordinaux)

Soient α et β deux ordinaux.

1. $\alpha \cup \beta$ est un ordinal et $\alpha \cup \beta = \max(\alpha, \beta)$.
2. $\alpha \cap \beta$ est un ordinal et $\alpha \cap \beta = \min(\alpha, \beta)$.

Soit X un ensemble dont tous les éléments sont des ordinaux.

3. $\bigcup X$ est un ordinal et $\bigcup X = \sup(X)$.
La notion de borne supérieure est à comprendre ici "*parmi les ordinaux*".
4. Si $X \neq \emptyset$ alors $\bigcap X$ est un ordinal et $\bigcap X = \min(X)$.



Démonstration

On a $\alpha \leq \beta$ ou $\beta < \alpha$ d'après le théorème 1 page 21.

1.

- Si $\alpha \leq \beta$ alors $\beta = \max(\alpha, \beta)$ par définition du maximum.
Or on a $\alpha \subseteq \beta$ par définition de \leq , donc $\alpha \cup \beta = \beta$.
En particulier $\alpha \cup \beta$ est un ordinal, et par ce qui précède $\alpha \cup \beta = \max(\alpha, \beta)$.
- Si $\beta < \alpha$ on a $\alpha = \max(\alpha, \beta)$ par définition du maximum.
On a en particulier $\beta \leq \alpha$ puisque c'est l'ordre large associé.
On a donc $\beta \subseteq \alpha$ par définition de \leq et donc $\alpha \cup \beta = \alpha$.
En particulier $\alpha \cup \beta$ est un ordinal, et par ce qui précède $\alpha \cup \beta = \max(\alpha, \beta)$.

Dans tous les cas $\boxed{\alpha \cup \beta \text{ est un ordinal et } \alpha \cup \beta = \max(\alpha, \beta)}$.

2. On a déjà vu lors de la proposition 7 page 18 que $\boxed{\alpha \cap \beta \text{ est un ordinal}}$.

- Si $\alpha \leq \beta$ alors $\alpha = \min(\alpha, \beta)$ par définition du minimum.
Or on a $\alpha \subseteq \beta$ par définition de \leq donc $\alpha \cap \beta = \alpha$.
On a donc $\alpha \cap \beta = \min(\alpha, \beta)$.
- Si $\beta < \alpha$ alors on a $\beta = \min(\alpha, \beta)$ par définition du minimum.
En particulier on a $\beta \leq \alpha$ puisque c'est l'ordre large associé.
On a donc $\beta \subseteq \alpha$ par définition de \leq .
On a donc $\alpha \cap \beta = \beta$ et donc $\alpha \cap \beta = \min(\alpha, \beta)$ par ce qui précède.

Dans tous les cas on a $\boxed{\alpha \cap \beta = \min(\alpha, \beta)}$.

3. Commençons par montrer que $\bigcup X$ est un ordinal.

- Montrons que $\bigcup X$ est transitif.

Soit $x \in \bigcup X$.

Par définition de la réunion, il existe $\alpha \in X$ tel que $x \in \alpha$.

Comme X est un ensemble d'ordinaux, α est un ordinal.

Donc α est transitif et donc $x \subseteq \alpha$.

Comme $\alpha \in X$, on a $\alpha \subseteq \bigcup X$ donc $x \subseteq \bigcup X$ par transitivité de l'inclusion.

Donc $\forall x \in \bigcup X, x \subseteq \bigcup X$ donc $\boxed{\bigcup X \text{ est transitif}}$.

- Montrons que \in est un bon ordre strict sur $\bigcup X$.

► \in est antiréflexive sur $\bigcup X$.

Soit $x \in \bigcup X$.

Par définition de la réunion, il existe $\alpha \in X$ tel que $x \in \alpha$.

Comme X est un ensemble d'ordinaux, α est un ordinal.

Donc x est un ordinal d'après la proposition 6 page 17.

En particulier $x \notin x$ par antiréflexivité de \in sur ON .

Donc $\forall x \in \bigcup X, x \notin x$ donc \in est antiréflexive sur $\bigcup X$.

► \in est transitive sur $\bigcup X$.

Soient x, y et z dans $\bigcup X$.

Il existe α, β et γ dans X tels que $x \in \alpha, y \in \beta$ et $z \in \gamma$.

Or tous les éléments de X sont des ordinaux donc α, β et γ sont des ordinaux.

En particulier d'après 1 $\alpha \cup \beta \cup \gamma$ est un ordinal, dont x, y et z sont des éléments.

Supposons que $x \in y \in z$.

On vient de dire que $\alpha \cup \beta \cup \gamma$ est un ordinal.

Donc $(\alpha \cup \beta \cup \gamma, \in)$ est strictement bien ordonné.

Donc \in est transitive sur $\alpha \cup \beta \cup \gamma$.

On a donc $x \in z$ par transitivité.

Donc si $x \in y \in z$ alors $x \in z$.

Donc \in est transitive sur $\bigcup X$.

Ainsi \in est un ordre strict sur $\bigcup X$.

► \in est un bon ordre strict sur $\bigcup X$.

Soit A une partie non vide de $\bigcup X$.

Soit $a \in A$.

Comme $A \subseteq \bigcup X$ on a $a \in \bigcup X$ par définition de l'inclusion.

Par définition de la réunion, il existe $\alpha \in X$ tel que $a \in \alpha$.

Or tous les éléments de X sont des ordinaux donc α est un ordinal.

Donc a est un ordinal d'après la proposition 6 page 17.

Donc tous les éléments de A sont des ordinaux.

Comme A est non vide, il possède un minimum d'après le théorème 1 page 21.

Donc toutes les parties non vides de $\bigcup X$ possèdent un minimum.

Donc \in est un bon ordre strict sur $\bigcup X$.

Donc $\boxed{\bigcup X \text{ est un ordinal}}$.

- Montrons que $\bigcup X = \sup(X)$.

Pour tout $\alpha \in X$, on a $\alpha \subseteq \bigcup X$ par définition de la réunion.

En particulier $\bigcup X$ est un majorant de X dans (ON, \subseteq) .

Soit β un majorant de X dans (ON, \subseteq) .

On a donc pour tout $\alpha \in X$, on a $\alpha \subseteq \beta$.

On a donc $\bigcup X \subseteq \beta$ par minimalité de la réunion pour l'inclusion.

Donc tout ordinal majorant de X dans (ON, \subseteq) est plus grand que ou égal à $\bigcup X$.

Ainsi, $\bigcup X$ est le plus petit ordinal majorant de X dans (ON, \subseteq) .

Donc $\boxed{\sup(X) = \bigcup X}$.

4. Supposons que X est non vide.

Commençons par montrer que $\bigcap X$ est un ordinal.

- $\bigcap X$ est transitif.

En effet, soit $x \in \bigcap X$.

Pour tout $\alpha \in X$, on a $x \in \alpha$.

Or tous les éléments de X sont des ordinaux donc sont transitifs.

Donc pour tout $\alpha \in X$, on a $x \subseteq \alpha$.

Donc $x \subseteq \bigcap X$ par maximalité de l'intersection pour l'inclusion.

Donc $\forall x \in \bigcap X, x \subseteq \bigcap X$.

Donc $\bigcap X$ est transitif.

- Comme X est non vide, il existe $\alpha \in X$.

On a alors $\bigcap X \subseteq \alpha$.

Or tous les éléments de X sont des ordinaux donc α est un ordinal.

Donc (α, \in) est strictement bien ordonné.

Donc $(\bigcap X, \in)$ est strictement bien ordonné d'après la proposition 3 page 11.

On en conclut que $\boxed{\bigcap X \text{ est un ordinal}}$.

- Montrons que $\bigcap X = \min(X)$.

Par définition X est un ensemble non vide d'ordinaux.

Donc X admet un minimum ξ d'après le théorème 1 page 21.

Ainsi $\forall \alpha \in X, \xi \leq \alpha$ par définition du minimum.

Donc $\forall \alpha \in X, \xi \subseteq \alpha$ par définition de \subseteq .

Donc $\xi \subseteq \bigcap X$ par maximalité de l'intersection pour l'inclusion.

Or on a $\xi \in X$ puisque ξ est le minimum de X .

On a donc $\bigcap X \subseteq \xi$ par définition de l'intersection.

On a donc $\bigcap X = \xi$ par antisymétrie de l'inclusion.

Or par définition $\xi = \min(X)$ donc $\boxed{\bigcap X = \min(X)}$.

CQFD.

Ainsi, on vient de voir que tout ensemble d'ordinaux admet une borne supérieure : c'est sa réunion. En particulier, tout ensemble d'ordinaux est majoré. En fait, cela fonctionne aussi dans l'autre sens : si une classe d'ordinaux est majorée, alors cette classe est (issue d')un ensemble !

Proposition 12 (Une classe majorée est un ensemble)

Soit C une classe d'ordinaux.

Supposons que C est majorée, c'est-à-dire qu'il existe un ordinal α tel que $\forall \beta \in C, \beta < \alpha$.

Alors C est un ensemble.

 *Démonstration*

Par hypothèse on a $\forall \beta \in C, \beta < \alpha$.

On a donc $\forall \beta \in C, \beta \in \alpha$ par définition de $<$.

En particulier on a $C = \{\beta \in C \mid \beta \in \alpha\} = \{\beta \in \alpha \mid \beta \in C\}$, qui est un ensemble d'après l'axiome de compréhension, car α est un ordinal donc un ensemble.

CQFD.

4 Successseurs, limites et entiers naturels

Ce qui nous a motivé à introduire la notion de bon ordre est le fait qu'étant donné un élément x , il est possible de répondre à la question « *quel élément suit directement x ?* ». En effet, nous avons dit que dans ce cas-là, il suffit de prendre l'ensemble des éléments strictement plus grands que x , et de considérer alors son minimum. On parle alors du **successeur** de x . Dans le cas particulier des ordinaux, ce successeur a une expression simple (que l'on peut quand-même définir en toute généralité).

Définition 10 (Successseur d'un ensemble)

Soit x un ensemble.

On appelle **successeur** de x l'ensemble $S(x) := x \cup \{x\}$.

Exemple :

Il est temps de définir nos premiers entiers naturels.

On pose $0 := \emptyset$ et $1 := S(0)$.

Plus précisément, on a $1 = S(0) = S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\}$.

Nous verrons un peu plus tard comment définir tous les autres entiers naturels.

Quand nous aurons défini les entiers naturels, nous verrons que $n + 1$ sera défini comme étant $S(n)$, ce qui correspond bien à l'intuition de l'entier naturel qui suit directement n .

Nous avons défini la notion de successeur d'un ensemble en toute généralité, mais cette notion devient intéressante dans le cas des ordinaux puisqu'elle répond bien à la question de l'ordinal qui suit directement.

Proposition 13 (Successseur d'un ordinal)

Soient α et β deux ordinaux.

1. $S(\alpha)$ est un ordinal tel que $\alpha < S(\alpha)$.

Ainsi $S(\alpha)$ est un ordinal strictement plus grand que α .

Ou encore, α est un ordinal strictement plus petit que $S(\alpha)$.

2. On a l'équivalence $\alpha < \beta \iff S(\alpha) \leq \beta$.

Ainsi $S(\alpha)$ est le plus petit des ordinaux strictement plus grands qu' α .

3. On a l'équivalence $\beta < S(\alpha) \iff \beta \leq \alpha$.

Ainsi α est le plus grand des ordinaux strictement plus petits que $S(\alpha)$.

Ainsi $S(\alpha)$ porte bien son nom : c'est l'ordinal qui vient « *juste après* » α .

4. On a l'implication $S(\alpha) = S(\beta) \implies \alpha = \beta$.

Ainsi le passage au successeur est injectif.

 *Démonstration*

1. Commençons par montrer que $S(\alpha)$ est un ordinal.

Pour cela, montrons que $S(\alpha)$ est transitif.

Soit $x \in S(\alpha)$.

On a donc $x \in \alpha \cup \{\alpha\}$ par définition du successeur.

On a donc $x \in \alpha$ ou $x \in \{\alpha\}$.

► Plaçons-nous dans le cas où $x \in \alpha$.

On a alors $x \subseteq \alpha$ car α est transitif car ordinal.

► Plaçons-nous dans le cas où $x = \alpha$.

On a alors $x \subseteq \alpha$ par réflexivité de l'inclusion.

Dans les deux cas on a donc $x \subseteq \alpha$.

En particulier on a $x \subseteq \alpha \cup \{\alpha\}$ et donc $x \subseteq S(\alpha)$.

Donc pour tout $x \in S(\alpha)$, on a $x \subseteq S(\alpha)$.

Donc $S(\alpha)$ est transitif.

Montrons que $(S(\alpha), \in)$ est strictement bien ordonné.

On a $S(\alpha) = \alpha \cup \{\alpha\}$.

Donc chaque élément de $S(\alpha)$ est soit un élément de α , soit α lui-même.

Or α est un ordinal.

Donc tous les éléments de α sont des ordinaux d'après la prop. 6 p. 17.

Donc tous les éléments de $S(\alpha)$ sont des ordinaux.

Donc \in est transitive et antiréflexive sur $S(\alpha)$ d'après le théorème 1 page 21.

De même, toute partie non vide de $S(\alpha)$ est alors un ensemble non vide d'ordinaux.

Donc toute partie non vide de $S(\alpha)$ admet un minimum d'après ce même théorème.

On en conclut que \in est un bon ordre strict sur $S(\alpha)$.

Ainsi, $S(\alpha)$ est transitif et \in est un bon ordre strict sur $S(\alpha)$.

Donc $S(\alpha)$ est un ordinal.

Par définition on a $\alpha \in \{\alpha\}$.

On a donc $\alpha \in \alpha \cup \{\alpha\}$.

On a donc $\alpha \in S(\alpha)$ par définition du successeur.

Autrement dit on a $\alpha < S(\alpha)$.

2. Raisonnons par double implications.



Supposons que $\alpha < \beta$.

En particulier on a $\alpha \leq \beta$ donc $\alpha \subseteq \beta$.

De plus comme $\alpha < \beta$, on a $\alpha \in \beta$ et donc $\{\alpha\} \subseteq \beta$.

Comme $\alpha \subseteq \beta$ et $\{\alpha\} \subseteq \beta$, on a $S(\alpha) = \alpha \cup \{\alpha\} \subseteq \beta$.

On a donc $S(\alpha) \leq \beta$.

Donc si $\alpha < \beta$ alors $S(\alpha) \leq \beta$.



Supposons que $S(\alpha) \leq \beta$.

On a donc $S(\alpha) \subseteq \beta$ et donc $\alpha \cup \{\alpha\} \subseteq \beta$.

En particulier $\{\alpha\} \subseteq \beta$ donc $\alpha \in \beta$ et donc $\alpha < \beta$.

Donc si $S(\alpha) \leq \beta$ alors $\alpha < \beta$.

Finalement, $\boxed{\alpha < \beta \iff S(\alpha) \leq \beta}$.

3. On a les équivalences suivantes

$$\begin{aligned}\beta < S(\alpha) &\iff \beta \in S(\alpha) \\ &\iff \beta \in \alpha \cup \{\alpha\} \\ &\iff \beta \in \alpha \text{ ou } \beta = \alpha \\ &\iff \beta \subseteq \alpha \text{ d'après la prop. 8 p. 19} \\ &\iff \beta \leq \alpha\end{aligned}$$

On a donc bien l'équivalence $\boxed{\beta < S(\alpha) \iff \beta \leq \alpha}$.

4. Supposons que $S(\alpha) = S(\beta)$.

D'après 1 on a $\beta < S(\beta)$ donc $\beta < S(\alpha)$ et donc $\beta \leq \alpha$ d'après 3.

De même, on a $\alpha < S(\alpha)$ donc $\alpha < S(\beta)$ et donc $\alpha \leq \beta$ d'après 3.

On a donc bien $\boxed{\alpha = \beta}$ par antisymétrie de \leq .

CQFD.

Remarque :

Comme on a défini 0 comme étant égal à \emptyset , on a déjà vu que 0 est le plus petit des ordinaux. En particulier il n'est le successeur d'aucun ordinal puisque sinon celui-ci serait strictement plus petit que 0.

Nous allons enfin définir la notion d'**entiers naturels** : il s'agit des premiers ordinaux. En effet 0 est le plus petit des ordinaux, 1 est son successeur, 2 est le successeur de 1 et ainsi de suite. On pourrait penser qu'en partant de 0 et en enchaînant l'opération de successeur suffisamment de fois, on finirait par avoir parcouru tous les ordinaux. Il n'en est rien : il existe des ordinaux qui ne seront jamais atteints de cette manière. Le plus petit de ces ordinaux est noté ω : oui, il s'agit tout simplement de l'ensemble des entiers naturels, aussi noté \mathbb{N} . Tout ordinal plus petit que lui va donc lui appartenir, et donc être un entier naturel. Il n'a donc pas de prédécesseur direct : en effet si $\alpha < \omega$ alors $\alpha \in \omega$ donc α est un entier naturel par définition donc $S(\alpha) = \alpha + 1$ est aussi un entier naturel donc n'est pas ω (puisque sinon $\omega \in \omega$, ce qui est impossible chez les ordinaux).

Ainsi il existe des ordinaux qui ne sont pas successeurs d'aucun ordinal : il y a 0 bien sûr, mais aussi ω comme nous venons de le voir. Nous allons les appeler ordinaux **limites**, car il y a en quelque sorte une limite à franchir pour les atteindre, on ne peut pas simplement partir d'un ordinal et enchaîner des opérations de successeurs.

Définition 11 (Ordinaux successeurs, limites et entiers naturels)

Soit β un ordinal.

1. On dit que β est **successeur** si et seulement s'il existe un ordinal α tel que $\beta = S(\alpha)$.
2. On dit que β est **limite** si et seulement si β n'est pas successeur.
3. On dit que β est un **entier naturel** si et seulement si pour tout ordinal $\alpha \leq \beta$,
 - ou bien $\alpha = 0$
 - ou bien α est un ordinal successeur.

On dit aussi que β est **fini**.

Dans le cas contraire on dit que β est **infini**, ou encore **transfini**.

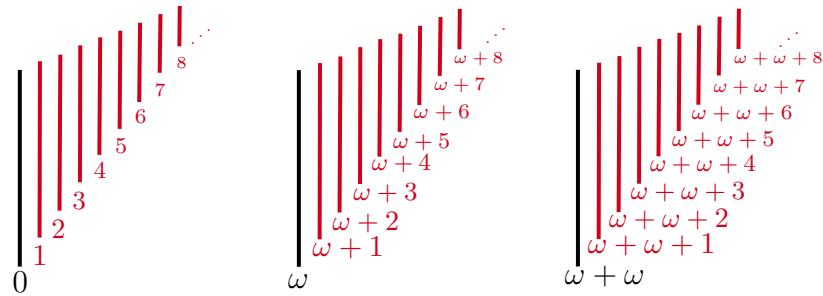
Exemple :

1. Comme $0 = \emptyset$, on a déjà vu que 0 est un ordinal.
1 en tant que successeur de 0 est aussi un ordinal.
1 est un successeur (de 0 donc) et 0 est limite car non successeur car vide.
0 et 1 sont tous les deux des entiers naturels.
2. Quand nous l'aurons défini, nous verrons que $\mathbb{N} = \omega$ l'ensemble des entiers naturels est un ordinal limite.
3. De même, nous définirons plus tard l'ordinal $\omega + 1 = S(\omega)$, qui est lui bien successeur de ω donc n'est pas limite, mais n'est pas entier naturel non plus puisque $\omega \leq \omega + 1$ alors que ω n'est ni 0 ni successeur.

Remarque :

La plupart des ouvrages sur le sujet considère que 0 n'est pas un ordinal limite. En effet, dans ce cas-là il n'y a pas eu de limite à franchir pour l'atteindre. Pour des soucis de simplification d'énoncés, nous considérons ici bien qu'il l'est : au contraire l'exclure demande souvent de l'exclure artificiellement de beaucoup d'énoncés de résultats sur les ordinaux limites.

Pour aider à visualiser tout cela, on peut proposer l'illustration suivante :



Une représentation visuelle des ordinaux.

Il faut ici voir la disposition des bâtons comme s'étendant à l'infini à l'horizon, l'ordre des bâtons étant rangés de la gauche vers la droite : au début on a un bâton pour chaque entier naturel, puis après tous les entiers naturels vient le bâton associé à ω . Ensuite vient le bâton associé à $\omega + 1$, puis $\omega + 2$ et ainsi de suite pour tous les ordinaux de la forme $\omega + n$ où n est un entier naturel, donc une infinité de bâtons sont disposés après celui de ω . Mais après tous ceux-là se trouve un bâton associé à l'ordinal $\omega + \omega$, et ainsi de suite. Nous aurons bien entendu tout le temps de définir proprement chacun de ces ordinaux, l'idée est ici simplement de comprendre intuitivement ce que nous sommes en train de construire. Gardons bien en tête que la taille des bâtons n'a aucune importance, seul leur agencement horizontal importe. Le fait de représenter des bâtons de plus en plus petits est seulement une astuce pour en faire tenir une infinité.

On peut voir sur l'illustration que les ordinaux limites sont les bâtons qui n'ont pas de prédécesseur direct : nous les avons représentés en **noir**. En **rouge** sont représentés les ordinaux successeurs.

Proposition 14 (successeur et ordinal limite)

Soit α un ordinal.

Les assertions suivantes sont équivalentes :

1. α est un ordinal limite.
2. Pour tout ordinal $\beta < \alpha$ on a $S(\beta) < \alpha$.

Autrement dit, un ordinal est limite si et seulement si partant d'un ordinal strictement plus petit, on reste strictement plus petit même en passant au successeur.

Démonstration

$1 \Rightarrow 2$

Supposons que α est un ordinal limite.

Soit β un ordinal tel que $\beta < \alpha$.

On a alors $S(\beta) \leq \alpha$ d'après la proposition 13 page 33.

On a donc $S(\beta) < \alpha$ ou $S(\beta) = \alpha$.

Or $S(\beta) = \alpha$ est impossible car par définition α est un ordinal limite.

On a donc nécessairement $S(\beta) < \alpha$.

Donc pour tout ordinal $\beta < \alpha$ on a $S(\beta) < \alpha$.

Donc si α est un ordinal limite alors pour tout ordinal $\beta < \alpha$, on a $S(\beta) < \alpha$.

2⇒1

Supposons que pour tout ordinal $\beta < \alpha$ on a $S(\beta) < \alpha$.

Supposons par l'absurde que α n'est pas limite.

Par définition, α est donc successeur.

Il existe donc un ordinal β tel que $\alpha = S(\beta)$.

Or on a $\beta < S(\beta)$ d'après la proposition 13 page 33 donc $\beta < \alpha$.

On a donc $S(\beta) < \alpha$ par l'hypothèse.

Or on a dit que $\alpha = S(\beta)$, si bien que $\alpha < \alpha$.

C'est absurde par antiréflexivité de $<$.

Par l'absurde, on vient de montrer que α est limite.

Donc si pour tout ordinal $\beta < \alpha$ on a $S(\beta) < \alpha$ alors α est limite.

CQFD.

Proposition 15 (Successeur d'un entier naturel)

Soit n un entier naturel.

1. $S(n)$ est un entier naturel.
2. Tout ordinal $\alpha \leq n$ est aussi un entier naturel.

Démonstration

1. Par définition n est un entier naturel donc est un ordinal.

Donc $S(n)$ est un ordinal d'après la proposition 13 page 33.

Soit un ordinal $\alpha \leq S(n)$.

On a donc ou bien $\alpha < S(n)$ ou bien $\alpha = S(n)$.

- Supposons que $\alpha < S(n)$.

On a alors $\alpha \leq n$ d'après la proposition 13 page 33.

Or n est un entier naturel.

Donc ou bien $\alpha = 0$, ou bien α est un successeur.

- Supposons que $\alpha = S(n)$. Comme n est un entier naturel, n est un ordinal.

Donc $\alpha = S(n)$ est un successeur.

Dans les deux cas, ou bien $\alpha = 0$ ou bien α est un successeur.

Donc tous les ordinaux plus petit que $S(n)$ sont ou 0 ou bien un successeur.

Comme $S(n)$ est un ordinal, c'est donc par définition un entier naturel.

2. Soit un ordinal $\alpha \leq n$.

Soit un ordinal $\beta \leq \alpha$.

On a donc $\beta \leq n$ par transitivité de \leq .

Or n est un entier naturel et β un ordinal.

Donc ou bien $\beta = 0$ ou bien β est un successeur.

Donc tous les ordinaux plus petits que α sont ou bien 0 ou bien un successeur.

Comme α est un ordinal, par définition α est un entier naturel.

CQFD.

Bien souvent en mathématiques nous sommes amenés à mener un **raisonnement par récurrence** afin de prouver qu'une assertion P à paramètres est vraie pour tout entier naturel n . Pour cela on raisonne en deux étapes :

1. On prouve que $P(0)$ est vraie : c'est l'étape d'**initialisation**.
2. On prouve que pour tout entier naturel n , si $P(n)$ est vraie alors $P(n + 1)$ est aussi vraie. C'est l'étape d'**héritéité**.

Grâce à ces deux étapes, on en conclut que $P(n)$ est vraie pour tout entier naturel n . Qu'est-ce qui justifie la validité de ce raisonnement ? La réponse se cache dans le théorème suivant.

Théorème 3 (Principe d'induction chez les entiers naturels)

Soit X un ensemble tel que :

1. $0 \in X$
2. Pour tout $x \in X$ on a $S(x) \in X$.

Alors X contient tous les entiers naturels.

Démonstration

Voici rapidement l'idée de la preuve : on suppose par l'absurde que X ne contient pas un entier naturel donné n . On regarde alors l'ensemble Y des entiers naturels plus petits que n qui ne sont pas dans X . En particulier n est dedans donc Y n'est pas vide : il va avoir un plus petit entier naturel k . Le soucis vient alors du fait que k est le plus petit entier naturel à ne pas être dans X , et donc $k - 1$ va être dans X , ce qui contredit l'hypothèse 2 selon laquelle X est stable par passage au successeur.

Soit n un entier naturel.

Supposons par l'absurde que $n \notin X$.

Considérons $Y := S(n) \setminus X$.

Comme n est un entier naturel, $S(n)$ est un entier naturel d'après la prop. 15 p. 38.

Donc tous les éléments de $S(n)$ sont des entiers naturels d'après la prop. 15 p. 38.

Or $Y = S(n) \setminus X$ donc $Y \subseteq S(n)$.

Donc tous les éléments de Y sont des entiers naturels.

On a $n < S(n)$ d'après la proposition 13 page 33.

On a donc $n \in S(n)$ par définition de $<$, et $n \notin X$ par hypothèse.

On a donc $n \in Y$ puisque $Y = S(n) \setminus X$.

Donc Y est un ensemble non vide d'entiers naturels.

Il possède donc un entier naturel minimum k d'après le théorème 1 page 21.

On a $k \leq k$ par réflexivité de \leq .

Donc k est un ordinal plus petit qu'un entier naturel (lui-même).

Donc k est ou bien 0 ou bien un successeur par définition.

Or $k \in Y = S(n) \setminus X$ donc $k \notin X$. Comme $0 \in X$ par hypothèse, on a donc $k \neq 0$.

Donc k est un successeur : il existe un ordinal i tel que $k = S(i)$.

Or on a $i < S(i)$ d'après la proposition 13 page 33 donc $i < k$.

Donc $i \notin Y$ car k est le minimum de Y .

Mais comme $n \in Y$, on a aussi $k \leq n$, toujours par minimalité de k .

Ainsi on a $i < k \leq n < S(n)$ donc $i < S(n)$ par transitivité.

On a donc $i \in S(n)$ par définition de $<$.

Ainsi $i \notin Y$ et $i \in S(n)$ donc $i \in S(n) \setminus Y = X$.

Or par hypothèse X est stable par successeur donc $S(i) \in X$ et donc $k \in X$.

C'est absurde puisque $k \in Y$ et $Y = S(n) \setminus X$ donc $k \notin X$.

Donc par l'absurde on vient donc de montrer $\boxed{n \in X}$.

CQFD.

Nous pouvons donc justifier la validité du raisonnement par récurrence : imaginons avoir démontré l'étape d'initialisation et l'étape d'hérédité. On peut alors considérer l'ensemble $X := \{n \in \mathbb{N} \mid P(n)\}$ qui répond alors aux hypothèses du théorème ci-dessus : il contient tous les entiers naturels, ce qui prouve donc bien que $P(n)$ est vraie pour tout entier naturel.

En vérité, il manque quelque chose pour valider ce que nous venons d'affirmer : l'existence de \mathbb{N} lui-même. En effet on affirme depuis le début que \mathbb{N} , l'ensemble de tous les entiers naturels existe, et on l'a même aussi noté ω en affirmant qu'il s'agit d'un ordinal. Malheureusement, on ne peut le faire sans un axiome, que l'on va donc rajouter aux différents axiomes de ZFC du précédent livre : l'**axiome de l'infini**.

Axiome 1 (de l'infini)

Il existe au moins un ensemble X tel que

1. $0 \in X$
2. Pour tout $x \in X$ on a $S(x) \in X$.

Nous sommes à présent armés pour définir proprement \mathbb{N} .

Proposition 16 (Ensemble des entiers naturels)

Il existe un unique ensemble \mathbb{N} tel que pour tout ensemble n , on a l'équivalence

$$n \in \mathbb{N} \iff n \text{ est un entier naturel}$$

On dit donc que \mathbb{N} est l'**ensemble des entiers naturels**, et on le note aussi parfois ω .

Démonstration

Existence :

D'après l'**axiome de l'infini**, il existe un ensemble X tel que

1. $0 \in X$
2. Pour tout $x \in X$ on a $S(x) \in X$.

Posons alors $\mathbb{N} := \{x \in X \mid x \text{ est un entier naturel}\}$.

Montrons que \mathbb{N} ainsi défini vérifie l'équivalence de l'énoncé.

Soit n un ensemble.

\Rightarrow Si $n \in \mathbb{N}$ alors par définition n est un entier naturel.



Supposons que n est un entier naturel.

Alors $n \in X$ d'après le principe d'induction chez les entiers naturels.

Donc $n \in \mathbb{N}$ et n est un entier naturel.

Donc $n \in \mathbb{N}$ par définition de $n \in \mathbb{N}$.

Donc si n est un entier naturel alors \mathbb{N} .

Ainsi pour tout ensemble n , on a bien l'équivalence $n \in \mathbb{N} \iff n \text{ est un entier naturel}$.

Unicité :

L'unicité est garantie par le fait que cette équivalence caractérise l'appartenance à \mathbb{N} .

CQFD.

Avant de prouver que \mathbb{N} est un ordinal, intéressons-nous aux segments initiaux de ON .

Proposition 17 (Segment initiaux des ordinaux)

Soit X un ensemble.

Les assertions suivantes sont équivalentes :

1. X est un ordinal.
2. Tous les éléments de X sont des ordinaux et X est transitif.
3. X est un segment initial de ON .



Démonstration

Nous allons montrer $1 \Leftrightarrow 2 \Leftrightarrow 3$.

$1 \Rightarrow 2$

Supposons que X est un ordinal.

En particulier X est transitif par définition.

De plus, tous les éléments de X sont des ordinaux d'après la proposition 6 page 17.

$1 \Leftrightarrow 2$

Supposons que tous les éléments de X sont des ordinaux et que X est transitif.

Comme X est un ensemble d'ordinaux, (X, \in) est strictement bien ordonné d'après le théorème 1 page 21. Comme X est transitif, on en conclut que $[X \text{ est un ordinal}]$.

$2 \Rightarrow 3$

Supposons que tous les éléments de X sont des ordinaux et que X est transitif.

En particulier on sait déjà que $X \subseteq ON$ par définition de ON .

Soient α et β deux ordinaux.

Supposons que $\alpha \leq \beta \in X$.

Comme $\alpha \leq \beta$, on a ($\alpha < \beta$ ou $\alpha = \beta$).

► Plaçons-nous dans le cas où $\alpha < \beta$.

Par définition cela veut dire que $\alpha \in \beta$.

Ainsi on a $\alpha \in \beta \in X$.

Or X est transitif, donc par définition $\alpha \in X$.

► Plaçons-nous dans le cas où $\alpha = \beta$.

Par hypothèse on a $\beta \in X$ donc $\alpha \in X$.

Donc dans les deux cas on a $\alpha \in X$.

Donc si $\alpha \leq \beta \in X$ alors $\alpha \in X$.

Donc $[X \text{ est un segment initial de } ON]$.

$2 \Leftarrow 3$

Supposons que X est un segment initial de ON .

Par définition on a alors $X \subseteq ON$.

Autrement dit, tous les éléments de X sont des ordinaux.

Soit $\beta \in X$.

Comme on vient de le dire, tous les éléments de X sont des ordinaux.

Donc β est un ordinal.

Soit $\alpha \in \beta$.

Alors α est un ordinal en tant qu'élément d'un ordinal.

Comme $\alpha \in \beta$ on a $\alpha < \beta$ et en particulier $\alpha \leq \beta$.

Ainsi α et β sont deux ordinaux tels que $\alpha \leq \beta \in X$.

On a donc $\alpha \in X$ puisque X est un segment initial de ON .

Donc $\forall \alpha \in \beta, \alpha \in X$ et donc $\beta \subseteq X$ par définition de l'inclusion.

Donc $\forall \beta \in X, \beta \subseteq X$ et donc X est transitif.

CQFD.

Remarque :

On a vu grâce au théorème 1 page 21 que les ordinaux sont munis d'un bon ordre.

X est un segment initial des ordinaux est donc équivalent à l'existence d'un ordinal ξ tel que $X = ON_{<\xi}$ d'après la proposition 5 page 13. Ici ξ est tout trouvé : c'est X lui-même d'après cette proposition. C'est d'ailleurs assez logique puisque la relation d'ordre strict sur les ordinaux est l'appartenance et donc

$$X = \{\alpha \in ON \mid \alpha \in X\} = \{\alpha \in ON \mid \alpha < X\} = ON_{<X}$$

Nous pouvons désormais prouver que ω , autre nom donné à \mathbb{N} , est un ordinal. Comme nous l'avons dit plus tôt, c'est même un ordinal limite, c'est-à-dire qu'il n'est pas un successeur. C'est même le plus petit des ordinaux limites non nuls, c'est-à-dire le tout premier après 0.

Proposition 18 (omega est le plus petit ordinal limite non nul)

1. ω est un ordinal limite.
 2. ω est le plus petit des ordinaux limites non nuls.
- Autrement dit pour tout ordinal limite non nul α , on a $\omega \leq \alpha$.

Démonstration

1.

• Montrons que ω est un ordinal.

D'après la proposition 17 page 42, il suffit de montrer que ω est un segment initial de ON .

ω ne contient que des entiers naturels (et les contient tous) par définition.

En particulier ω est un ensemble d'ordinaux : on sait déjà que $\omega \subseteq ON$.

Soient n et m deux ordinaux.

Supposons que $n \leq m \in \omega$.

On a $m \in \omega$ donc m est un entier naturel par définition.

On a $n \leq m$ donc n est un entier naturel d'après la proposition 15 page 38.

On a donc $n \in \omega$ par définition.

Donc si $n \leq m \in \omega$ alors $n \in \omega$.

Donc pour tout n et m dans ON , si $n \leq m \in \omega$ alors $n \in \omega$.

Donc ω est un segment initial de ON .

Donc $\boxed{\omega \text{ est un ordinal}}$ d'après la proposition 17 page 42.

- Montrons que ω est un ordinal limite.

Supposons par l'absurde que ω est successeur.

Il existe donc un ordinal α tel que $\omega = S(\alpha)$.

On a $\alpha < S(\alpha)$ d'après la proposition 13 page 33.

On a donc $\alpha < \omega$, c'est-à-dire $\alpha \in \omega$ par définition de $<$.

Donc α est un entier naturel par définition de ω .

Donc $S(\alpha)$ est un entier naturel d'après la proposition 15 page 38.

Donc $S(\alpha) \in \omega$ par définition de ω , c'est-à-dire $\omega \in \omega$.

On a donc $\omega < \omega$: c'est absurde par antiréflexivité de $<$.

Donc ω n'est pas un successeur et donc $\boxed{\omega \text{ est limite}}$.

2. Montrons que ω est plus petit que tout ordinal limite non nul.

Soit α un ordinal limite non nul.

Soit n un ordinal tel que $n < \omega$.

On a donc $n \in \omega$ par définition de $<$.

Donc n est un entier naturel par définition de ω .

Or on a $n \leq n$ par réflexivité de \leq .

Donc n est un ordinal plus petit qu'un entier naturel (lui-même).

Donc $n = 0$ ou n est un successeur par définition des entiers naturels.

Donc tout ordinal strictement plus petit que ω est ou bien nul ou bien successeur.

Or α est limite non nul donc n'est ni nul ni successeur.

Donc α n'est pas strictement plus petit que ω .

On a donc $\boxed{\omega \leq \alpha}$ puisque \leq est total chez les ordinaux.

CQFD.

Nous l'avons dit quand nous avons évoqué le paradoxe de Burali-Forti : il n'est pas possible d'encapsuler tous les ordinaux dans un seul ensemble. En fait le résultat est même plus fort : tout ensemble d'ordinaux est majoré par d'autres ordinaux qui ne sont pas dans l'ensemble. En

particulier parmi tous ces majorants stricts se cache un plus petit majorant strict.

Proposition 19 (Plus petit majorant strict d'ordinaux)

Soit X un ensemble d'ordinaux.

Alors il existe un unique ordinal α tel que

1. α est un majorant strict de X .

Autrement dit $\forall \xi \in X, \xi < \alpha$.

2. α est plus petit que tout majorant strict de X .

Autrement dit pour tout ordinal β , si $\forall \xi \in X, \xi < \beta$ alors $\alpha \leq \beta$.

Autrement dit α est le plus petit de tous les majorants stricts de X .

Démonstration

D'après la proposition 11 page 29, $\bigcup X$ est un ordinal.

Le plus petit des majorants stricts de X va dépendre de si $\bigcup X$ appartient à X ou non.

$$\text{Posons alors } \alpha := \begin{cases} \bigcup X & \text{si } \bigcup X \notin X \\ S(\bigcup X) & \text{si } \bigcup X \in X \end{cases}.$$

1.

On a vu lors de la proposition 11 page 29 que $\bigcup X = \sup(X)$.

En particulier $\bigcup X$ est un majorant de X .

- Plaçons-nous dans le cas où $\bigcup X \notin X$.

Alors $\bigcup X$ est un majorant strict de X puisqu'il n'en est pas un élément.

Or dans ce cas-là on a $\alpha = \bigcup X$ donc α est un majorant strict de X .

- Supposons à présent que $\bigcup X \in X$.

On a donc $\alpha = S(\bigcup X)$.

Or $\bigcup X$ est un majorant de X donc $\forall \xi \in X, \xi \leq \bigcup X$.

Donc $\forall \xi \in X, \xi < S(\bigcup X)$ d'après la proposition 13 page 33.

On a donc $\forall \xi \in X, \xi < \alpha$ et donc α est un majorant strict de X .

Dans les deux cas, α est un majorant strict de X .

2. Soit β un ordinal majorant strict de X .

Comme α et β sont deux ordinaux, on a $\beta < \alpha$ ou $\alpha \leq \beta$.

Supposons par l'absurde que $\beta < \alpha$.

- Plaçons-nous dans le cas où $\bigcup X \notin X$.

Par définition de α on a alors $\alpha = \bigcup X$ donc $\beta < \bigcup X$.

Autrement dit on a $\beta \in \bigcup X$ par définition de $<$.

Par définition de la réunion, il existe donc $\xi \in X$ tel que $\beta \in \xi$.

Autrement dit il existe $\xi \in X$ tel que $\beta < \xi$ par définition de $<$.

Donc β n'est pas un majorant de X , ce qui est absurde.

► Plaçons-nous dans le cas où $\bigcup X \in X$.

Par définition de α on a alors $\alpha = S(\bigcup X)$ donc $\beta \in S(\bigcup X)$.

On a donc $\beta < S(\bigcup X)$ par définition de $<$.

Donc $\beta \leq \bigcup X$ d'après la proposition 13 page 33.

Or $\bigcup X \in X$ par hypothèse donc β n'est pas un majorant strict de X .

C'est absurde.

Dans les deux cas on aboutit à une absurdité.

Donc par l'absurde on a montré que l'on n'a pas $\beta < \alpha$, et donc $\boxed{\alpha \leq \beta}$.

CQFD.

Dans la preuve qui précède, nous avons discuté du fait que pour un ensemble d'ordinaux X , sa borne supérieure $\sup(X) = \bigcup X$ est un élément de X ou non. Si ce n'est pas le cas, on est en fait assuré que $\sup(X)$ est un ordinal limite.

Proposition 20 (Borne supérieure qui n'est pas un maximum)

Soit X un ensemble d'ordinaux.

Si $\sup(X) \notin X$ alors $\sup(X)$ est un ordinal limite.



Démonstration

Montrons le résultat par contraposition.

Supposons que $\sup(X)$ n'est pas un ordinal limite.

Donc $\sup(X)$ est un successeur par définition.

Il existe donc un ordinal α tel que $\sup(X) = S(\alpha)$.

Par définition $\sup(X)$ est un majorant de X .

Donc $S(\alpha)$ est un majorant de X : on a $\forall \xi \in X, \xi \leq S(\alpha)$.

Supposons par l'absurde que $\sup(X) \notin X$.

On a donc $S(\alpha) \notin X$ donc $\forall \xi \in X, \xi \neq S(\alpha)$.

Donc $\forall \xi \in X, \xi < S(\alpha)$ d'après ce qui précède.

On a donc $\forall \xi \in X, \xi \leq \alpha$ d'après la proposition 10 page 33.

Donc α est un majorant de X .

Or on a $\alpha < S(\alpha)$ d'après la proposition 10 page 33.

Donc $S(\alpha)$ n'est pas le plus petit des majorants de X .

C'est absurde : cela veut dire que $S(\alpha)$ n'est pas la borne supérieure de X .

Par l'absurde, on vient de montrer que $\sup(X) \in X$.

Donc si $\sup(X)$ n'est pas limite alors $\sup(X) \in X$.

Par contraposition, on a $\boxed{\text{si } \sup(X) \notin X \text{ alors } \sup(X) \text{ est limite}}$.

CQFD.

Dans le cas où X est lui-même un ordinal, cette proposition se précise. Cela nous fournit même une autre caractérisation d'être un ordinal limite, en plus de celle donnée par la proposition 14 page 37.

Proposition 21 (Ordinal limite et borne supérieure)

Soit α un ordinal.

Les assertions suivantes sont équivalentes :

1. α est un ordinal limite.
2. $\sup(\alpha) = \alpha$

Démonstration

$1 \Rightarrow 2$ Supposons que α est un ordinal limite.

On a donc $\forall \beta \in \alpha, S(\beta) < \alpha$ d'après la proposition 14 page 37.

Notons (\star) cette affirmation.

Par définition de $<$ on a $\forall \beta \in \alpha, \beta < \alpha$.

En particulier $\forall \beta \in \alpha, \beta \leq \alpha$ donc α est un majorant de lui-même.

On a donc $\sup(\alpha) \leq \alpha$ par minimalité de la borne supérieure.

Supposons par l'absurde que $\sup(\alpha) \neq \alpha$.

On a donc $\sup(\alpha) < \alpha$ par ce qui précède.

On a donc $S(\sup(\alpha)) < \alpha$ d'après (\star) , donc $S(\sup(\alpha)) \in \alpha$ par définition de $<$.

Or on a $\sup(\alpha) < S(\sup(\alpha))$ d'après la proposition 13 page 33.

Donc $\sup(\alpha)$ n'est pas un majorant de α .

C'est absurde par définition de la borne supérieure.

Par l'absurde, on vient de montrer que $\boxed{\sup(\alpha) = \alpha}$.

$2 \Rightarrow 1$ Supposons que $\sup(\alpha) = \alpha$.

On n'a pas $\alpha < \alpha$ par antiréflexivité de $<$.

On n'a donc pas $\sup(\alpha) < \alpha$ par hypothèse, et donc $\sup(\alpha) \notin \alpha$ par définition de $<$.

Donc $\sup(\alpha)$ est un ordinal limite d'après la proposition 20 page 46.

Donc comme $\sup(\alpha) = \alpha$, on en déduit que $\boxed{\alpha \text{ est un ordinal limite}}$.

CQFD.

5 Isomorphisme avec les ordinaux

Jusqu'à présent, nous avons définis, construits et étudiés les ordinaux pour eux-mêmes. Or à la base nous les avons introduits dans l'optique d'en faire des représentants de classes d'isomorphie. Il est donc temps d'étudier d'un peu plus près les isomorphismes. Commençons par constater quelques propriétés de base que conservent les isomorphismes : c'est en cela que l'on peut dire que si deux ensembles ordonnés sont isomorphes, alors ils représentent en fait la même structure d'ensemble ordonné.

Proposition 22 (Isomorphismes et propriétés conservées)

Soient E et F deux ensembles ordonnés.

Supposons que E et F sont **isomorphes**.

1. E est totalement ordonné si et seulement si F est totalement ordonné.
2. L'ordre sur E est bien fondé si et seulement si l'ordre sur F est bien fondé.
3. E est bien ordonné si et seulement si F est bien ordonné.



Démonstration

Notons \preccurlyeq l'ordre sur E et \sqsubseteq l'ordre sur F .

Soit $f : E \longrightarrow F$ un isomorphisme d'ordres.

1. \Rightarrow

Supposons que E est totalement ordonné.

Soient y_1 et y_2 dans F .

Par définition f est un isomorphisme d'ordres.

En particulier f est surjective dans F .

Il existe donc x_1 et x_2 dans E tels que $y_1 = f(x_1)$ et $y_2 = f(x_2)$.

Or E est totalement ordonné donc $x_1 \preccurlyeq x_2$ ou $x_2 \preccurlyeq x_1$.

On a donc $f(x_1) \sqsubseteq f(x_2)$ ou $f(x_2) \sqsubseteq f(x_1)$ car f est un isomorphisme d'ordres. On a donc $y_1 \sqsubseteq y_2$ ou $y_2 \sqsubseteq y_1$.

Donc F est totalement ordonné.

Donc si E est totalement ordonné alors F est totalement ordonné.

\Leftarrow

Supposons que F est totalement ordonné.

Par définition $f : E \longrightarrow F$ est un isomorphisme d'ordres.

Donc f est inversible et $f^{-1} : F \longrightarrow E$ est un isomorphisme d'ordres.

Donc E est totalement ordonné d'après le sens \Rightarrow .

Donc si F est totalement ordonné alors E est totalement ordonné.

Finalement, E est totalement ordonné si et seulement si F est totalement ordonné.

2. \Rightarrow

Supposons que l'ordre sur E est bien fondé.

Soit B une partie non vide de F .

Considérons $A := f^\leftarrow(B)$.

Comme f est un isomorphisme d'ordres, A est une partie non vide de E .

Or l'ordre sur E est bien fondé donc A admet un élément minimal a_0 .

Donc $f(a_0)$ est un élément minimal de $f^\rightarrow(A)$.

Or f est un isomorphisme d'ordre, donc en particulier

$$f^\rightarrow(A) = f^\rightarrow(f^\leftarrow(B)) = B$$

Donc B admet un élément minimal.

Donc toute partie non vide de F admet un élément minimal.

Donc l'ordre sur F est bien fondé.

Donc si l'ordre sur E est bien fondé alors l'ordre sur F est bien fondé.

 \Leftarrow

Supposons que l'ordre sur F est bien fondé.

Par définition $f : E \longrightarrow F$ est un isomorphisme d'ordres.

Donc f est inversible et $f^{-1} : F \longrightarrow E$ est un isomorphisme d'ordres.

Donc l'ordre sur E est bien fondé d'après le sens \Rightarrow .

Donc si l'ordre sur F est bien fondé alors l'ordre sur E est bien fondé.

Finalement, L'ordre sur E est bien fondé si et seulement si l'ordre sur F est bien fondé.

11. On a les équivalences suivantes :

$$\begin{aligned} E \text{ est bien ordonné} &\iff \text{L'ordre sur } E \text{ est total et bien fondé d'après la prop. 2 p. 10} \\ &\iff \text{L'ordre sur } F \text{ est total et bien fondé d'après 1. et 2.} \\ &\iff F \text{ est bien ordonné d'après la prop. 2 p. 10} \end{aligned}$$

D'où l'équivalence recherchée.

CQFD.

Rentrons dans le cœur de ce qui nous intéresse : constatons la proposition suivante, qui n'est au fond pas étonnante dans la mesure où l'on a dit lors de la proposition 5 page 13 qu'étant donné un ensemble ordonné (E, \preccurlyeq) , tout segment initial propre est de la forme E_{\prec_x} .

Proposition 23 (Ensemble des segments initiaux)

Soient (E, \preccurlyeq) un ensemble **bien ordonné** et \prec l'ordre strict associé.

Soit X l'ensemble des segments initiaux propres de E .

On munit X de la relation d'ordre \subseteq .

$$\text{Soit } f := \begin{pmatrix} E & \longrightarrow & X \\ x & \longmapsto & E_{\prec x} \end{pmatrix}.$$

On a alors :

1. f est un isomorphisme d'ordres de E vers X .
2. Si de plus E est un ordinal alors $f = \text{id}_E$ et en particulier $E = X$.

Démonstration

1.

- Montrons que f est strictement croissante.

Rappelons que la relation d'ordre strict sur X est l'inclusion stricte \subsetneq .

Soient x et y dans E .

Supposons que $x \prec y$.

On a alors $x \in E_{\prec y}$ par définition.

Or on n'a pas $x \prec x$ par antiréflexivité de \prec donc $x \notin E_{\prec x}$.

Comme $x \in E_{\prec y}$ et $x \notin E_{\prec x}$ on a $E_{\prec x} \neq E_{\prec y}$.

Soit $z \in E_{\prec x}$.

On a alors $z \prec x$ par définition.

Donc $z \prec y$ par transitivité de \prec .

Donc $z \in E_{\prec y}$ par définition.

Donc $E_{\prec x} \subseteq E_{\prec y}$ par définition de l'inclusion.

Comme $E_{\prec x} \neq E_{\prec y}$, on a donc $E_{\prec x} \subsetneq E_{\prec y}$

Ainsi on a $f(x) \subsetneq f(y)$ par définition de f .

Donc si $x \prec y$ alors $f(x) \subsetneq f(y)$.

Donc f est strictement croissante.

Or E est bien ordonné donc en particulier totalement ordonné d'après la prop. 2 p. 10.

Donc le domaine de f est totalement ordonné et donc f est croissante et injective.

- Montrons que f est surjective dans X .

Par définition de f on sait déjà que $\text{im}(f) \subseteq X$.

Soit $A \in X$.

Alors A est un segment initial propre de E par définition de X .

Or E est bien ordonné donc il existe $x \in E$ tel que $A = E_{\prec x}$ d'après la prop. 5 p. 13.

On a donc $A = f(x)$ et donc $A \in \text{im}(f)$.

Donc $\text{im}(f) \supseteq X$ et donc $\text{im}(f) = X$.

Ainsi f est surjective dans X .

- Ainsi f est croissante, injective et surjective dans X .

Or on a dit que E le domaine de f est totalement ordonné.

Donc f est un isomorphisme de E vers X .

2. Supposons que E est un ordinal.

Dans ce cas particulier, l'ordre strict \prec est l'appartenance \in .

Ainsi pour tout $\alpha \in E$ on a $E_{\prec\alpha} = \{\beta \in E \mid \beta \prec \alpha\} = \{\beta \in E \mid \beta \in \alpha\} = E \cap \alpha$.

Remarquons pour commencer que f et id_E ont le même domaine E .

Soit $\alpha \in E$.

Comme E est ordinal, E est transitif donc $\alpha \subseteq E$ et donc $E \cap \alpha = \alpha$.

Or on a vu que $E_{\prec\alpha} = E \cap \alpha$ donc $E_{\prec\alpha} = \alpha$.

En particulier $f(\alpha) = E_{\prec\alpha} = \alpha = \text{id}_E(\alpha)$.

Donc $\forall \alpha \in E, f(\alpha) = \text{id}_E(\alpha)$.

Donc $f = \text{id}_E$.

En particulier $E = \text{im}(\text{id}_E) = \text{im}(f) = X$.

CQFD.

Remarque :

1. On peut remarquer que $g := \begin{pmatrix} X & \longrightarrow & E \\ A & \longmapsto & \min(E \setminus A) \end{pmatrix}$ est la réciproque de f .

2. Le cas où E est un ordinal n'est pas non plus étonnant : on a déjà vu lors de la proposition 17 page 42 que les ordinaux sont eux-mêmes les segments initiaux de ON .

Proposition 24 (Isomorphismes entre bons ordres)

Soient E et F deux ensembles **bien ordonnés**.

Il y a au plus un isomorphisme d'ordres de E vers F .

Démonstration

Notons \preccurlyeq l'ordre sur E et \prec la relation d'ordre strict associée.

Notons \trianglelefteq l'ordre sur F et \triangleleft la relation d'ordre strict associée.

Supposons qu'il existe un isomorphisme d'ordres $f : E \longrightarrow F$.

Soit $g : E \longrightarrow F$ un autre isomorphisme d'ordres.

Montrons que $f = g$.

Supposons par l'absurde que $f \neq g$.

Considérons alors l'ensemble $A := \{x \in E \mid f(x) \neq g(x)\}$.

Par hypothèse A est donc une partie non vide de E .

Or E est bien ordonné donc A admet un minimum a .

Comme $a = \min(A)$, pour tout $b \in E$ tel que $b \prec a$ on a $b \notin A$ donc $f(b) = g(b)$.

De plus $a \in A$ donc $f(a) \neq g(a)$. Or F est bien ordonné donc totalement ordonné d'après la proposition 2 page 10, donc $f(a) \triangleleft g(a)$ ou $g(a) \triangleleft f(a)$.

► Plaçons-nous dans le cas où $f(a) \triangleleft g(a)$.

Soit $b \in E$.

Comme E est bien ordonné, il est en particulier totalement ordonné d'après la proposition 2 page 10.

On a donc $b \prec a$ ou $a \preccurlyeq b$.

Si $b \prec a$ alors $g(b) = f(b) \triangleleft f(a)$ par stricte croissance de f .

Si $a \preccurlyeq b$ alors $f(a) \triangleleft g(a) \preccurlyeq g(b)$ par croissance de g .

Dans les deux cas on a $g(b) \neq f(a)$.

Ainsi pour tout $b \in E$ on a $g(b) \neq f(a)$.

Ainsi $f(a)$ est un élément de F que g n'atteint pas.

C'est absurde puisque g est surjective dans F .

► Plaçons-nous dans le cas où $g(a) \triangleleft f(a)$.

On montre de la même manière que dans ce cas-là $g(a)$ est un élément de F que f n'atteint pas. C'est absurde puisque f est surjective dans F .

Dans les deux cas on aboutit une absurdité concernant la surjectivité d'une des deux applications.

Par l'absurde on vient de montrer que $f = g$, d'où l'unicité.

CQFD.

Remarque :

En particulier pour E un ensemble bien ordonné quelconque, il n'a que l'identité comme isomorphisme d'ordres de E vers E .

Quand nous avons dit que les ordinaux fournissaient un représentant de chaque classe d'isomorphie pour les bons ordres, nous avons aussi affirmé qu'il n'y en avait qu'un seul par classe. Autrement dit, si deux ordinaux sont isomorphes, alors nécessairement il s'agit d'un même ordinal.

Proposition 25 (Isomorphisme entre ordinaux)

Soient α et β deux ordinaux et $f : \alpha \longrightarrow \beta$.

Si f est un isomorphisme de α vers β alors $f = \text{id}_\alpha$ et donc $\alpha = \beta$.



Démonstration

Supposons que f est un isomorphisme de α vers β .

En particulier f est injective, surjective sur β et croissante.

Étant injective et croissante, f est strictement croissante.

Remarquons que comme α est un ordinal, α est transitif.

Donc pour tout $\xi \in \alpha$, on aussi $\xi \subseteq \alpha$.

Autrement dit tout élément de α est aussi une partie de α .

Donc pour tout $\xi \in \alpha$, on peut s'intéresser à la fois à $f(\xi)$ et $f^\rightarrow(\xi)$.

- Montrons que pour tout $\xi \in \alpha$, on a $f(\xi) = f^\rightarrow(\xi)$.

Soit $\xi \in \alpha$.

Montrons que $f(\xi) = f^\rightarrow(\xi)$.



Soit $\gamma \in f(\xi)$.

Comme f est surjective sur β on a $\text{im}(f) = \beta$ donc $f(\xi) \in \beta$.

Comme β est un ordinal, β est transitif donc $f(\xi) \subseteq \beta$.

On a donc $\gamma \in \beta$ par définition de l'inclusion.

Comme $\text{im}(f) = \beta$ on a $\gamma \in \text{im}(f)$ donc il existe $\mu \in \alpha$ tel que $\gamma = f(\mu)$.

Comme α est un ordinal, μ et ξ sont des ordinaux d'après la prop. 6 p. 17.

On a donc $\mu < \xi$ ou $\mu = \xi$ ou $\xi < \mu$ d'après le théorème 1 page 21.

- Si $\mu = \xi$ alors $\gamma = f(\mu) = f(\xi)$.

Or par définition on a $\gamma \in f(\xi)$, donc $\gamma \in \gamma$ et donc $\gamma < \gamma$.

- Si $\xi < \mu$ alors $f(\xi) < f(\mu)$ par stricte croissance de f .

Comme $f(\mu) = \gamma$ par définition de μ , on a donc $f(\xi) < \gamma$.

Or par définition on a $\gamma \in f(\xi)$ donc $\gamma < f(\xi)$.

On a donc $\gamma < \gamma$ par transitivité de $<$.

Dans ces deux cas-là on a donc nécessairement $\gamma < \gamma$.

C'est absurde par antiréflexivité de $<$.

On a donc nécessairement le troisième cas $\mu < \xi$.

On a donc $\mu \in \xi$ par définition de $<$.

Donc $f(\mu) \in f^\rightarrow(\xi)$ par définition de l'image directe.

Comme $\gamma = f(\mu)$, on a donc $\gamma \in f^\rightarrow(\xi)$.

Donc $f(\xi) \subseteq f^\rightarrow(\xi)$.

□

Soit $\gamma \in f^\rightarrow(\xi)$.

Il existe donc $\mu \in \xi$ tel que $\gamma = f(\mu)$.

Comme $\mu \in \xi$, on a $\mu < \xi$ par définition de $<$.

On a donc $f(\mu) < f(\xi)$ par stricte croissance de f .

Or $\gamma = f(\mu)$ donc $\gamma < f(\xi)$.

On a donc $\gamma \in f(\xi)$ par définition de $<$.

Donc $f(\xi) \supseteq f^\rightarrow(\xi)$.

Finalement on a bien $f(\xi) = f^\rightarrow(\xi)$.

Donc pour tout $\xi \in \alpha$, on a $f(\xi) = f^\rightarrow(\xi)$ $(*)$.

- On veut montrer que $f = \text{id}_\alpha$.

Comme elles ont le même domaine α , cela revient à montrer que $\forall \xi \in \alpha, f(\xi) = \xi$.

Pour cela, considérons $X := \{\xi \in \alpha \mid f(\xi) \neq \xi\}$ et montrons que $X = \emptyset$.

Supposons par l'absurde que X est non vide.

Par définition X est donc une partie non vide de l'ordinal α .

Or tous les éléments de α sont des ordinaux d'après la proposition 6 page 17.

Donc X est un ensemble non vide dont les éléments sont tous des ordinaux.

Il possède donc un ordinal minimum ξ d'après le théorème 1 page 21.

Soit $\mu \in \xi$.

On a donc $\mu < \xi$ par définition de $<$.

Comme ξ est minimum de X , on a $\mu \notin X$.

Or $\xi \in X$ et $X \subseteq \alpha$ par définitions, donc $\xi \in \alpha$ par définition de l'inclusion.

On a donc $\xi < \alpha$ par définition de $<$.

Ainsi on a $\mu < \xi < \alpha$ donc $\mu < \alpha$ par transitivité de $<$.

On a donc $\mu \in \alpha$ par définition de $<$, et on a vu que $\mu \notin X$.

On a donc $f(\mu) = \mu$ par définition de X .

Donc $\forall \mu \in \xi, f(\mu) = \mu$.

Donc $f(\xi) = f^\rightarrow(\xi) = \{f(\mu) \mid \mu \in \xi\} = \{\mu \mid \mu \in \xi\} = \xi$.

Donc $f(\xi) = \xi$ donc $\xi \notin X$ par définition de X : c'est absurde.

Par l'absurde, on vient de montrer que X est vide.

Donc $\forall \xi \in \alpha, f(\xi) = \xi$ par définition de X .

Finalement, on a donc $f = \text{id}_\alpha$.

On a en particulier $\alpha = \text{im}(\text{id}_\alpha) = \text{im}(f) = \beta$ par surjectivité de f sur β .

CQFD.

Ainsi, on vient de montrer qu'au sein d'une classe d'isomorphie, il ne peut y avoir au maximum qu'un seul ordinal. Précisions ce que l'on entend par là.

Proposition 26 (Au plus un ordinal associé à un bon ordre)

Soit (E, \preccurlyeq) un ensemble ordonné.

Supposons qu'il existe au moins un ordinal α tel que (E, \preccurlyeq) et (α, \leq) sont isomorphes.

Alors un tel α est unique, et l'isomorphisme de (E, \preccurlyeq) vers (α, \leq) est unique.

Démonstration

- **Unicité de l'ordinal**

Soit $f : E \longrightarrow \alpha$ un isomorphisme.

Soit β un ordinal tel que (E, \preccurlyeq) et (β, \leq) sont isomorphes.

Il existe donc un isomorphisme $g : E \longrightarrow \beta$.

Comme $f : E \longrightarrow \alpha$ un isomorphisme, $f^{-1} : \alpha \longrightarrow E$ est un isomorphisme.

Donc $g \circ f^{-1} : \alpha \longrightarrow \beta$ est un isomorphisme.

Donc $\alpha = \beta$ d'après la proposition 25 page 53.

On a donc $\boxed{\text{unicité de l'ordinal } \alpha \text{ isomorphe à } E}$.

- **Unicité de l'isomorphisme.**

Soit $g : E \longrightarrow \alpha$ un isomorphisme.

Alors $g^{-1} : \alpha \longrightarrow E$ est un isomorphisme.

Donc $f \circ g^{-1} : \alpha \longrightarrow \alpha$ est un isomorphisme.

Donc $f \circ g^{-1} = \text{id}_\alpha$ d'après la proposition 25 page 53.

Donc $f = f \circ \text{id}_E = f \circ (g^{-1} \circ g) = (f \circ g^{-1}) \circ g = \text{id}_\alpha \circ g = g$.

On a donc $\boxed{\text{unicité de l'isomorphisme de } E \text{ vers } \alpha}$.

CQFD.

Venons-en finalement à ce qui nous intéressait depuis le début : utiliser les ordinaux pour représenter n'importe quel bon ordre, à isomorphisme près. On vient déjà de voir l'unicité, mais formulons quand-même complètement un théorème digne de ce nom !

Pour le démontrer, nous allons utiliser l'idée proposée par la proposition 23 page 50, qui affirme qu'à isomorphisme près, un ensemble bien ordonné se comporte comme l'ensemble de ses segments initiaux propres. Autrement dit, on peut tout à fait raisonner sur les segments initiaux propres plutôt que sur l'ensemble directement.

Théorème 4 (Unique ordinal associé à un bon ordre)

Soit (E, \preccurlyeq) un ensemble **bien ordonné**.

Alors il existe un unique ordinal α tel que (E, \preccurlyeq) et (α, \leq) sont isomorphes.

On dit alors que α est le **type** de (E, \preccurlyeq) et on note $\text{type}(E, \preccurlyeq) := \alpha$.



Démonstration

- Soit \prec l'ordre strict sur E associé à \preccurlyeq .

Soit $x \in E$.

Rappelons-nous que $E_{\prec x} = \{y \in E \mid y \prec x\}$.

Ainsi $E_{\prec x}$ est une partie de E et (E, \preccurlyeq) est bien ordonné.

Donc $(E_{\prec x}, \preccurlyeq)$ est bien ordonné d'après la prop. 3 p. 11.

Pour la suite de cette démonstration, on dira que x est **bon** si et seulement s'il existe au moins un ordinal ξ tel que $(E_{\prec x}, \preccurlyeq)$ est isomorphe à (ξ, \leq) .

Dans ce cas-là, un tel ordinal ξ est unique d'après la proposition 26 page 55.

Nous noterons par la suite $f(x) := \xi$ cet unique ordinal.

Posons $G := \{x \in E \mid x \text{ est bon}\}$.

On a donc l'application $f : G \longrightarrow ?$ qui à $x \in G$ associe l'unique ordinal $f(x)$ tel que $(E_{\prec x}, \preccurlyeq)$ est isomorphe à $(f(x), \leq)$.

Pour tout $x \in G$, l'isomorphisme de $(E_{\prec x}, \preccurlyeq)$ vers $(f(x), \leq)$ est unique d'après la proposition 26 page 55 : on le notera h_x . Ainsi on a $h_x : E_{\prec x} \longrightarrow f(x)$.

Avant d'avancer, remarquons que par définition pour tout $x \in G$, son image $f(x)$ est un ordinal. En particulier $\text{im}(f)$ est un ensemble d'ordinaux et est donc naturellement muni de l'ordre induit \leq .

Voici à présent les différentes étapes de la preuve :

1. On prouve que G est un segment initial de E .
2. On montre que f est un isomorphisme de (G, \preccurlyeq) dans $(\text{im}(f), \leq)$.
3. On montre que $\text{im}(f)$ est un ordinal : G est donc lui-même isomorphe à un ordinal.
4. On montre qu'en fait $G = E$, ce qui permet de conclure.

1. Montrons que G est un segment initial de E .

Autrement dit, montrons que pour tout x et y dans E , si $y \prec x \in G$ alors $y \in G$.

Montrons aussi au passage que dans ce cas-là on a $f(y) = h_x(y)$.

Soient x et y dans E tels que $y \prec x \in G$.

Alors pour tout $z \in E_{\prec y}$, on a $z \prec y$ donc $z \prec x$ par transitivité de \prec et donc $z \in E_{\prec x}$. Ainsi $E_{\prec y} \subseteq E_{\prec x}$ donc on peut considérer la restriction de h_x à $E_{\prec y}$.

Nous allons montrer que $(h_x)_{|E_{\prec y}}$ est un isomorphisme de $E_{\prec y}$ vers $h_x(y)$.

Par définition on sait déjà que $(h_x)_{|E_{\prec y}}$ a pour domaine $E_{\prec y}$.

Par définition h_x est un isomorphisme d'ordre, donc est injectif et croissant.

On en déduit déjà que $(h_x)_{|E_{\prec y}}$ est lui aussi injectif et croissant.

De plus, on en déduit aussi que h_x est strictement croissant.

Montrons que $(h_x)_{|E_{\prec y}}$ est surjectif sur $h_x(y)$, c'est-à-dire $\text{im}((h_x)_{|E_{\prec y}}) = h_x(y)$.



Soit $u \in \text{im}((h_x)_{|E_{\prec y}})$.

Il existe donc $z \in E_{\prec y}$ tel que $u = (h_x)_{|E_{\prec y}}(z) = h_x(z)$.

Comme $z \in E_{\prec y}$, on a $z \prec y$ donc $h_x(z) < h_x(y)$ par stricte croissance de h_x .

Ainsi on a $h_x(z) \in h_x(y)$ par définition de $<$, et donc $u \in h_x(y)$.

Donc $\text{im}((h_x)_{|E_{\prec y}}) \subseteq h_x(y)$.



Soit $\beta \in h_x(y)$.

Comme $h_x : E_{\prec x} \longrightarrow f(x)$ on a $h_x(y) \in f(x)$.

Ainsi on a $\beta \in h_x(y) \in f(x)$.

Or $f(x)$ est un ordinal par définition de f donc $f(x)$ est transitif.

On en déduit donc que $\beta \in f(x)$.

Or par définition h_x est surjectif dans $f(x)$.

Il existe donc $b \in E_{\prec x}$ tel que $h_x(b) = \beta$.

Or (E, \preccurlyeq) est bien ordonné par définition.

Donc (E, \preccurlyeq) est totalement ordonné d'après la proposition 2 page 10.

On a donc $b \prec y$ ou $y \preccurlyeq b$.

Supposons par l'absurde que $y \preccurlyeq b$.

On a alors $h_x(y) \leq h_x(b)$ par croissance de h_x .

Comme $h_x(b) = \beta$ par définition de b , on a $h_x(y) \leq \beta$.

Or on a $\beta \in h_x(y)$ par définition de β , c'est-à-dire $\beta < h_x(y)$.

On vient donc de montrer $\beta < h_x(y) \leq \beta$, ce qui est absurde.

On a donc nécessairement l'autre option $b \prec y$ c'est-à-dire $b \in E_{\prec y}$.

Comme $\beta = h_x(b)$, on a donc $\beta \in h_x^-(E_{\prec y})$.

Autrement dit, on a $\beta \in \text{im}((h_x)_{|E_{\prec y}})$.

On a donc $\text{im}((h_x)_{|E_{\prec y}}) \supseteq h_x(y)$ et donc $\text{im}((h_x)_{|E_{\prec y}}) = h_x(y)$.

Ainsi $(h_x)_{|E_{\prec y}}$ est surjective sur $h_x(y)$.

On a donc $(h_x)_{|E_{\prec y}}$ est croissante, injective et surjective sur $h_x(y)$.

Or on a dit que \preccurlyeq est total sur E , donc sur $E_{\prec y}$ le domaine de $(h_x)_{|E_{\prec y}}$.

Donc $(h_x)_{|E_{\prec y}}$ est un isomorphisme de $E_{\prec y}$ vers $h_x(y)$.

Ainsi $E_{\prec y}$ et $h_x(y)$ sont isomorphes.

Or on a dit que $h_x(y)$ est un ordinal puisqu'élément de $f(x)$.

Donc $E_{\prec y}$ est isomorphe à un ordinal, et donc $y \in G$.

On note au passage que par unicité de l'ordinal on a $f(y) = h_x(y)$.

Donc pour tout x et y dans E , si $y \prec x \in G$ alors $y \in G$ avec $f(y) = h_x(y)$ (\star) .

2. Montrons que f est un isomorphisme de (G, \preccurlyeq) dans $(\text{im}(f), \leq)$.

Pour cela, montrons que f est strictement croissante.

Soient x et y dans G .

Supposons que $y \prec x$.

D'après (\star) on a alors $f(y) = h_x(y)$.

Or par définition h_x est à valeurs dans $f(x)$.

On a donc $h_x(y) \in f(x)$ et donc $f(y) \in f(x)$.

On a donc $f(y) < f(x)$ par définition de $<$.

Donc si $y \prec x$ alors $f(y) < f(x)$.

Donc f est strictement croissante.

En particulier f est croissante et injective, donc f est croissante et bijective sur $\text{im}(f)$.

Or on a dit que \preccurlyeq est total sur E donc \preccurlyeq est total sur G puisque $G \subseteq E$.

Donc f est un isomorphisme de (G, \preccurlyeq) dans $(\text{im}(f), \leq)$.

3. Montrons que $\text{im}(f)$ est un ordinal.

► Montrons que $\text{im}(f)$ est transitif.

Soit $a \in \text{im}(f)$.

Il existe donc $x \in G$ tel que $a = f(x)$.

Soit $b \in a$.

Par définition h_x est un isomorphisme d'ordres de $E_{\prec x}$ dans $f(x)$.

En particulier h_x est surjectif sur $f(x)$.

Comme $a = f(x)$, on en déduit que h_x est surjectif sur a .

On a donc $\text{im}(h_x) = a$ et donc $b \in \text{im}(h_x)$.

Il existe donc $y \in E_{\prec x}$ tel que $b = h_x(y)$.

Comme $y \in E_{\prec x}$ on a $y \prec x$ et donc $h_x(y) = f(y)$ d'après (\star) .

Donc $b = f(y)$ et donc $b \in \text{im}(f)$.

Donc $a \subseteq \text{im}(f)$ par définition de l'inclusion.

Donc $\forall a \in \text{im}(f), a \subseteq \text{im}(f)$.

Donc $\text{im}(f)$ est transitif.

► Par définition de f , tous les éléments de $\text{im}(f)$ sont des ordinaux.

Donc \in est un bon ordre strict sur $\text{im}(f)$ d'après le théorème 1 page 21.

Ainsi, $\text{im}(f)$ est transitif et \in est un bon ordre strict sur $\text{im}(f)$.

Donc $\boxed{\text{im}(f) \text{ est un ordinal}}$.

4. Ainsi f est un isomorphisme de G vers l'ordinal $\text{im}(f)$.

Il ne reste plus qu'à prouver que $G = E$ pour conclure.

Supposons par l'absurde que $G \subsetneq E$.

Alors $E \setminus G$ est une partie non vide de l'ensemble bien ordonné E .

Elle admet donc un minimum e .

Montrons que $E_{\prec e} = G$.

\subseteq

Soit $x \in E_{\prec e}$.

Par définition on a $x \prec e$ donc $x \notin E \setminus G$ car e en est minimum.

On a donc $x \in G$.

Donc $E_{\prec e} \subseteq G$ par définition de l'inclusion.

\supseteq

Soit $x \in G$.

On a dit que \preccurlyeq est total sur E donc $x \prec e$ ou $x = e$ ou $e \prec x$.

► Si $x = e$ alors $e \in G$ puisque $x \in G$.

► Si $e \prec x$ alors $e \in G$ d'après $(*)$.

Dans ces deux cas-là on a donc $e \in G$ ce qui est absurde puisque $e \in E \setminus G$.

On a donc nécessairement $x < e$ et donc $x \in E_{\prec e}$.

Donc $E_{\prec e} \supseteq G$ et donc $E_{\prec e} = G$.

Or G est isomorphe à l'ordinal $\text{im}(f)$ d'après ce qui précède.

Donc $E_{\prec e}$ est isomorphe à l'ordinal $\text{im}(f)$.

Donc $e \in G$ par définition de G , ce qui est absurde puisque $e \in E \setminus G$.

On a donc $E = G$.

Or G est isomorphe à l'ordinal $\text{im}(f)$ d'après ce qui précède.

Donc $\boxed{E \text{ est isomorphe à l'ordinal } \text{im}(f)}$.

L'unicité est garantie par la proposition 26 page 55.

CQFD.

Remarque :

1. Ainsi on note $\text{type}(E, \preccurlyeq)$ l'unique ordinal isomorphe à l'ensemble bien ordonné (E, \preccurlyeq) . Comme nous avons déjà eu l'occasion de le faire, on omet parfois d'écrire \preccurlyeq car l'ordre est sous-entendu, afin de simplifier et fluidifier le discours. On notera très donc très souvent $\text{type}(E)$.
2. Soit α un ordinal : comme α est nécessairement isomorphe à lui-même, on a donc $\text{type}(\alpha) = \alpha$.

Pour la proposition qui suit, rappelons qu'une partie d'un ensemble bien ordonné est elle aussi bien ordonnée en vertu de la proposition 3 page 11.

Proposition 27 (Ordinal associé et inclusion)

Soient (A, \preccurlyeq) un ensemble **bien ordonné** et X une partie de A .
On a $\text{type}(X, \preccurlyeq) \leq \text{type}(A, \preccurlyeq)$.



Démonstration

- Commençons par supposer que A est un ordinal : la relation \preccurlyeq est donc \leq .
Ainsi tous les éléments de A sont des ordinaux d'après la proposition 6 page 17.
Or X est une partie de A donc X est un ensemble d'ordinaux.
Posons alors $\delta := \text{type}(X, \leq)$ et $f : X \longrightarrow \delta$ l'isomorphisme associé.
Rappelons que δ étant un ordinal, X et δ sont munis tous deux de \leq .
En particulier les éléments de X et les éléments de δ sont comparables pour \leq .
En particulier pour tout $\xi \in X$, $f(\xi)$ et ξ sont comparables pour \leq .

Montrons que $\forall \xi \in X, f(\xi) \leq \xi$.

Pour cela posons $E := \{\xi \in X \mid f(\xi) \leq \xi\}$ et montrons que $E = X$.

Supposons par l'absurde que $E \subsetneq X$.

Alors $X \setminus E$ est un ensemble non vide d'ordinaux.

Il admet donc un minimum ξ d'après le théorème 1 page 21.

Comme $\xi \in X \setminus E$, on a $\xi \notin E$ et donc $f(\xi) \not\leq \xi$ par définition de E .

Or \leq est total chez les ordinaux donc on a $\xi < f(\xi)$.

Or $\text{im}(f) = \delta$ par définition de f donc $f(\xi) \in \delta$ et donc $f(\xi) < \delta$.

Ainsi $\xi < f(\xi) < \delta$ donc $\xi < \delta$ par transitivité de $<$.

Comme $\text{im}(f) = \delta$ on a donc $\xi < \text{im}(f)$ et donc $\xi \in \text{im}(f)$.

Il existe donc $\gamma \in E$ tel que $\xi = f(\gamma)$.

Comme $\xi < f(\xi)$ on a donc $f(\gamma) < f(\xi)$.

Comme f est un isomorphisme d'ordres, on a donc $\gamma < \xi$.

Comme ξ est le minimum de $X \setminus E$, on a donc $\gamma \notin X \setminus E$ donc $\gamma \in E$.

On a donc $f(\gamma) \leq \gamma$ par définition de E , c'est-à-dire $\xi \leq \gamma$ par définition de γ .

C'est absurde puisque l'on a dit que $\gamma < \xi$.

Par l'absurde, on a donc montré que $E = X$.

Ainsi, $\boxed{\forall \xi \in X, f(\xi) \leq \xi}$ (\star_1) .

Montrons que $\delta \leq A$.

Soit $\varepsilon \in \delta$.

Par définition de f on a $\text{im}(f) = \delta$ donc $\varepsilon \in \text{im}(f)$.

Il existe donc $\xi \in X$ tel que $\varepsilon = f(\xi)$.

D'après (\star_1) on a $f(\xi) \leq \xi$ donc $\varepsilon \leq \xi$.

Or on a $\xi \in X \subseteq A$ donc $\xi \in A$ par définition de l'inclusion.

On a donc $\xi < A$ par définition de $<$, donc $\varepsilon \leq \xi < A$ et donc $\varepsilon < A$ par transitivité.

Ainsi on a $\varepsilon \in A$.

Donc $\delta \subseteq A$ par définition de l'inclusion, et donc $\delta \leq A$.

Or par définition $\delta = \text{type}(X, \leq)$ donc $\text{type}(X, \leq) \leq A$.

Or A est un ordinal donc en particulier est l'unique ordinal isomorphe à lui-même.

Autrement dit on a $A = \text{type}(A, \leq)$.

On a donc bien $\boxed{\text{type}(X, \leq) \leq \text{type}(A, \leq)}$ (\star_2) .

- Plus généralement on ne suppose plus spécialement que A est un ordinal.

Soient alors $\alpha := \text{type}(A, \preccurlyeq)$ et $g : A \longrightarrow \alpha$ l'isomorphisme associé.

Considérons $Y := g^\rightarrow(X)$, de telle sorte que Y est une partie de α .

On se retrouve dans la situation précédente : d'après (\star_2) on a $\text{type}(Y, \leq) \leq \text{type}(\alpha, \leq)$.

Or g est un isomorphisme d'ordres.

Donc en particulier g est croissant et injectif.

Donc $g|_X$ est croissant et injectif.

Donc $g|_X$ est croissant et une bijection de X dans $g^\rightarrow(X) = Y$.

Or X est bien ordonné donc en particulier est totalement ordonné d'après la proposition 2 page 10. Donc $g|_X$ est un isomorphisme d'ordres de X dans Y .

En particulier X et Y sont isomorphes.

Donc X et $\text{type}(Y, \leq)$ sont isomorphes par transitivité de l'isomorphie.

Donc $\text{type}(X, \preccurlyeq) = \text{type}(Y, \leq)$ par unicité de l'ordinal associé.

On a donc $\boxed{\text{type}(X, \preccurlyeq) \leq \text{type}(A, \preccurlyeq)}$.

CQFD.

Proposition 28 (Segments initiaux et isomorphisme)

Soient E et F deux ensembles **totalement ordonnés**.

Notons \prec l'ordre strict sur E et \triangleleft l'ordre strict sur F .

- Supposons qu'il existe $f : E \rightarrow F$ un isomorphisme d'ordres.

Alors pour tout $x \in E$, $f|_{E_{\prec x}} : E_{\prec x} \rightarrow F_{\triangleleft f(x)}$ est un isomorphisme d'ordres.

- En particulier supposons que E est un ensemble **bien ordonné**.

Notons $f : E \rightarrow \text{type}(E)$ l'isomorphisme associé.

Alors pour tout $x \in E$, on a $f(x) = \text{type}(E_{\prec x})$ et $f|_{E_{\prec x}}$ est l'isomorphisme associé.



Démonstration

Notons \preccurlyeq l'ordre sur E et \leq l'ordre sur F .

- Soit $x \in E$.

Par hypothèse $f : E \rightarrow F$ est un isomorphisme d'ordres.

En particulier f est injective donc $f|_{E_{\prec x}}$ est injective.

De même f est croissante donc $f|_{E_{\prec x}}$ est croissante.

Il reste à montrer que $\text{im}(f|_{E_{\prec x}}) = F_{\triangleleft f(x)}$.



Soit $y \in \text{im}(f|_{E_{\prec x}})$.

Il existe donc $z \in E_{\prec x}$ tel que $y = f|_{E_{\prec x}}(z) = f(z)$.

Comme $f : E \rightarrow F$ est un isomorphisme d'ordres, f est strictement croissante.

Comme $z \in E_{\prec x}$, on a $z \prec x$ et donc $f(z) \triangleleft f(x)$.

Autrement dit on a $y \triangleleft f(x)$ donc $y \in F_{\triangleleft f(x)}$.

Ainsi on a $\text{im}(f|_{E_{\prec x}}) \subseteq F_{\triangleleft f(x)}$.



Soit $y \in F_{\triangleleft f(x)}$.

En particulier on a $y \in F$.

Or $f : E \rightarrow F$ est un isomorphisme d'ordre donc f est surjective dans F .

Il existe donc $z \in E$ tel que $y = f(z)$.

Par hypothèse E est totalement ordonné donc on a $z \prec x$ ou $x \preccurlyeq z$.

Supposons par l'absurde que $x \preccurlyeq z$.

Comme f est croissante on a $f(x) \preccurlyeq f(z)$, donc $f(x) \leq y$.

C'est absurde puisque par définition $y \in F_{\triangleleft f(x)}$ donc $y \triangleleft f(x)$.

On a donc nécessairement $z \prec x$, donc $z \in E_{\prec x}$.

Comme $y = f(z)$, on a $y \in f^{\rightarrow}(E_{\prec x}) = \text{im}(f|_{E_{\prec x}})$.

Ainsi on a $\text{im}(f|_{E_{\prec x}}) \supseteq F_{\triangleleft f(x)}$ et donc $\text{im}(f|_{E_{\prec x}}) = F_{\triangleleft f(x)}$.

Ainsi $f|_{E_{\prec x}}$ est surjective dans $F_{\triangleleft f(x)}$.

Ainsi $f|_{E_{\prec x}}$ est croissante, injective et surjective dans $F_{\triangleleft f(x)}$.

Or par hypothèse E est totalement ordonné, donc $E_{\prec x}$ son domaine aussi.

Donc $f|_{E_{\prec x}} : E_{\prec x} \longrightarrow F_{\triangleleft f(x)}$ est un isomorphisme d'ordres.

2. Soit $x \in E$.

Un ensemble bien ordonné est totalement ordonné d'après la proposition 2 page 10.

On peut donc appliquer 1.

On a donc $f|_{E_{\prec x}} : E_{\prec x} \longrightarrow \text{type}(E)_{\prec f(x)}$ est un isomorphisme d'ordres.

Or par définition $\text{type}(E)$ est un ordinal, donc $f(x)$ aussi.

On a aussi $\text{type}(E)_{\prec f(x)} = f(x)$ d'après la proposition 17 page 42.

Ainsi $f|_{E_{\prec x}} : E_{\prec x} \longrightarrow f(x)$ est un isomorphisme d'ordres et $f(x)$ un ordinal.

On a donc $f(x) = \text{type}(E_{\prec x})$.

CQFD.

6 Récurrence : induction et récursion

Au lycée, nous découvrons la notion de récurrence. On la retrouve notamment à travers le **raisonnement par récurrence** qui, comme nous l'avons déjà explicité, permet de prouver qu'une propriété est vraie pour tous les entiers naturels :

$$\text{Supposons } \begin{cases} P(0) \\ \forall n \in \mathbb{N}, [P(n) \implies P(n+1)] \end{cases}$$

Alors on a $\forall n \in \mathbb{N}, P(n)$.

Dans le même temps, on retrouve aussi la récurrence à travers les **définitions par récurrence**, qui permettent de définir une suite à partir d'une donnée initiale et d'une règle pour passer d'un entier au suivant :

$$\begin{cases} u_0 := a \\ \forall n \in \mathbb{N}, u_{n+1} := f(u_n) \end{cases}$$

Ces deux incarnations de la récurrence portent chacune un nom : le raisonnement par récurrence est aussi appelé **induction**, et la définition par récurrence est aussi appelée **récursion**.

6.1 Induction

L'induction chez les ordinaux est donc une généralisation de l'induction chez les nombres entiers : le principe est le même que pour le raisonnement par récurrence classique, c'est-à-dire prouver qu'une assertion est vraie à un certain ordinal et qu'elle se transmet de proche en proche par opération de successeur :

$$\text{Supposons } \begin{cases} P(0) \\ \forall \alpha \in ON, [P(\alpha) \implies P(S(\alpha))] \end{cases}$$

Alors on a $\forall \alpha \in ON, P(\alpha)$.

Cependant, nous l'avons dit : certains ordinaux ne sont le successeur de personne et donc il est impossible que l'assertion leur parvienne de cette façon (sauf pour 0 qui est déjà atteint au début). Autrement dit, la formulation qui précède n'est pas correcte.

Pour palier ce problème, on peut commencer par reformuler le raisonnement par récurrence sur les entiers naturels d'une autre manière. Pour démontrer qu'une assertion à paramètres P est vraie pour tout entier naturel n , on peut plutôt montrer :

$$\text{Supposons } \begin{cases} P(0) \\ \forall n \in \mathbb{N}, [\forall m < n, P(m)] \implies P(n) \end{cases}$$

Alors on a $\forall n \in \mathbb{N}, P(n)$.

On retrouve ce que l'on appelle usuellement le raisonnement par récurrence **forte**. En réalité, il ne s'agit pas d'un raisonnement plus fort que le raisonnement par récurrence classique. Pour s'en convaincre, il suffit de poser $Q(n) : \forall m < n, P(m)$. On peut alors remarquer que faire l'hypothèse de $Q(n)$, c'est faire l'hypothèse que P est vraie pour tout entier de 0 à $n - 1$, et dire que cela implique alors $P(n)$ signifie désormais que P est vraie pour un entier de plus,

c'est-à-dire pour tout entier précédent $n + 1$, et donc que $Q(n + 1)$ est vraie. Il s'agit donc tout simplement de l'implication $Q(n) \implies Q(n + 1)$. C'est donc bel et bien une hérédité classique.

Remarquons au passage qu'on peut enfouir l'initialisation $P(0)$ dans l'implication

$\left[\forall m < 0, P(m) \right] \implies P(0)$ puisque la prémissse étant toujours vraie, cette implication est équivalente à $P(0)$. Autrement dit, on peut reformuler le raisonnement par récurrence sur les entiers de la façon suivante :

Supposons $\forall n \in \mathbb{N}, \left[\forall m < n, P(m) \right] \implies P(n)$.

Alors on a $\forall n \in \mathbb{N}, P(n)$.

Or cette fois-ci il n'est pas question de successeur : cette formulation se généralise très bien aux ordinaux ! C'est l'objet du théorème qui suit.

Théorème 5 (Principe d'induction transfinie)

Soit P une assertion à paramètres.

Supposons que pour tout ordinal α , on a

$$\left[\forall \beta < \alpha, P(\beta) \right] \implies P(\alpha)$$

Alors pour tout ordinal α , on a $P(\alpha)$.

Démonstration

Supposons que pour tout ordinal α , on a $\left[\forall \beta < \alpha, P(\beta) \right] \implies P(\alpha)$.

Supposons par l'absurde qu'il existe au moins un ordinal α tel que l'on n'a pas $P(\alpha)$.

Soit X l'ensemble des ordinaux plus petit ou égaux à α et qui ne vérifient pas P .

Par définition on a $\alpha \in X$ donc X est un ensemble non vide d'ordinaux.

Il admet donc un ordinal minimum ξ d'après le théorème 1 page 21.

Soit μ un ordinal tel que $\mu < \xi$.

Alors $\mu \notin X$ car ξ est minimum de X .

Or $\xi \in X$ donc $\xi \leq \alpha$ et donc $\mu < \xi \leq \alpha$.

On a donc $\mu \leq \alpha$ par transitivité.

Ainsi on a $\mu \notin X$ alors que $\mu \leq \alpha$.

Donc nécessairement on a $P(\mu)$ par définition de X .

Donc $\forall \mu < \xi, P(\mu)$.

On a donc $P(\xi)$ par hypothèse du théorème.

C'est absurde puisque $\xi \in X$ et donc $P(\xi)$ est faux.

Donc par l'absurde on a montré que $\boxed{\text{pour tout ordinal } \alpha \text{ on a } P(\alpha)}$.

CQFD.

On est cependant en droit de se demander : la formulation classique avec le passage de n à $n + 1$ a-t-elle une généralisation chez les ordinaux ? La réponse est oui, à condition de traiter séparément le cas des ordinaux limites puisqu'ils ne sont pas successeurs. C'est donc au fond un mélange des deux formulations. À la manière de la récurrence classique et de la récurrence forte chez les entiers naturels, c'est une formulation équivalente à la précédente, on ne dit au fond rien de moins même si naïvement on peut en avoir l'impression.

Proposition 29 (Principe faible d'induction transfinie)

Soit P une assertion à paramètres.

Supposons que :

1. On a $P(0)$.
2. Pour tout ordinal α , si $P(\alpha)$ alors $P(S(\alpha))$.
3. Pour tout ordinal limite α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Alors pour tout ordinal α on a $P(\alpha)$.

Démonstration

Appliquons le théorème 5 page 65.

Soit α un ordinal.

Supposons que $\forall \beta < \alpha, P(\beta)$ (\star).

► Si $\alpha = 0$, alors d'après l'hypothèse 1 on a $P(\alpha)$.

► Supposons que α est un successeur.

Par définition il existe un ordinal β tel que $\alpha = S(\beta)$.

Alors $\beta < \alpha$ d'après la proposition 13 page 33.

On a donc $P(\beta)$ d'après l'hypothèse (\star).

On a donc $P(\alpha)$ d'après l'hypothèse 2.

► Supposons que α est un ordinal limite.

On alors $P(\alpha)$ d'après les hypothèses (\star) et 3.

Dans tous les cas on a donc $P(\alpha)$.

Donc si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Donc pour tout ordinal α , on l'implication $(\forall \beta < \alpha, P(\beta)) \implies P(\alpha)$.

Donc pour tout ordinal α , on a $P(\alpha)$ d'après le théorème 5 page 65.

CQFD.

Remarque :

0 étant un ordinal limite, il entre à la fois dans le cas 1 et le cas 3, mais comme $\forall \beta < 0, P(\beta)$ est nécessairement vraie, l'implication $(\forall \beta < 0, P(\beta)) \implies P(0)$ est équivalente à $P(0)$ et donc il y a seulement une redondance, pas de contradiction.

6.2 Récursion

Nous l'avons dit, la récursion chez les entiers naturels est aussi connue sous le nom de définition par récurrence, pour définir une suite : on définit la valeur de cette suite en un entier puis l'on se donne une règle pour déterminer la valeur de la suite sur l'entier suivant à partir du précédent, ce qui permet de proche en proche de définir la suite sur chaque entier. Par exemple on pourrait être amenés à définir la suite suivante :

$$\begin{cases} u_0 := 1 \\ \forall n \in \mathbb{N}^*, u_n := 3u_{n-1} \end{cases}$$

qui va alors donner la suite des puissances de 3. On peut se retrouver dans le cas où l'étape de propagation nécessite en fait les deux termes précédents (auquel cas il faut déterminer la valeur de la suite sur deux entiers au début), comme c'est le cas avec **la suite de Fibonacci** :

$$\begin{cases} u_0 := 1 \\ u_1 := 1 \\ \forall n \geq 2, u_n := u_{n-1} + u_{n-2} \end{cases}$$

Plus généralement, on peut même vouloir définir un terme à partir de toutes les valeurs précédentes. La notion qui va permettre de donner une "*règle de construction*" en toute généralité pour se servir des valeurs précédentes est celle d'**assertion fonctionnelle** que nous avons rappelée au début de ce chapitre. Ainsi, si H est une assertion fonctionnelle, la forme la plus générale qu'on pourrait être amenés à utiliser pour définir une suite par récurrence sur les entiers est

$$\forall n \in \mathbb{N}, u_n := H(u_{\llbracket 0, n \rrbracket})$$

où $u_{\llbracket 0, n \rrbracket}$ désigne la restriction de la suite u à tous les entiers de 0 à $n - 1$. Ainsi, on tient bien compte des valeurs de u jusqu'à n (non compris). Remarquons bien que comme H est très générale, elle peut en particulier ne regarder que quelques valeurs parmi les précédentes et non toutes (par exemple seulement les deux précédentes comme dans le cas de Fibonacci), et aussi être constante en quelques entiers pour s'assurer d'avoir fixé les premières valeurs de la suite. Ainsi dans le cas de la suite de Fibonacci, H serait définie de telle sorte à avoir

$$\begin{cases} H(u_{\llbracket 0, n \rrbracket}) := 1 & \text{si } n = 0 \\ H(u_{\llbracket 0, n \rrbracket}) := 1 & \text{si } n = 1 \\ H(u_{\llbracket 0, n \rrbracket}) := u_{n-1} + u_{n-2} & \text{sinon} \end{cases}$$

C'est ce cadre-là que nous allons désormais définir proprement pour le généraliser encore plus, c'est-à-dire à présent sur tous les ordinaux. Le principe va cependant rester le même : se donner une règle de propagation via les assertions fonctionnelles, et l'utiliser pour définir la valeur d'une suite à un ordinal à partir de la restriction de la suite aux ordinaux précédents.

On remarque que pour que u puisse être définie, il faut que ses restrictions respectives soient bien dans le domaine de H pour que l'étape de propagation ait du sens. C'est à travers la notion d'application inductive (une suite étant une application particulière) que nous allons faire cela.

Définition 12 (Application inductive)

Soient H une assertion fonctionnelle et u une application.

On dit que u est **H -inductive** si et seulement si :

1. $\text{dom}(u)$ est un ordinal.
2. Pour tout $\beta \in \text{dom}(u)$, on a $u|_{\beta} \in \text{dom}(H)$ et $u(\beta) = H(u|_{\beta})$.

Ne perdons pas de vue qu'un ordinal est lui-même l'ensemble de tous les ordinaux qui le précédent, autrement dit $u|_{\beta}$ est bien la restriction de u à tous les ordinaux qui viennent avant β , au même titre que $u|_{[0,n]}$ est la restriction de u à tous les entiers qui précèdent n . Dans le cadre conceptuel des ordinaux, on a d'ailleurs bien l'égalité : $n = [0, n]$, donc on retombe bien sur nos pieds.

À ce stade, nous avons déjà défini les notions utiles pour construire les différentes suites que nous avons évoquées : il suffit pour cela de bien choisir le H en question et les applications u qui sont H -inductives et concernées seront celles telles que $\text{dom}(u) = \omega$ l'ensemble des entiers naturels.

Cependant nous avons exprimé le souhait d'aller au delà de ω à travers la théorie plus générale des ordinaux. On pourrait tout à fait se contenter pour cela de la définition que nous venons d'énoncer : si l'on souhaite se rendre jusqu'à un ordinal α , même très grand, il suffit de demander $\text{dom}(u) = \alpha$ pour les applications H -inductives qui nous intéressent.

Il y a néanmoins des cas où nous ne voudrions pas particulièrement limiter l'ordinal jusqu'où construire l'application u . Typiquement, étant donnés deux ordinaux α et β , nous serons amenés à définir l'addition $\alpha + \beta$. Nous le ferons à l'aide d'une assertion fonctionnelle H bien choisie. En passant par une application u qui est H -inductive, on pourra faire en sorte que

$$\forall \beta \in \text{dom}(u), \alpha + \beta := u(\beta)$$

en ayant fixé α au préalable. Le problème vient alors de savoir le sens à donner à $\alpha + \text{dom}(u)$. En effet, $\text{dom}(u)$ est lui-même un ordinal, que l'on devra en plus choisir arbitrairement. On peut se contenter de se limiter à $\text{dom}(u)$ en l'ayant pris très grand, mais cela présente une inélégance que l'on peut corriger.

On aimerait pour cela ne pas limiter le domaine des applications : comment faire pour que le domaine soit ON la classe de tous les ordinaux ? En fait, il nous suffit de passer par la généralisation des applications dont nous avons déjà tant parlée : les assertions fonctionnelles. Ainsi, nous allons simplement étendre la définition précédente aux assertions fonctionnelles.

Définition 13 (Assertion fonctionnelle inductive)

Soient H et F deux assertions fonctionnelles.

On dit que F est **H -inductive** si et seulement si :

1. $\text{dom}(F) \in ON$ ou $\text{dom}(F) = ON$.
2. Pour tout $\beta \in \text{dom}(F)$, on a $F|_{\beta} \in \text{dom}(H)$ et $F(\beta) = H(F|_{\beta})$.

Le seul véritable cas nouveau est celui pour lequel $\text{dom}(F) = ON$. En effet, si $\text{dom}(F) \in ON$ alors $\text{dom}(F)$ est un ordinal donc F est alors associée à une application et donc on confond sans problème les deux. On peut cependant se demander pourquoi le seul cas nouveau que l'on rajoute est celui où $\text{dom}(F) = ON$. Au fond, tant qu'on est à généraliser, on pourrait demander plus largement $\text{dom}(F) \subseteq ON$, non ? La réponse nous a déjà été fournie par la proposition 10 page 26 : ON est en quelque sorte la seule classe propre légitime à généraliser les ordinaux.

On peut remarquer la chose suivante : restreindre une assertion fonctionnelle (ou donc une application) qui est H -inductive à un ordinal de son domaine va nécessairement produire une application qui est encore H -inductive. En effet : *qui peut le plus peut le moins*.

Proposition 30 (Restriction d'une application inductive)

Soient H et F deux assertions fonctionnelles.

Si F est H -inductive alors pour tout $\beta \in \text{dom}(F)$, l'application $F|_\beta$ est H -inductive.

Démonstration

Supposons que F est H -inductive.

Alors on a $\text{dom}(F) \in ON$ ou $\text{dom}(F) = ON$ par définition.

On a donc $\text{dom}(F) \subseteq ON$ d'après la proposition 6 page 17.

Soit $\beta \in \text{dom}(F)$.

Comme $\text{dom}(F) \subseteq ON$, on en déduit que $\text{dom}(F|_\beta) = \beta$ est un ordinal.

Soit $\gamma \in \beta$.

On a alors $\gamma \in \beta \in \text{dom}(F)$ donc $\gamma \in \text{dom}(F)$ par transitivité :

Si $\text{dom}(F) \in ON$ alors c'est la transitivité de \in sur ON qu'on applique.

Si $\text{dom}(F) = ON$, c'est la transitivité de ON qu'on applique.

Or F est H -inductive par hypothèse.

Donc $F|_\gamma$ est dans le domaine de H et $F(\gamma) = H(F|_\gamma)$.

Or $\gamma \in \beta$ donc $\gamma \subseteq \beta$ par transitivité.

Donc $(F|_\beta)|_\gamma = F|_\gamma$ donc en particulier $(F|_\beta)|_\gamma$ est dans le domaine de H .

De plus $F|_\beta(\gamma) = F(\gamma) = H(F|_\gamma) = H((F|_\beta)|_\gamma)$.

Donc pour tout $\gamma \in \beta$, $(F|_\beta)|_\gamma$ est dans le domaine de H et $F|_\beta(\gamma) = H((F|_\beta)|_\gamma)$.

Donc $F|_\beta$ est H -inductive.

Donc pour tout $\beta \in \text{dom}(F)$, $F|_\beta$ est H -inductive.

CQFD.

Pour pouvoir dire qu'on souhaite définir une assertion fonctionnelle F par la relation

$$\forall \beta \in \text{dom}(F), F(\beta) = H(F|_\beta)$$

il faut s'assurer qu'une telle relation ne convient pas pour plusieurs assertions fonctionnelles : c'est l'objet de la proposition suivante.

Proposition 31 (Au plus une application inductive)

Soit H une assertion fonctionnelle.

Soit C une classe telle que $C \in ON$ ou $C = ON$.

Il existe au plus une assertion fonctionnelle de domaine C qui est H -inductive.



Démonstration

Soient F et G deux assertions fonctionnelles qui sont toutes deux H -inductives et telles que $\text{dom}(F) = C = \text{dom}(G)$.

Montrons que $F = G$ par l'absurde.

Supposons par l'absurde que $F \neq G$.

Comme elles ont le même domaine C , il existe $\alpha \in C$ tel que $F(\alpha) \neq G(\alpha)$.

Considérons alors la classe $X := \{\beta \in C \mid F(\beta) \neq G(\beta)\}$.

Par définition on a $\alpha \in X$ donc X est non vide.

On a aussi $X \subseteq C$ par définition de X .

Or on a ($C \in ON$ ou $C = ON$) par définition de C , donc $C \subseteq ON$ et donc $X \subseteq ON$.

Ainsi X est classe non vide telle que $X \subseteq ON$.

Donc X possède un ordinal minimum ξ d'après la proposition 9 page 24.

Remarquons que l'on a $\xi \in X$ et $X \subseteq C$ donc $\xi \in C$ par définition de l'inclusion.

Soit $\gamma \in \xi$.

On a dit que $\xi \in C$ donc $\gamma \in C$ par transitivité :

Si $C \in ON$ alors c'est la transitivité de \in sur ON qu'on applique.

Si $C = ON$, c'est la transitivité de ON qu'on applique.

On a aussi $\gamma < \xi$ par définition de $<$.

Comme ξ est le minimum de X , on a donc $\gamma \notin X$.

Ainsi on a $\gamma \notin X$ alors que $\gamma \in C$.

Nécessairement on a donc $F(\gamma) = G(\gamma)$ par définition de X .

Donc $\forall \gamma \in \xi, F(\gamma) = G(\gamma)$ et donc $F|_\xi = G|_\xi$.

Or F et G sont H -inductives donc $F(\xi) = H(F|_\xi) = H(G|_\xi) = G(\xi)$.

C'est absurde puisque $\xi \in X$ donc $F(\xi) \neq G(\xi)$.

Donc par l'absurde on a $F = G$, d'où l'unicité.

CQFD.

Remarque :

En particulier si F est H -inductive, alors pour tout $\alpha \in \text{dom}(F)$, l'application $F|_\alpha$ est l'unique application H -inductive de domaine α .

Nous venons de voir qu'à domaine fixé, il existe au plus une assertion fonctionnelle qui est H -inductive. Mais en existe-t-il au moins une ? La réponse est oui, mais à une certaine condition sur H . En fait l'idée est de construire notre assertion fonctionnelle de proche en proche, donc

par récurrence (plus précisément par récursion), puisque la valeur en un ordinal est fonction des valeurs en les ordinaux précédents.

Imaginons que l'on ait défini notre application/assertion fonctionnelle jusqu'à l'ordinal α . Cela revient à dire que nous avons à notre disposition une application $v : \alpha \longrightarrow ?$ qui est H -inductive. Pour pouvoir poursuivre à nouveau la construction, et faire en sorte que v ne soit en fait que la restriction à α de notre assertion fonctionnelle finale, il faut simplement s'assurer que v elle-même est dans le domaine de H , et non pas seulement ses restrictions. C'est pour cela que dans le théorème suivant, on a rajouté cette condition.

Théorème 6 (Principe de récursion transfinie)

Soit H une assertion fonctionnelle.

Soit C une classe telle que $C \in ON$ ou $C = ON$.

Supposons que pour tout $\alpha \in C$ et toute application $v : \alpha \longrightarrow ?$ on a

si v est H -inductive alors v est dans le domaine de H

Alors il existe une unique assertion fonctionnelle de domaine C qui est H -inductive.

Démonstration

Unicité

C'est exactement l'objet de la proposition 31 page 70.

Existence

Considérons T la classe des $\alpha \in C$ tel qu'il existe une application $\alpha \longrightarrow ?$ qui est H -inductive et dans le domaine de H . Une telle application est alors unique d'après la proposition 31 page 70. Ainsi T représente en quelque sorte le domaine maximal auquel on peut construire notre assertion fonctionnelle. Notre but va donc simplement être de montrer que $T = C$: on peut en fait aller jusqu'au bout.

Voici les différentes étapes de notre preuve :

1. On montre que comme pour C , on a $T \in ON$ ou $T = ON$

Autrement dit T est un ordinal ou est la généralisation d'un ordinal.

On s'assure ainsi que le domaine est pertinent pour une construction par récursion.

- (a) Cela passe d'abord par montrer que T est transitive.
- (b) Puis on conclut avec le fait que (T, \in) est strictement bien ordonné.

2. On construit l'assertion fonctionnelle $U_T : T \longrightarrow ?$ qui est H -inductive.

La construction va se faire par récursion : c'est justement notre objectif.

3. On montre que l'existence de U_T implique nécessairement que $T = C$.

1.(a) Montrons que T est transitive.

Soit $\alpha \in T$.

Comme $\alpha \in T$, par définition de T il existe une application $u : \alpha \longrightarrow ?$ qui est H -inductive et dans le domaine de H .

Soit $\beta \in \alpha$.

Comme u est H -inductive, $u|_\beta$ est dans le domaine de H .

De plus $u|_\beta : \beta \longrightarrow ?$ est H -inductive d'après la proposition 30 page 69.

Ainsi il existe une application $\beta \longrightarrow ?$ qui est H -inductive et dans le domaine de H . Pour prouver que $\beta \in T$, il reste donc à prouver que $\beta \in C$.

Or on a $\beta \in \alpha \in C$ donc $\beta \in C$ par transitivité.

Si $C \in ON$ alors c'est la transitivité de \in sur ON qu'on applique.

Si $C = ON$, c'est la transitivité de ON qu'on applique.

Donc $\beta \in T$ par définition de T .

Donc pour tout $\beta \in \alpha$, on a $\beta \in T$, et donc $\alpha \subseteq T$ par définition de l'inclusion.

Ainsi $\forall \alpha \in T, \alpha \subseteq T$ donc T est transitive.

1.(b) Montrons que $T \in ON$ ou $T = ON$.

On a $(C \in ON \text{ ou } C = ON)$ donc $C \subseteq ON$.

Comme $T \subseteq C$ on a donc $T \subseteq ON$.

Ainsi tous les éléments de T sont des ordinaux.

Donc (T, \in) est strictement bien ordonnée d'après le théorème 1 page 21.

Or on vient de montrer que T est transitive.

- Si T est issue d'un ensemble alors T est un ordinal par définition d'être un ordinal.
- Si T est une classe propre, alors $T = ON$ d'après la proposition 10 page 26.

On a donc nécessairement $T \in ON$ ou $T = ON$.

2. Pour tout $\alpha \in T$, posons u_α l'unique application $\alpha \longrightarrow ?$ qui est H -inductive et dans le domaine de H . Construisons maintenant U_T l'unique assertion fonctionnelle $T \longrightarrow ?$ qui est H -inductive.

Pour cela, on peut raisonner par analyse synthèse : imaginons qu'on ait déjà à notre disposition l'unique assertion fonctionnelle $U_T : T \longrightarrow ?$ qui est H -inductive.

Le fait qu'elle est H -inductive signifie en particulier que pour tout $\alpha \in T$, on a

$$U_T(\alpha) = H((U_T)|_\alpha)$$

Autrement dit, pour connaître la valeur de $U_T(\alpha)$, il nous faut connaître $(U_T)|_\alpha$.

Mais comme U_T est H -inductive, on a forcément que $(U_T)_{|\alpha} : \alpha \longrightarrow ?$ est elle-même H -inductive d'après la proposition 30 page 69. Or justement on sait qu'une telle application est nécessairement u_α par définition de u_α . Autrement dit, on a nécessairement $(U_T)_{|\alpha} = u_\alpha$ et donc on a nécessairement $U_T(\alpha) = H(u_\alpha)$. Ainsi, notre candidat pour U_T est donné par

$$U_T := \begin{pmatrix} T & \longrightarrow & ? \\ \alpha & \longmapsto & H(u_\alpha) \end{pmatrix}$$

Montrons que U_T ainsi définie est H -inductive.

On sait déjà que $\text{dom}(U_T) = T$ est ou bien un ordinal ou bien ON toute entière.

Soit $\alpha \in T$.

Montrons que $(U_T)_{|\alpha} = u_\alpha$.

Soit $\beta \in \alpha$.

► On a alors $\beta \in \alpha \in T$ donc $\beta \in T$ par transitivité.

Si $T \in ON$ alors c'est la transitivité de \in sur ON qu'on applique.

Si $T = ON$, c'est la transitivité de ON qu'on applique.

On a $(U_T)_{|\alpha}(\beta) = U_T(\beta)$ par définition d'une restriction.

Or $U_T(\beta) = H(u_\beta)$ par définition de U_T , donc $(U_T)_{|\alpha}(\beta) = H(u_\beta)$.

► D'un autre côté, on sait que u_α est par définition H -inductive.

Donc $u_\alpha(\beta) = H((u_\alpha)_{|\beta})$ par définition de la H -inductivité.

► Enfin $(u_\alpha)_{|\beta}$ est H -inductive d'après la proposition 30 page 69.

Donc $(u_\alpha)_{|\beta} = u_\beta$ par unicité de l'application de domaine β qui est H -inductive.

$$\text{On a donc montré que } \begin{cases} (U_T)_{|\alpha}(\beta) = H(u_\beta) \\ u_\alpha(\beta) = H((u_\alpha)_{|\beta}) \\ (u_\alpha)_{|\beta} = u_\beta \end{cases}$$

Les deux dernières lignes nous disent que $u_\alpha(\beta) = H(u_\beta)$.

Combiné à la première ligne, on en déduit que $(U_T)_{|\alpha}(\beta) = u_\alpha(\beta)$.

Ainsi $\forall \beta \in \alpha, (U_T)_{|\alpha}(\beta) = u_\alpha(\beta)$.

Or $(U_T)_{|\alpha}$ et u_α ont le même domaine α donc $(U_T)_{|\alpha} = u_\alpha$.

Or par définitions u_α est dans le domaine de H et $U_T(\alpha) = H(u_\alpha)$.

Donc $(U_T)_{|\alpha}$ est dans le domaine de H et $U_T(\alpha) = H((U_T)_{|\alpha})$.

Donc pour tout $\alpha \in T$, $(U_T)_{|\alpha}$ est dans le domaine de H et $U_T(\alpha) = H((U_T)_{|\alpha})$.

Donc U_T est H -inductive.

3. Enfin, montrons que $T = C$ pour conclure.

Supposons par l'absurde que $T \neq C$.

Par définition de T on a $T \subseteq C$ donc on a $T \subsetneq C$.

On a dit que $(T \in ON \text{ ou } T = ON)$ et $(C \in ON \text{ ou } C = ON)$.

- Si $T \in ON$ et $C \in ON$ alors $T \in C$ d'après la prop. 8 p. 19.
- Si $T \in ON$ et $C = ON$ alors immédiatement $T \in C$.
- Si $T = ON$ et $C \in ON$ c'est absurde puisque $T \subsetneq C$.
- Si $T = ON$ et $C = ON$ c'est absurde puisque $T \subsetneq C$.

On a donc nécessairement $T \in C$.

Comme $(C \in ON \text{ ou } C = ON)$ on a $C \subseteq ON$.

On a donc $T \in ON$ par définition de l'inclusion.

Donc T est un ordinal donc en particulier est un ensemble.

Donc U_T est une assertion fonctionnelle de domaine un ensemble.

Donc U_T est une application.

Ainsi U_T est une application H -inductive dont le domaine appartient à C .

Donc U_T est dans le domaine de H par hypothèse du théorème.

Ainsi T est un élément de C et le domaine d'une application H -inductive qui est elle-même dans le domaine de H .

Donc $T \in T$ par définition de T .

C'est absurde par antiréflexivité de \in sur ON .

On a donc montré par l'absurde que $T = C$.

Or U_T est une assertion fonctionnelle de domaine T qui est H -inductive.

Donc U_T est une assertion fonctionnelle de domaine C qui est H -inductive.

CQFD.

Exemple :

Reprendons les exemples du début et éclairons-les de ce que l'on vient d'apprendre.

1. Si l'on souhaite définir proprement l'unique suite $(u_n)_{n \in \mathbb{N}}$ vérifiant

$$\begin{cases} u_0 = 1 \\ \forall n \in \mathbb{N}, u_{n+1} = 3u_n \end{cases}$$

il suffit de considérer l'assertion fonctionnelle H définie pour toute application f

telle que $\text{dom}(f) \in \mathbb{N}$ par

$$\begin{cases} H(f) := 1 & \text{si } \text{dom}(f) = 0 \\ H(f) := 3f(n) & \text{si } \text{dom}(f) = n + 1 \text{ avec } n \in \mathbb{N} \end{cases}$$

En effet, on considère alors $C = \mathbb{N}$, qui vérifie bien $C \in ON$ d'après la proposition 18 page 43. Prenons alors $n \in \mathbb{N}$ et $f : n \longrightarrow ?$ quelconque : par définition de H , f est nécessairement dans le domaine de H . Cela reste donc vrai si en particulier f est H -inductive. Ainsi H et C vérifient les hypothèses du théorème précédent : il existe une unique suite $u = (u_n)_{n \in \mathbb{N}}$ qui est H -inductive. Cette suite vérifie alors

$$\begin{cases} u_0 = u(0) = H(u|_0) = 1 \\ \forall n \in \mathbb{N}, u_{n+1} = u(n+1) = H(u|_{n+1}) = 3u(n) = 3u_n \end{cases}$$

2. Si l'on souhaite définir proprement l'unique suite $(u_n)_{n \in \mathbb{N}}$ vérifiant

$$\begin{cases} u_0 = 1 \\ u_1 = 1 \\ \forall n \in \mathbb{N}, u_{n+2} = u_{n+1} + u_n \end{cases}$$

il suffit de considérer l'assertion fonctionnelle H définie pour toute application f telle que $\text{dom}(f) \in \mathbb{N}$ par

$$\begin{cases} H(f) := 1 & \text{si } \text{dom}(f) = 0 \\ H(f) := 1 & \text{si } \text{dom}(f) = 1 \\ H(f) := f(n+1) + f(n) & \text{si } \text{dom}(f) = n + 2 \text{ pour } n \in \mathbb{N} \end{cases}$$

On considère ici encore $C = \mathbb{N}$. Pour les mêmes raisons que l'exemple précédent, H et C vérifient les hypothèses du théorème précédent : il existe une unique suite $u = (u_n)_{n \in \mathbb{N}}$ qui est H -inductive. Cette suite vérifie alors

$$\begin{cases} u_0 = u(0) = H(u|_0) = 1 \\ u_1 = u(1) = H(u|_1) = 1 \\ \forall n \in \mathbb{N}, u_{n+2} = u(n+2) = H(u|_{n+2}) = u(n+1) + u(n) = u_{n+1} + u_n \end{cases}$$

C'est la célèbre **suite de Fibonacci**.

Pour la petite histoire



Leonardo Fibonacci (vers 1170 – vers 1250), de son vrai nom Léonard De Pise, est le fils d'un commerçant toscan. Ce dernier émigre avec toute sa famille à Béjaïa dans l'actuelle Algérie et Leonardo est encouragé à maîtriser les comptes pour l'aider. Par la suite, Fibonacci parcourt l'Égypte, la Sicile, la Grèce et la Syrie. Il entre ainsi en contact avec les mathématiques arabes et grecques.

Convaincu de la supériorité du système d'écriture des nombres par les chiffres arabes, il écrit *Liber abaci* à son retour en Europe en 1202, ce qui les introduira en occident. Dans cet ouvrage, il explique la notation de position, les méthodes de calcul des opérations élémentaires, et des méthodes de résolutions d'équations.

Si la suite de Fibonacci était déjà connue au moins depuis 200 avant JC en Inde, c'est Fibonacci qui la rendra célèbre en occident dans *Liber abaci*.

6.3 Suites

Les exemples précédents parlent de suites, mais nous n'avons pas encore eu l'occasion de définir proprement ce que c'est. Intuitivement, une suite est une liste d'objets mathématiques, un pour chaque entier naturel. Autrement dit, c'est simplement une application $\mathbb{N} \longrightarrow ?$.

Définition 14 (Suite)

On appelle **suite** toute application $u : \mathbb{N} \longrightarrow ?$.

Pour tout ensemble E , on appelle **suite à valeurs dans E** toute suite $u : \mathbb{N} \longrightarrow E$.

Remarque :

1. Nous avons vu dans le livre précédent que les familles ne sont qu'un autre nom donné aux applications, mais qu'il est associé à d'autres notations et usages. Les suites rentrent plutôt dans le cadre des familles : c'est pourquoi une suite u est souvent notée $(u_n)_{n \in \mathbb{N}}$.
2. L'ensemble des suites à valeurs dans E étant l'ensemble $\mathcal{F}(\mathbb{N} \longrightarrow E)$ des applications de \mathbb{N} dans E , il est aussi souvent noté $E^{\mathbb{N}}$.
3. Pour la fin de ce chapitre, notons $n + 1$ à la place de $S(n)$, puisque c'est de toute manière l'intuition qui se cache derrière $S(n)$. Bien heureusement, au prochain

chapitre nous définirons proprement l'addition, et nous verrons que l'on a bien $n + 1 = S(n)$.

Dans les exemples précédents nous avons défini des suites par récurrence, plus précisément par récursion. Pour cela, nous avons invoqué une assertion fonctionnelle particulière et utilisé le théorème 6 page 71. Nous aimerais ne plus avoir à passer par une assertion fonctionnelle à chaque fois et pouvoir directement définir la suite en question. La proposition suivante se propose simplement de le faire une bonne fois pour toute.

Proposition 32 (Suites définies par récurrence)

Soient E un ensemble, $a \in E$ et $f : E \longrightarrow E$.

Alors il existe une unique suite $(u_n)_{n \in \mathbb{N}}$ à valeurs dans E telle que

$$\begin{cases} u_0 = a \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

Démonstration

Pour tout $m \in \mathbb{N}$ et tout $g : m \longrightarrow E$, posons

$$\begin{cases} H(g) = a & \text{si } m = 0 \\ H(g) = f(g(n)) & \text{si } m = n + 1 \text{ avec } n \in \mathbb{N} \end{cases}$$

On vient ainsi de définir une assertion fonctionnelle H . Par définition, un ensemble g est dans le domaine de H si et seulement s'il existe $m \in \mathbb{N}$ tel que $g : m \longrightarrow E$.

Nous allons utiliser le théorème 6 page 71 avec $C := \mathbb{N}$. On sait déjà que C est un ordinal d'après la proposition 18 page 43, donc vérifie $C \in ON$. Il reste donc à montrer que H et C vérifie l'hypothèse du théorème.

Soient $m \in \mathbb{N}$ et $g : m \longrightarrow ?$ qui est H -inductive.

Montrons que $g : m \longrightarrow E$, c'est-à-dire $\text{im}(g) \subseteq E$.

Soit $y \in \text{im}(g)$.

Par définition il existe $n \in \text{dom}(g)$ tel que $y = g(n)$.

Or g est H -inductive par définition.

On a donc $g|_n \in \text{dom}(H)$ et $y = g(n) = H(g|_n)$.

Remarquons deux choses : tout d'abord on a $g|_n : n \longrightarrow ?$ mais on vient de voir que $g|_n \in \text{dom}(H)$ donc $g|_n : n \longrightarrow E$ par définition de H . Or $f : E \longrightarrow E$ donc pour tout $p < n$, on a $g|_n(p) \in \text{dom}(f)$, si bien que l'on peut tout à fait

parler de $f(g|_n(p))$.

Deuxièmement, $n \in \text{dom}(g)$ et $g : m \longrightarrow ?$ donc $n \in m$, c'est-à-dire $n < m$ par définition de $<$. Comme $m \in \mathbb{N}$, on a $n \in \mathbb{N}$ d'après la proposition 15 page 38. Donc n est ou bien 0, ou bien le successeur d'un entier naturel p par définition d'être un entier naturel.

► Plaçons-nous dans le cas où $n = 0$.

Alors $\text{dom}(g|_0) = 0$ donc $y = H(g|_0) = a$ par définition de H .

Comme $a \in E$ par définition, on a $y \in E$.

► Plaçons-nous dans le cas où $n = p+1$ avec $p \in \mathbb{N}$. Alors $\text{dom}(g|_n) = p+1$ donc $y = H(g|_n) = f(g|_n(p))$ par définition de H . Or par définition on a $f : E \longrightarrow E$ donc $\text{im}(f) \subseteq E$ et donc $y \in E$.

Dans les deux cas on a $y \in E$.

Ainsi on a $\text{im}(g) \subseteq E$ par définition de l'inclusion, et donc $g : m \longrightarrow E$.

Donc pour tout $m \in \mathbb{N}$ et toute application $g : m \longrightarrow ?$, si g est H -inductive alors $g : m \longrightarrow E$ et donc g est dans le domaine de H .

Ainsi $C = \mathbb{N}$ et H vérifient les hypothèses du théorème 6 page 71.

Il existe donc une unique application $u : \mathbb{N} \longrightarrow ?$ qui est H -inductive.

En particulier u vérifie

$$\begin{cases} u_0 = u(0) = H(u|_0) = a \\ \forall n \in \mathbb{N}, u_{n+1} = u(n+1) = H(u|_{n+1}) = f(u(n)) = f(u_n) \end{cases}$$

CQFD.

Remarque :

Dans la proposition précédente, nous avons explicité le cas où la valeur de u en un entier ne dépend que de la valeur de u à l'entier précédent, nous permettant de ne plus avoir à passer par une assertion fonctionnelle à chaque fois qu'une suite est définie par une récurrence. Cependant, nous n'avons pas décrit le cas où la valeur de u en un entier dépend des valeurs de u aux deux, trois, voire tous les entiers précédents.

Nous n'allons évidemment pas le faire : tenter de traiter tous les cas possibles reviendrait précisément à réénoncer le théorème 6 page 71. Autrement dit, nous allons désormais simplement considérer que le lecteur comprend comment le faire en toute généralité. Nous avons par exemple déjà donné la recette de cuisine pour la suite de Fibonacci.

Remarquons que cela nous permet par exemple de définir des suites vérifiant $u_{n+1} = \sum_{k=0}^n u_k$ (en ayant fixée la valeur en 0) puisque l'on a dit qu'on peut utiliser toutes les valeurs aux

entiers précédents !

En tant qu'ordinal, $\mathbb{N} = \omega$ est muni d'une relation d'ordre. À ce titre, pour tout ensemble ordonné E , on a déjà donné du sens dans le livre précédent à la notion de suite (strictement) croissante, (strictement) décroissante et constante. Rajoutons une petite dernière : celle de stationnarité.

Définition 15 (Suite stationnaire)

Soient E un ensemble et $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans E .

On dit que $(u_n)_{n \in \mathbb{N}}$ est **stationnaire** si et seulement s'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, u_n = u_{n_0}$$

Remarque :

Une suite constante est un cas particulier de suite stationnaire : elle l'est simplement depuis le début, c'est-à-dire qu'on peut prendre $n_0 = 0$.

La notion de successeurs chez les ordinaux et donc chez les entiers naturels offre une caractérisation intéressante de ces propriétés dont on a parlé juste avant : il suffit qu'elles soient vérifiées entre un entier et le suivant pour se propager en fait à tous les entiers. C'est l'objet de la proposition suivante. On utilise ici la notation $n + 1$ pour désigner $S(n)$, bien que nous verrons qu'elle coïncide bien avec l'addition quand nous l'aurons définie dans le prochain chapitre.

Proposition 33 (Suites croissantes, décroissantes, constantes)

Soient (E, \preccurlyeq) un ensemble ordonné et $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans E .

Soit \prec l'ordre strict associé à \preccurlyeq .

1. $(u_n)_{n \in \mathbb{N}}$ est croissante si et seulement si $\forall n \in \mathbb{N}, u_n \preccurlyeq u_{n+1}$.
2. $(u_n)_{n \in \mathbb{N}}$ est strictement croissante si et seulement si $\forall n \in \mathbb{N}, u_n \prec u_{n+1}$.
3. $(u_n)_{n \in \mathbb{N}}$ est décroissante si et seulement si $\forall n \in \mathbb{N}, u_{n+1} \preccurlyeq u_n$.
4. $(u_n)_{n \in \mathbb{N}}$ est strictement décroissante si et seulement si $\forall n \in \mathbb{N}, u_{n+1} \prec u_n$.
5. $(u_n)_{n \in \mathbb{N}}$ est constante si et seulement si $\forall n \in \mathbb{N}, u_n = u_{n+1}$.
6. $(u_n)_{n \in \mathbb{N}}$ est stationnaire si et seulement si $\exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n = u_{n+1}$.

Démonstration

1. \Rightarrow

Supposons que $(u_n)_{n \in \mathbb{N}}$ est croissante.

Soit $n \in \mathbb{N}$.

On sait que $n < n + 1$ d'après la proposition 13 page 33.

En particulier on a $n \leq n + 1$.

On a donc $u_n \preccurlyeq u_{n+1}$ par croissance de u .

Donc $\boxed{\forall n \in \mathbb{N}, u_n \preccurlyeq u_{n+1}}$.



Supposons que $\forall n \in \mathbb{N}, u_n \preccurlyeq u_{n+1}$.

Soit $m \in \mathbb{N}$.

Pour tout $n \in \mathbb{N}$, posons $P(n)$ l'assertion $m \leq n \implies u_m \preccurlyeq u_n$.

Montrons par récurrence que pour tout $n \in \mathbb{N}$, on a $P(n)$.

Initialisation

► Plaçons-nous dans le cas où $m = 0$.

Dans ce cas-là on a $u_m = u_0$ donc $u_m \preccurlyeq u_0$ par réflexivité de \preccurlyeq .

L'implication $m \leq 0 \implies u_m \preccurlyeq u_0$ est donc vraie.

► Plaçons-nous dans le cas où $m \neq 0$.

L'implication $m \leq 0 \implies u_m \preccurlyeq u_0$ est alors vraie car sa prémissse est fausse.

Dans les deux cas, on a $P(0)$.

Héritéité

Soit $n \in \mathbb{N}$ tel que $P(n)$.

Supposons que $m \leq n + 1$.

On a donc $m < n + 1$ ou $m = n + 1$.

► Plaçons-nous dans le cas où $m < n + 1$.

On a alors $m \leq n$ d'après la proposition 13 page 33.

On a donc $u_m \preccurlyeq u_n$ par $P(n)$.

Or on a $u_n \preccurlyeq u_{n+1}$ par hypothèse.

On a donc $u_m \preccurlyeq u_{n+1}$ par transitivité de \preccurlyeq .

► Plaçons-nous dans le cas où $m = n + 1$.

On a alors $u_m = u_{n+1}$ donc $u_m \preccurlyeq u_{n+1}$ par réflexivité de \preccurlyeq .

Dans les deux cas on a $u_m \preccurlyeq u_{n+1}$.

Donc si $m \leq n + 1$ alors $u_m \preccurlyeq u_{n+1}$.

On a donc $P(n + 1)$.

Donc pour tout $n \in \mathbb{N}$, on a $P(n) \implies P(n + 1)$.

D'après le principe d'induction sur les entiers naturels, on a donc $\forall n \in \mathbb{N}, P(n)$.

Autrement dit, $\forall n \in \mathbb{N}, (m \leq n \Rightarrow u_m \preccurlyeq u_n)$.

Donc $\forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m \leq n \Rightarrow u_m \preceq u_n)$.

Donc $(u_n)_{n \in \mathbb{N}}$ est croissante.

2. \Rightarrow

Supposons que $(u_n)_{n \in \mathbb{N}}$ est strictement croissante.

Soit $n \in \mathbb{N}$.

On sait que $n < n + 1$ d'après la proposition 13 page 33.

On a donc $u_n \prec u_{n+1}$ par stricte croissance de u .

Donc $\forall n \in \mathbb{N}, u_n \prec u_{n+1}$.

\Leftarrow

Supposons que $\forall n \in \mathbb{N}, u_n \prec u_{n+1}$.

Soit $m \in \mathbb{N}$.

Pour tout $n \in \mathbb{N}$, posons $P(n)$ l'assertion $m < n \implies u_m \prec u_n$.

Montrons par récurrence que pour tout $n \in \mathbb{N}$, on a $P(n)$.

Initialisation

L'implication $m < 0 \implies u_m \prec u_0$ est vraie car sa prémissse est fausse.

On a donc $P(0)$.

Héritéité

Soit $n \in \mathbb{N}$ tel qu'on a $P(n)$.

Supposons que $m < n + 1$.

On a donc $m \leq n$ d'après la proposition 13 page 33.

On a donc $m < n$ ou $m = n$.

► Plaçons-nous dans le cas où $m < n$.

On a alors $u_m \prec u_n$ d'après $P(n)$.

Or on a $u_n \prec u_{n+1}$ par hypothèse.

On a donc $u_m \prec u_{n+1}$ par transitivité de \prec .

► Plaçons-nous dans le cas où $m = n$.

On a donc $u_m = u_n$.

Or on a $u_n \prec u_{n+1}$ par hypothèse.

On a donc $u_m \prec u_{n+1}$.

Dans les deux cas on a $u_m \prec u_{n+1}$.

Donc si $m < n + 1$ alors $u_m \prec u_{n+1}$.

On a donc $P(n + 1)$.

Donc pour tout $n \in \mathbb{N}$, on a $P(n) \implies P(n+1)$.

D'après le principe d'induction sur les entiers naturels, on a donc $\forall n \in \mathbb{N}, P(n)$.

Autrement dit $\forall n \in \mathbb{N}, (m < n \implies u_m \prec u_n)$.

Donc $\forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m < n \implies u_m \prec u_n)$.

Donc $\boxed{(u_n)_{n \in \mathbb{N}} \text{ est strictement croissante}}$.

3. Considérons la relation \succsim symétrique de \preccurlyeq , c'est-à-dire que pour tout x et y dans E , on a

$$x \succsim y \iff y \preccurlyeq x$$

On peut montrer que \succsim est une relation d'ordre sur E .

On a alors les équivalences :

$$\begin{aligned} (u_n)_{n \in \mathbb{N}} \text{ est décroissante pour } \preccurlyeq &\iff \forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m \leq n \implies u_n \preccurlyeq u_m) \\ &\iff \forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m \leq n \implies u_m \succsim u_n) \\ &\iff (u_n)_{n \in \mathbb{N}} \text{ est croissante pour } \succsim \\ &\iff \forall n \in \mathbb{N}, u_n \succsim u_{n+1} \text{ d'après 1.} \\ &\iff \forall n \in \mathbb{N}, u_{n+1} \preccurlyeq u_n \end{aligned}$$

D'où $\boxed{(u_n)_{n \in \mathbb{N}} \text{ est décroissante si et seulement si } \forall n \in \mathbb{N}, u_{n+1} \preccurlyeq u_n}$.

4. Considérons la relation \succ symétrique de \prec , c'est-à-dire que pour tout x et y dans E , on a

$$x \succ y \iff y \prec x$$

On peut montrer que \succ est une relation d'ordre strict sur E .

On a alors les équivalences :

$$\begin{aligned} (u_n)_{n \in \mathbb{N}} \text{ est strictement décroissante pour } \prec &\iff \forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m < n \implies u_n \prec u_m) \\ &\iff \forall m \in \mathbb{N}, \forall n \in \mathbb{N}, (m < n \implies u_m \succ u_n) \\ &\iff (u_n)_{n \in \mathbb{N}} \text{ est strictement croissante pour } \succ \\ &\iff \forall n \in \mathbb{N}, u_n \succ u_{n+1} \text{ d'après 2.} \\ &\iff \forall n \in \mathbb{N}, u_{n+1} \prec u_n \end{aligned}$$

D'où $\boxed{(u_n)_{n \in \mathbb{N}} \text{ est strictement décroissante si et seulement si } \forall n \in \mathbb{N}, u_{n+1} \prec u_n}$.

5. On a les équivalences suivantes :

$$\begin{aligned}
 (u_n)_{n \in \mathbb{N}} \text{ est constante} &\iff \forall n \in \mathbb{N}, \forall m \in \mathbb{N}, u_n = u_m \\
 &\iff \forall n \in \mathbb{N}, \forall m \in \mathbb{N}, (n \leq m \Rightarrow u_n = u_m) \\
 &\iff \forall n \in \mathbb{N}, \forall m \in \mathbb{N}, (n \leq m \Rightarrow (u_n \preccurlyeq u_m \text{ et } u_m \preccurlyeq u_n)) \\
 &\iff (u_n)_{n \in \mathbb{N}} \text{ est croissante et décroissante} \\
 &\iff (\forall n \in \mathbb{N}, u_n \preccurlyeq u_{n+1}) \text{ et } (\forall n \in \mathbb{N}, u_{n+1} \preccurlyeq u_n) \text{ d'après 1. et 3.} \\
 &\iff \forall n \in \mathbb{N}, (u_n \preccurlyeq u_{n+1} \text{ et } u_{n+1} \preccurlyeq u_n) \\
 &\iff \forall n \in \mathbb{N}, u_n = u_{n+1} \text{ par antisymétrie et réflexivité de } \preccurlyeq
 \end{aligned}$$

D'où $(u_n)_{n \in \mathbb{N}}$ est constante si et seulement si $\forall n \in \mathbb{N}, u_n = u_{n+1}$.

6. \Rightarrow

Supposons que $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

Il existe donc $n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0, u_n = u_{n_0}$.

Soit $n \geq n_0$.

On a donc $u_{n_0} = u_n$ par hypothèse.

Comme $n_0 \leq n$ on a aussi $n_0 < n + 1$ d'après la proposition 13 page 33.

En particulier $n_0 \leq n + 1$ donc $u_{n_0} = u_{n+1}$ par hypothèse.

On a donc $u_n = u_{n+1}$.

Donc $\forall n \geq n_0, u_n = u_{n+1}$.

\Leftarrow

Supposons qu'il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0, u_n = u_{n+1}$.

Pour tout $n \in \mathbb{N}$, posons $P(n)$ l'assertion $n_0 \leq n \implies u_{n_0} = u_n$.

Montrons par récurrence que pour tout $n \in \mathbb{N}$ on a $P(n)$.

Initialisation

Supposons que $n_0 \leq 0$.

On a donc $n_0 \subseteq 0$ par définition de \leq .

On a donc $n_0 = 0$ puisque par définition $0 = \emptyset$.

Donc $u_{n_0} = u_0$.

On a donc $n_0 \leq 0 \implies u_{n_0} = u_0$, c'est-à-dire $P(0)$.

Hérédité

Soit $n \in \mathbb{N}$ tel que $P(n)$.

Supposons que $n_0 \leq n + 1$.

On a donc $n_0 < n + 1$ ou $n_0 = n + 1$.

► Plaçons-nous dans le cas où $n_0 < n + 1$.

On a donc $n_0 \leq n$ d'après la proposition 13 page 33.

On a donc $u_{n_0} = u_n$ d'après $P(n)$.

Or on a $u_n = u_{n+1}$ par hypothèse.

Donc $u_{n_0} = u_{n+1}$.

► Plaçons-nous dans le cas où $n_0 = n + 1$.

On donc $u_{n_0} = u_{n+1}$.

Dans les deux cas, on a donc $u_{n_0} = u_{n+1}$.

Donc si $n_0 \leq n + 1$ alors $u_{n_0} = u_{n+1}$, donc on a $P(n + 1)$.

Donc pour tout $n \in \mathbb{N}$, si $P(n)$ alors $P(n + 1)$.

D'après le principe d'induction chez les entiers, on a donc $\forall n \in \mathbb{N}, P(n)$.

Autrement dit $\forall n \in \mathbb{N}, (n_0 \leq n \Rightarrow u_n = u_{n_0})$.

Dit encore autrement, $\forall n \geq n_0, u_n = u_{n_0}$.

Donc $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

CQFD.

Observons à présent un phénomène qui peut parfois se produire : pour certains ensembles ordonnés, une suite décroissante est forcément stationnaire. Autrement dit, c'est comme si toute suite décroissante était forcée de s'arrêter à un moment. On parle de la **condition de la chaîne descendante**.

Définition 16 (Condition de la chaîne descendante)

Soit E un ensemble ordonné.

On dit que E vérifie la **condition de la chaîne descendante** si et seulement si pour toute suite $(u_n)_{n \in \mathbb{N}}$ à valeurs dans E , si $(u_n)_{n \in \mathbb{N}}$ est décroissante alors $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

Venons-en à une caractérisation très puissante des ensembles bien ordonnés : parmi les ensembles totalement ordonnés, ce sont précisément ceux qui vérifient la condition de la chaîne descendante.

Proposition 34 (Cond. de la chaîne descendante et bon ordre)

Soit E un ensemble ordonné.

Les assertions suivantes sont équivalentes :

1. E est bien ordonné.
2. E est totalement ordonné et vérifie la condition de la chaîne descendante.

 *Démonstration*

Notons \preccurlyeq la relation d'ordre sur E et \prec l'ordre strict associé.



Supposons que E est bien ordonné.

On sait déjà que E est totalement ordonné d'après la proposition 2 page 10.

Montrons que E vérifie la condition de la chaîne descendante.

Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans E .

Supposons que $(u_n)_{n \in \mathbb{N}}$ est décroissante.

Considérons $X := \{u_n \mid n \in \mathbb{N}\}$, qui est donc une partie non vide de E .

Or E est **bien ordonné** par hypothèse.

Donc X admet un minimum x : on a donc $\forall n \in \mathbb{N}, x \preccurlyeq u_n$.

Mais on sait aussi que $x \in X$ donc il existe $m \in \mathbb{N}$ tel que $x = u_m$.

Ainsi pour tout $n \geq m$, on a donc $u_m \preccurlyeq u_n$ par ce qui précède.

Mais on sait aussi que $(u_n)_{n \in \mathbb{N}}$ est décroissante.

Donc pour tout $n \geq m$, on a $u_n \preccurlyeq u_m$ et donc $u_n = u_m$ par antisymétrie de \preccurlyeq .

Ainsi $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

Donc si $(u_n)_{n \in \mathbb{N}}$ est décroissante alors $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

Donc E vérifie la condition de la chaîne descendante.



Supposons que E est totalement ordonné et vérifie la condition de la chaîne descendante.

Supposons par l'absurde que E n'est pas bien ordonné.

Il existe donc une partie X de E qui est non vide mais n'admet pas de minimum.

Pour tout $x \in X$, posons $X_{\prec x}$ l'ensemble $\{y \in X \mid y \prec x\}$.

Supposons par l'absurde qu'il existe $x \in X$ tel que $X_{\prec x}$ est vide.

Autrement dit pour tout $y \in X$, on n'a pas $y \prec x$.

Or E est **totalement ordonné** par hypothèse donc pour tout $y \in X$ on a $x \preccurlyeq y$.

Autrement dit x est le minimum de X , ce qui est absurde car X n'en a pas.

Ainsi pour tout $x \in X$, on a $X_{\prec x} \neq \emptyset$.

D'après l'**axiome du choix**, il existe une application $g : \mathcal{P}(X) \setminus \{\emptyset\} \longrightarrow ?$ telle que $\forall A \in \mathcal{P}(X) \setminus \{\emptyset\}, g(A) \in A$.

En particulier pour tout $x \in X$, on a $g(X_{\prec x}) \in X_{\prec x}$ donc $g(X_{\prec x}) \prec x$.

Pour tout $x \in X$, posons $f(x) := g(X_{\prec x})$ de sorte que $f(x) \prec x$.

On a dit que X est non vide : soit $a \in X$.

Considérons alors la suite $(u_n)_{n \in \mathbb{N}}$ définie par récurrence par

$$\begin{cases} u_0 := a \\ \forall n \in \mathbb{N}, u_{n+1} := f(u_n) \end{cases}$$

Ainsi $(u_n)_{n \in \mathbb{N}}$ est une suite à valeurs dans X donc dans E car $X \subseteq E$.

De plus on a $\forall n \in \mathbb{N}, u_{n+1} = f(u_n) \prec u_n$ donc $(u_n)_{n \in \mathbb{N}}$ est strictement décroissante.

En particulier $(u_n)_{n \in \mathbb{N}}$ est décroissante.

Or E vérifie **la condition de la chaîne descendante** donc $(u_n)_{n \in \mathbb{N}}$ est stationnaire.

C'est absurde puisqu'on vient de dire qu'elle est strictement décroissante.

Par l'absurde, on vient de montrer que E est bien ordonné.

CQFD.

Pour la petite histoire



Emmy Noether (23 mars 1882 – 14 avril 1935) est une mathématicienne allemande spécialiste d'algèbre abstraite et de physique théorique. Elle a révolutionné les théories des anneaux, des corps et des algèbres, notamment en développant la notion d'idéal. En physique, le théorème de Noether explique le lien fondamental entre la symétrie et les lois de conservation et est considéré comme aussi important que la théorie de la relativité.

Noether introduit la condition de la chaîne descendante dans son article de 1921 *Idealtheorie in Ringbereichen* mais précise que ce concept avait déjà été introduit précédemment par Dedekind (dans le cas des corps de nombres) et par Lasker (dans le cas des polynômes). Elle est la première à l'introduire dans un cadre aussi général que celui de son article : celui des anneaux commutatifs dont chaque idéal est finiment engendré.

Bien souvent, on souhaite itérer une application, c'est-à-dire considérer la suite $x, f(x), f(f(x)), f(f(f(x)))$, et ainsi de suite : on prend un élément, on le donne à manger à f , on redonne le résultat à manger à f , et ainsi de suite autant de fois que nécessaire. Cela conduit à la définition suivante, qui est donc tout naturellement récursive.

Définition 17 (Itérées d'une application)

Soient E un ensemble et $f : E \rightarrow E$.

On pose alors par récursion :

$$\begin{cases} f^0 := \text{id}_E \\ f^{n+1} := f^n \circ f \end{cases}$$

Pour tout $n \in \mathbb{N}$, on dit que f^n est la $n^{\text{ème}}$ **itérée** de f .

Exemple :

- On peut montrer que pour tout $n \in \mathbb{N}$, on a $\text{id}_E^n = \text{id}_E$.
- On a $f^1 = f^{0+1} = f^0 \circ f = \text{id}_E \circ f = f$.
- On a $f^2 = f^{1+1} = f^1 \circ f = f \circ f$.

Proposition 35 (Itérées d'une application injective)

Soient E un ensemble et $f : E \rightarrow E$.

Si f est injective alors pour tout $n \in \mathbb{N}$, f^n est injective.



Démonstration

Supposons que f est injective.

Pour tout $n \in \mathbb{N}$, on pose $P(n)$ l'assertion « f^n est injective ».

Initialisation

$f^0 = \text{id}_E$ par définition, qui est injective, donc on a $P(0)$.

Hérédité

Soit $n \in \mathbb{N}$ tel qu'on a $P(n)$, c'est-à-dire que f^n est injective.

Par hypothèse f est injective, si bien que $f^n \circ f$ est injective.

Or $f^{n+1} = f^n \circ f$, donc f^{n+1} est injective, et donc on a $P(n + 1)$.

Ainsi pour tout $n \in \mathbb{N}$, si on a $P(n)$ alors on a $P(n + 1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout $n \in \mathbb{N}$, on a $P(n)$, c'est-à-dire que f^n est injective.

CQFD.

Chapitre 2

Opérations sur les ordinaux



Note de l'auteur

Nous avons défini et étudié les ordinaux lors du chapitre précédent, en se munissant de toute une batterie d'outils pour cela. Il est temps maintenant d'agir sur ceux-ci via l'introduction d'opérations entre ordinaux. Nous allons voir comment les additionner, les multiplier et les éléver à une puissance. L'outil de récursion développé à la fin du chapitre précédent nous sera pour cela d'une grande aide.

Notez que ce chapitre est un peu sec : il va consister concrètement à énumérer plein de propriétés sur les ordinaux et à les démontrer principalement avec des récurrences. Ce n'est pas forcément la grande joie. L'auteur conseille au lecteur de passer la plupart des démonstrations.

Sommaire

1	Généralités	90
2	Addition d'ordinaux	94
2.1	Définition et propriétés	94
2.2	Interprétation graphique : la concaténation	111
3	Multiplication d'ordinaux	126
3.1	Définition et propriétés	126
3.2	Interprétation graphique : le produit cartésien	141
4	Exponentiation d'ordinaux	148
4.1	Définition et propriétés	148
4.2	Applications à support fini	159
5	Forme normale de Cantor et ε_0	174
5.1	Logarithme ordinal et forme normale de Cantor	174
5.2	L'ordinal ε_0 et la classe des points fixes	180

1 Généralités

Si nous avons déployé l’artillerie lourde avec la notion d’assertion fonctionnelle inductive, c’est pour avoir les mains libres au moment de la définition de trois opérations importantes chez les ordinaux : l’addition, la multiplication et l’exponentiation. Prenons pour exemple l’addition des ordinaux : nous aimerais donner du sens à l’addition $\alpha + \beta$ pour α et β deux ordinaux. On peut pour cela s’inspirer de l’addition chez les entiers naturels.

Comment allons-nous définir l’addition $2 + 7$ par exemple ? On considère que $2 + 6$ a déjà été définie et on pose simplement que $2 + 7$ est l’entier qui vient juste après $2 + 6$, c’est-à-dire $2 + 7 := S(2 + 6)$. Autrement dit, on pose $2 + S(6) := S(2 + 6)$. Pour définir l’addition de 2 par tous les entiers, on le fait simplement de la manière suivante en initialisant la valeur en 0 :

$$\begin{cases} 2 + 0 := 2 \\ 2 + S(m) := S(2 + m) \text{ pour tout entier naturel } m \end{cases}$$

Pour donner du sens à $2 + S(m)$, on considère que $2 + m$ a déjà du sens : c’est bien là une récursion. Cependant, comme ω n’est pas un successeur, quel sens donner alors à $2 + \omega$? On va simplement dire que c’est l’ordinal qui vient juste après tous les $2 + n$ avec $n < \omega$. Autrement dit, on va dire que $2 + \omega$ est la borne supérieure de l’ensemble $\{2 + n \mid n < \omega\}$. Plus généralement, on va poser

$$\begin{cases} 2 + 0 := 2 \\ 2 + S(\beta) := S(2 + \beta) \text{ pour tout ordinal } \beta \\ 2 + \gamma := \sup_{\delta < \gamma} (2 + \delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Encore plus généralement, pour α un ordinal quelconque fixé à l’avance, on va poser

$$\begin{cases} \alpha + 0 := \alpha \\ \alpha + S(\beta) := S(\alpha + \beta) \text{ pour tout ordinal } \beta \\ \alpha + \gamma := \sup_{\delta < \gamma} (\alpha + \delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Comment allons-nous procéder pour s’assurer qu’il s’agit d’une définition rigoureuse ? On souhaite en fait définir par récursion une assertion fonctionnelle F_α vérifiant :

$$\begin{cases} F_\alpha(0) := \alpha \\ F_\alpha(S(\beta)) := S(F_\alpha(\beta)) \text{ pour tout ordinal } \beta \\ F_\alpha(\gamma) := \sup_{\delta < \gamma} F_\alpha(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

et on pose alors $\alpha + \beta := F_\alpha(\beta)$. Pour pouvoir justifier proprement qu’une telle construction est possible, et pouvoir de même définir la multiplication et l’exponentiation, énonçons la proposition suivante. L’ordinal μ_0 joue le rôle du résultat de l’initialisation, et l’assertion fonctionnelle G est là pour généraliser S .

Proposition 36 (Justification des opérations sur les ordinaux)

Soient μ_0 un ordinal et $G : ON \longrightarrow ON$ une assertion fonctionnelle.
Alors il existe une unique assertion fonctionnelle $F : ON \longrightarrow ON$ telle que

$$\begin{cases} F(0) = \mu_0 \\ F(S(\beta)) = G(F(\beta)) \text{ pour tout ordinal } \beta \\ F(\gamma) := \sup_{\delta < \gamma} F(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Démonstration

Existence

Pour toute application f telle que $\text{dom}(f)$ est un ordinal et tel que $\text{im}(f) \subseteq ON$, on pose

$$\begin{cases} H(f) := \mu_0 & \text{si } \text{dom}(f) = 0 \\ H(f) := G(f(\beta)) & \text{si } \text{dom}(f) = S(\beta) \text{ avec } \beta \text{ un ordinal} \\ H(f) := \sup_{\delta < \gamma} f(\delta) & \text{si } \text{dom}(f) = \gamma \text{ est un ordinal limite non nul} \end{cases}$$

On obtient alors H une assertion fonctionnelle.

Remarquons que par définition de H , pour tout ensemble f on a

$$f \in \text{dom}(H) \iff f \text{ est une application telle que } \text{dom}(f) \in ON \text{ et } \text{im}(f) \subseteq ON$$

- Montrons que $\text{im}(H) \subseteq ON$.

Soit $y \in \text{im}(H)$.

Il existe donc $f \in \text{dom}(H)$ tel que $y = H(f)$.

Par définition de H , f est une application telle que $\text{dom}(f) \in ON$ et $\text{im}(f) \subseteq ON$.

- Si $\text{dom}(f) = 0$ alors $y = H(f) = \mu_0 \in ON$.
- Si $\text{dom}(f) = S(\beta)$ avec β un ordinal, alors $y = H(f) = G(f(\beta)) \in \text{im}(G)$.
Or par définition $G : ON \longrightarrow ON$ donc $\text{im}(G) \subseteq ON$ et donc $y \in ON$.
- Si $\text{dom}(f)$ est un ordinal limite non nul γ alors $y = H(f) = \sup_{\delta < \gamma} f(\delta)$.

Or $\text{im}(f) \subseteq ON$ donc $\text{im}(f) = \{f(\delta) \mid \delta \in \gamma\}$ est un ensemble d'ordinaux.

Donc $\{f(\delta) \mid \delta < \gamma\}$ est un ensemble d'ordinaux par définition de $<$.

Donc y sa borne supérieure est un ordinal.

Dans tous les cas, on a bien $y \in ON$.

Donc $\forall y \in \text{im}(H), y \in ON$, si bien que $\text{im}(H) \subseteq ON$.

- Montrons que H vérifie l'hypothèse du théorème 6 page 71.

Soient α un ordinal et $f : \alpha \longrightarrow ?$ une application H -inductive.

On sait déjà que $\text{dom}(f) = \alpha$ est un ordinal.

Il suffit donc de montrer que $\text{im}(f) \subseteq ON$.

Soit $y \in \text{im}(f)$.

Il existe donc $\beta \in \alpha$ tel que $y = f(\beta)$.

Or f est H -inductive par définition.

On a donc $f|_{\beta} \in \text{dom}(H)$ et $y = f(\beta) = H(f|_{\beta}) \in \text{im}(H)$.

Or on a dit que $\text{im}(H) \subseteq ON$ donc $y \in ON$.

On a donc $\forall y \in \text{im}(f), y \in ON$ donc $\text{im}(f) \subseteq ON$.

Ainsi on a $\text{dom}(f) \in ON$ et $\text{im}(f) \subseteq ON$ donc $f \in \text{dom}(H)$.

Ainsi pour tout $\alpha \in ON$ et toute application $f : \alpha \longrightarrow ?,$ si f est H -inductive alors $f \in \text{dom}(H)$. Donc H et ON vérifient l'hypothèse du théorème 6 page 71.

- Il existe donc une unique assertion fonctionnelle $F : ON \longrightarrow ?$ qui est H -inductive.

Par définition de la H -inductivité et par définition de H , on a alors

$$\left\{ \begin{array}{l} F(0) = H(F|_0) = \mu_0 \\ F(S(\beta)) = H(F|_{S(\beta)}) = G(F|_{S(\beta)}(\beta)) = G(F(\beta)) \text{ pour tout ordinal } \beta \\ F(\gamma) = H(F|_{\gamma}) = \sup_{\delta < \gamma} F|_{\gamma}(\delta) = \sup_{\delta < \gamma} F(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

Ainsi F vérifie les conditions de l'énoncé.

Remarquons que comme $\text{im}(H) \subseteq ON$ et comme par définition on a $\text{im}(F) \subseteq \text{im}(H)$, on a donc $\text{im}(F) \subseteq ON$ et donc on a bien $F : ON \longrightarrow ON$.

Unicité

Soit $F' : ON \longrightarrow ON$ une assertion fonctionnelle vérifiant les conditions de l'énoncé.

Montrons par induction transfinie que $F = F'$.

Considérons alors l'assertion à paramètres P définie pour tout ordinal α par

$$P(\alpha) \iff F(\alpha) = F'(\alpha)$$

► Initialisation

On a $F(0) = \mu_0 = F'(0)$ donc $P(0)$ est vraie.

► Héritéité

Soit α un ordinal tel que $P(\alpha)$.

On a donc $F(\alpha) = F'(\alpha)$.

Donc $F(S(\alpha)) = G(F(\alpha)) = G(F'(\alpha)) = F'(S(\alpha))$.

On a donc $P(S(\alpha))$.

Donc pour tout ordinal α , on a $P(\alpha) \implies P(S(\alpha))$.

► *Héritage limite*

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

On a donc $\forall \beta < \alpha, F(\beta) = F'(\beta)$.

Donc $F(\alpha) = \sup_{\beta < \alpha} F(\beta) = \sup_{\beta < \alpha} F'(\beta) = F'(\alpha)$.

On a donc $P(\alpha)$.

Donc pour tout ordinal limite non nul α , on a $(\forall \beta < \alpha, P(\beta)) \implies P(\alpha)$.

Ainsi P vérifie les trois hypothèses du principe faible d'induction transfinie.

Donc pour tout ordinal α , on a $P(\alpha)$, c'est-à-dire $F(\alpha) = F'(\alpha)$.

Ainsi on a $F = F'$, d'où l'unicité.

CQFD.

2 Addition d'ordinaux

2.1 Définition et propriétés

Nous pouvons enfin définir l'addition sur les ordinaux.

Définition 18 (Addition d'ordinaux)

Soit α un ordinal.

On pose

$$\left\{ \begin{array}{l} \alpha + 0 := \alpha \\ \alpha + S(\beta) := S(\alpha + \beta) \text{ pour tout ordinal } \beta \\ \alpha + \gamma := \sup_{\delta < \gamma} (\alpha + \delta) \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

Remarque :

Pour justifier proprement que cette définition a du sens, on utilise simplement la proposition 36 page 91 qui précède, en posant $\mu_0 := \alpha$ et $G(\xi) := S(\xi)$ pour tout ordinal ξ . La proposition nous donne alors une assertion fonctionnelle F_α telle que

$$\left\{ \begin{array}{l} F_\alpha(0) := \alpha \\ F_\alpha(S(\beta)) := S(F_\alpha(\beta)) \text{ pour tout ordinal } \beta \\ F_\alpha(\gamma) := \sup_{\delta < \gamma} F_\alpha(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

et on pose alors $\alpha + \beta := F_\alpha(\beta)$ pour tout ordinal β .

Exemple :

1. Nous allons voir juste après que pour tout ordinal α , on a $\alpha + 1 = S(\alpha)$.
Ainsi par exemple $\omega + 1$ est le successeur de ω .
2. On a

$$3 + \omega = \sup_{n < \omega} (3 + n) = \sup\{3 + 0, 3 + 1, 3 + 2, \dots\} = \sup\{3, 4, 5, 6, \dots\} = \omega$$

En fait plus généralement pour tout entier naturel m , on a

$$m + \omega = \sup\{m, m + 1, m + 2, \dots\} = \omega$$

En particulier $1 + \omega = \omega$. Et pourtant, nous venons de dire que $\omega + 1 = S(\omega)$, si bien que $1 + \omega < \omega + 1$ et donc $1 + \omega \neq \omega + 1$. Et oui, l'addition des ordinaux n'est pas commutative en général. Heureusement nous verrons qu'elle l'est quand on se restreint aux entiers naturels !

Nous affirmons depuis des pages et des pages que pour un entier naturel n , on va définir $n + 1$ comme étant $S(n)$. En fait comme n et 1 sont des cas particuliers d'ordinaux, on a déjà donné du sens à $n + 1$, et on retombe bien sur $S(n)$. Plus généralement, on a la proposition suivante.

Proposition 37 (Successeur et plus un)

Soit α un ordinal.

On a $S(\alpha) = \alpha + 1$.



Démonstration

On a les égalités suivantes :

$$\begin{aligned}\alpha + 1 &= \alpha + S(0) \text{ par définition de 1} \\ &= S(\alpha + 0) \text{ par définition de l'addition} \\ &= S(\alpha) \text{ puisque } \alpha + 0 = \alpha\end{aligned}$$

On a donc $\boxed{\alpha + 1 = S(\alpha)}$

CQFD.

Remarque :

De plus en plus, nous remplacerons la notation $S(\alpha)$ par $\alpha + 1$ car cela coïncide plus facilement avec notre intuition. Par exemple dans la définition de l'addition, l'étape intermédiaire peut se réécrire

$$\alpha + (\beta + 1) = (\alpha + \beta) + 1 \text{ pour tout ordinal } \beta$$

De même, la condition d'hérédité du principe faible d'induction peut se réécrire

Pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

Cependant, nous allons continuer pour l'instant à utiliser $S(\alpha)$ car nous travaillons encore sur l'addition et que des confusions entre définition et propriétés risquent d'émerger. Le changement de notation opérera surtout à partir de la multiplication.

Proposition 38 (0 est neutre pour l'addition des ordinaux)

Pour tout ordinal α , on a $\alpha + 0 = \alpha = 0 + \alpha$.

On dit que 0 est **neutre** pour l'addition des ordinaux.



Démonstration

Par définition de l'addition, on sait déjà que pour tout ordinal α on a $\alpha + 0 = \alpha$.

Montrons l'autre égalité par induction.

Considérons P l'assertion à paramètre définie pour tout ordinal α par

$$P(\alpha) \iff 0 + \alpha = \alpha$$

► *Initialisation*

Par définition de l'addition on a $0 + 0 = 0$ donc on a $P(0)$.

► *Héritage*

Soit α un ordinal tel que $P(\alpha)$.

Autrement dit on a $0 + \alpha = \alpha$.

On a alors par définition de l'addition $0 + S(\alpha) = S(0 + \alpha) = S(\alpha)$.

On a donc $P(S(\alpha))$.

Donc pour tout ordinal α , si $P(\alpha)$ alors $P(S(\alpha))$.

► *Héritage limite*

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

Autrement dit pour tout $\beta < \alpha$, on a $0 + \beta = \beta$.

On a alors par définition de l'addition $0 + \alpha = \sup_{\beta < \alpha} (0 + \beta) = \sup_{\beta < \alpha} \beta = \sup_{\beta \in \alpha} \beta$.

Or α est un ordinal limite donc $\sup_{\beta \in \alpha} \beta = \sup(\alpha) = \alpha$ d'après la prop. 21 p. 47.

On a donc $0 + \alpha = \alpha$ et donc $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $P(\alpha)$.

Autrement dit pour tout ordinal α on a $[0 + \alpha = \alpha]$.

CQFD.

On vient de définir une addition sur tous les ordinaux, et donc en particulier sur tous les entiers naturels. Comme on l'a vu en introduction de ce chapitre, il s'agit normalement de l'addition avec laquelle on est familière dans la vie de tous les jours. Une des règles de cette addition est évidemment qu'ajouter deux entiers naturels donne toujours un entier naturel.

Proposition 39 (Addition de deux entiers naturels)

Pour tout entiers naturels n et m , l'ordinal $n + m$ est un entier naturel.

On dit que $\mathbb{N} = \omega$ est **stable par addition**.

 *Démonstration*

Fixons n un entier naturel.

Soit P l'assertion à paramètre définie pour tout entier naturel m par

$$P(m) \iff n + m \in \mathbb{N}$$

Raisonnons par induction sur les entiers naturels.

► *Initialisation*

Par définition de l'addition on a $n + 0 = n$.

Or n est un entier naturel par définition.

Donc $n + 0$ est un entier naturel et donc $P(0)$.

► *Héritéité*

Soit m un entier naturel tel que $P(m)$.

Autrement dit $n + m$ est un entier naturel.

Donc $S(n + m)$ est un entier naturel d'après la proposition 15 page 38.

Or $n + S(m) = S(n + m)$ par définition de l'addition.

Donc $n + S(m)$ est un entier naturel : autrement dit on a $P(S(m))$.

Donc pour tout entier naturel m , si $P(m)$ alors $P(S(m))$.

Ainsi P vérifie les deux conditions de l'induction chez les entiers naturels.

Donc pour tout entier naturel m , on a $P(m)$.

Autrement dit, pour tout entier naturel m , $n + m$ est un entier naturel.

CQFD.

Proposition 40 (Croissance de l'addition des ordinaux)

Soient α , β et γ trois ordinaux.

1. Si $\beta < \gamma$ alors $\alpha + \beta < \alpha + \gamma$.

On dit que l'addition à gauche est **strictement croissante**.

2. Si $\beta \leq \gamma$ alors $\alpha + \beta \leq \alpha + \gamma$.

On dit que l'addition à gauche est **croissante**.

3. Si $\beta \leq \gamma$ alors $\beta + \alpha \leq \gamma + \alpha$.

On dit que l'addition à droite est **croissante**.



Démonstration

1. Fixons α et β .

Posons $P_{\alpha,\beta}$ l'assertion à paramètre définie pour tout ordinal γ par

$$P_{\alpha,\beta}(\gamma) \iff (\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma)$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

Il est toujours faux de dire $\beta \in \emptyset$ donc il est toujours faux de dire $\beta \in 0$.

Autrement dit la prémissse $\beta < 0$ est fausse, donc on a l'implication $\beta < 0 \Rightarrow \alpha + \beta < \alpha + 0$.

Ainsi on a $P_{\alpha,\beta}(0)$.

► *Héritéité*

Soit γ un ordinal tel que $P_{\alpha,\beta}(\gamma)$.

Ainsi on a $\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$.

Supposons que $\beta < S(\gamma)$.

On a alors $\beta \leq \gamma$ d'après la proposition 13 page 33, donc $\beta < \gamma$ ou $\beta = \gamma$.

- Plaçons-nous dans le cas où $\beta < \gamma$.

On a alors $\alpha + \beta < \alpha + \gamma$ d'après $P_{\alpha,\beta}(\gamma)$.

Or on a $\alpha + \gamma < S(\alpha + \gamma)$ d'après la proposition 13 page 33.

On a donc $\alpha + \beta < S(\alpha + \gamma)$ par transitivité de $<$.

- Plaçons-nous dans le cas où $\beta = \gamma$.

On a donc $\alpha + \beta = \alpha + \gamma$.

Or on a $\alpha + \gamma < S(\alpha + \gamma)$ d'après la proposition 13 page 33.

On a donc $\alpha + \beta < S(\alpha + \gamma)$.

Ainsi dans les deux cas on a $\alpha + \beta < S(\alpha + \gamma)$.

Or par définition de l'addition on a $S(\alpha + \gamma) = \alpha + S(\gamma)$.

On a donc $\alpha + \beta < \alpha + S(\gamma)$.

Donc si $\beta < S(\gamma)$ alors $\alpha + \beta < \alpha + S(\gamma)$.

Ainsi on a $P_{\alpha,\beta}(S(\gamma))$.

Donc pour tout ordinal γ on a $P_{\alpha,\beta}(\gamma) \implies P_{\alpha,\beta}(S(\gamma))$.

► *Héritéité limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

Supposons que $\beta < \gamma$.

On a alors $S(\beta) < \gamma$ d'après la proposition 14 page 37 car γ est limite.

On a donc $\alpha + S(\beta) \leq \sup_{\delta < \gamma} (\alpha + \delta)$ par définition de la borne supérieure.

Comme γ est limite non nul, on a $\alpha + \gamma = \sup_{\delta < \gamma} (\alpha + \delta)$ et donc $\alpha + S(\beta) \leq \alpha + \gamma$.

De plus par hypothèse on a $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

Or on a dit que $S(\beta) < \gamma$ donc $P_{\alpha,\beta}(S(\beta))$.

Autrement dit on a $\beta < S(\beta) \implies \alpha + \beta < \alpha + S(\beta)$.

Or on a $\beta < S(\beta)$ d'après la proposition 13 page 33.

On a donc $\alpha + \beta < \alpha + S(\beta)$ par modus ponens.

Ainsi on a $\alpha + \beta < \alpha + S(\beta) \leq \alpha + \gamma$.

On a donc $\alpha + \beta < \alpha + \gamma$.

Donc si $\beta < \gamma$ alors $\alpha + \beta < \alpha + \gamma$.

Autrement dit on a $P_{\alpha,\beta}(\gamma)$.

Donc pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$ alors $P_{\alpha,\beta}(\gamma)$.

Ainsi $P_{\alpha,\beta}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $P_{\alpha,\beta}(\gamma)$.

Autrement dit, pour tout ordinal γ on a $\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$.

2. Supposons que $\beta \leq \gamma$.

On a donc $\beta < \gamma$ ou $\beta = \gamma$.

► Plaçons-nous dans le cas où $\beta < \gamma$.

On a donc $\alpha + \beta < \alpha + \gamma$ d'après 1.

On a en particulier $\alpha + \beta \leq \alpha + \gamma$ d'après la proposition 8 page 19.

► Plaçons-nous dans le cas où $\beta = \gamma$.

On a alors $\alpha + \beta = \alpha + \gamma$.

En particulier on a $\alpha + \beta \leq \alpha + \gamma$ par réflexivité de \leq .

Dans les deux cas on a donc $\alpha + \beta \leq \alpha + \gamma$.

3. Fixons β et γ deux ordinaux tels que $\beta \leq \gamma$.

Posons $Q_{\beta,\gamma}$ l'assertion à paramètre définie pour tout ordinal α par

$$Q_{\beta,\gamma}(\alpha) \iff \beta + \alpha \leq \gamma + \alpha$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

On a $\beta + 0 = \beta$ et $\gamma + 0 = \gamma$ par définition de l'addition.

Or on a $\beta \leq \gamma$ par hypothèse, donc $\beta + 0 \leq \gamma + 0$.

Autrement dit on a $Q_{\beta,\gamma}(0)$.

► *Héritéité*

Soit α un ordinal tel que $Q_{\beta,\gamma}(\alpha)$, c'est-à-dire $\beta + \alpha \leq \gamma + \alpha$.

On a donc $\beta + \alpha < \gamma + \alpha$ ou $\beta + \alpha = \gamma + \alpha$.

- Plaçons-nous dans le cas où $\beta + \alpha < \gamma + \alpha$.

Alors $S(\beta + \alpha) \leq \gamma + \alpha < S(\gamma + \alpha)$ d'après la proposition 13 page 33.

On a donc $S(\beta + \alpha) < S(\gamma + \alpha)$ par transitivité.

En particulier on a $S(\beta + \alpha) \leq S(\gamma + \alpha)$.

- Plaçons-nous dans le cas où $\beta + \alpha = \gamma + \alpha$.

Alors $S(\beta + \alpha) = S(\gamma + \alpha)$.

En particulier on a $S(\beta + \alpha) \leq S(\gamma + \alpha)$ par réflexivité de \leq .

Dans les deux cas on a $S(\beta + \alpha) \leq S(\gamma + \alpha)$.

On a donc $\beta + S(\alpha) \leq \gamma + S(\alpha)$ par définition de l'addition.

Autrement dit on a $Q_{\beta, \gamma}(S(\alpha))$.

Donc pour tout ordinal α , si $Q_{\beta, \gamma}(\alpha)$ alors $Q_{\beta, \gamma}(S(\alpha))$.

► Héritage limite

Soit α un ordinal limite non nul tel que $\forall \delta < \alpha, Q_{\beta, \gamma}(\delta)$.

Autrement dit pour tout $\delta < \alpha$ on a $\beta + \delta \leq \gamma + \delta$.

Soit δ un ordinal tel que $\delta < \alpha$.

On a alors

$\beta + \delta \leq \gamma + \delta$ d'après ce qui précède

$\leq \sup_{\varepsilon < \alpha} (\gamma + \varepsilon)$ car la borne supérieure est un majorant

$= \gamma + \alpha$ car α est limite non nul

On a donc $\beta + \delta \leq \gamma + \alpha$ par transitivité de \leq .

Donc pour tout $\delta < \alpha$ on a $\beta + \delta \leq \gamma + \alpha$.

On a donc $\sup_{\delta < \alpha} (\beta + \delta) \leq \gamma + \alpha$ par minimalité de la borne supérieure.

On a donc $\beta + \alpha \leq \gamma + \alpha$ car α est limite non nul. Autrement dit, on a $Q_{\beta, \gamma}(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \delta < \alpha, Q_{\beta, \gamma}(\delta)$ alors $Q_{\beta, \gamma}(\alpha)$.

Ainsi $Q_{\beta, \gamma}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $Q_{\beta, \gamma}(\alpha)$.

Autrement dit pour tout ordinal α on a $\boxed{\beta + \alpha \leq \gamma + \alpha}$.

CQFD.

Remarque :

1. En particulier supposons que $\beta \leq \gamma$.

Alors $\beta + 1 \leq \gamma + 1$ donc $S(\beta) \leq S(\gamma)$ d'après la proposition 37 page 95.

2. Soit $n \in \omega$.

Pour tout $m \in \omega$, on a $n + m \in \omega$ par stabilité de ω pour l'addition.

Donc pour tout ordinal $m < \omega$, on a $n + m < \omega$ par définition de $<$.
 On a donc $\sup_{m < \omega} (n + m) \leq \omega$ par minimalité de la borne supérieure.
 On a donc :

$$\begin{aligned}\omega &= 0 + \omega \text{ par neutralité de } 0 \text{ pour l'addition} \\ &\leq n + \omega \text{ par croissance de l'addition à droite} \\ &= \sup_{m < \omega} (n + m) \text{ par définition de l'addition} \\ &\leq \omega \text{ par ce qui précède}\end{aligned}$$

Ainsi on a $\omega \leq n + \omega \leq \omega$ et donc $n + \omega = \omega$.

On vient donc de montrer ce que nous avions constaté dans un exemple précédent.

Proposition 41 (Régularité de l'addition des ordinaux)

Soient α, β et γ trois ordinaux.

Si $\alpha + \beta = \alpha + \gamma$ alors $\beta = \gamma$.

On dit que l'addition à gauche des ordinaux est **régulière**.

Démonstration

Montrons-le par contraposition.

Supposons que $\beta \neq \gamma$.

On a donc $\beta < \gamma$ ou $\gamma < \beta$ d'après le théorème 1 page 21.

Si $\beta < \gamma$ alors $\alpha + \beta < \alpha + \gamma$ par stricte croissance de l'addition à gauche.

Si $\gamma < \beta$ alors $\alpha + \gamma < \alpha + \beta$ par stricte croissance de l'addition à gauche.

Dans les deux cas on a $\alpha + \beta \neq \alpha + \gamma$ par antiréflexivité de $<$.

Donc si $\beta \neq \gamma$ alors $\alpha + \beta \neq \alpha + \gamma$.

Par contraposition si $\alpha + \beta = \alpha + \gamma$ alors $\beta = \gamma$.

CQFD.

Remarque :

Malheureusement l'addition à droite n'est pas régulière.

En effet on a montré lors d'une précédente remarque que

$$1 + \omega = \omega = 2 + \omega$$

alors que l'on n'a pas $1 = 2$.

Dans la définition de l'addition, si γ est un ordinal limite non nul alors

$$\alpha + \gamma = \sup_{\delta < \gamma} (\alpha + \delta)$$

Or un ordinal limite est lui-même sa propre borne supérieure d'après la proposition 21 page 47

$$\gamma = \sup_{\delta \in \gamma} \delta = \sup_{\delta < \gamma} \delta$$

si bien que l'on a en fait

$$\alpha + \sup_{\delta < \gamma} \delta = \sup_{\delta < \gamma} (\alpha + \delta)$$

Ainsi l'addition à gauche commute avec la borne supérieure. Nous verrons que c'est vrai même pour la borne supérieure d'un ensemble qui n'est pas lui-même un ordinal. Pour l'heure, généralisons ce concept avec la définition qui suit.

Définition 19 (Assertion fonctionnelle croissante continue)

Soient C et D deux classes d'**ordinaux**.

Soit $F : C \rightarrow D$ une assertion fonctionnelle.

1. On dit que F est **croissante** si et seulement si pour α et β dans C on a ,

$$\alpha \leq \beta \implies F(\alpha) \leq F(\beta)$$

2. On se place dans le cas où $C = ON = D$, de sorte que $F : ON \rightarrow ON$.

Supposons que F est **croissante**.

On dit que F est **continue** si et seulement si pour tout X ensemble **non vide** d'ordinaux, on a

$$F(\sup(X)) = \sup(F^\rightarrow(X))$$

Remarque :

Le point 2 de cette définition est bien une généralisation de ce que nous avons vu juste avant :

$$F\left(\sup_{\xi \in X} \xi\right) = \sup_{\xi \in X} F(\xi)$$

Pourquoi cette propriété s'appelle-t-elle *continuité*? Parce qu'elle rappelle ce qu'il se passe dans le cadre de l'analyse : par exemple pour $f : \mathbb{R} \rightarrow \mathbb{R}$ une application **croissante**, f est *continue* (à gauche) en $a \in \mathbb{R}$ si et seulement si

$$f(a) = f\left(\sup_{x < a} x\right) = f\left(\lim_{\substack{x \rightarrow a \\ x < a}} x\right) = \lim_{\substack{x \rightarrow a \\ x < a}} f(x) = \sup_{x < a} f(x)$$

Il s'avère qu'en fait on peut affaiblir cette condition et quand-même retrouver la continuité en question : en la demandant seulement sur les ordinaux limites, on la retrouve partout.

Proposition 42 (Caractérisation de continuité)

Soit $F : ON \longrightarrow ON$ une assertion fonctionnelle **croissante**.

Les assertions suivantes sont équivalentes :

1. F est continue.
2. Pour tout ordinal limite non nul γ , on a $F(\gamma) = \sup_{\delta < \gamma} F(\delta)$.

Démonstration

$1 \Rightarrow 2$

Supposons que F est continue.

Par définition pour tout ensemble non vide d'ordinaux X , on a $F\left(\sup_{\xi \in X} \xi\right) = \sup_{\xi \in X} F(\xi)$.

Soit γ un ordinal limite non nul.

On a $\gamma = \sup_{\delta \in \gamma} \delta$ d'après la proposition 21 page 47.

En prenant $X := \gamma$ on a donc $F(\gamma) = F\left(\sup_{\delta \in \gamma} \delta\right) = \sup_{\delta \in \gamma} F(\delta) = \sup_{\delta < \gamma} F(\delta)$.

Donc $\boxed{\text{pour tout ordinal limite non nul } \gamma, \text{ on a } F(\gamma) = \sup_{\delta < \gamma} F(\delta)}$.

$1 \Leftarrow 2$

Supposons que pour tout ordinal limite non nul γ on a $F(\gamma) = \sup_{\delta < \gamma} F(\delta)$.

Soit X un ensemble non vide d'ordinaux.

Montrons que $F(\sup(X)) = \sup(F^\rightarrow(X))$.

Rappelons-nous la chose suivante : comme X est un ensemble, $F^\rightarrow(X) = \{F(\xi) \mid \xi \in X\}$ est aussi un ensemble d'après le schéma d'axiome de remplacement.

C'est bien un ensemble d'ordinaux car F est à valeurs dans ON .

Raisonnons par double inégalités.

\leq

- Plaçons-nous dans le cas où $\sup(X) \in X$.

Alors $F(\sup(X)) \in F^\rightarrow(X)$ par définition de l'image directe.

On a donc $F(\sup(X)) \leq \sup(F^\rightarrow(X))$ car la borne supérieure est un majorant.

- Plaçons-nous dans le cas où $\sup(X) \notin X$.

Alors $\sup(X)$ est un ordinal limite d'après la proposition 20 page 46.

Supposons par l'absurde que $\sup(X) = 0$.

Comme $\sup(X)$ est un majorant de X , on a $\forall \xi \in X, \xi \leq \sup(X)$.

Autrement dit on a $\forall \xi \in X, \xi \subseteq \sup(X)$.

Comme $\sup(X) = 0 = \emptyset$, on a $\forall \xi \in X, \xi = \sup(X)$.

Comme X est non vide, on a donc $X = \{\sup(X)\}$ et donc $\sup(X) \in X$.

C'est absurde puisqu'on a justement supposé que $\sup(X) \notin X$.

Donc $\sup(X)$ est un ordinal limite non nul.

Donc par hypothèse on a $F(\sup(X)) = \sup_{\delta < \sup(X)} F(\delta)$.

Montrons donc que $\sup_{\delta < \sup(X)} F(\delta) \leq \sup(F^\rightarrow(X))$.

Soit δ un ordinal tel que $\delta < \sup(X)$.

Par définition $\sup(X)$ est le plus petit majorant de X .

Donc δ n'est pas un majorant de X .

Il existe donc $\xi \in X$ tel que $\delta \leq \xi$.

Par croissance de F on a donc $F(\delta) \leq F(\xi)$.

Or $\xi \in X$ donc $F(\xi) \in F^\rightarrow(X)$ et donc $F(\xi) \leq \sup(F^\rightarrow(X))$.

On a donc $F(\delta) \leq \sup(F^\rightarrow(X))$ par transitivité de \leq .

Donc $\forall \delta < \sup(X), F(\delta) \leq \sup(F^\rightarrow(X))$.

Donc $\sup_{\delta < \sup(X)} F(\delta) \leq \sup(F^\rightarrow(X))$ par minimalité de la borne supérieure.

On a donc $F(\sup(X)) \leq \sup(F^\rightarrow(X))$ d'après ce qui précède.

Ainsi dans les deux cas on a $F(\sup(X)) \leq \sup(F^\rightarrow(X))$.

\geq

Soit $\mu \in F^\rightarrow(X)$.

Par définition il existe $\xi \in X$ tel que $\mu = F(\xi)$.

On a $\xi \leq \sup(X)$ car la borne supérieure est un majorant.

On a donc $F(\xi) \leq F(\sup(X))$ par croissance de F .

On a donc $\mu \leq F(\sup(X))$ par définition de ξ .

Donc pour tout $\mu \in F^\rightarrow(X)$, on a $\mu \leq F(\sup(X))$.

Donc $\sup(F^\rightarrow(X)) \leq F(\sup(X))$ par minimalité de la borne supérieure.

Finalement on a bien $F(\sup(X)) = \sup(F^\rightarrow(X))$ par antiréflexivité de \leq .

CQFD.

Ce que l'on vient de dire s'applique en particulier à l'addition qui vérifie bien la condition sur les ordinaux limites non vides. Ainsi l'addition à gauche est continue.

Proposition 43 (Continuité de l'addition des ordinaux)

Soient α un ordinal et X un ensemble **non vide** d'ordinaux.

On a

$$\sup_{\xi \in X} (\alpha + \xi) = \alpha + \sup_{\xi \in X} \xi$$

Autrement dit l'addition à gauche est continue.



Démonstration

Par définition de l'addition, pour tout ordinal limite non nul γ , on a

$$\alpha + \gamma = \sup_{\delta < \gamma} (\alpha + \delta)$$

On peut alors appliquer la proposition 42 page 103 pour conclure.

CQFD.

Remarque :

Malheureusement l'addition à droite n'est pas continue.

En effet, prenons $X = \omega = \alpha$.

On sait que ω est un ordinal limite d'après la proposition 18 page 43.

On a donc $\sup_{n \in \omega} n = \omega$ d'après la proposition 21 page 47.

Or on a montré dans une précédente remarque que pour tout $n \in \omega$, on a $n + \omega = \omega$.

On a donc $\sup_{n \in \omega} (n + \omega) = \sup_{n \in \omega} \omega = \omega$ tandis que $\left(\sup_{n \in \omega} n \right) + \omega = \omega + \omega$.

Or $\omega < S(\omega)$ 13 p. 33 $= \omega + 1 \leq \sup_{n \in \omega} (\omega + n) = \omega + \omega$ 37 p. 95.

On a donc $\omega \neq \omega + \omega$ et donc $\sup_{n \in \omega} (n + \omega) \neq \left(\sup_{n \in \omega} n \right) + \omega$.

Proposition 44 (Associativité de l'addition des ordinaux)

Pour tout ordinaux α , β et γ , on a l'égalité

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$$

On dit que l'addition des ordinaux est **associative**.



Démonstration

Montrons-le à l'aide du principe faible d'induction transfinie.

Fixons α et β deux ordinaux.

Posons $P_{\alpha, \beta}$ l'assertion à paramètre définie pour tout ordinal γ par

$$P_{\alpha, \beta}(\gamma) \iff (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$$

► *Initialisation*

On a les égalités suivantes :

$$\begin{aligned} (\alpha + \beta) + 0 &= \alpha + \beta \text{ car } 0 \text{ est neutre pour l'addition} \\ &= \alpha + (\beta + 0) \text{ car } 0 \text{ est neutre pour l'addition} \end{aligned}$$

On a donc $(\alpha + \beta) + 0 = \alpha + (\beta + 0)$ et donc $P_{\alpha,\beta}(0)$.

► *Héritéité*

Soit γ un ordinal tel que $P_{\alpha,\beta}(\gamma)$.

On a alors

$$\begin{aligned} (\alpha + \beta) + S(\gamma) &= S((\alpha + \beta) + \gamma) \text{ par définition de l'addition} \\ &= S(\alpha + (\beta + \gamma)) \text{ puisqu'on a } P_{\alpha,\beta}(\gamma) \\ &= \alpha + S(\beta + \gamma) \text{ par définition de l'addition} \\ &= \alpha + (\beta + S(\gamma)) \text{ par définition de l'addition} \end{aligned}$$

Ainsi on a $(\alpha + \beta) + S(\gamma) = \alpha + (\beta + S(\gamma))$, c'est-à-dire $P_{\alpha,\beta}(S(\gamma))$.

Donc pour tout ordinal γ , on a $P_{\alpha,\beta}(\gamma) \implies P_{\alpha,\beta}(S(\gamma))$.

► *Héritéité limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

Posons $X := \{\beta + \delta \mid \delta < \gamma\}$.

On a alors $\{\alpha + (\beta + \delta) \mid \delta < \gamma\} = \{\alpha + \xi \mid \xi \in X\}$.

En particulier $\sup_{\delta < \gamma} (\alpha + (\beta + \delta)) = \sup_{\xi \in X} (\alpha + \xi) \quad (\star)$.

On en déduit donc que :

$$\begin{aligned} (\alpha + \beta) + \gamma &= \sup_{\delta < \gamma} ((\alpha + \beta) + \delta) \text{ par définition de l'addition} \\ &= \sup_{\delta < \gamma} (\alpha + (\beta + \delta)) \text{ puisque } \forall \delta < \gamma, P_{\alpha,\beta}(\delta) \\ &= \sup_{\xi \in X} (\alpha + \xi) \text{ par } (\star) \\ &= \alpha + \sup_{\xi \in X} \xi \text{ par continuité de l'addition à gauche} \\ &= \alpha + \sup_{\delta < \gamma} (\beta + \delta) \text{ par définition de } X \\ &= \alpha + (\beta + \gamma) \text{ par définition de l'addition} \end{aligned}$$

On a donc $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Autrement dit on a $P_{\alpha,\beta}(\gamma)$.

Ainsi pour tout ordinal limite non nul γ , on a $(\forall \delta < \gamma, P_{\alpha,\beta}(\delta)) \implies P_{\alpha,\beta}(\gamma)$.

Ainsi $P_{\alpha,\beta}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ , on a $P_{\alpha,\beta}(\gamma)$, c'est-à-dire $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

CQFD.

Remarque :

Désormais pour α, β et γ trois ordinaux, on notera $\alpha + \beta + \gamma$ pour désigner indifféremment $(\alpha + \beta) + \gamma$ et $\alpha + (\beta + \gamma)$, puisqu'il y a égalité.

On a vu plus tôt que l'addition chez les ordinaux n'est malheureusement pas commutative. Heureusement, elle l'est chez les entiers naturels : on retrouve donc bien un résultat qui nous semble évident avec l'addition de tous les jours. Ouf !

Pour le montrer, on commence par montrer que 1 commute avec tous les entiers naturels.

Proposition 45 (Commutativité de l'addition des entiers naturels)

Soient m et n deux entiers naturels.

Alors $m + n = n + m$.

On dit que l'addition des entiers naturels est **commutative**.

Démonstration

- On sait déjà que $m + 1 = S(m)$ d'après la proposition 37 page 95.

Montrons que l'on a aussi $S(m) = 1 + m$, de sorte que $m + 1 = 1 + m$.

Pour tout entier naturel m , posons $P(m)$ l'assertion « $S(m) = 1 + m$ ».

Initialisation

On a $S(0) = 1$ par définition de 1.

Or $1 = 1 + 0$ par neutralité de 0 pour l'addition.

On a donc $S(0) = 1 + 0$ et donc on a $P(0)$.

Hérédité

Soit m un entier naturel tel que $P(m)$.

On a alors

$$\begin{aligned} S(S(m)) &= S(m) + 1 \text{ d'après la proposition 37 page 95} \\ &= (1 + m) + 1 \text{ d'après } P(m) \end{aligned}$$

$$\begin{aligned}
 &= 1 + (m + 1) \text{ par associativité de l'addition} \\
 &= 1 + S(m) \text{ d'après la proposition 37 page 95}
 \end{aligned}$$

Ainsi $S(S(m)) = 1 + S(m)$ et donc $P(S(m))$.

Ainsi pour tout entier naturel m , si $P(m)$ alors $P(S(m))$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel m , on a $P(m)$, c'est-à-dire $S(m) = 1 + m$.

- Montrons le résultat attendu.

Fixons m un entier naturel.

Pour tout entier naturel n , posons $Q(n)$ l'assertion « $m + n = n + m$ ».

Initialisation

On a $m + 0 = m = 0 + m$ car 0 est neutre pour l'addition.

On a donc $Q(0)$.

Hérédité

Soit n un entier naturel tel que $Q(n)$.

On a alors

$$\begin{aligned}
 m + S(n) &= m + (n + 1) \text{ d'après la proposition 37 page 95} \\
 &= (m + n) + 1 \text{ par associativité de l'addition} \\
 &= 1 + (m + n) \text{ d'après ce que l'on a montré plus haut} \\
 &= 1 + (n + m) \text{ d'après } Q(n) \\
 &= (1 + n) + m \text{ par associativité de l'addition} \\
 &= S(n) + m \text{ d'après ce que l'on a montré plus haut}
 \end{aligned}$$

Ainsi $n + S(n) = S(n) + m$ et donc $Q(S(n))$.

Ainsi pour tout entier naturel n , si $Q(n)$ alors $Q(S(n))$.

Finalement Q vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on $Q(n)$, c'est-à-dire $m + n = n + m$.

CQFD.

À la toute fin du précédent chapitre, nous avons défini par récursion les itérées d'une application, c'est-à-dire le fait de prendre un objet, de le donner à manger à une application, de donner le résultat à manger à l'application, et ainsi de suite autant de fois que désiré. Il s'avère bien heureusement que le faire $n + m$ fois, c'est là même chose que le faire m fois puis n fois : encore une fois l'intuition est préservée !

Proposition 46 (Itérées d'une application et addition)

Soient E un ensemble et $f : E \longrightarrow E$.

Pour tout entiers naturels n et m , on a $f^{n+m} = f^n \circ f^m$.



Démonstration

Fixons n un entier naturel.

Pour tout entier naturel m , posons $P(m)$ l'assertion « $f^{n+m} = f^n \circ f^m$ ».

Initialisation

On a $f^{n+0} = f^n = f^n \circ \text{id}_E = f^n \circ \text{id}_E$ et donc $P(0)$.

Hérédité

Soit m un entier naturel tel que $P(m)$, c'est-à-dire $f^{n+m} = f^n \circ f^m$.

On a alors

$$\begin{aligned} f^{n+(m+1)} &= f^{(n+m)+1} \text{ par associativité de l'addition} \\ &= f^{n+m} \circ f \text{ par définition des itérées de } f \\ &= (f^n \circ f^m) \circ f \text{ d'après } P(m) \\ &= f^n \circ (f^m \circ f) \text{ par associativité de la composition} \\ &= f^n \circ f^{m+1} \text{ par définition des itérées de } f \end{aligned}$$

Ainsi on a $f^{n+(m+1)} = f^n \circ f^{m+1}$ et donc $P(m+1)$.

Ainsi pour tout entier naturel m , si $P(m)$ alors $P(m+1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

On a donc pour tout $m \in \mathbb{N}$, $P(m)$, c'est-à-dire $f^{n+m} = f^n \circ f^m$.

CQFD.

Remarque :

En particulier pour tout entier naturel n , on a $f^n \circ f = f^{n+1} = f^{1+n} = f^1 \circ f^n = f \circ f^n$.

Plus généralement, comme l'addition des entiers naturels est commutative, les itérées de f commutent.

Que peut-on dire de l'image directe et de l'image réciproque par une itérée d'une application ? On a vu lors du précédent livre que $(g \circ f)^{\rightarrow} = g^{\rightarrow} \circ f^{\rightarrow}$ pour deux applications f et g . Autrement dit, on va par exemple avoir $(f^3)^{\rightarrow} = (f \circ f \circ f)^{\rightarrow} = f^{\rightarrow} \circ f^{\rightarrow} \circ f^{\rightarrow} = (f^{\rightarrow})^3$. Plus généralement, on a le résultat suivant.

Proposition 47 (Images directes et réciproques d'une itérée)

Soient E un ensemble, $f : E \longrightarrow E$ et $n \in \mathbb{N}$.

1. On a $(f^n)^\rightarrow = (f^\rightarrow)^n$.
2. On a $(f^n)^\leftarrow = (f^\leftarrow)^n$.



Démonstration

1. Pour tout entier naturel n , posons $P(n)$ l'assertion « $(f^n)^\rightarrow = (f^\rightarrow)^n$ ».

Initialisation

On a $(f^0)^\rightarrow = \text{id}_E^\rightarrow = \text{id}_{\mathcal{P}(E)} = (f^\rightarrow)^0$ donc on a $P(0)$.

Hérédité

Soit n un entier naturel tel que $P(n)$.

On a alors

$$\begin{aligned} (f^{n+1})^\rightarrow &= (f^n \circ f)^\rightarrow \text{ par définition des itérées de } f \\ &= (f^n)^\rightarrow \circ f^\rightarrow \text{ d'après le précédent livre} \\ &= (f^\rightarrow)^n \circ f^\rightarrow \text{ d'après } P(n) \\ &= (f^\rightarrow)^{n+1} \text{ par définition des itérées de } f^\rightarrow \end{aligned}$$

On a donc $(f^{n+1})^\rightarrow = (f^\rightarrow)^{n+1}$, c'est-à-dire $P(n+1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n+1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

On a donc pour tout n entier naturel $P(n)$, c'est-à-dire $\boxed{(f^n)^\rightarrow = (f^\rightarrow)^n}$.

2. Pour tout entier naturel n , posons $Q(n)$ l'assertion « $(f^n)^\leftarrow = (f^\leftarrow)^n$ ».

Initialisation

On a $(f^0)^\leftarrow = \text{id}_E^\leftarrow = \text{id}_{\mathcal{P}(E)} = (f^\leftarrow)^0$ donc on a $Q(0)$.

Hérédité

Soit n un entier naturel tel que $Q(n)$.

On a alors

$$(f^{n+1})^\leftarrow = (f^n \circ f)^\leftarrow \text{ par définition des itérées de } f$$

$$\begin{aligned}
 &= (f \circ f^n)^\leftarrow \text{ car les itérées de } f \text{ commutent} \\
 &= (f^n)^\leftarrow \circ f^\leftarrow \text{ d'après le précédent livre} \\
 &= (f^\leftarrow)^n \circ f^\leftarrow \text{ d'après } Q(n) \\
 &= (f^\leftarrow)^{n+1} \text{ par définition des itérées de } f^\leftarrow
 \end{aligned}$$

On a donc $(f^{n+1})^\leftarrow = (f^\leftarrow)^{n+1}$, c'est-à-dire $Q(n+1)$.

Ainsi pour tout entier naturel n , si $Q(n)$ alors $Q(n+1)$.

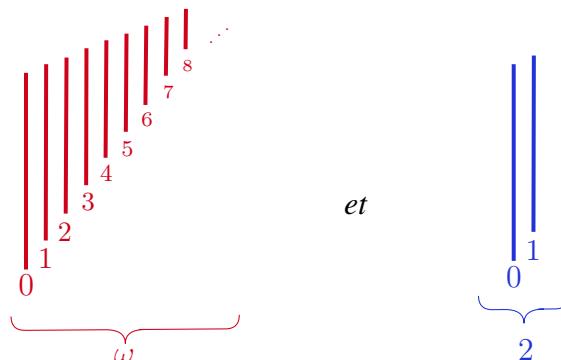
Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

On a donc pour tout n entier naturel $Q(n)$, c'est-à-dire $\boxed{(f^n)^\leftarrow = (f^\leftarrow)^n}$.

CQFD.

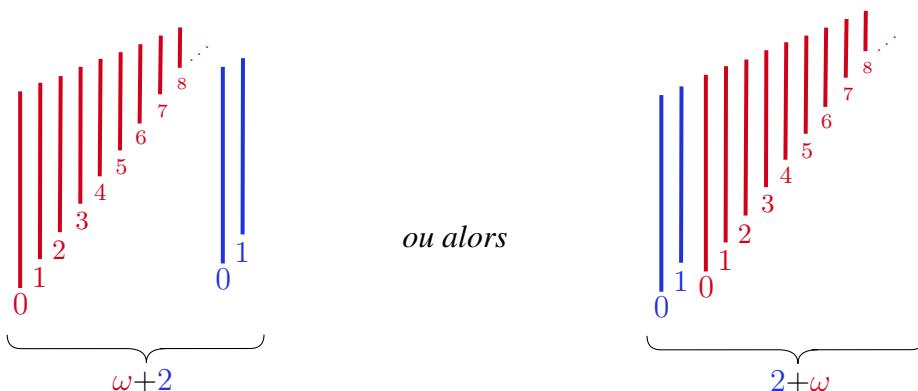
2.2 Interprétation graphique : la concaténation

L'addition des ordinaux a une interprétation graphique : visualisons par exemple l'addition des ordinaux ω et 2. Commençons par les représenter tous les deux avec des bâtons indépendamment l'un de l'autre, ω et ses éléments étant en **rouge** et 2 et ses éléments étant en **bleu** :

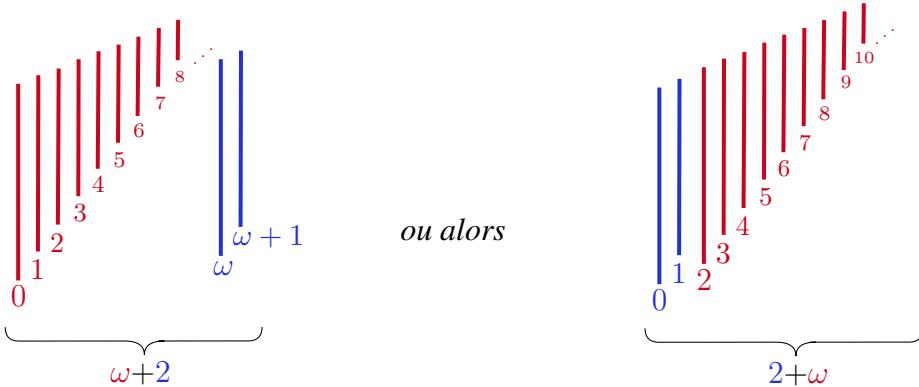


Rappelons qu'ici la taille des bâtons n'a aucune importance, seul leur agencement horizontal importe. Le fait de représenter des bâtons de plus en plus petits est seulement une astuce pour en faire tenir une infinité. L'interprétation visuelle consiste alors à concaténer les deux ordinaux à additionner, c'est-à-dire à les placer l'un derrière l'autre. Ainsi, on peut les disposer de deux manières :

- d'abord ω puis 2 à sa droite, ce qui donne la représentation graphique de $\omega + 2$
- d'abord 2 puis ω à sa droite, ce qui donne la représentation graphique de $2 + \omega$



On renumérote alors les bâtons en fonction de l'ordre dans lequel ils arrivent, de la gauche vers la droite,



ce qui permet ainsi de voir que

- $\omega + 2$ est égal à $\{0, 1, 2, \dots, \omega, \omega + 1\}$ donc est l'ordinal qui vient juste après $\omega + 1$, c'est-à-dire son successeur.
- $2 + \omega$ est égal à $\{0, 1, 2, 3, \dots\}$, c'est-à-dire tout simplement ω lui-même.

On retrouve bien le fait démontré précédemment que $2 + \omega = \omega$ et que $2 + \omega \neq \omega + 2$.

Pour l'heure, cette illustration est là pour nous faire comprendre l'intuition derrière l'addition des ordinaux : cette idée de concaténation va se traduire formellement par la notion d'**union disjointe**. On va en fait associer les éléments de A aux couples de la forme $(0, a)$ avec $a \in A$, et les éléments de B aux couples de la forme $(1, b)$ avec $b \in B$. De cette façon, on pourra dire "*parmi ces couples-là, ceux avec une première composante nulle viennent avant ceux avec une première composante qui vaut 1*", et donc s'assurer que les éléments de A viennent avant les éléments de B .

Définition 20 (Union disjointe de deux ensembles)

Soient A et B deux ensembles.

On appelle **union disjointe** de A et B l'ensemble

$$A \amalg B := (\{0\} \times A) \cup (\{1\} \times B)$$

Dans le cas où A et B sont munit d'un ordre, voyons comment s'en servir pour construire un ordre sur $A \amalg B$.

- on souhaite que tout élément de A vienne nécessairement avant tout élément de B . On va donc dire que tout élément de première composante nulle est plus petit que tout élément de première composante 1.
- si les deux couples à comparer sont de première composante nulle, alors les deuxièmes composantes sont dans A et on peut donc les comparer dans A .
- si les deux couples à comparer sont de première composante 1, alors les deuxièmes composantes sont dans B et on peut donc les comparer dans B .

C'est en quelque sorte comme l'ordre lexicographique : on compare les premières composantes et si éventuellement elles sont égales, on compare les secondes. C'est l'objet de la définition qui suit.

Définition 21 (Ordre sur l'union disjointe de deux ensembles)

Soient (A, \preccurlyeq) et (B, \sqsubseteq) deux ensembles ordonnés.

On appelle **ordre de concaténation** sur $A \amalg B$ la relation binaire \trianglelefteq définie pour tout (i, x) et (j, y) dans $A \amalg B$ par

$$(i, x) \trianglelefteq (j, y) \iff \begin{cases} i = 0 \text{ et } j = 1 \\ \text{ou} \\ i = 0 = j \text{ et } x \preccurlyeq y \\ \text{ou} \\ i = 1 = j \text{ et } x \sqsubseteq y \end{cases}$$

Bien évidemment, comme son nom l'indique, cette relation est une relation d'ordre.

Proposition 48 (Ordre de concaténation)

Soient (A, \preccurlyeq) et (B, \sqsubseteq) deux ensembles ordonnés.

Soit \trianglelefteq l'ordre de concaténation associé sur $A \amalg B$.

Alors \trianglelefteq est une relation d'ordre sur $A \amalg B$.

Démonstration

Réflexivité

Soit $(i, x) \in A \amalg B$.

► Supposons que $i = 0$.

Par définition de $A \amalg B$, on a alors $x \in A$.

Par réflexivité de \preccurlyeq sur A , on a $x \preccurlyeq x$.

Ainsi $(i = 0 = i \text{ et } x \preccurlyeq x)$ donc $(i, x) \trianglelefteq (i, x)$ par définition de \trianglelefteq .

► On raisonne de la même manière si $i = 1$.

On a donc $(i, x) \trianglelefteq (i, x)$.

Donc \trianglelefteq est réflexive sur $A \amalg B$.

Antisymétrie

Soient (i, x) et (j, y) dans $A \amalg B$.

Supposons que $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (i, x)$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

C'est impossible puisqu'on a $(j, y) \trianglelefteq (i, x)$.

► Plaçons-nous dans le cas où $i = 1$ et $j = 0$.

C'est impossible puisqu'on a $(i, x) \trianglelefteq (j, y)$.

- Plaçons-nous dans le cas où $i = 0 = j$.

Par définition de $A \amalg B$, on a alors $x \in A$ et $y \in A$.

Comme $(i, x) \trianglelefteq (j, y)$, on a $x \preccurlyeq y$.

Comme $(j, y) \trianglelefteq (i, x)$, on a $y \preccurlyeq x$.

On a donc $x = y$ par antisymétrie de \preccurlyeq .

Ainsi $i = j$ et $x = y$ donc $(i, x) = (j, y)$.

- Le cas $i = 1 = j$ se traite de la même manière.

Ainsi dans les deux cas possibles on a $(i, x) = (j, y)$.

Donc si $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (i, x)$ alors $(i, x) = (j, y)$.

Donc \trianglelefteq est antisymétrique.

Transitivité

Soient (i, x) , (j, y) et (k, z) dans $A \amalg B$.

Supposons que $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (k, z)$.

- Plaçons-nous dans le cas où $i = j = k = 0$.

Par définition de $A \amalg B$, on a alors $x \in A$, $y \in A$ et $z \in A$.

Comme $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (k, z)$, on a $x \preccurlyeq y$ et $y \preccurlyeq z$.

On a donc $x \preccurlyeq z$ par transitivité de \preccurlyeq .

Comme $i = 0 = k$ et $x \preccurlyeq z$, on a $(i, x) \trianglelefteq (k, z)$.

- Le cas $i = j = k = 1$ se traite de la même manière.

- Si $i = 0 = j$ et $k = 1$ alors on a automatiquement $(i, x) \trianglelefteq (k, z)$.

- Si $i = 0$ et $j = 1 = k$ alors on a automatiquement $(i, x) \trianglelefteq (k, z)$.

- Les autres cas sur les valeurs de i , j et k sont impossibles puisque l'on a $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (k, z)$.

Dans tous les cas, on a nécessairement $(i, x) \trianglelefteq (k, z)$.

Donc si $(i, x) \trianglelefteq (j, y)$ et $(j, y) \trianglelefteq (k, z)$ alors $(i, x) \trianglelefteq (k, z)$.

Donc \trianglelefteq est transitive.

Ainsi \trianglelefteq est réflexive sur $A \amalg B$, est antisymétrique et transitive.

Donc \trianglelefteq est une relation d'ordre sur $A \amalg B$.

CQFD.

Nous avons formalisé ce qu'était l'opération de concaténation de deux ensembles ordonnés : passer par leur union disjointe et associer à celle-ci l'ordre de concaténation. Il s'avère que si les deux ensembles en question sont bien ordonnés, il en va de même pour leur union disjointe.

Proposition 49 (Concaténation de bons ordres)

Soient (A, \preccurlyeq) et (B, \sqsubseteq) deux ensembles ordonnés.

Soit \trianglelefteq l'ordre de concaténation associé sur $A \amalg B$.

Si (A, \preccurlyeq) et (B, \sqsubseteq) sont bien ordonnés alors $(A \amalg B, \trianglelefteq)$ est bien ordonné.



Démonstration

Supposons que (A, \preccurlyeq) et (B, \sqsubseteq) sont bien ordonnés.

Montrons que toute partie non vide de $A \amalg B$ admet un minimum.

Soit C une partie non vide $A \amalg B$.

- Supposons dans un premier temps qu'il existe $a \in A$ tel que $(0, a) \in C$.

Considérons alors $E := \{a \in A \mid (0, a) \in C\}$.

Alors E est donc une partie non vide de A .

Comme A est bien ordonné, E admet un minimum a_0 .

Ainsi on a $(0, a_0) \in C$ par définition de E .

Montrons que $(0, a_0)$ est le minimum de C .

Soit $(i, x) \in C$.

- Plaçons-nous dans le cas où $i = 0$.

Dans ce cas-là $x \in A$ par définition de $A \amalg B$.

Donc $x \in A$ est tel que $(0, x) = (i, x) \in C$.

Alors $x \in E$ par définition de E .

Donc $a_0 \preccurlyeq x$ car a_0 est le minimum de E .

Ainsi ($i = 0$ et $a_0 \preccurlyeq x$) donc $(0, a_0) \trianglelefteq (i, x)$ par définition de \trianglelefteq .

- Plaçons-nous dans le cas où $i = 1$.

On a alors $(0, a_0) \trianglelefteq (i, x)$ par définition de \trianglelefteq .

Dans tous les cas on a $(0, a_0) \trianglelefteq (i, x)$.

Donc $(0, a_0)$ est le minimum de C .

- Supposons à présent qu'il n'existe pas de $a \in A$ tel que $(0, a) \in C$.

Donc pour tout $(i, x) \in C$, on a $i = 1$ et $x \in B$ par définition de $A \amalg B$.

Posons alors $F := \{b \in B \mid (1, b) \in C\}$.

Comme C est non vide, F est une partie non vide de B .

Or B est bien ordonné donc F admet un minimum b_0 .

Ainsi on a $(1, b_0) \in C$ par définition de F .

Montrons que $(1, b_0)$ est le minimum de C .

Soit $(i, x) \in C$.

D'après ce qui précède, on a nécessairement $i = 1$ et $x \in B$.

Ainsi $x \in B$ est tel que $(1, x) = (i, x) \in C$.

Donc $x \in F$ par définition de F .

Donc $b_0 \sqsubseteq x$ car b_0 est le minimum de F .

Ainsi $(i = 1 \text{ et } b_0 \sqsubseteq x)$ donc $(1, b_0) \trianglelefteq (i, x)$ par définition de \trianglelefteq .

Donc $(1, b_0)$ est le minimum de C .

Dans les deux cas C admet un minimum.

Donc toute partie non vide de $A \amalg B$ admet un minimum.

Donc $A \amalg B$ est bien ordonné.

CQFD.

Notation :

Soient E et F deux ensembles ordonnés.

On note $E \cong F$ si et seulement si E et F sont isomorphes.

La première partie de la proposition qui suit nous indique que la concaténation se comporte bien vis à vis de l'isomorphie d'ordres. La deuxième partie nous indique qu'étant donnés un ordinal α et un ordinal plus petit β , il est possible de simplement concaténer β et $\alpha \setminus \beta$ pour retrouver α . Le lecteur avisé pourra reconnaître les premiers jalons de la soustraction d'ordinaux !

Proposition 50 (Union disjointe et isomorphismes)

1. Soient A_0, A_1, B_0 et B_1 quatre ensembles ordonnés.

Si $A_0 \cong B_0$ et $A_1 \cong B_1$ alors $A_0 \amalg A_1 \cong B_0 \amalg B_1$.

2. Soient α et β deux ordinaux.

Si $\beta \leq \alpha$ alors $\alpha \cong \beta \amalg (\alpha \setminus \beta)$.

Démonstration

1. Supposons que $A_0 \cong B_0$ et $A_1 \cong B_1$.

Pour tout $i \in \{0, 1\}$, il existe donc $f_i : A_i \longrightarrow B_i$ un isomorphisme d'ordre.

Considérons alors $\varphi := \begin{pmatrix} A_0 \amalg A_1 & \longrightarrow & B_0 \amalg B_1 \\ (i, x) & \longmapsto & (i, f_i(x)) \end{pmatrix}$.

Montrons que φ est un isomorphisme d'ordre de $A_0 \amalg A_1$ vers $B_0 \amalg B_1$.

- Montrons que φ est bijective de $A_0 \amalg A_1$ vers $B_0 \amalg B_1$.

Pour tout $i \in \{0, 1\}$, f_i est un isomorphisme d'ordre de A_i vers B_i .

Donc tout $i \in \{0, 1\}$, f_i est inversible, avec $f_i^{-1} : B_i \longrightarrow A_i$.

Considérons alors $\psi := \begin{pmatrix} B_0 \amalg B_1 & \longrightarrow & A_0 \amalg A_1 \\ (j, y) & \longmapsto & (j, f_j^{-1}(y)) \end{pmatrix}$.

Montrons que φ et ψ sont réciproques l'une de l'autre.

En effet, pour tout $(i, x) \in A_0 \amalg A_1$ on a

$$\begin{aligned} (\psi \circ \varphi)(i, x) &= \psi(\varphi(i, x)) = \psi(i, f_i(x)) = \left(i, f_i^{-1}(f_i(x)) \right) = (i, x) \\ &= \text{id}_{A_0 \amalg A_1}(i, x) \end{aligned}$$

si bien que $\psi \circ \varphi = \text{id}_{A_0 \amalg A_1}$.

De même pour tout $(j, y) \in B_0 \amalg B_1$ on a

$$\begin{aligned} (\varphi \circ \psi)(j, y) &= \varphi(\psi(j, y)) = \varphi(j, f_j^{-1}(y)) = \left(j, f_j(f_j^{-1}(y)) \right) = (j, y) \\ &= \text{id}_{B_0 \amalg B_1}(j, y) \end{aligned}$$

si bien que $\varphi \circ \psi = \text{id}_{B_0 \amalg B_1}$.

Ainsi φ et ψ sont réciproques l'une de l'autre.

En particulier φ est bijective de $A_0 \amalg A_1$ vers $B_0 \amalg B_1$.

• Montrons que φ est croissante.

Pour tout $i \in \{0, 1\}$, f_i est un isomorphisme de A_i vers B_i .

Donc pour tout $i \in \{0, 1\}$, f_i est croissante.

Pour tout $i \in \{0, 1\}$, notons \preccurlyeq_i l'ordre sur A_i .

Notons \trianglelefteq_A et \trianglelefteq_B les ordres de concaténation associés sur $A_0 \amalg A_1$ et $B_0 \amalg B_1$.

Soient (i, x) et (j, y) dans $A_0 \amalg A_1$.

Supposons que $(i, x) \trianglelefteq_A (j, y)$.

► Plaçons-nous dans le cas où $i = j$.

On a $(i, x) \trianglelefteq_A (i, y)$ donc $x \preccurlyeq_i y$ par définition de \trianglelefteq_A .

Donc $f_i(x) \preccurlyeq_i f_i(y)$ par croissance de f_i .

Donc $(i, f_i(x)) \trianglelefteq_B (i, f_i(y))$ par définition de \trianglelefteq_B .

Donc $\varphi(i, x) \trianglelefteq_B \varphi(i, y)$ par définition de φ .

On a donc $\varphi(i, x) \trianglelefteq_B \varphi(j, y)$ puisque $i = j$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

On a $(0, f_0(x)) \trianglelefteq_B (1, f_1(y))$ par définition de \trianglelefteq_B .

Donc $\varphi(0, x) \trianglelefteq_B \varphi(1, y)$ par définition de φ .

On a donc $\varphi(i, x) \trianglelefteq_B \varphi(j, y)$ puisque $i = 0$ et $j = 1$.

► Plaçons-nous dans le cas où $i = 1$ et $j = 0$.

C'est absurde puisque par hypothèse on a $(i, x) \leq_A (j, y)$.

Donc dans tous les cas possibles, on a $\varphi(i, x) \leq_B \varphi(j, y)$.

Donc si $(i, x) \leq_A (j, y)$ alors $\varphi(i, x) \leq_B \varphi(j, y)$.

Donc φ est croissante.

On montre de la même manière que ψ est croissante.

- Ainsi φ est bijective de $A_0 \amalg A_1$ vers $B_0 \amalg B_1$, et φ et sa réciproque ψ sont croissantes.

Donc φ est un isomorphisme d'ordre de $A_0 \amalg A_1$ vers $B_0 \amalg B_1$.

Donc $[A_0 \amalg A_1 \cong B_0 \amalg B_1]$.

2. Supposons que $\beta \subseteq \alpha$.

On a donc $\beta \subseteq \alpha$ par définition de \leq .

Pour tout $\gamma \in \beta$, on a donc $\gamma \in \alpha$ par définition de l'inclusion.

De même on a $\alpha \setminus \beta \subseteq \alpha$ donc pour tout $\gamma \in \alpha \setminus \beta$ on a $\gamma \in \alpha$.

Ainsi toute deuxième composante d'un élément de $\beta \amalg (\alpha \setminus \beta)$ est un élément de α .

On peut donc définir $\varphi := \begin{pmatrix} \beta \amalg (\alpha \setminus \beta) & \longrightarrow & \alpha \\ (i, \gamma) & \longmapsto & \gamma \end{pmatrix}$.

- Montrons que φ est injective.

Soient (i, γ) et (j, δ) dans $\beta \amalg (\alpha \setminus \beta)$ tels que $\varphi(i, \gamma) = \varphi(j, \delta)$.

Par définition de φ on a alors $\gamma = \delta$.

En particulier on a à la fois $(i, \gamma) \in \beta \amalg (\alpha \setminus \beta)$ et à la fois $(j, \gamma) \in \beta \amalg (\alpha \setminus \beta)$.

Supposons par l'absurde que $i \neq j$.

On vient de voir que $(i, \gamma) \in \beta \amalg (\alpha \setminus \beta)$ et $(j, \gamma) \in \beta \amalg (\alpha \setminus \beta)$.

Donc $\gamma \in \beta$ et $\gamma \in \alpha \setminus \beta$ par définition de $\beta \amalg (\alpha \setminus \beta)$.

C'est absurde puisque β et $\alpha \setminus \beta$ sont disjoints par définition.

Par l'absurde on vient de montrer que $i = j$.

Comme $\gamma = \delta$ on a donc $(i, \gamma) = (j, \delta)$.

Donc pour tout (i, γ) et (j, δ) dans $\beta \amalg (\alpha \setminus \beta)$, si $\varphi(i, \gamma) = \varphi(j, \delta)$ alors $(i, \gamma) = (j, \delta)$.

Donc φ est injective.

- Montrons que φ est surjective dans α .

Soit $\gamma \in \alpha$.

Comme $\beta \subseteq \alpha$, on a $\alpha = \beta \cup (\alpha \setminus \beta)$.

On a donc $\gamma \in \beta \cup (\alpha \setminus \beta)$ et donc ($\gamma \in \beta$ ou $\gamma \in \alpha \setminus \beta$).

Si $\gamma \in \beta$ alors $\gamma = \varphi(0, \gamma)$ et si $\gamma \in \alpha \setminus \beta$ alors $\gamma = \varphi(1, \gamma)$ par définition de φ .

Dans les deux cas on a $\gamma \in \text{im}(\varphi)$.

On a donc $\alpha \subseteq \text{im}(\varphi)$.

Or par définition de φ on a $\alpha \supseteq \text{im}(\varphi)$, et donc $\alpha = \text{im}(\varphi)$.

Ainsi φ est surjective dans α .

On en conclut donc que φ est bijective de $\beta \amalg (\alpha \setminus \beta)$ vers α .

- Montrons que φ est croissante.

Considérons \trianglelefteq l'ordre de concaténation associé sur $\beta \amalg (\alpha \setminus \beta)$.

Soient (i, γ) et (j, δ) dans $\beta \amalg (\alpha \setminus \beta)$.

Supposons que $(i, \gamma) \trianglelefteq (j, \delta)$.

- Plaçons-nous dans le cas où $i = j$.

On a donc $\gamma \leq \delta$ par définition de \trianglelefteq .

- Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

On a donc $\gamma \in \beta$ et $\delta \in \alpha \setminus \beta$ par définition de $\beta \amalg (\alpha \setminus \beta)$.

Comme $\delta \in \alpha \setminus \beta$, on a $\delta \notin \beta$ donc $\delta \not\prec \beta$ par définition de $<$.

Or \leq est total chez les ordinaux donc $\beta \leq \delta$.

Comme $\gamma \in \beta$, on a $\gamma < \beta$ donc $\gamma < \delta$ et donc $\gamma \leq \delta$.

- Plaçons-nous dans le cas où $i = 1$ et $j = 0$.

C'est absurde puisqu'on a fait l'hypothèse que $(i, \gamma) \trianglelefteq (j, \delta)$.

Donc dans les seuls cas possibles, on a nécessairement $\gamma \leq \delta$.

Or on a $\varphi(i, \gamma) = \gamma$ et $\varphi(j, \delta) = \delta$ par définition de φ , donc $\varphi(i, \gamma) \leq \varphi(j, \delta)$.

Donc si $(i, \gamma) \trianglelefteq (j, \delta)$ alors $\varphi(i, \gamma) \leq \varphi(j, \delta)$.

Donc φ est croissante.

Ainsi φ est bijective de $\beta \amalg (\alpha \setminus \beta)$ vers α et est croissante.

Or β et $\alpha \setminus \beta$ sont bien ordonnés d'après le théorème 1 page 21.

Donc $\beta \amalg (\alpha \setminus \beta)$ est bien ordonné d'après la proposition 49 page 115.

En particulier $\beta \amalg (\alpha \setminus \beta)$ est totalement ordonné d'après la proposition 2 page 10.

Donc φ est un isomorphisme d'ordre $\beta \amalg (\alpha \setminus \beta)$ vers α .

En particulier on a $\boxed{\alpha \cong \beta \amalg (\alpha \setminus \beta)}$.

CQFD.

Nous y voilà ! Prenons deux ordinaux α et β : ils sont bien ordonnés par \leq donc $\alpha \amalg \beta$ est aussi bien ordonné par l'ordre de concaténation associé. Donc d'après le théorème 4 page 56, il existe un unique ordinal isomorphe à $\alpha \amalg \beta$, que l'on a noté type($\alpha \amalg \beta$). L'intuition est confirmée par le théorème suivant : cet unique ordinal est en fait $\alpha + \beta$!

Au passage, remarquons que le fait de passer de $\alpha \amalg \beta$ à type($\alpha \amalg \beta$) correspond à la renumérotation que l'on a fait dans l'exemple visuel de $2 + \omega$: c'était une étape nécessaire pour s'assurer d'avoir un ordinal à la fin.

Théorème 7 (Addition d'ordinaux et concaténation)

Soient α et β deux ordinaux.

On munit $\alpha \amalg \beta$ de l'ordre de concaténation associé.

Alors $\alpha + \beta = \text{type}(\alpha \amalg \beta)$.

Démonstration

Notons \leq l'ordre de concaténation associé à $\alpha \amalg \beta$.

Construisons un isomorphisme d'ordre entre $\alpha \amalg \beta$ et $\alpha + \beta$.

- Construction de l'application.

Remarquons que pour tout $(i, \gamma) \in \alpha \amalg \beta$, on a :

► Plaçons-nous dans le cas où $i = 0$.

Alors par définition de $\alpha \amalg \beta$ on a $\gamma \in \alpha$.

On sait que $0 \subseteq \beta$ car le vide est inclus dans tout ensemble, donc $0 \leq \beta$.

Donc $\alpha + 0 \leq \alpha + \beta$ par croissance de l'addition à gauche.

Or $\alpha = \alpha + 0$ par définition de l'addition donc $\alpha \leq \alpha + \beta$.

On a donc $\alpha \subseteq \alpha + \beta$ par définition de \leq .

On a donc $\gamma \in \alpha + \beta$ par définition de l'inclusion.

► Plaçons-nous dans le cas où $i = 1$.

Alors par définition de $\alpha \amalg \beta$ on a $\gamma \in \beta$ et donc $\gamma < \beta$.

On a donc $\alpha + \gamma < \alpha + \beta$ par stricte croissance de l'addition à gauche.

Donc $\alpha + \gamma \in \alpha + \beta$ par définition de $<$.

$$\text{Ainsi, on peut poser } \varphi_\beta := \begin{cases} \alpha \amalg \beta & \longrightarrow \alpha + \beta \\ (i, \gamma) & \longmapsto \begin{cases} \gamma & \text{si } i = 0 \\ \alpha + \gamma & \text{si } i = 1 \end{cases} \end{cases}$$

Le fait d'avoir mis β en indice nous servira pour une preuve par induction sur β .

- Montrons que φ_β est croissante.

Soient (i, γ) et (j, δ) dans $\alpha \amalg \beta$.

Supposons que $(i, \gamma) \leq (j, \delta)$.

► Plaçons-nous dans le cas où $i = 0 = j$.

Alors on a $\varphi_\beta(i, \gamma) = \gamma$ et $\varphi_\beta(j, \delta) = \delta$ par définition de φ_β .

Comme $(i, \gamma) \leq (j, \delta)$ on a $\gamma \leq \delta$ par définition de \leq .

On a donc $\varphi_\beta(i, \gamma) \leq \varphi_\beta(j, \delta)$.

► Plaçons-nous dans le cas où $i = 1 = j$.

Alors $\varphi_\beta(i, \gamma) = \alpha + \gamma$ et $\varphi_\beta(j, \delta) = \alpha + \delta$ par définition de φ_β .

Comme $(i, \gamma) \trianglelefteq (j, \delta)$ on a $\gamma \leq \delta$ par définition de \trianglelefteq .

On a donc $\alpha + \gamma \leq \alpha + \delta$ par croissance de l'addition à gauche.

On a donc $\varphi_\beta(i, \gamma) \leq \varphi_\beta(j, \delta)$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

Alors $\gamma \in \alpha$ (et $\delta \in \beta$) par définition de $\alpha \amalg \beta$, donc $\gamma < \alpha$.

De plus $\varphi_\beta(i, \gamma) = \gamma$ et $\varphi_\beta(j, \delta) = \alpha + \delta$ par définition de φ_β .

Or on a $\gamma < \alpha = \alpha + 0 \leq \alpha + \delta$.

On a donc $\gamma \leq \alpha + \delta$ par transitivité et donc $\varphi_\beta(i, \gamma) \leq \varphi_\beta(j, \delta)$.

► Plaçons-nous dans le cas où $i = 1$ et $j = 0$.

C'est absurde puisqu'on a fait l'hypothèse que $(i, \gamma) \trianglelefteq (j, \delta)$.

Dans tous les cas possibles on a $\varphi_\beta(i, \gamma) \leq \varphi_\beta(j, \delta)$.

Donc si $(i, \gamma) \trianglelefteq (j, \delta)$ alors $\varphi_\beta(i, \gamma) \leq \varphi_\beta(j, \delta)$.

Donc φ_β est croissante.

• Montrons que φ_β est injective.

Soit (i, γ) et (j, δ) dans $\alpha \amalg \beta$.

Supposons que $\varphi_\beta(i, \gamma) = \varphi_\beta(j, \delta)$.

► Plaçons-nous dans le cas où $i = 0 = j$.

On a alors $\varphi_\beta(i, \gamma) = \gamma$ et $\varphi_\beta(j, \delta) = \delta$ par définition de φ_β .

Comme $\varphi_\beta(i, \gamma) = \varphi_\beta(j, \delta)$ on a donc $\gamma = \delta$.

Comme $i = j$ et $\gamma = \delta$ on a $(i, \gamma) = (j, \delta)$.

► Plaçons-nous dans le cas où $i = 1 = j$.

On a alors $\varphi_\beta(i, \gamma) = \alpha + \gamma$ et $\varphi_\beta(j, \delta) = \alpha + \delta$ par définition de φ_β .

Comme $\varphi_\beta(i, \gamma) = \varphi_\beta(j, \delta)$ on a donc $\alpha + \gamma = \alpha + \delta$.

On en déduit que $\gamma = \delta$ par régularité de l'addition à gauche.

Comme $i = j$ et $\gamma = \delta$ on a $(i, \gamma) = (j, \delta)$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

On a alors $\gamma \in \alpha$ (et $\delta \in \beta$) par définition de $\alpha \amalg \beta$, donc $\gamma < \alpha$.

On a aussi $\varphi_\beta(i, \gamma) = \gamma$ et $\varphi_\beta(j, \delta) = \alpha + \delta$ par définition de φ_β .

Or on a $\gamma < \alpha = \alpha + 0 \leq \alpha + \delta$.

On a donc $\gamma < \alpha + \delta$ par transitivité, donc $\varphi_\beta(i, \gamma) < \varphi_\beta(j, \delta)$.

Or on a fait l'hypothèse que $\varphi_\beta(i, \gamma) = \varphi_\beta(j, \delta)$.

C'est absurde par antiréflexivité de $<$.

► Le cas où $i = 1$ et $j = 0$ est absurde pour la même raison.

Les deux seuls cas possibles conduisent alors à $(i, \gamma) = (j, \delta)$.

Donc si $\varphi_\beta(i, \gamma) = \varphi_\beta(j, \delta)$ alors $(i, \gamma) = (j, \delta)$.

Donc $\boxed{\varphi_\beta \text{ est injective}}$.

- Montrons que φ_β est surjective sur $\alpha + \beta$.

On sait déjà par définition de φ_β que $\text{im}(\varphi_\beta) \subseteq \alpha + \beta$.

Fixons un ordinal α et posons P l'assertion à paramètre définie pour tout ordinal β par

$$P(\beta) \iff \text{im}(\varphi_\beta) = \alpha + \beta$$

Remarquons la chose suivante : soient β et β' deux ordinaux tels que $\beta \leq \beta'$.

On a alors $\beta \subseteq \beta'$ donc $(\{1\} \times \beta) \subseteq (\{1\} \times \beta')$ et donc $\alpha \amalg \beta \subseteq \alpha \amalg \beta'$.

Or pour tout $(i, \gamma) \in \alpha \amalg \beta$ on a $\varphi_\beta(i, \gamma) = \varphi_{\beta'}(i, \gamma)$ par définition de φ_β et $\varphi_{\beta'}$.

Autrement dit on a alors $\varphi_\beta = (\varphi_{\beta'})|_{\alpha \amalg \beta}$. Notons $(*)$ ce fait.

► Initialisation

On a alors $\alpha + 0 = \alpha$ par définition de l'addition.

Montrons que $\text{im}(\varphi_0) \supseteq \alpha$.

Soit $\gamma \in \alpha$.

On a alors $\varphi_0(0, \gamma) = \gamma$ par définition de φ_0 .

Donc $\gamma \in \text{im}(\varphi_0)$.

On a donc $\text{im}(\varphi_0) \supseteq \alpha$ et donc $\text{im}(\varphi_0) = \alpha$.

Ainsi on a $\text{im}(\varphi_0) = \alpha + 0$ et donc on a $P(0)$.

► Héritéité

Soit β un ordinal tel que $P(\beta)$, c'est-à-dire $\text{im}(\varphi_\beta) = \alpha + \beta$.

On a $\alpha + S(\beta) = S(\alpha + \beta)$ par définition de l'addition.

Montrons que $\text{im}(\varphi_{S(\beta)}) \supseteq S(\alpha + \beta)$.

Soit $\gamma \in S(\alpha + \beta)$.

On a donc $\gamma < S(\alpha + \beta)$ par définition de $<$.

On a donc $\gamma \leq \alpha + \beta$ d'après la proposition 13 page 33.

On a donc $\gamma < \alpha + \beta$ ou $\gamma = \alpha + \beta$.

Plaçons-nous dans le cas où $\gamma < \alpha + \beta$.

On a donc $\gamma \in \alpha + \beta$ par définition de $<$.

Or par hypothèse on a $\text{im}(\varphi_\beta) = \alpha + \beta$ donc on a $\gamma \in \text{im}(\varphi_\beta)$.

Or $\beta < S(\beta)$ d'après la proposition 13 page 33 donc $\beta \leq S(\beta)$.

On a donc $\varphi_\beta = (\varphi_{S(\beta)})_{|\alpha \amalg \beta}$ d'après (\star) et donc $\text{im}(\varphi_\beta) \subseteq \text{im}(\varphi_{S(\beta)})$.
 Ainsi on a $\gamma \in \text{im}(\varphi_{S(\beta)})$ par définition de l'inclusion.

Plaçons-nous dans le cas où $\gamma = \alpha + \beta$.

On a $\beta < S(\beta)$ d'après la proposition 13 page 33 donc $\beta \in S(\beta)$.

Ainsi $\gamma = \alpha + \beta = \varphi_{S(\beta)}(1, \beta)$ par définition de $\varphi_{S(\beta)}$.

On a donc $\gamma \in \text{im}(\varphi_{S(\beta)})$.

Ainsi dans les deux cas on a $\gamma \in \text{im}(\varphi_{S(\beta)})$.

On a donc $\text{im}(\varphi_{S(\beta)}) \supseteq S(\alpha + \beta)$ et donc $\text{im}(\varphi_{S(\beta)}) = S(\alpha + \beta)$.

Autrement dit on a $P(S(\beta))$.

Donc pour tout ordinal β , si $P(\beta)$ alors $P(S(\beta))$.

► Héritage limite

Soit β un ordinal limite non nul tel que $\forall \delta < \beta, P(\delta)$.

Montrons que $\text{im}(\varphi_\beta) \supseteq \alpha + \beta$.

Soit $\gamma \in \alpha + \beta$.

On a donc $\gamma < \alpha + \beta$ par définition de $<$.

Par définition β est un ordinal limite non nul.

On a donc $\alpha + \beta = \sup_{\delta < \beta} (\alpha + \delta)$ par définition de l'addition.

On a donc $\gamma < \sup_{\delta < \beta} (\alpha + \delta)$ donc γ n'est pas un majorant de $\{\alpha + \delta \mid \delta < \beta\}$.

Il existe donc un ordinal $\delta < \beta$ tel que l'on a n'a pas $\alpha + \delta \leq \gamma$.

Comme \leq est total chez les ordinaux, on a donc $\gamma < \alpha + \delta$ et donc $\gamma \in \alpha + \delta$.

Or $\delta < \beta$ donc par hypothèse on a $P(\delta)$ et donc $\alpha + \delta = \text{im}(\varphi_\delta)$.

Comme $\gamma \in \alpha + \delta$ on a donc $\gamma \in \text{im}(\varphi_\delta)$.

Comme $\delta < \beta$ on a $\delta \leq \beta$ donc $\varphi_\delta = (\varphi_\beta)|_{\alpha \amalg \delta}$ d'après (\star) .

En particulier on a $\text{im}(\varphi_\delta) \subseteq \text{im}(\varphi_\beta)$ et donc $\gamma \in \text{im}(\varphi_\beta)$.

On a donc $\text{im}(\varphi_\beta) \supseteq \alpha + \beta$ et donc $\text{im}(\varphi_\beta) = \alpha + \beta$.

Autrement dit on a $P(\beta)$.

Donc pour tout ordinal limite non nul β , si $\forall \delta < \beta, P(\delta)$ alors $P(\beta)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal β on a $P(\beta)$.

Autrement dit pour tout ordinal β on a $\text{im}(\varphi_\beta) = \alpha + \beta$.

Autrement dit pour tout ordinal β , φ_β est surjective dans $\alpha + \beta$.

• Conclusion.

Fixons à nouveau un ordinal β .

Alors l'application $\varphi_\beta : \alpha \amalg \beta \longrightarrow \alpha + \beta$ est croissante, injective et surjective dans $\alpha + \beta$.

Or α et β sont bien ordonnés d'après le théorème 1 page 21.

Donc $\alpha \amalg \beta$ est bien ordonné d'après la proposition 49 page 115.

Donc $\alpha \amalg \beta$ est totalement ordonné d'après la proposition 2 page 10.

Donc φ_β est un isomorphisme d'ordres de $\alpha \amalg \beta$ dans $\alpha + \beta$.

Or $\alpha + \beta$ est un ordinal par définition.

Donc $\boxed{\text{type}(\alpha \amalg \beta) = \alpha + \beta}$ par définition du type d'un ensemble bien ordonné.

CQFD.

À l'école primaire, après l'addition vient rapidement la soustraction. Nous apprenons à cet âge-là qu'il n'est pas possible de soustraire un nombre par un nombre plus grand : la raison est simplement que cela produit un nombre négatif, notion qui n'est à ce moment-là pas abordée. C'est un peu le cas ici : il n'est pas encore temps de définir les nombres négatifs, et donc nous nous contenterons de ne pouvoir soustraire que les ordinaux plus grands par des ordinaux plus petits.

Comme remarqué plus haut, nous allons bien évidemment nous servir de la proposition 50 page 116 qui nous fournit au fond tout ce dont nous avons besoin.

Proposition 51 (Soustraction d'ordinaux)

Soient α et β deux ordinaux.

Les assertions suivantes sont équivalentes :

1. $\beta \leq \alpha$
2. Il existe un ordinal σ tel que $\alpha = \beta + \sigma$.

Dans ce cas-là, un tel ordinal σ est unique, et vérifie $\sigma \leq \alpha$.

Plus précisément on a $\sigma = \text{type}(\alpha \setminus \beta)$, et donc $\alpha = \beta + \text{type}(\alpha \setminus \beta)$.

Démonstration

Commençons par montrer l'équivalence.

$1 \Rightarrow 2$

Supposons que $\beta \leq \alpha$.

On a alors $\alpha \cong \beta \amalg (\alpha \setminus \beta)$ d'après la proposition 50 page 116.

Par définition du type on a $\alpha \setminus \beta \cong \text{type}(\alpha \setminus \beta)$.

On a aussi $\beta \cong \beta$ par réflexivité de l'isomorphie d'ordres.

On a donc $\beta \amalg (\alpha \setminus \beta) \cong \beta \amalg \text{type}(\alpha \setminus \beta)$ d'après la proposition 50 page 116.

On a donc les isomorphies d'ordres suivantes :

$$\begin{aligned}\alpha &\cong \beta \amalg (\alpha \setminus \beta) \\ &\cong \beta \amalg \text{type}(\alpha \setminus \beta) \text{ par ce qui précède} \\ &\cong \text{type}(\beta \amalg \text{type}(\alpha \setminus \beta)) \text{ par définition du type} \\ &\cong \beta + \text{type}(\alpha \setminus \beta) \text{ d'après le théorème 7 page 120}\end{aligned}$$

Ainsi $\alpha \cong \beta + \text{type}(\alpha \setminus \beta)$ par transitivité de l'isomorphie d'ordres.

Or deux ordinaux isomorphes sont nécessairement égaux d'après la proposition 25 page 53.

On a donc $\boxed{\alpha = \beta + \text{type}(\alpha \setminus \beta)}$.

Remarquons que $\alpha \setminus \beta \subseteq \alpha$ par définition de la différence ensembliste.

On a donc $\text{type}(\alpha \setminus \beta) \leq \text{type}(\alpha)$ d'après la proposition 27 page 60.

Or α est un ordinal donc $\text{type}(\alpha) = \alpha$, si bien que $\text{type}(\alpha \setminus \beta) \leq \alpha$.

Autrement dit en notant $\sigma := \text{type}(\alpha \setminus \beta)$, on a $\boxed{\alpha = \beta + \sigma \text{ et } \sigma \leq \alpha}$.

$\boxed{1 \Leftarrow 2}$

Supposons qu'il existe un ordinal σ tel que $\alpha = \beta + \sigma$.

Ainsi on a $\beta = \beta + 0 \leq \beta + \sigma = \alpha$ donc $\boxed{\beta \leq \alpha}$.

Montrons l'unicité

On se place à présent dans le cas où effectivement $\beta \leq \alpha$.

Soit σ' un ordinal tel que $\alpha = \beta + \sigma'$.

On a $\sigma < \sigma'$ ou $\sigma' < \sigma$ ou $\sigma = \sigma'$ d'après le théorème 1 page 21.

Si $\sigma < \sigma'$ alors $\beta + \sigma < \beta + \sigma'$ par stricte croissance de l'addition à gauche.

De même si $\sigma' < \sigma$ alors $\beta + \sigma' < \beta + \sigma$.

Dans ces deux cas on a donc $\alpha < \alpha$, ce qui est absurde par antiréflexivité de $<$.

On a donc nécessairement $\sigma = \sigma'$.

On a donc $\boxed{\text{unicité d'un tel ordinal } \sigma}$.

CQFD.

3 Multiplication d'ordinaux

3.1 Définition et propriétés

Pour définir la multiplication chez les ordinaux, on peut à nouveau s'inspirer de la multiplication chez les entiers naturels. Comment définir 5×3 ? Intuitivement il s'agit de $5 + 5 + 5$, c'est-à-dire une répétition d'additions de 5, où le nombre 5 est répété 3 fois. Ce n'est cependant pas une définition très pratique ici, car on n'a pas particulièrement donné de sens à "*répéter 3 fois une addition*". Il nous faudrait plutôt une définition par récursion, puisqu'on a développé tous les outils pour cela précédemment. On peut remarquer que $5 \times 2 = 5 + 5$, si bien que $5 \times 3 = 5 + 5 + 5 = (5 + 5) + 5 = (5 \times 2) + 5$. En voilà une définition par récursion!

Ainsi, pour définir 5×3 on considère que 5×2 est déjà défini, puis on pose $5 \times 3 := (5 \times 2) + 5$. Autrement dit, on a posé $5 \times (2 + 1) := (5 \times 2) + 5$. Cela nous guide donc vers la définition suivante.

Définition 22 (Multiplication d'ordinaux)

Soit α un ordinal.

On pose

$$\left\{ \begin{array}{l} \alpha \cdot 0 := 0 \\ \alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha \text{ pour tout ordinal } \beta \\ \alpha \cdot \gamma := \sup_{\delta < \gamma} (\alpha \cdot \delta) \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

Remarque :

1. Comme on peut le voir dans la définition, pour deux ordinaux α et β on note généralement $\alpha \cdot \beta$ plutôt que $\alpha \times \beta$. Cela évitera au passage la confusion avec le produit cartésien, même si nous verrons qu'il y a un lien entre les deux. Parfois-même on n'écrira aucun symbole entre les deux : on notera $\alpha\beta$ à la place de $\alpha \cdot \beta$.
2. Afin de simplifier les expressions, on considère désormais que la multiplication des ordinaux est prioritaire sur l'addition des ordinaux. Ainsi, l'expression $\alpha \cdot \beta + \gamma$ désigne $(\alpha \cdot \beta) + \gamma$ et non $\alpha \cdot (\beta + \gamma)$.
3. Pour justifier proprement cette définition, on utilise simplement la proposition 36 page 91, en posant $\mu_0 := 0$, et $G(\xi) := \xi + \alpha$ pour tout ordinal ξ . La proposition nous donne alors une unique assertion fonctionnelle F_α telle que

$$\left\{ \begin{array}{l} F_\alpha(0) = 0 \\ F_\alpha(\beta + 1) = F_\alpha(\beta) + \alpha \text{ pour tout ordinal } \beta \\ F_\alpha(\gamma) = \sup_{\delta < \gamma} F_\alpha(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

et on pose alors $\alpha \cdot \beta := F_\alpha(\beta)$ pour tout ordinal β .

Exemple :

On a $3\omega = \sup_{n<\omega} (3n) = \sup\{0, 3, 6, 9, 12, \dots\} = \omega$.

En fait plus généralement pour tout entier naturel m on a

$$m\omega = \sup\{m \cdot 0, m \cdot 1, m \cdot 2, \dots\} = \omega$$

En particulier $2\omega = \omega$. Pourtant, on a $\omega \cdot 2 = \omega \cdot 1 + \omega = \omega + \omega$. Or on a vu dans un exemple précédent que $\omega < \omega + \omega$, si bien que $2\omega < \omega \cdot 2$ et donc $2\omega \neq \omega \cdot 2$. Et oui, la multiplication des ordinaux n'est elle non plus pas commutative ! Heureusement nous verrons qu'elle l'est quand on se restreint aux entiers naturels !

Proposition 52 (0 est absorbant pour la multiplication)

Pour tout ordinal α , on a $\alpha \cdot 0 = 0 = 0 \cdot \alpha$.

On dit que 0 est **absorbant** pour la multiplication des ordinaux.

Démonstration

On sait déjà que pour tout ordinal α , on a $\alpha \cdot 0 = \alpha$ par définition de la multiplication.

Montrons l'autre égalité par induction.

Considérons P l'assertion à paramètres définie pour tout ordinal α par

$$P(\alpha) \iff 0 \cdot \alpha = 0$$

► Initialisation

On a $0 \cdot 0 = 0$ par définition de la multiplication, et donc $P(0)$.

► Hérédité

Soit α un ordinal tel que $P(\alpha)$.

On a donc

$$\begin{aligned} 0 \cdot (\alpha + 1) &= 0 \cdot \alpha + 0 \text{ par définition de la multiplication} \\ &= 0 \cdot \alpha \text{ par neutralité de } 0 \text{ pour l'addition} \\ &= 0 \text{ puisqu'on a } P(\alpha) \end{aligned}$$

Ainsi on a $0 \cdot (\alpha + 1) = 0$ et donc $P(\alpha + 1)$.

Ainsi pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

► Hérédité limite

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

Autrement dit pour tout $\beta < \alpha$, on a $0 \cdot \beta = 0$.

Par définition de la multiplication on a donc $0 \cdot \alpha = \sup_{\beta < \alpha} (0 \cdot \beta) = \sup_{\beta < \alpha} 0 = 0$.

Autrement dit on a $P(\alpha)$.

Ainsi pour tout ordinal limite non nul α , si $\forall \beta \in \alpha, P(\beta)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α , on a $P(\alpha)$.

Autrement dit pour tout ordinal α on a $0 \cdot \alpha = 0$.

CQFD.

Proposition 53 (1 est neutre pour la multiplication)

Pour tout ordinal α , on a $\alpha \cdot 1 = \alpha = 1 \cdot \alpha$.

On dit que 1 est **neutre** pour la multiplication des ordinaux.



Démonstration

Pour tout ordinal α on a

$$\begin{aligned}\alpha \cdot 1 &= \alpha \cdot (0 + 1) \text{ par définition de 1} \\ &= \alpha \cdot 0 + \alpha \text{ par définition de la multiplication} \\ &= 0 + \alpha \text{ par définition de la multiplication} \\ &= \alpha \text{ par neutralité de 0 pour l'addition}\end{aligned}$$

Ainsi on sait déjà que pour tout ordinal α on a $\alpha \cdot 1 = \alpha$.

Montrons l'autre égalité par induction.

Considérons P l'assertion à paramètre définie pour tout ordinal α par

$$P(\alpha) \iff 1 \cdot \alpha = \alpha$$

► Initialisation

On a $1 \cdot 0 = 0$ par définition de la multiplication, on a donc $P(0)$.

► Hérédité

Soit α un ordinal tel que $P(\alpha)$.

On a alors

$$\begin{aligned}1 \cdot (\alpha + 1) &= 1 \cdot \alpha + 1 \text{ par définition de l'addition} \\ &= \alpha + 1 \text{ puisque l'on a } P(\alpha)\end{aligned}$$

Ainsi on a $1 \cdot (\alpha + 1) = \alpha + 1$.

Autrement dit on a $P(\alpha + 1)$.

Ainsi pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

► *Héritage limite*

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

Autrement dit pour tout $\beta < \alpha$ on a $1 \cdot \beta = \beta$.

Or α est un ordinal limite donc on a $\sup_{\beta < \alpha} \beta = \sup_{\beta < \alpha} (\alpha) = \alpha$ d'après la prop. 21 p. 47.

On a donc $1 \cdot \alpha = \sup_{\beta < \alpha} (1 \cdot \beta) = \sup_{\beta < \alpha} \beta = \alpha$ par définition de la multiplication.

Ainsi on a $1 \cdot \alpha = \alpha$, et donc on a $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α , on a $P(\alpha)$.

Autrement dit pour tout ordinal α on a $1 \cdot \alpha = \alpha$.

CQFD.

De même que l'addition de deux entiers naturels donne toujours un entier naturel, nous sommes bien heureux qu'il en soit de même pour la multiplication.

Proposition 54 (Multiplication de deux entiers naturels)

Pour tout entiers naturels n et m , l'ordinal nm est un entier naturel.

On dit que $\mathbb{N} = \omega$ est **stable** par multiplication.

 *Démonstration*

Fixons n un entier naturel.

Soit P l'assertion à paramètre définie pour tout entier naturel m par

$$P(m) \iff nm \in \mathbb{N}$$

Raisonnons par induction sur les entiers naturels.

► *Initialisation*

On a $n \cdot 0 = 0$ par définition de la multiplication.

Or 0 est un entier naturel donc $n \cdot 0$ est un entier naturel : on a donc $P(0)$.

► *Héritéité*

Soit m un entier naturel tel que $P(m)$.

Autrement dit nm est un entier naturel.

Donc $nm + n$ est un entier naturel d'après la proposition 39 page 96.

Or on a $n(m + 1) = nm + n$ par définition de la multiplication.

Donc $n(m + 1)$ est un entier naturel : on a donc $P(m + 1)$.

Donc pour tout entier naturel m , si $P(m)$ alors $P(m + 1)$.

Ainsi P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel m on a $P(m)$.

Autrement dit [pour tout entier naturel m , l'ordinal nm est un entier naturel].

CQFD.

Dans la proposition qui suit, la stricte croissance nécessite forcément que α soit non nul car 0 est absorbant pour la multiplication.

Proposition 55 (Croissance de la multiplication des ordinaux)

Soient α, β et γ trois ordinaux.

1. Supposons ici que α est **non nul**.

Si $\beta < \gamma$ alors $\alpha\beta < \alpha\gamma$.

On dit que la multiplication à gauche est **strictement croissante**.

2. Si $\beta \leq \gamma$ alors $\alpha\beta \leq \alpha\gamma$.

On dit que la multiplication à gauche est **croissante**.

3. Si $\beta \leq \gamma$ alors $\beta\alpha \leq \gamma\alpha$.

On dit que la multiplication à droite est **croissante**.



Démonstration

1. Fixons deux ordinaux α et β , avec α non nul.

Posons $P_{\alpha,\beta}$ l'assertion à paramètre définie pour tout ordinal γ par

$$P_{\alpha,\beta}(\gamma) \iff (\beta < \gamma \Rightarrow \alpha\beta < \alpha\gamma)$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

Il est faux de dire que $\beta \in \emptyset$ donc il est faux de dire que $\beta \in 0$ et donc de dire $\beta < 0$.

La prémissse $\beta < 0$ étant fausse, on a l'implication $\beta < 0 \Rightarrow \alpha\beta < \alpha \cdot 0$.

Autrement dit on a $P_{\alpha,\beta}(0)$.

► *Héritage*

Soit γ un ordinal tel que $P_{\alpha,\beta}(\gamma)$.

Commençons par remarquer que l'on a :

$$\alpha\gamma = \alpha\gamma + 0 \text{ car } 0 \text{ est neutre pour l'addition}$$

$$\begin{aligned} &< \alpha\gamma + \alpha \text{ par stricte croissance de l'addition à gauche et } \alpha \neq 0 \\ &= \alpha(\gamma + 1) \text{ par définition de la multiplication} \end{aligned}$$

Ainsi on a $\alpha\gamma < \alpha(\gamma + 1)$.

Supposons que $\beta < \gamma + 1$.

On a alors $\beta \leq \gamma$ d'après la proposition 13 page 33.

On a donc $\beta < \gamma$ ou $\beta = \gamma$.

- Plaçons-nous dans le cas où $\beta < \gamma$.

On a alors $\alpha\beta < \alpha\gamma$ d'après $P_{\alpha,\beta}(\gamma)$.

Or on a dit que $\alpha\gamma < \alpha(\gamma + 1)$, donc $\alpha\beta < \alpha(\gamma + 1)$ par transitivité de $<$.

- Plaçons-nous dans le cas où $\beta = \gamma$.

On a donc $\alpha\beta = \alpha\gamma$.

Or on a dit que $\alpha\gamma < \alpha(\gamma + 1)$, donc $\alpha\beta < \alpha(\gamma + 1)$.

Dans les deux cas on a $\alpha\beta < \alpha(\gamma + 1)$.

Donc si $\beta < \gamma + 1$ alors $\alpha\beta < \alpha(\gamma + 1)$.

Autrement dit on a $P_{\alpha,\beta}(\gamma + 1)$.

Donc pour tout ordinal γ , si $P_{\alpha,\beta}(\gamma)$ alors $P_{\alpha,\beta}(\gamma + 1)$.

► *Héritage limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

Supposons que $\beta < \gamma$.

On a donc $\beta + 1 < \gamma$ d'après la proposition 14 page 37 car γ est limite.

On a donc $\alpha(\beta + 1) \leq \sup_{\delta < \gamma} (\alpha\delta)$ car la borne supérieure est un majorant.

Comme γ est limite, on a $\alpha\gamma = \sup_{\delta < \gamma} (\alpha\delta)$ donc $\alpha(\beta + 1) \leq \alpha\gamma$.

De plus par hypothèse on a $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

Or on a dit que $\beta + 1 < \gamma$ donc $P_{\alpha,\beta}(\beta + 1)$.

Autrement dit on a $\beta < \beta + 1 \Rightarrow \alpha\beta < \alpha(\beta + 1)$.

Or on a $\beta < \beta + 1$ d'après la proposition 13 page 33.

On a donc $\alpha\beta < \alpha(\beta + 1)$ par modus ponens.

Ainsi on a $\alpha\beta < \alpha(\beta + 1) \leq \alpha\gamma$ donc $\alpha\beta < \alpha\gamma$.

Donc si $\beta < \gamma$ alors $\alpha\beta < \alpha\gamma$.

Autrement dit on a $P_{\alpha,\beta}(\gamma)$.

Donc pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$ alors $P_{\alpha,\beta}(\gamma)$.

Ainsi $P_{\alpha,\beta}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ , on a $P_{\alpha,\beta}(\gamma)$.

Autrement dit pour tout ordinal γ , si $\beta < \gamma$ alors $\alpha\beta < \alpha\gamma$.

2. Fixons trois ordinaux α, β et γ .

Plaçons-nous dans un premier temps dans le cas où $\alpha = 0$.

On a alors $\alpha\beta = 0 = \alpha\gamma$ car 0 est absorbant pour la multiplication.

On a donc $\alpha\beta \leq \alpha\gamma$ par réflexivité de \leq .

En particulier on a l'implication $\beta \leq \gamma \Rightarrow \alpha\beta \leq \alpha\gamma$.

Plaçons-nous à présent dans le cas où α est non nul.

Supposons que $\beta \leq \gamma$.

On a donc $\beta < \gamma$ ou $\beta = \gamma$.

► Plaçons-nous dans le cas où $\beta < \gamma$.

Comme $\alpha \neq 0$, on a $\alpha\beta < \alpha\gamma$ d'après 1, donc $\alpha\beta \leq \alpha\gamma$.

► Plaçons-nous dans le cas où $\beta = \gamma$.

On a donc $\alpha\beta = \alpha\gamma$, donc $\alpha\beta \leq \alpha\gamma$ par réflexivité de \leq .

Dans les deux cas on a $\alpha\beta \leq \alpha\gamma$.

Donc si $\beta \leq \gamma$ alors $\alpha\beta \leq \alpha\gamma$.

3. Fixons deux ordinaux β et γ .

Supposons que $\beta \leq \gamma$.

Posons $Q_{\beta,\gamma}$ l'assertion à paramètre définie pour tout ordinal α par

$$Q_{\beta,\gamma}(\alpha) \iff \beta\alpha \leq \gamma\alpha$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

Par définition de la multiplication on a $\beta \cdot 0 = 0 = \gamma \cdot 0$.

En particulier on a $\beta \cdot 0 \leq \gamma \cdot 0$, c'est-à-dire $Q_{\beta,\gamma}(0)$.

► *Hérédité*

Soit α un ordinal tel que $Q_{\beta,\gamma}(\alpha)$.

Autrement dit on a $\beta\alpha \leq \gamma\alpha$.

On a alors

$$\begin{aligned}\beta(\alpha + 1) &= \beta\alpha + \beta \text{ par définition de la multiplication} \\ &\leq \gamma\alpha + \beta \text{ par croissance de l'addition à droite} \\ &\leq \gamma\alpha + \gamma \text{ par croissance de l'addition à gauche} \\ &= \gamma(\alpha + 1) \text{ par définition de la multiplication}\end{aligned}$$

On a donc $\beta(\alpha + 1) \leq \gamma(\alpha + 1)$, donc $Q_{\beta,\gamma}(\alpha + 1)$.

Donc pour tout ordinal α , si $Q_{\beta,\gamma}(\alpha)$ alors $Q_{\beta,\gamma}(\alpha + 1)$..

► *Hérédité limite*

Soit α un ordinal limite non nul tel que $\forall \delta < \alpha, Q_{\beta,\gamma}(\delta)$.

Soit δ un ordinal tel que $\delta < \alpha$. On a alors

$$\begin{aligned}\beta\delta &\leq \gamma\delta \text{ d'après } Q_{\beta,\gamma}(\delta) \\ &\leq \sup_{\varepsilon < \alpha} (\gamma\varepsilon) \text{ car la borne supérieure est un majorant} \\ &= \gamma\alpha \text{ par définition de la multiplication}\end{aligned}$$

On a donc $\beta\delta \leq \gamma\alpha$.

Donc pour tout ordinal δ tel que $\delta < \alpha$ on a $\beta\delta \leq \gamma\alpha$.

Donc $\sup_{\delta < \alpha} (\beta\delta) \leq \gamma\alpha$ par minimalité de la borne supérieure.

Donc $\beta\alpha \leq \gamma\alpha$ par définition de la multiplication, donc $Q_{\beta,\gamma}(\alpha)$.

Donc pour tout ordinal limite non nul γ , si $\forall \delta < \alpha, Q_{\beta,\gamma}(\delta)$ alors $Q_{\beta,\gamma}(\alpha)$.

Ainsi $Q_{\beta,\gamma}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $Q_{\beta,\gamma}(\alpha)$.

Autrement dit pour tout ordinal α , on a $\beta\alpha \leq \gamma\alpha$.

Donc pour tout ordinal α , si $\beta \leq \gamma$ alors $\beta\alpha \leq \gamma\alpha$.

CQFD.

Remarque :

Nous pouvons désormais montrer que $2\omega = \omega$ et que $2\omega \neq \omega \cdot 2$.

Pour tout $n \in \omega$ on a $2n \in \omega$ d'après la proposition 54 page 129.

Donc pour tout ordinal $n < \omega$ on a $2n < \omega$ par définition de $<$.

On a donc $2\omega = \sup_{n<\omega} (2n) \leq \omega$ par minimalité de la borne supérieure, donc $2\omega \leq \omega$.

Par définition de 2 on a $2 = S(1)$.

On a donc $1 < 2$ d'après la proposition 13 page 33, donc $1 \leq 2$.

On a donc $1\omega \leq 2\omega$ par croissance de la multiplication à droite.

Or on a $1\omega = \omega$ car 1 est neutre pour la multiplication.

On a donc $\omega \leq 2\omega$ et donc finalement $\boxed{\omega = 2\omega}$.

Mais on a dit que $1 < 2$ donc $\omega \cdot 1 < \omega \cdot 2$ par stricte croissance de la multiplication.

Or on a $\omega \cdot 1 = \omega$ car 1 est neutre pour la multiplication.

On a donc $\omega < \omega \cdot 2$, donc en particulier $\omega \neq \omega \cdot 2$.

On en déduit que $\boxed{2\omega \neq \omega \cdot 2}$ par ce qui précède.

Proposition 56 (Régularité de la multiplication à gauche)

Soient α, β et γ trois ordinaux, avec α **non nul**.

Si $\alpha\beta = \alpha\gamma$ alors $\beta = \gamma$.

On dit que la multiplication à gauche des ordinaux est **régulière**.



Démonstration

Montrons-le par contraposition.

Supposons que $\beta \neq \gamma$.

On a donc $\beta < \gamma$ ou $\gamma < \beta$ d'après le théorème 1 page 21.

Si $\beta < \gamma$ alors $\alpha\beta < \alpha\gamma$ par stricte croissance de la multiplication à gauche.

Si $\gamma < \beta$ alors $\alpha\gamma < \alpha\beta$ par stricte croissance de la multiplication à gauche.

Dans les deux cas on a $\alpha\beta \neq \alpha\gamma$ par antiréflexivité de $<$.

Donc si $\beta \neq \gamma$ alors $\alpha\beta \neq \alpha\gamma$.

Donc par contraposition, $\boxed{\text{si } \alpha\beta = \alpha\gamma \text{ alors } \beta = \gamma}$.

CQFD.

Remarque :

Malheureusement la multiplication à droite n'est pas régulière.

En effet, on a vu que $1\omega = \omega = 2\omega$ alors que $1 \neq 2$.

Proposition 57 (Continuité à droite de la multiplication)

Soient α un ordinal et X un ensemble **non vide** d'ordinaux.

On a $\sup_{\xi \in X} (\alpha\xi) = \alpha \sup_{\xi \in X} \xi$.

Autrement dit la multiplication à gauche est continue.

 *Démonstration*

Par définition de la multiplication des ordinaux, pour tout ordinal limite non nul γ , on a

$$\alpha\gamma = \sup_{\delta \in \gamma} (\alpha\delta)$$

On peut alors appliquer la proposition 42 page 103 pour conclure.

CQFD.

Remarque :

Malheureusement, la multiplication à droite des ordinaux n'est pas continue.

En effet, prenons $X = \omega = \alpha$. À la manière de la remarque où l'on a montré que $2\omega = \omega$, on peut montrer que pour tout entier naturel non nul n , on a $n\omega = \omega$.

On a donc $\sup_{n < \omega} (n\omega) = \sup_{n < \omega} \omega = \omega$.

Mais comme ω est un ordinal limite, on a $\sup_{n < \omega} n = \omega$ donc $\left(\sup_{n < \omega} n\right)\omega = \omega\omega$.

Or par définition de la multiplication on a $\omega\omega = \sup_{n < \omega} (\omega n)$.

En particulier on a $\omega \cdot 2 \leq \sup_{n < \omega} (\omega n)$ donc $\omega \cdot 2 \leq \omega\omega$.

Or on a montré plus tôt que $\omega < \omega \cdot 2$, donc $\omega < \omega\omega$ par transitivité de $<$.

En particulier $\omega \neq \omega\omega$ et donc $\sup_{n < \omega} (n\omega) \neq \left(\sup_{n < \omega} n\right)\omega$.

Proposition 58 (Distributivité de la multiplication sur l'addition)

Pour tout ordinaux α, β et γ , on a l'égalité

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

On dit que la multiplication à gauche est **distributive** sur l'addition.

 *Démonstration*

Fixons deux ordinaux α et β .

Posons $P_{\alpha,\beta}$ l'assertion à paramètre définie pour tout γ par

$$P_{\alpha,\beta}(\gamma) \iff \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

On a

$$\begin{aligned}
 \alpha(\beta + 0) &= \alpha\beta \text{ car } 0 \text{ est neutre pour l'addition} \\
 &= \alpha\beta + 0 \text{ car } 0 \text{ est neutre pour l'addition} \\
 &= \alpha\beta + \alpha \cdot 0 \text{ car } 0 \text{ est absorbant pour la multiplication}
 \end{aligned}$$

On a donc $\alpha(\beta + 0) = \alpha\beta + \alpha \cdot 0$, c'est-à-dire $P_{\alpha,\beta}(0)$.

► *Héritéité*

Soit γ un ordinal tel que $P_{\alpha,\beta}(\gamma)$.

On a alors

$$\begin{aligned}
 \alpha(\beta + \gamma + 1) &= \alpha(\beta + \gamma) + \alpha \text{ par définition de la multiplication} \\
 &= \alpha\beta + \alpha\gamma + \alpha \text{ par } P_{\alpha,\beta}(\gamma) \\
 &= \alpha\beta + \alpha(\gamma + 1) \text{ par définition de la multiplication}
 \end{aligned}$$

Ainsi on a $\alpha(\beta + \gamma + 1) = \alpha\beta + \alpha(\gamma + 1)$, c'est-à-dire $P_{\alpha,\beta}(\gamma + 1)$.

Donc pour tout ordinal γ , si $P_{\alpha,\beta}(\gamma)$ alors $P_{\alpha,\beta}(\gamma + 1)$.

► *Héritéité limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

On a alors

$$\begin{aligned}
 \alpha(\beta + \gamma) &= \alpha \sup_{\delta < \gamma} (\beta + \delta) \text{ par définition de l'addition} \\
 &= \sup_{\delta < \gamma} (\alpha(\beta + \delta)) \text{ par continuité de la multiplication à gauche} \\
 &= \sup_{\delta < \gamma} (\alpha\beta + \alpha\delta) \text{ puisque } \forall \delta < \gamma, P_{\alpha,\beta}(\delta) \\
 &= \alpha\beta + \sup_{\delta < \gamma} (\alpha\delta) \text{ par continuité de l'addition à gauche} \\
 &= \alpha\beta + \alpha\gamma \text{ par définition de la multiplication}
 \end{aligned}$$

Ainsi on a $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Autrement dit on a $P_{\alpha,\beta}(\gamma)$.

Ainsi pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$ alors $P_{\alpha,\beta}(\gamma)$.

Ainsi $P_{\alpha,\beta}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $P_{\alpha,\beta}(\gamma)$, c'est-à-dire $\boxed{\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma}$.

CQFD.

Remarque :

Malheureusement la multiplication à droite n'est pas distributive.
 En effet, on a déjà vu que $(1 + 1)\omega = 2\omega = \omega$.
 On a aussi vu que $1\omega + 1\omega = \omega + \omega$.
 Comme $\omega \neq \omega + \omega$, on a $(1 + 1)\omega \neq 1\omega + 1\omega$.

Proposition 59 (Associativité de la multiplication)

Pour tout ordinaux α , β et γ , on a l'égalité

$$(\alpha\beta)\gamma = \alpha(\beta\gamma)$$

On dit que la multiplication des ordinaux est **associative**.

Démonstration

Fixons deux ordinaux α et β .

Posons $P_{\alpha,\beta}$ l'assertion à paramètre définie pour tout ordinal γ par

$$P_{\alpha,\beta}(\gamma) \iff (\alpha\beta)\gamma = \alpha(\beta\gamma)$$

Montrons le résultat par le principe faible d'induction.

► Initialisation

On a $(\alpha\beta) \cdot 0 = 0 = \alpha \cdot 0 = \alpha(\beta \cdot 0)$ car 0 est absorbant pour la multiplication des ordinaux.

On a donc $P_{\alpha,\beta}(0)$.

► Héritéité

Soit γ un ordinal tel que $P_{\alpha,\beta}(\gamma)$.

On a alors

$$\begin{aligned} (\alpha\beta)(\gamma + 1) &= (\alpha\beta)\gamma + \alpha\beta \text{ par définition de la multiplication} \\ &= \alpha(\beta\gamma) + \alpha\beta \text{ par } P_{\alpha,\beta}(\gamma) \\ &= \alpha(\beta\gamma + \beta) \text{ par distributivité} \\ &= \alpha(\beta(\gamma + 1)) \text{ par définition de la multiplication} \end{aligned}$$

Ainsi on a $(\alpha\beta)(\gamma + 1) = \alpha(\beta(\gamma + 1))$.

Autrement dit on a $P_{\alpha,\beta}(\gamma + 1)$.

Donc pour tout ordinal γ , si $P_{\alpha,\beta}(\gamma)$ alors $P_{\alpha,\beta}(\gamma + 1)$.

► *Héritéité limite*

Soit γ un ordinal limite tel que $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$.

On a alors

$$\begin{aligned}
 (\alpha\beta)\gamma &= \sup_{\delta < \gamma} ((\alpha\beta)\delta) \text{ par définition de la multiplication} \\
 &= \sup_{\delta < \gamma} (\alpha(\beta\delta)) \text{ car } \forall \delta < \gamma, P_{\alpha,\beta}(\delta) \\
 &= \alpha \sup_{\delta < \gamma} (\beta\delta) \text{ par continuité de la multiplication à gauche} \\
 &= \alpha(\beta\gamma) \text{ par définition de la multiplication}
 \end{aligned}$$

Ainsi on a $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

Autrement dit on a $P_{\alpha,\beta}(\gamma)$.

Ainsi pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, P_{\alpha,\beta}(\delta)$ alors $P_{\alpha,\beta}(\gamma)$.

Ainsi $P_{\alpha,\beta}$ vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $P_{\alpha,\beta}(\gamma)$.

Autrement dit pour tout ordinal γ , on a $\boxed{(\alpha\beta)\gamma = \alpha(\beta\gamma)}$.

CQFD.

Remarque :

Désormais pour trois ordinaux α , β et γ trois ordinaux on notera $\alpha\beta\gamma$ pour désigner indifféremment $(\alpha\beta)\gamma$ et $\alpha(\beta\gamma)$, puisque l'on vient de voir qu'il s'agit du même ordinal.

On a vu lors de la proposition 52 page 127 que 0 est absorbant pour la multiplication.

Autrement dit pour deux ordinaux, si au moins l'un des deux est nul alors leur produit sera nul. Mais la réciproque est-elle vraie ? Autrement dit, si le produit de deux ordinaux est nul, cela veut-il dire qu'au moins un des deux est nul ? La réponse est oui : on parle cette fois d'intégrité.

Proposition 60 (Intégrité de la multiplication des ordinaux)

Soient α et β deux ordinaux.

Si $\alpha\beta = 0$ alors ($\alpha = 0$ ou $\beta = 0$).

On dit que la multiplication des ordinaux est **intègre**.

Démonstration

Supposons que $\alpha\beta = 0$ et que $\alpha \neq 0$.

On sait que $\alpha \cdot 0 = 0$ car 0 est absorbant pour la multiplication.

Ainsi on a $\alpha\beta = \alpha \cdot 0$.

Or α est non nul par hypothèse.

On a donc $\beta = 0$ par régularité de la multiplication à gauche.

Donc si $(\alpha\beta = 0 \text{ et } \alpha \neq 0)$ alors $\beta = 0$.

Donc $\boxed{\text{si } \alpha\beta = 0 \text{ alors } (\alpha = 0 \text{ ou } \beta = 0)}$.

CQFD.

On a vu lors d'exemple précédents que la multiplication entre ordinaux n'est pas commutative. Heureusement, elle l'est chez les entiers naturels.

Proposition 61 (Commutativité de la multiplication des entiers)

Soient n et m deux entiers naturels.

Alors $mn = nm$.

On dit que la multiplication des entiers naturels est **commutative**.



Démonstration

- On a vu que la multiplication à gauche est distributive sur l'addition.

Montrons que chez les entiers naturels, la multiplication à droite l'est aussi.

Fixons p et q deux entiers naturels.

Pour tout entier naturel n , posons $P(n)$ l'assertion « $(p + q)n = pn + qn$ ».

Initialisation

On a

$$\begin{aligned} (p + q) \cdot 0 &= 0 \text{ car } 0 \text{ est absorbant pour la multiplication} \\ &= 0 + 0 \text{ car } 0 \text{ est neutre pour l'addition} \\ &= p \cdot 0 + q \cdot 0 \text{ car } 0 \text{ est absorbant pour la multiplication} \end{aligned}$$

Ainsi on a $(p + q) \cdot 0 = p \cdot 0 + q \cdot 0$ et donc $P(0)$.

Hérédité

Soit n un entier naturel tel que $P(n)$.

On a alors

$$\begin{aligned} (p + q)(n + 1) &= (p + q)n + (p + q) \cdot 1 \text{ par distributivité de la multiplication à gauche} \\ &= (p + q)n + (p + q) \text{ car } 1 \text{ est neutre pour la multiplication} \\ &= (pn + qn) + (p + q) \text{ d'après } P(n) \\ &= pn + qn + p + q \text{ par associativité de l'addition} \end{aligned}$$

$$\begin{aligned}
 &= pn + p + qn + q \text{ par commutativité de l'addition chez les entiers naturels} \\
 &= pn + p \cdot 1 + qn + q \cdot 1 \text{ car } 1 \text{ est neutre pour la multiplication} \\
 &= p(n+1) + q(n+1) \text{ par distributivité de la multiplication à gauche}
 \end{aligned}$$

Ainsi $(p+q)(n+1) = p(n+1) + q(n+1)$ et donc $P(n+1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n+1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$, c'est-à-dire $(p+q)n = pn + qn$.

- Montrons le résultat attendu.

Fixons un entier naturel m .

Pour tout entier naturel n , posons $Q(n)$ l'assertion « $mn = nm$ ».

Initialisation

On a $m \cdot 0 = 0 = 0m$ car 0 est absorbant pour la multiplication.

On a donc $Q(0)$.

Hérité

Soit n un entier naturel tel que $Q(n)$.

On a alors

$$\begin{aligned}
 m(n+1) &= mn + m \cdot 1 \text{ par distributivité de la multiplication à gauche} \\
 &= nm + m \cdot 1 \text{ d'après } Q(n) \\
 &= nm + 1m \text{ car } 1 \text{ est neutre pour la multiplication} \\
 &= (n+1)m \text{ par distributivité de la multiplication à droite}
 \end{aligned}$$

Ainsi $m(n+1) = (n+1)m$ et donc $Q(n+1)$.

Ainsi pour tout entier naturel n , si $Q(n)$ alors $Q(n+1)$.

Finalement Q vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $Q(n)$, c'est-à-dire $mn = nm$.

CQFD.

On peut remarquer que $f^{5 \cdot 3} = f^{5+5+5} = f^5 \circ f^5 \circ f^5 = (f^5)^3$.

Plus généralement, on a la propriété suivante.

Proposition 62 (Itérées d'une application et multiplication)

Soient E un ensemble et $f : E \longrightarrow E$.

Pour tout entiers naturels n et m , on a $f^{nm} = (f^n)^m$.

Démonstration

Fixons n un entier naturel.

Pour tout entier naturel m , on pose $P(m)$ l'assertion « $f^{nm} = (f^n)^m$ ».

Initialisation

On a $f^{n \cdot 0} = f^0 = \text{id}_E = (f^n)^0$ et donc on a $P(0)$.

Hérédité

Soit m un entier naturel tel que $P(m)$, c'est-à-dire $f^{nm} = (f^n)^m$.

On a alors

$$\begin{aligned} f^{n(m+1)} &= f^{nm+n} \text{ par définition de la multiplication} \\ &= f^{nm} \circ f^n \text{ d'après la prop. 46 p. 109} \\ &= (f^n)^m \circ f^n \text{ d'après } P(m) \\ &= (f^n)^{m+1} \text{ par définition des itérées de } f^n \end{aligned}$$

On a donc $f^{n(m+1)} = (f^n)^{m+1}$ et donc $P(m+1)$.

Ainsi pour tout entier naturel m , si $P(m)$ alors $P(m+1)$.

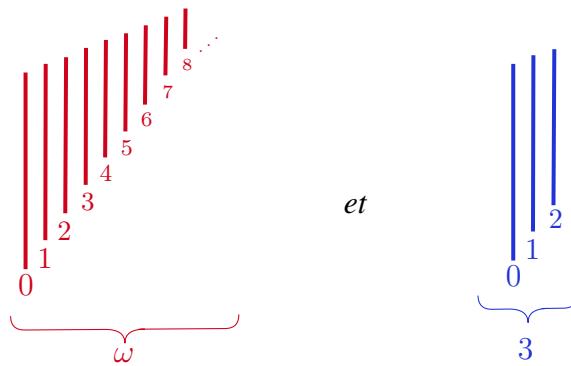
Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Pour tout entier naturel m , on a donc $P(m)$, c'est-à-dire $\boxed{f^{nm} = (f^n)^m}$.

CQFD.

3.2 Interprétation graphique : le produit cartésien

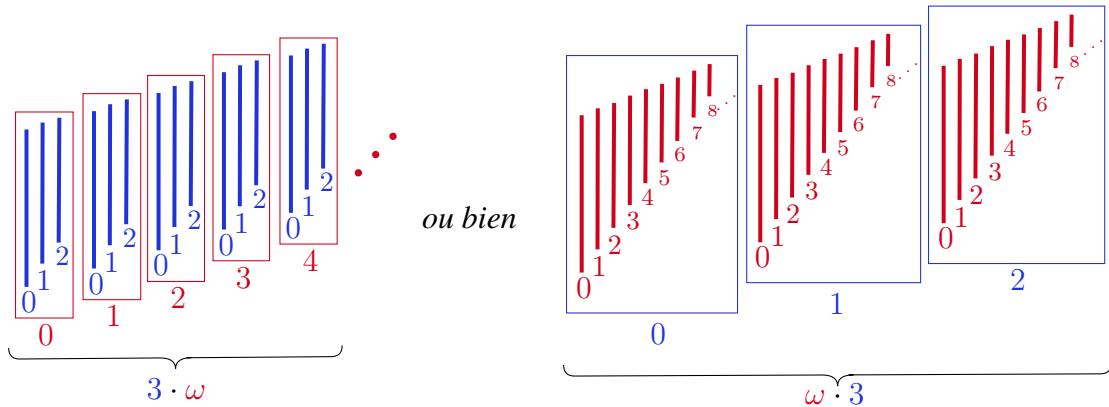
La multiplication des ordinaux a aussi une interprétation graphique : visualisons par exemple la multiplication des ordinaux ω et 3. Comme précédemment, commençons par les représenter tous les deux avec des bâtons indépendamment l'un de l'autre, ω et ses éléments étant en **rouge** et 3 est ses éléments étant en **bleu**.



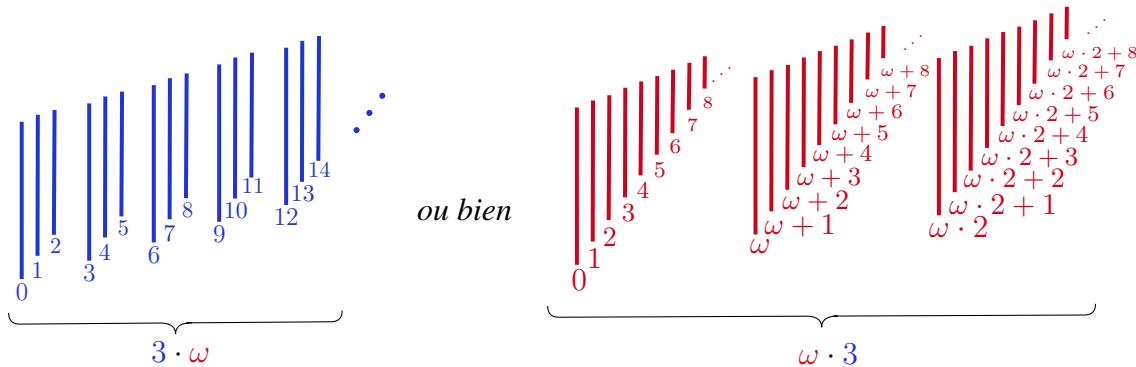
Rappelons à nouveau qu'ici seul l'ordre d'arrivée de la gauche vers la droite importe (pour

traduire l'ordre sur les ordinaux), la taille verticale des bâtons n'est qu'une astuce pour visualiser une infinité de bâtons.

Pour représenter visuellement $\alpha\beta$, on vient remplacer chaque bâton de β par une copie de l'ensemble des bâtons constitutifs de α , ce qui dans notre cas donne

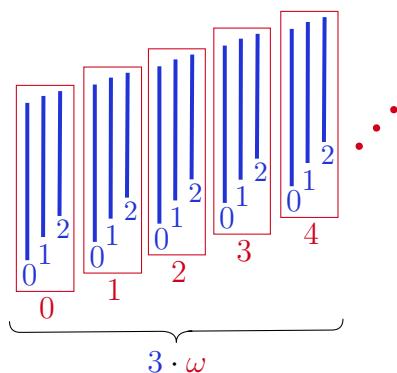


Enfin, comme pour l'addition on renumérote les bâtons en fonction de leur ordre d'arrivée de la gauche vers la droite, ce qui nous donne



On constate bien par exemple que $3\omega = \{0, 1, 2, 3, 4, 5, \dots\} = \omega$, comme nous l'avons expliqué plus tôt.

Quand il a fallu formaliser l'interprétation graphique de l'addition, nous avons dû introduire une nouvelle notion : celle de l'union disjointe ainsi que de l'ordre de concaténation. En ce qui concerne la multiplication, nous avons déjà tous les outils pour cela, puisqu'il s'agit du **produit cartésien**. En effet, si l'on reprend ce dessin d'interprétation de 3ω :



Le premier rectangle rouge tout à gauche peut être vu comme l'ensemble $\{(0, 0), (0, 1), (0, 2)\}$, le deuxième comme l'ensemble $\{(1, 0), (1, 1), (1, 2)\}$, le troisième comme l'ensemble

$\{(2,0), (2,1), (2,2)\}$, le quatrième comme l'ensemble $\{(3,0), (3,1), (3,2)\}$, et ainsi de suite, avec un rectangle pour chaque $n \in \omega$ de la forme $\{(n,0), (n,1), (n,2)\}$. En réunissant tous ces rectangles, on reconnaît précisément le produit cartésien $\omega \times \{0,1,2\} = \omega \times 3$. Oui, il faudra prendre garde au fait que la multiplication 3ω est associée au produit cartésien $\omega \times 3$, **il y a une inversion du sens**.

De quel ordre munir alors $\omega \times 3$? En fait, nous l'avons évoqué dans le premier livre et le premier chapitre : il s'agit de l'ordre lexicographique (en référence à l'ordre avec lequel fonctionne un dictionnaire). En effet, toujours avec le même dessin, on peut voir que tous les éléments du rectangle 1 sont situés avant tous les éléments du rectangle 2 : pour comparer les éléments du produit cartésien, il faut d'abord comparer les premières composantes de chaque couple et éventuellement si elles sont égales comparer les deuxièmes composantes. Nous avons déjà vu lors de la proposition 4 page 12 que si A et B sont munis de deux bons ordres, alors l'ordre lexicographique sur $A \times B$ est aussi un bon ordre. En particulier, si α et β sont deux ordinaux, alors $\alpha \times \beta$ est bien ordonné donc est associé à un unique ordinal (qu'on appellé son **type**). On retombe sur nos pieds : cet unique ordinal est précisément $\beta\alpha$.

Quel isomorphisme allons-nous construire? Il s'agit sur l'illustration de trouver, étant donné un rectangle et la position du bâton au sein de ce rectangle, quel sera le numéro du bâton après renumérotation. Pour un couple (γ, δ) de $\alpha \times \beta$, on trouve le bon rectangle en prenant la $\gamma^{\text{ème}}$ copie de β , puis on se déplace de δ bâtons vers la droite, c'est-à-dire $\beta\gamma + \delta$. Par exemple le bâton 1 au sein du rectangle 2 a bien été renuméroté en $3 \cdot 2 + 1 = 7$ à la fin.

Théorème 8 (Multiplication d'ordinaux et produit cartésien)

Soient α et β deux ordinaux.

On munit $\alpha \times \beta$ de l'ordre lexicographique.

On a alors $\beta\alpha = \text{type}(\alpha \times \beta)$.

Démonstration

- Construction de l'isomorphisme

Soit $(\gamma, \delta) \in \alpha \times \beta$.

On a alors $\gamma \in \alpha$ et $\delta \in \beta$.

Comme $\gamma \in \alpha$, on a $\gamma < \alpha$ donc $\gamma + 1 \leq \alpha$ d'après la proposition 13 page 33.

De même $\delta \in \beta$ on a $\delta < \beta$, donc on a :

$$\begin{aligned} \beta\gamma + \delta &< \beta\gamma + \beta \text{ par stricte croissance de l'addition à gauche} \\ &= \beta\gamma + \beta \cdot 1 \text{ car } 1 \text{ est neutre pour la multiplication} \\ &= \beta(\gamma + 1) \text{ par distributivité} \\ &\leq \beta\alpha \text{ par croissance de la multiplication à gauche} \end{aligned}$$

On a donc $\beta\gamma + \delta < \beta\alpha$ et donc $\beta\gamma + \delta \in \beta\alpha$ par définition de $<$.

Ainsi pour tout $(\gamma, \delta) \in \alpha \times \beta$, on a $\beta\gamma + \delta \in \beta\alpha$.

On peut donc considérer l'application $\varphi_\alpha := \begin{pmatrix} \alpha \times \beta & \longrightarrow & \beta\alpha \\ (\gamma, \delta) & \longmapsto & \beta\gamma + \delta \end{pmatrix}$.

Le fait d'avoir mis α en indice nous servira pour une preuve par induction sur α .

Montrons que φ_α est un isomorphisme d'ordres.

• Surjectivité

Pour cette partie fixons β .

Montrons que pour tout ordinal α , l'application φ_α est surjective sur $\beta\alpha$.

Posons P l'assertion à paramètre définie pour tout ordinal α par

$$P(\alpha) \iff \text{im}(\varphi_\alpha) = \beta\alpha$$

Remarquons que pour α et α' deux ordinaux, si $\alpha \leq \alpha'$ alors $\alpha \subseteq \alpha'$ donc $\alpha \times \beta \subseteq \alpha' \times \beta$ et pour tout $(\gamma, \delta) \in \alpha \times \beta$ on a $\varphi_\alpha(\gamma, \delta) = \beta \cdot \gamma + \delta = \varphi_{\alpha'}(\gamma, \delta)$, si bien que $\varphi_\alpha = (\varphi_{\alpha'})|_{\alpha \times \beta}$. En particulier dans ce cas-là on a $\text{im}(\varphi_\alpha) \subseteq \text{im}(\varphi_{\alpha'})$. Notons cela $(*)$

Remarquons aussi que pour tout ordinal α on a $\text{im}(\varphi_\alpha) \subseteq \beta\alpha$ donc on doit seulement montrer l'inclusion réciproque.

► Initialisation

Par définition de la multiplication on a $\beta \cdot 0 = 0$.

On a donc $\text{im}(\varphi_0) \subseteq 0$ et comme $0 = \emptyset$ on a donc $\text{im}(\varphi_0) = 0$, c'est-à-dire $P(0)$.

► Héritéité

Soit α un ordinal tel que $P(\alpha)$, c'est-à-dire $\beta\alpha = \text{im}(\varphi_\alpha)$.

Soit $\varepsilon \in \beta(\alpha + 1)$.

On a donc $\varepsilon < \beta(\alpha + 1)$ par définition de $<$.

Comme \leq est total sur les ordinaux, on a $\varepsilon < \beta\alpha$ ou $\beta\alpha \leq \varepsilon$.

► Plaçons-nous dans le cas où $\varepsilon < \beta\alpha$.

On a donc $\varepsilon \in \beta\alpha$, et comme $\beta\alpha = \text{im}(\varphi_\alpha) \underset{(*)}{\subseteq} \text{im}(\varphi_{\alpha+1})$ on a $\varepsilon \in \text{im}(\varphi_{\alpha+1})$.

► Plaçons-nous désormais dans le cas où $\beta\alpha \leq \varepsilon$.

Il existe un ordinal μ tel que $\varepsilon = \beta\alpha + \mu$ d'après la proposition 51 page 124.

Comme \leq est total sur les ordinaux, on a $\mu < \beta$ ou $\beta \leq \mu$.

Supposons par l'absurde que $\beta \leq \mu$.

On a alors

$$\begin{aligned}
 \beta(\alpha + 1) &= \beta\alpha + \beta \cdot 1 \text{ par distributivité} \\
 &= \beta\alpha + \beta \text{ car } 1 \text{ est neutre pour la multiplication} \\
 &\leq \beta\alpha + \mu \text{ par croissance de l'addition à gauche} \\
 &= \varepsilon \text{ par définition de } \mu
 \end{aligned}$$

On a donc $\beta(\alpha + 1) \leq \varepsilon$.

C'est absurde puisqu'on a dit que $\varepsilon < \beta(\alpha + 1)$.

Par l'absurde on vient de montrer que l'on n'a pas $\beta \leq \mu$, donc on a $\mu < \beta$.

Or on a $\alpha < \alpha + 1$ d'après la proposition 13 page 33.

Ainsi on a $\mu \in \beta$ et $\alpha \in \alpha + 1$ donc $(\alpha, \mu) \in (\alpha + 1) \times \beta$.

Donc $\varepsilon = \beta\alpha + \mu = \varphi_{\alpha+1}(\alpha, \mu)$ par définition de $\varphi_{\alpha+1}$.

On a donc $\varepsilon \in \text{im}(\varphi_{\alpha+1})$.

On a donc $\text{im}(\varphi_{\alpha+1}) \supseteq \beta(\alpha + 1)$ et donc $\text{im}(\varphi_{\alpha+1}) = \beta(\alpha + 1)$.

Autrement dit on a $P(\alpha + 1)$.

Donc pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

► Héritage limite

Soit α un ordinal limite non nul tel que $\forall \mu < \alpha, P(\mu)$.

Soit $\varepsilon \in \beta\alpha$.

On a donc $\varepsilon < \beta\alpha$ par définition de $<$.

Par définition α est un ordinal limite non nul.

Par définition de la multiplication on a donc $\beta\alpha = \sup_{\mu < \alpha} (\beta\mu)$.

On a donc $\varepsilon < \sup_{\mu < \alpha} (\beta\mu)$ donc ε n'est pas un majorant de $\{\beta\mu \mid \mu < \alpha\}$.

Il existe donc un ordinal $\mu < \alpha$ tel que l'on n'a pas $\beta\mu \leq \varepsilon$.

Comme \leq est total chez les ordinaux, on a $\varepsilon < \beta\mu$ donc $\varepsilon \in \beta\mu$.

Or $\mu < \alpha$ donc par hypothèse on a $P(\mu)$ et donc $\text{im}(\varphi_\mu) = \beta\mu$.

On a donc $\varepsilon \in \text{im}(\varphi_\mu)$.

Or $\mu < \alpha$ donc on a $\text{im}(\varphi_\mu) \subseteq \text{im}(\varphi_\alpha)$ d'après (\star) .

On a donc $\varepsilon \in \text{im}(\varphi_\alpha)$.

On a donc $\text{im}(\varphi_\alpha) \supseteq \beta\alpha$ et donc $\text{im}(\varphi_\alpha) = \beta\alpha$.

Autrement dit on a $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \mu < \alpha, P(\mu)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $P(\alpha)$.

Autrement dit pour tout ordinal α , on a $\text{im}(\varphi_\alpha) = \beta\alpha$.

Autrement dit pour tout ordinal α , $\boxed{\varphi_\alpha \text{ est surjective dans } \beta\alpha}$.

• Stricte croissance

Fixons à nouveau α et β .

Notons \triangleleft l'ordre lexicographique strict associé sur $\alpha \times \beta$.

Soient (γ, δ) et (γ', δ') dans $\alpha \times \beta$ tels que $(\gamma, \delta) \triangleleft (\gamma', \delta')$.

On a donc $\gamma < \gamma'$ ou $(\gamma = \gamma' \text{ et } \delta < \delta')$ par définition de \triangleleft .

► Plaçons-nous dans le cas où $\gamma < \gamma'$.

On a donc $\gamma + 1 \leq \gamma'$ d'après la proposition 13 page 33.

On a donc

$$\begin{aligned} \beta\gamma + \delta &< \beta\gamma + \beta \text{ par stricte croissance de l'addition à gauche} \\ &= \beta\gamma + \beta \cdot 1 \text{ car } 1 \text{ est neutre pour la multiplication} \\ &= \beta(\gamma + 1) \text{ par distributivité} \\ &\leq \beta\gamma' \text{ par croissance de la multiplication à gauche} \\ &= \beta\gamma' + 0 \text{ car } 0 \text{ est neutre pour l'addition} \\ &\leq \beta\gamma' + \delta' \text{ par croissance de l'addition à gauche} \end{aligned}$$

On a donc $\beta\gamma + \delta < \beta\gamma' + \delta'$ donc $\varphi_\alpha(\gamma, \delta) < \varphi_\alpha(\gamma', \delta')$.

► Plaçons-nous dans le cas où $\gamma = \gamma'$ et $\delta < \delta'$.

On a alors $\beta\gamma + \delta < \beta\gamma + \delta'$ par stricte croissance de l'addition à gauche.

Comme $\gamma = \gamma'$, on a donc $\beta\gamma + \delta < \beta\gamma' + \delta'$ et donc $\varphi_\alpha(\gamma, \delta) < \varphi_\alpha(\gamma', \delta')$.

Dans les deux cas on a donc $\varphi_\alpha(\gamma, \delta) < \varphi_\alpha(\gamma', \delta')$.

Donc pour tout (γ, δ) et (γ', δ') dans $\alpha \times \beta$, si $(\gamma, \delta) \triangleleft (\gamma', \delta')$ alors $\varphi_\alpha(\gamma, \delta) < \varphi_\alpha(\gamma', \delta')$.

Donc φ_α est strictement croissante.

Or $\text{dom}(\varphi_\alpha) = \alpha \times \beta$ est bien ordonné d'après la proposition 4 page 12.

En particulier $\text{dom}(\varphi_\alpha)$ est totalement ordonné d'après la proposition 2 page 10.

Donc $\boxed{\varphi_\alpha \text{ est injective et croissante}}$.

• Conclusion

Ainsi φ_α est croissante, injective et surjective sur $\beta\alpha$.

Or on vient de dire que $\text{dom}(\varphi_\alpha) = \alpha \times \beta$ est totalement ordonné.

Donc φ_α est un isomorphisme d'ordres de $\alpha \times \beta$ vers $\beta\alpha$.

Ainsi $\alpha \times \beta$ et $\beta\alpha$ sont isomorphes, et donc $\boxed{\beta\alpha = \text{type}(\alpha \times \beta)}$.
CQFD.

On a vu à la suite de l'addition comment faire une espèce de soustraction via la proposition 51 page 124. Celle-ci nous explique que si $\beta \leq \alpha$, alors il existe un unique ordinal σ tel que $\alpha = \beta + \sigma$. On retrouve une propriété similaire avec le principe de division ordinaire. Tout comme la division habituelle, celle-ci requiert qu'on ne divise pas par zéro, c'est-à-dire que β soit non nul.

Proposition 63 (Division ordinaire)

Soient α et β deux ordinaux, avec β **non nul**.

Il existe deux ordinaux δ et σ tels que $\alpha = \beta\delta + \sigma$ avec $\delta \leq \alpha$ et $\sigma < \beta$.

De plus, de tels ordinaux sont uniques.

Démonstration

Par définition β est non nul donc $0 < \beta$.

On a donc $1 = 0 + 1 \leq \beta$ d'après la proposition 13 page 33.

De même on a $\alpha < \alpha + 1$ d'après la proposition 13 page 33.

On a donc

$$\begin{aligned}\alpha &= 1\alpha \text{ car } 1 \text{ est neutre pour la multiplication} \\ &\leq \beta\alpha \text{ par croissance de la multiplication à droite} \\ &< \beta(\alpha + 1) \text{ par stricte croissance de la multiplication à gauche}\end{aligned}$$

Ainsi on a $\alpha < \beta(\alpha + 1)$ et donc $\alpha \in \beta(\alpha + 1)$ par définition de $<$.

Or on a vu dans la preuve du théorème 8 page 143 que l'application

$$\varphi := \begin{pmatrix} (\alpha + 1) \times \beta & \longrightarrow & \beta(\alpha + 1) \\ (\delta, \sigma) & \longmapsto & \beta\delta + \sigma \end{pmatrix} \text{ est bijective de } (\alpha + 1) \times \beta \text{ dans } \beta(\alpha + 1).$$

Il existe donc un unique $(\delta, \sigma) \in (\alpha + 1) \times \beta$ tel que $\alpha = \varphi(\delta, \sigma) = \beta\delta + \sigma$.

Remarquons pour conclure que comme $(\delta, \sigma) \in (\alpha + 1) \times \beta$, on a $\delta \in (\alpha + 1)$ et $\sigma \in \beta$.

Autrement dit on a $\delta < \alpha + 1$ donc $\delta \leq \alpha$ d'après la proposition 13 page 33.

De même on a $\sigma < \beta$.

CQFD.

Remarque :

Bien que cette division ressemble à la division euclidienne, il faut prendre garde au fait que le quotient δ peut être égal au dividende α . Par exemple, si l'on prend $\alpha = \omega$ et $\beta = 2$, on a $\omega = 2\omega + 0$.

4 Exponentiation d'ordinaux

4.1 Définition et propriétés

Attaquons-nous à la dernière des trois opérations ordinaires : l'exponentiation. Encore une fois, inspirons-nous de notre intuition sur les entiers. Intuitivement, 5^3 c'est $5 \cdot 5 \cdot 5$, avec le nombre 5 répété 3 fois, mais encore une fois le fait de répéter une opération un certain nombre de fois n'a pas été pleinement défini jusqu'à présent. Rattachons-nous une fois de plus aux définitions par récursion. On remarque simplement que $5^2 = 5 \cdot 5$ et donc $5^3 = (5 \cdot 5) \cdot 5 = 5^2 \cdot 5$. La récursion nous est donc encore donnée sur un plateau.

Ainsi pour définir 5^3 on considère que 5^2 est déjà défini puis on pose $5^3 := 5^2 \cdot 5$. Autrement dit on a posé $5^{2+1} := 5^2 \cdot 5$. Cela nous guide vers la définition suivante.

Définition 23 (Exponentiation d'ordinaux)

Soit α un ordinal **non nul**.

On pose

$$\begin{cases} \alpha^0 := 1 \\ \alpha^{\beta+1} := \alpha^\beta \alpha \text{ pour tout ordinal } \beta \\ \alpha^\gamma := \sup_{\delta < \gamma} \alpha^\delta \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Enfin, on pose $0^0 := 1$ et pour tout ordinal β non nul on pose $0^\beta := 0$.

Remarque :

1. Dans le livre précédent, nous avons indiqué que pour deux ensembles A et B , l'ensemble des applications de A vers B est noté $\mathcal{F}(A \rightarrow B)$ ou encore B^A . Nous allons temporairement éviter cette dernière notation afin d'éviter la confusion avec l'exponentiation d'ordinaux. Cela dit, le fait qu'il y ait la même notation pour ces deux concepts n'est pas une coïncidence, comme nous aurons l'occasion de le voir dans le dernier chapitre, où nous pourrons la reprendre sereinement pour la suite de nos aventures.
2. Pour justifier proprement cette définition, on utilise simplement la proposition 36 page 91, en posant $\mu_0 := 1$ et $G(\xi) := \xi \cdot \alpha$ pour tout ordinal ξ . La proposition nous donne alors une unique assertion fonctionnelle $F_\alpha : ON \longrightarrow ON$ telle que

$$\begin{cases} F_\alpha(0) = 1 \\ F_\alpha(\beta + 1) = F_\alpha(\beta) \cdot \alpha \text{ pour tout ordinal } \beta \\ F_\alpha(\gamma) = \sup_{\delta < \gamma} F_\alpha(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

et on pose alors $\alpha^\beta := F_\alpha(\beta)$ pour tout ordinal β .

3. Pourquoi ce traitement à part pour $\alpha = 0$? Imaginons un instant ne pas avoir fait d'exception pour $\alpha = 0$. On a donc $0^0 = 1$ par définition, puis $0^1 = 0^0 \cdot 0 = 1 \cdot 0 = 0$ puis pour tout entier naturel n on a $0^n = 0$. On aimerait qu'à partir de là, $0^\beta = 0$

dès que β est non nul. Le soucis vient pour $\beta = \omega$ car $0^\beta = \sup_{n<\omega} 0^n = 1$ puisqu'on a aussi $n = 0$ dans le lot. Mais au fond pourquoi vouloir que $0^\omega = 0$ plutôt que $0^\omega = \sup_{n<\omega} 0^n = 1$? Cela vient du fait que l'assertion fonctionnelle F_0 n'est pas croissante, et donc cette histoire de borne supérieure et de continuité ne tient plus. En fait, à partir de 1 elle est bien croissante, et donc il semble naturel de simplement exclure 0 de la prise en compte dans la borne supérieure, afin de refaire fonctionner le fait que la borne supérieure représente *le passage à la limite au voisinage d'un ordinal limite*. On peut donc tout aussi bien poser $0^\gamma = \sup_{0<\delta<\gamma} 0^\delta$ pour tout ordinal limite non nul γ , et cela revient au même.

Proposition 64 (Exponentiations avec 1)

Pour tout ordinal α , on a $\alpha^1 = \alpha$ et $1^\alpha = 1$.



Démonstration

- Pour tout ordinal α , on a

$$\begin{aligned}\alpha^1 &= \alpha^{0+1} \text{ par définition de } 1 \\ &= \alpha^0 \alpha \text{ par définition de l'exponentiation} \\ &= 1 \cdot \alpha \text{ par définition de l'exponentiation} \\ &= \alpha \text{ car } 1 \text{ est neutre pour la multiplication}\end{aligned}$$

et donc $\boxed{\alpha^1 = \alpha}$.

- Posons P l'assertion à paramètres définie pour tout ordinal α par

$$P(\alpha) \iff 1^\alpha = 1$$

Montrons le résultat par le principe faible d'induction.

► Initialisation

Par définition de l'exponentiation on a $1^0 = 1$.

Autrement dit on a $P(0)$.

► Hérédité

Soit α un ordinal tel que $P(\alpha)$.

On a alors

$$\begin{aligned} 1^{\alpha+1} &= 1^\alpha \cdot 1 \text{ par définition de l'exponentiation} \\ &= 1 \cdot 1 \text{ puisqu'on a } P(\alpha) \\ &= 1 \text{ car } 1 \text{ est neutre pour la multiplication} \end{aligned}$$

On a donc $1^{\alpha+1} = 1$.

Autrement dit on a $P(\alpha + 1)$.

Donc pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

► *Hérité limite*

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

Autrement dit $\forall \beta < \alpha, 1^\beta = 1$.

Par définition de l'exponentiation on a donc $1^\alpha = \sup_{\beta < \alpha} 1^\beta = \sup_{\beta < \alpha} 1 = 1$.

Autrement dit on a $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $P(\alpha)$.

Autrement dit pour tout ordinal α on a $1^\alpha = 1$.

CQFD.

Une fois n'est pas coutume, l'exponentiation de deux entiers naturels est encore un entier naturel.

Proposition 65 (Exponentiation de deux entiers naturels)

Pour tout entiers naturels n et m , l'ordinal n^m est un entier naturel.

On dit que $\mathbb{N} = \omega$ est **stable** par exponentiation.

 *Démonstration*

Fixons n un entier naturel.

Posons P l'assertion à paramètre définie pour tout entier naturel m par

$$P(m) \iff n^m \in \mathbb{N}$$

Montrons le résultat par induction sur les entiers naturels.

► *Initialisation*

Par définition de l'exponentiation on a $n^0 = 1$ qui est un entier naturel.

On a donc $P(0)$.

► *Héritéité*

Soit m un entier naturel tel que $P(m)$.

On a alors $n^{m+1} = n^m n$ par définition de l'exponentiation.

Or n^m est un entier naturel d'après $P(m)$ et n l'est par définition.

Donc $n^m n$ est un entier naturel d'après la proposition 54 page 129 .

Donc n^{m+1} est un entier naturel, et donc $P(m + 1)$.

Donc pour tout entier naturel m , si $P(m)$ alors $P(m + 1)$.

Ainsi P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel m on a $P(m)$.

Autrement dit pour tout entier naturel m , l'ordinal n^m est un entier naturel.

CQFD.

Pour la proposition qui suit, qui nous dit encore que l'opération du jour est croissante,

- on a exclut 0 dans les deux premiers cas $0^0 = 1$ est plus grand que tout $0^m = 0$ quand bien même $0 < m$,
- on a exclut 1 du premier cas car $1^\beta = 1$ est constant en β et donc on n'a pas de stricte croissance.

Proposition 66 (Croissance de l'exponentiation des ordinaux)

Soient α , β et γ trois ordinaux.

1. Supposons que $1 < \alpha$.

Si $\beta < \gamma$ alors $\alpha^\beta < \alpha^\gamma$.

On dit que l'exponentiation à gauche est **strictement croissante**.

2. Supposons que $0 < \alpha$.

Si $\beta \leq \gamma$ alors $\alpha^\beta \leq \alpha^\gamma$.

On dit que l'exponentiation à gauche est **croissante**.

3. Si $\beta \leq \gamma$ alors $\beta^\alpha \leq \gamma^\alpha$.

On dit que l'exponentiation à droite est **croissante**.

 *Démonstration*

1. Fixons α et β deux ordinaux tels que $1 < \alpha$.

Posons P l'assertion à paramètre définie pour tout ordinal γ par

$$P(\gamma) \iff (\beta < \gamma \Rightarrow \alpha^\beta < \alpha^\gamma)$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

Il est faux de dire que $\beta \in \emptyset$ donc il est faux de dire que $\beta \in 0$ et donc de dire que $\beta < 0$.

Ainsi, la prémissse $\beta < 0$ étant fausse, on a nécessairement l'implication $\beta < 0 \Rightarrow \alpha^\beta < \alpha^0$.

Autrement dit on a $P(0)$.

► *Hérédité*

Soit γ un ordinal tel que $P(\gamma)$.

Supposons que $\beta < \gamma + 1$.

On a donc $\beta \leq \gamma$ d'après la proposition 13 page 33, donc $\beta < \gamma$ ou $\beta = \gamma$.

Rappelons que par hypothèse on a $1 < \alpha$, et donc on a

$$\begin{aligned} \alpha^\gamma &= \alpha^\gamma \cdot 1 \text{ car } 1 \text{ est neutre pour la multiplication} \\ &< \alpha^\gamma \alpha \text{ par stricte croissance de la multiplication à gauche} \\ &= \alpha^{\gamma+1} \text{ par définition de l'exponentiation} \end{aligned}$$

Ainsi on a $\alpha^\gamma < \alpha^{\gamma+1}$.

► Plaçons-nous dans le cas où $\beta < \gamma$.

On a alors $\alpha^\beta < \alpha^\gamma$ d'après $P(\gamma)$.

On a donc $\alpha^\beta < \alpha^{\gamma+1}$ par transitivité de $<$.

► Plaçons-nous dans le cas où $\beta = \gamma$.

On a donc $\alpha^\beta = \alpha^\gamma$ et donc $\alpha^\beta < \alpha^{\gamma+1}$.

Dans les deux cas on a donc $\alpha^\beta < \alpha^{\gamma+1}$.

Donc si $\beta < \gamma + 1$ alors $\alpha^\beta < \alpha^{\gamma+1}$, donc on a $P(\gamma + 1)$.

Donc pour tout ordinal γ , si $P(\gamma)$ alors $P(\gamma + 1)$.

► *Hérédité limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P(\delta)$.

Supposons que $\beta < \gamma$.

On a donc $\beta + 1 < \gamma$ d'après la proposition 14 page 37 car γ est limite.

On a donc $P(\beta + 1)$ par hypothèse, c'est-à-dire $\beta < \beta + 1 \Rightarrow \alpha^\beta < \alpha^{\beta+1}$.

Or $\beta < \beta + 1$ d'après la proposition 13 page 33.

On a donc $\alpha^\beta < \alpha^{\beta+1}$ par modus ponens.

On a dit que $\beta + 1 < \gamma$.

On a donc $\alpha^{\beta+1} \leq \sup_{\delta < \gamma} \alpha^\delta$ car la borne supérieure est un majorant.

Or on a $\alpha^\gamma = \sup_{\delta < \gamma} \alpha^\delta$ par définition de l'exponentiation.

On a donc $\alpha^{\beta+1} \leq \alpha^\gamma$ et donc $\alpha^\beta < \alpha^\gamma$ par transitivité.

Donc si $\beta < \gamma$ alors $\alpha^\beta < \alpha^\gamma$, donc on a $P(\gamma)$.

Donc pour tout ordinal limite γ , si $\forall \delta < \gamma, P(\delta)$ alors $P(\gamma)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $P(\gamma)$.

Autrement dit pour tout ordinal γ , si $\beta < \gamma$ alors $\alpha^\beta < \alpha^\gamma$.

2. Fixons α, β et γ trois ordinaux, avec $0 < \alpha$.

Supposons que $\beta \leq \gamma$.

On a $0 < \alpha$ par hypothèse donc $0 + 1 \leq \alpha$ d'après la proposition 13 page 33.

Comme $1 = 0 + 1$ on a $1 \leq \alpha$ donc ($1 = \alpha$ ou $1 < \alpha$).

Plaçons-nous dans le cas où $1 = \alpha$.

On a alors $\alpha^\beta = 1^\beta = 1 = 1^\gamma = \alpha^\gamma$ d'après la proposition 64 page 149.

En particulier on a $\alpha^\beta \leq \alpha^\gamma$ par réflexivité de \leq .

On se place à présent dans le cas où $1 < \alpha$.

On a fait l'hypothèse que $\beta \leq \gamma$ donc ($\beta < \gamma$ ou $\beta = \gamma$).

- Plaçons-nous dans le cas où $\beta < \gamma$.

Comme $1 < \alpha$ on a $\alpha^\beta < \alpha^\gamma$ d'après 1, donc en particulier $\alpha^\beta \leq \alpha^\gamma$.

- Plaçons-nous dans le cas où $\beta = \gamma$.

On a alors $\alpha^\beta = \alpha^\gamma$ et en particulier $\alpha^\beta \leq \alpha^\gamma$ par réflexivité de \leq .

Dans tous les cas on a donc $\alpha^\beta \leq \alpha^\gamma$.

Donc si $\beta \leq \gamma$ alors $\alpha^\beta \leq \alpha^\gamma$.

3. Fixons β et γ deux ordinaux.

Supposons que $\beta \leq \gamma$.

Posons Q l'assertion à paramètre définie pour tout ordinal α par

$$Q(\alpha) \iff \beta^\alpha \leq \gamma^\alpha$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

On a $\beta^0 = 1 = \gamma^0$ par définition de l'exponentiation.

On a donc $\beta^0 \leq \gamma^0$ par réflexivité de \leq , et donc $Q(0)$.

► *Hérédité*

Soit α un ordinal tel que $Q(\alpha)$, c'est-à-dire $\beta^\alpha \leq \gamma^\alpha$.

On a alors

$$\begin{aligned} \beta^{\alpha+1} &= \beta^\alpha \beta \text{ par définition de l'exponentiation} \\ &\leq \gamma^\alpha \beta \text{ par croissance de la multiplication à droite} \\ &\leq \gamma^\alpha \gamma \text{ par croissance de la multiplication à gauche} \\ &= \gamma^{\alpha+1} \text{ par définition de l'exponentiation} \end{aligned}$$

On a donc $\beta^{\alpha+1} \leq \gamma^{\alpha+1}$ et donc $Q(\alpha + 1)$.

Donc pour tout ordinal α , si $Q(\alpha)$ alors $Q(\alpha + 1)$.

► *Hérédité limite*

Soit α un ordinal limite non nul tel que $\forall \delta < \alpha, Q(\delta)$.

Soit δ un ordinal tel que $\delta < \alpha$.

On a alors

$$\begin{aligned} \beta^\delta &\leq \gamma^\delta \text{ d'après } Q(\delta) \\ &\leq \sup_{\varepsilon < \alpha} \gamma^\varepsilon \text{ car la borne supérieure est un majorant} \\ &= \gamma^\alpha \text{ par définition de l'exponentiation} \end{aligned}$$

On a donc $\beta^\delta \leq \gamma^\alpha$.

On a donc $\forall \delta < \alpha, \beta^\delta \leq \gamma^\alpha$.

Donc $\sup_{\delta < \alpha} \beta^\delta \leq \gamma^\alpha$ par minimalité de la borne supérieure.

Ainsi on a $\beta^\alpha \leq \gamma^\alpha$ par définition de l'exponentiation, et donc $Q(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \delta < \alpha, Q(\delta)$ alors $Q(\alpha)$.

Ainsi Q vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $Q(\alpha)$.

Autrement dit pour tout ordinal α on a $\beta^\alpha \leq \gamma^\alpha$.

Donc pour tout ordinal α , si $\beta \leq \gamma$ alors $\beta^\alpha \leq \gamma^\alpha$.

CQFD.

Remarque :

Prenons α un ordinal non nul et β un ordinal quelconque non nul.

On a alors $0 < \alpha^\beta$. En effet, comme \leq est total chez les ordinaux, on a $\beta \leq 0$ ou $0 < \beta$.

Mais $\beta \leq 0 \iff \beta \subseteq \emptyset \iff \beta = \emptyset \iff \beta = 0$.

Donc comme β est non nul, on a $0 < \beta$ donc $0 + 1 \leq \beta$ d'après la proposition 13 page 33.

Comme $0 + 1 = 1$, on a donc $1 \leq \beta$.

En particulier par croissance de l'exponentiation on a $0 < \alpha = \alpha^1 \leq \alpha^\beta$.

Proposition 67 (Régularité de l'exponentiation des ordinaux)

Soient α, β et γ des ordinaux, avec $1 < \alpha$.

Si $\alpha^\beta = \alpha^\gamma$ alors $\beta = \gamma$.

On dit que l'exponentiation à gauche est **régulière**.

Démonstration

Montrons-le par contraposition.

Supposons que $\beta \neq \gamma$.

On a donc $\beta < \gamma$ ou $\gamma < \beta$ d'après le théorème 1 page 21.

Si $\beta < \gamma$ alors $\alpha^\beta < \alpha^\gamma$ par stricte croissance de l'exponentiation à gauche.

Si $\gamma < \beta$ alors $\alpha^\gamma < \alpha^\beta$ par stricte croissance de l'exponentiation à gauche.

Dans les deux cas on a $\alpha^\gamma \neq \alpha^\beta$ par antiréflexivité de $<$.

Donc si $\beta \neq \gamma$ alors $\alpha^\gamma \neq \alpha^\beta$.

Donc par contraposition, si $\alpha^\beta = \alpha^\gamma$ alors $\beta = \gamma$.

CQFD.

Proposition 68 (Continuité de l'exponentiation des ordinaux)

Soient α un ordinal **non nul** et X un ensemble **non vide** d'ordinaux.

On a $\sup_{\xi \in X} \alpha^\xi = \alpha^{\sup(X)}$.

Autrement dit l'exponentiation à gauche est continue.

Démonstration

Par définition de l'exponentiation, pour tout ordinal limite non nul γ , on a

$$\alpha^\gamma = \sup_{\delta < \gamma} \alpha^\delta$$

On peut donc appliquer la proposition 42 page 103 pour conclure.

Notons que cette définition ne tient que parce que α est non nul.

CQFD.

Proposition 69 (Exponentiation, addition et multiplication)

Soient trois ordinaux α , β et γ .

1. On a $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$.
2. On a $(\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$

Démonstration

1.

Réglons tout d'abord le cas $\alpha = 0$.

On a $0^{0+0} = 0^0 = 1 = 1 \cdot 1 = 0^0 \cdot 0^0$.

Pour tout ordinal non nul γ on a $0^{0+\gamma} = 0^\gamma = 0 = 1 \cdot 0 = 0^0 \cdot 0^\gamma$.

Pour tout ordinal non nul γ on a $0^{\gamma+0} = 0^\gamma = 0 = 0 \cdot 1 = 0^\gamma \cdot 0^0$.

Pour tout ordinaux non nuls γ et δ on a $0^{\gamma+\delta} = 0 = 0 \cdot 0 = 0^\gamma \cdot 0^\delta$.

Fixons deux ordinaux α et β , avec α **non nul**.

Posons P l'assertion à paramètre définie pour tout ordinal γ par

$$P(\gamma) \iff \alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

On a

$$\begin{aligned} \alpha^{\beta+0} &= \alpha^\beta \text{ car } 0 \text{ est neutre pour l'addition} \\ &= \alpha^\beta \cdot 1 \text{ car } 1 \text{ est neutre pour la multiplication} \\ &= \alpha^\beta \alpha^0 \text{ par définition de l'exponentiation} \end{aligned}$$

Ainsi on a $\alpha^{\beta+0} = \alpha^\beta \alpha^0$, c'est-à-dire $P(0)$.

► *Héritage*

Soit γ un ordinal tel que $P(\gamma)$.

On a alors

$$\begin{aligned}\alpha^{\beta+\gamma+1} &= \alpha^{\beta+\gamma}\alpha \text{ par définition de l'exponentiation} \\ &= \alpha^\beta\alpha^\gamma\alpha \text{ par } P(\gamma) \\ &= \alpha^\beta\alpha^{\gamma+1} \text{ par définition de l'exponentiation}\end{aligned}$$

Ainsi on a $\alpha^{\beta+\gamma+1} = \alpha^\beta\alpha^{\gamma+1}$ et donc $P(\gamma + 1)$.

Donc pour tout ordinal γ , si $P(\gamma)$ alors $P(\gamma + 1)$.

► *Héritage limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P(\delta)$.

On a alors

$$\begin{aligned}\alpha^{\beta+\gamma} &= \alpha^{\sup_{\delta < \gamma}(\beta+\delta)} \text{ par définition de l'addition} \\ &= \sup_{\delta < \gamma} \alpha^{\beta+\delta} \text{ par continuité de l'exponentiation à gauche car } \alpha \neq 0 \\ &= \sup_{\delta < \gamma} \alpha^\beta\alpha^\delta \text{ puisque } \forall \delta < \gamma, P(\delta) \\ &= \alpha^\beta \sup_{\delta < \gamma} \alpha^\delta \text{ par continuité de la multiplication à gauche} \\ &= \alpha^\beta\alpha^\gamma \text{ par définition de l'exponentiation car } \alpha \neq 0\end{aligned}$$

On a donc $\alpha^{\beta+\gamma} = \alpha^\beta\alpha^\gamma$, c'est-à-dire $P(\gamma)$.

Donc pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, P(\delta)$ alors $P(\gamma)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $P(\gamma)$.

Autrement dit pour tout ordinal γ , on a $\boxed{\alpha^{\beta+\gamma} = \alpha^\beta\alpha^\gamma}$.

2.

Commençons par traiter le cas $\alpha = 0$.

Pour tout ordinal γ on a $(0^0)^\gamma = 1^\gamma = 1 = 0^0 = 0^{0\cdot\gamma}$.

Pour tout ordinal β non nul on a $(0^\beta)^0 = 0^0 = 0^{\beta\cdot 0}$.

Pour tout ordinal β et γ non nuls, $\beta\gamma$ est aussi non nul car la multiplication des ordinaux est intègre. On a donc $(0^\beta)^\gamma = 0^\gamma = 0 = 0^{\beta\gamma}$.

Fixons α et β deux ordinaux, avec α **non nul**.

Notons qu'alors α^β est non nul car $\alpha \neq 0$.

Posons Q l'assertion à paramètre définie pour tout ordinal γ par

$$Q(\gamma) \iff (\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$$

Montrons le résultat par le principe faible d'induction.

► *Initialisation*

On a

$$\begin{aligned} (\alpha^\beta)^0 &= 1 \text{ par définition de l'exponentiation} \\ &= \alpha^0 \text{ par définition de l'exponentiation} \\ &= \alpha^{\beta \cdot 0} \text{ car } 0 \text{ est absorbant pour la multiplication} \end{aligned}$$

On a donc $(\alpha^\beta)^0 = \alpha^{\beta \cdot 0}$.

Autrement dit on a $Q(0)$.

► *Hérédité*

Soit γ un ordinal tel que $Q(\gamma)$.

On a alors

$$\begin{aligned} (\alpha^\beta)^{\gamma+1} &= (\alpha^\beta)^\gamma \alpha^\beta \text{ par définition de l'exponentiation} \\ &= \alpha^{\beta\gamma} \alpha^\beta \text{ puisqu'on a } Q(\gamma) \\ &= \alpha^{\beta\gamma+\beta} \text{ d'après 1} \\ &= \alpha^{\beta(\gamma+1)} \text{ par définition de la multiplication} \end{aligned}$$

On a donc $(\alpha^\beta)^{\gamma+1} = \alpha^{\beta(\gamma+1)}$, c'est-à-dire $Q(\gamma + 1)$.

Donc pour tout ordinal γ , si $Q(\gamma)$ alors $Q(\gamma + 1)$.

► *Hérédité limite*

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, Q(\delta)$.

On a alors

$$\begin{aligned} (\alpha^\beta)^\gamma &= \sup_{\delta < \gamma} (\alpha^\beta)^\delta \text{ par définition de l'exponentiation car } \alpha^\beta \neq 0 \\ &= \sup_{\delta < \gamma} (\alpha^\beta \alpha^\delta) \text{ puisque } \forall \delta < \gamma, Q(\delta) \end{aligned}$$

$$\begin{aligned}
 &= \alpha^\beta \sup_{\delta < \gamma} \alpha^\delta \text{ par continuité de la multiplication à gauche} \\
 &= \alpha^\beta \alpha^\gamma \text{ par définition de l'exponentiation car } \alpha \neq 0
 \end{aligned}$$

On a donc $(\alpha^\beta)^\gamma = \alpha^\beta \alpha^\gamma$, c'est-à-dire $Q(\gamma)$.

Donc pour tout ordinal limite non nul γ , si $\forall \delta < \gamma, Q(\delta)$ alors $Q(\gamma)$.

Ainsi Q vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal γ on a $Q(\gamma)$.

Autrement dit pour tout ordinal γ on a $(\alpha^\beta)^\gamma = \alpha^\beta \alpha^\gamma$.

CQFD.

Quand nous avons vu l'addition des ordinaux, nous avons au passage vu la soustraction : étant donnés deux ordinaux α et β tels que $\beta \leq \alpha$, il existe un ordinal γ tel que $\alpha = \beta + \gamma$, et cet ordinal est unique. De même, quand nous avons vu la multiplication des ordinaux, nous avons au passage vu la division : étant donnés deux ordinaux α et β , avec β non nul, il existe deux ordinaux γ et δ tels que $\alpha = \beta\gamma + \delta$ avec $\gamma \leq \alpha$ et $\delta < \beta$, et ces deux ordinaux sont uniques. L'exponentiation ne déroge pas à la règle : il y a une généralisation du logarithme chez les ordinaux. Nous n'allons cependant pas la voir tout de suite, mais dans la prochaine section sur la forme normale de Cantor.

4.2 Applications à support fini

Nous n'allons pas ici fournir à proprement parler d'interprétation géométrique dans le cas de l'exponentiation, mais établir plutôt un isomorphisme d'ordre entre l'ordinal α^β et un autre ensemble bien ordonné, avec l'idée que cela pourra nous éclairer un peu mieux sur l'ordre fourni par cet ordinal.

Pour comprendre la direction dans laquelle nous partons, revenons à la définition de α^β . En ayant en tête que $\alpha^2 = \alpha\alpha$, ou encore que $\alpha^3 = \alpha\alpha\alpha$, d'une manière générale on a intuitivement l'égalité α^β est $\alpha\alpha\dots\alpha$ où α est répété β fois. Or l'interprétation graphique de $\alpha\alpha$ est le produit cartésien $\alpha \times \alpha$ muni de l'ordre lexicographique. Ne peut-on pas partir de cet ordre lexicographique, mais le généraliser à 3, voir une infinité de composantes (par exemple si β n'est pas un entier naturel) ? Pour trois composantes, ce n'est pas compliqué : on compare les premières, si elles sont égales on compare les deuxièmes, et si elles sont aussi égales on compare alors les troisièmes. Comment faire dans le cas d'une infinité de composantes ?

Rappelons-nous que dans le précédent livre :

1. Nous avons défini le produit cartésien de deux ensembles A et B comme étant l'ensemble des couples (a, b) avec $a \in A$ et $b \in B$.
2. Nous avons ensuite remarqué que le couple (a, b) peut être vu comme l'application $f : \{0, 1\} \longrightarrow ?$ définie par $f(0) = a \in A$ et $f(1) = b \in B$. En particulier en notant $A_0 := A$ et $A_1 := B$, on a $f(i) \in A_i$ pour tout $i \in \{0, 1\}$.
3. Cette approche nous a permis de généraliser la notion de produit cartésien à toute une famille d'ensembles $(A_i)_{i \in I}$. Nous avons donc défini $\prod_{i \in I} A_i$ comme l'ensemble de toutes

les applications $f : I \longrightarrow ?$ telles que $\forall i \in I, f(i) \in A_i$.

4. En particulier, en prenant $\forall i \in I, A_i := E$ pour E un ensemble quelconque fixé à l'avance, on obtient $\prod_{i \in I} E$ qui est l'ensemble des applications $f : I \longrightarrow ?$ telles que $\forall i \in I, f(i) \in E$, c'est-à-dire simplement l'ensemble $\mathcal{F}(I \rightarrow E)$ des applications $I \longrightarrow E$, que l'on a d'ailleurs aussi noté E^I .

Autrement dit, pour généraliser le produit cartésien de α par lui-même β fois, il suffit de regarder l'ensemble $\mathcal{F}(\beta \rightarrow \alpha)$. C'est d'ailleurs la raison pour laquelle on a choisi la notation exponentielle pour l'ensemble des applications : il semble intuitivement que $\mathcal{F}(\beta \rightarrow \alpha)$ et α^β sont isomorphes. **Ce n'est en général pas vrai**, mais pour comprendre cela réfléchissons tout d'abord à voir comment généraliser l'ordre lexicographique.

Si l'on se souvient bien, pour γ et δ deux ordinaux, le produit ordinal $\gamma\delta$ est isomorphe au produit cartésien $\delta \times \gamma$ muni de l'ordre lexicographique. On a ainsi une inversion entre la gauche et la droite. Cela a des conséquences sur l'exponentiation : en effet on a $\gamma^3 = (\gamma\gamma)\gamma$ qui est donc isomorphe à $\gamma \times (\gamma \times \gamma)$ muni de l'ordre lexicographique. Là où les γ supplémentaires s'accumulent sur la droite pour la multiplication, ils s'accumulent sur la gauche pour le produit cartésien.

Si l'on souhaite conserver la gauche et la droite, il faut donc non pas munir $\gamma \times \delta$ de l'ordre lexicographique, mais de l'ordre anti-lexicographique, c'est-à-dire que pour (ε, ζ) et (ε', ζ') dans $\gamma \times \delta$, on pose

$$(\varepsilon, \zeta) \sqsubseteq (\varepsilon', \zeta') \iff \begin{cases} \zeta < \zeta' \\ \text{ou} \\ \zeta = \zeta' \text{ et } \varepsilon \leq \varepsilon' \end{cases}$$

Dans ce cas-là, on peut montrer que $\gamma\delta$ est bien isomorphe à $\gamma \times \delta$ muni de cet ordre anti-lexicographique, le sens est donc bien conservé. Pour reprendre l'analogie avec le dictionnaire, cela veut dire que cette fois on compare les mots en commençant par les dernières lettres : dès qu'il y a une lettre de différence (en partant de la fin du mot), celle-ci nous dit quel est mot se trouve avant quel autre. La preuve n'est pas difficile : $\gamma \times \delta$ et $\delta \times \gamma$ sont isomorphes munit respectivement de l'ordre anti-lexicographique et lexicographique (et ça se voit tout de suite si on écrit les définitions des deux !).

Généralisons donc l'ordre anti-lexicographique à l'ensemble $\mathcal{F}(\beta \rightarrow \alpha)$.

Proposition 70 (Ordre anti-lexicographique sur les applications)

Soient α et β deux ordinaux.

Pour tout $f : \beta \longrightarrow \alpha$ et $g : \beta \longrightarrow \alpha$, posons

$$f \prec g \iff \exists \gamma \in \beta, \begin{cases} \forall \delta \in \beta, (\gamma < \delta \Rightarrow f(\delta) = g(\delta)) \\ \text{et } f(\gamma) < g(\gamma) \end{cases}$$

Alors \prec est un ordre strict sur $\mathcal{F}(\beta \rightarrow \alpha)$.

Son ordre (large) associé est appelé **ordre anti-lexicographique** sur $\mathcal{F}(\beta \rightarrow \alpha)$.

 *Démonstration*

Antiréflexivité

Soit $f : \beta \longrightarrow \alpha$.

Pour tout $\gamma \in \beta$, on ne peut pas avoir $f(\gamma) < f(\gamma)$ par antiréflexivité de $<$.

Il n'existe donc pas de $\gamma \in \beta$ qui remplit la deuxième condition de la définition de $f \prec f$, donc en particulier on n'a pas $f \prec f$.

Donc pour tout $f : \beta \longrightarrow \alpha$, on n'a pas $f \prec f$.

Donc \prec est antiréflexive sur $\mathcal{F}(\beta \rightarrow \alpha)$.

Transitivité

Soient $f : \beta \longrightarrow \alpha$, $g : \beta \longrightarrow \alpha$ et $h : \beta \longrightarrow \alpha$.

Supposons que $f \prec g \prec h$.

Il existe donc γ_1 et γ_2 dans β tels que

$$\left\{ \begin{array}{l} \forall \delta \in \beta, (\gamma_1 < \delta \Rightarrow f(\delta) = g(\delta)) \\ \text{et } f(\gamma_1) < g(\gamma_1) \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} \forall \delta \in \beta, (\gamma_2 < \delta \Rightarrow g(\delta) = h(\delta)) \\ \text{et } g(\gamma_2) < h(\gamma_2) \end{array} \right.$$

Comme \leq est total chez les ordinaux, on a $\gamma_1 < \gamma_2$ ou $\gamma_1 = \gamma_2$ ou $\gamma_2 < \gamma_1$.

► Plaçons-nous dans le cas où $\gamma_1 < \gamma_2$.

Soit $\delta \in \beta$.

Supposons que $\gamma_2 < \delta$.

Par définition de γ_2 on a $g(\delta) = h(\delta)$.

Mais on a aussi $\gamma_1 < \delta$ par transitivité de $<$.

Donc par définition de γ_1 on a $f(\delta) = g(\delta)$.

On a donc $f(\delta) = h(\delta)$.

Donc si $\gamma_2 < \delta$ alors $f(\delta) = h(\delta)$.

Donc $\forall \delta \in \beta, (\gamma_2 < \delta \Rightarrow f(\delta) = h(\delta))$.

Comme $\gamma_1 < \gamma_2$, on a $f(\gamma_2) = g(\gamma_2)$ par définition de γ_1 .

De plus $g(\gamma_2) < h(\gamma_2)$ par définition de γ_2 , et donc $f(\gamma_2) < h(\gamma_2)$.

Ainsi γ_2 vérifie les deux conditions de la définition de $f \prec h$.

► Plaçons-nous dans le cas où $\gamma_1 = \gamma_2$.

Alors $\forall \delta \in \beta, (\gamma_1 < \delta \Rightarrow f(\delta) = g(\delta) = h(\delta))$ par définition de γ_1 et γ_2 .

On a donc $\forall \delta \in \beta, (\gamma_1 < \delta \Rightarrow f(\delta) = h(\delta))$.

De même, on a $f(\gamma_1) < g(\gamma_1) < h(\gamma_1)$ par définition de γ_1 et γ_2 .

On a donc $f(\gamma_1) < h(\gamma_1)$ par transitivité de $<$.

Ainsi γ_1 vérifie les deux conditions de la définition de $f \prec h$.

► Plaçons-nous dans le cas où $\gamma_2 < \gamma_1$.

Soit $\delta \in \beta$.

Supposons que $\gamma_1 < \delta$.

Par définition de γ_1 on a $f(\delta) = g(\delta)$.

Mais on a aussi $\gamma_2 < \delta$ par transitivité de $<$.

Donc par définition de γ_2 on a $g(\delta) = h(\delta)$.

On a donc $f(\delta) = h(\delta)$.

Donc si $\gamma_1 < \delta$ alors $f(\delta) = h(\delta)$.

Donc $\forall \delta \in \beta, (\gamma_1 < \delta \Rightarrow f(\delta) = g(\delta))$.

Comme $\gamma_2 < \gamma_1$, on a $g(\gamma_1) = h(\gamma_1)$ par définition de γ_2 .

De plus $f(\gamma_1) < g(\gamma_1)$ par définition de γ_1 , et donc $f(\gamma_1) < h(\gamma_1)$.

Ainsi γ_1 vérifie les deux conditions de la définition de $f \prec h$.

Dans les trois cas, on a donc $f \prec h$.

Donc si $f \prec g \prec h$ alors $f \prec h$.

Donc \prec est transitive.

Finalement $\boxed{\prec \text{ est une relation d'ordre strict sur } \mathcal{F}(\beta \rightarrow \alpha)}$.

CQFD.

Malheureusement, bien que α et β soient bien ordonnés car des ordinaux, cela ne suffit pas à rendre $\mathcal{F}(\beta \rightarrow \alpha)$ bien ordonné, muni de l'ordre anti-lexicographique. En effet, cet ordre n'a même pas de raison d'être total ! Par exemple, prenons $\alpha := 2 = \{0, 1\}$ et $\beta := \omega$ puis $f : \omega \longrightarrow \{0, 1\}$ qui vaut 0 sur les entiers pairs et 1 sur les entiers impairs, et enfin $g : \omega \longrightarrow \{0, 1\}$ qui fait le contraire, c'est-à-dire qui vaut 1 sur les entiers pairs et 0 sur les entiers impairs.

$$f = (0, 1, 0, 1, 0, 1, 0, 1, \dots) \text{ et } g = (1, 0, 1, 0, 1, 0, 1, 0, \dots).$$

En partant de la droite, à partir de quand peut-on comparer f et g ?

Il n'existe alors pas d'entier n tel que pour tout $m \in \omega$, on ait $n < m \Rightarrow f(m) = g(m)$ si bien que l'on ne peut avoir ni $f \preccurlyeq g$ ni $g \preccurlyeq f$: ainsi f et g ne sont pas comparables. Autrement dit, comme 2^ω est un ordinal, il est totalement ordonné donc **n'est pas isomorphe** à $\mathcal{F}(\omega \rightarrow 2)$.

C'est un peu comme vouloir comparer des mots infinis en commençant par la fin (puisque on utilise l'ordre anti-lexicographique), cela n'a pas de sens. Une solution est en fait de ne pas chercher à comparer toutes les applications $\mathcal{F}(\beta \rightarrow \alpha)$, mais seulement celles qui finissent par « *s'arrêter* » à un moment, c'est-à-dire qui finissent par prendre uniquement la valeur 0 à partir d'un moment. Autrement dit, il s'agit des applications $f : \beta \longrightarrow \alpha$ telles que l'ensemble $\{\gamma \in \beta \mid f(\gamma) \neq 0\}$ admet un maximum, qui jouera alors le rôle de « *dernière position* ». Cet ensemble porte d'ailleurs un nom : c'est le **support** de f .

Définition 24 (Support d'une application ordinaire)

Soient β et α deux ordinaux, et $f : \beta \longrightarrow \alpha$.

On appelle **support** de f l'ensemble $\text{supp}(f) := \{\gamma \in \beta \mid f(\gamma) \neq 0\}$.

Ainsi si l'on ne regarde que les applications dont le support admet un maximum, les deux applications de tout à l'heure ne sont pas concernées, car elles alternent sans arrêt entre 0 et 1.

Malheureusement, cela ne suffit toujours pas. Prenons encore $\alpha := 2 = \{0, 1\}$ mais cette fois $\beta := \omega + 1$, et reprenons les mêmes applications f et g que tout à l'heure en rajoutant la valeur 0 en ω . Les supports $\text{supp}(f)$ et $\text{supp}(g)$ admettent bien ω pour maximum, mais pourtant il n'existe toujours pas d'élément de β en lequel f et g diffèrent mais au delà duquel f et g sont égales : ce n'est pas ω puisque $f(\omega) = 0 = g(\omega)$, et pour tous les autres éléments de β on retombe sur le même problème que tout à l'heure puisque ω est limite.

En vérité nous allons restreindre encore plus les applications de notre intérêt. Nous n'allons pas demander uniquement au support d'avoir un maximum, mais carrément à toutes ses parties non vides ! Qu'est-ce que cela implique sur le support ? La proposition qui suit va nous éclairer, en se rappelant que pour un ordinal, être fini veut dire la même chose qu'être un entier naturel.

Proposition 71 (Ordinaux finis et maximum)

Soit α un ordinal.

Les assertions suivantes sont équivalentes :

1. α est un ordinal fini.
2. Toute partie non vide de α admet un maximum.

 *Démonstration*

$1 \Rightarrow 2$

Supposons que α est un ordinal fini.

Soit X une partie non vide de α .

On a évidemment $\forall \gamma \in \alpha, \gamma \in \alpha$ donc $\forall \gamma \in \alpha, \gamma < \alpha$ par définition de $<$.

Donc $\text{sup}(\alpha) \leq \alpha$ par minimalité de la borne supérieure.

Or $X \subseteq \alpha$ donc $\text{sup}(X) \leq \text{sup}(\alpha)$ et donc $\text{sup}(X) \leq \alpha$.

Or α est un ordinal fini par hypothèse.

Donc $\text{sup}(X) = 0$ ou $\text{sup}(X)$ est un successeur par définition.

Supposons par l'absurde que $\text{sup}(X) \notin X$.

Alors $\text{sup}(X)$ est un ordinal limite d'après la proposition 20 page 46.

On a donc nécessairement $\text{sup}(X) = 0$ par ce qui précède.

On a donc en particulier $\forall \xi \in X, \xi \leq 0$ donc $\forall \xi \in X, \xi \subseteq 0$.

Comme $0 = \emptyset$, on a donc $\forall \xi \in X, \xi = 0$.

Comme X est non vide, on a donc $0 \in X$ et donc $\sup(X) \in X$.

C'est absurde puisqu'on a justement supposé $\sup(X) \notin X$.

Par l'absurde on vient de montrer que $\sup(X) \in X$.

Autrement dit X admet un maximum.

Donc toutes les parties non vides de α admettent un maximum.

Donc si α est fini alors toutes les parties non vides de α admettent un maximum.

2⇒1

Supposons que toutes les parties non vides de α admettent un maximum.

Supposons par l'absurde que α n'est pas fini.

Il existe donc un ordinal $\beta \leq \alpha$ tel que $\beta \neq 0$ et β n'est pas un successeur.

Autrement dit β est un ordinal limite non nul donc non vide.

En particulier on a $\sup(\beta) = \beta$ d'après la proposition 21 page 47.

On a donc $\sup(\beta) \not\in \beta$ par antiréflexivité de $<$.

Autrement dit on a $\sup(\beta) \notin \beta$ par définition de $<$.

Donc β n'admet pas de maximum.

C'est absurde puisque β est une partie non vide de α .

Par l'absurde, on vient de montrer que α est fini.

Donc si toutes les parties non vides de α admettent un maximum alors α est fini.

CQFD.

Ainsi, chez les ordinaux avoir toutes ses parties non vides qui ont un maximum revient à être fini. Remarquons alors que pour $f : \beta \longrightarrow \alpha$, comme $\text{supp}(f)$ est une partie de l'ordinal β , c'est une partie d'un ensemble bien ordonné, si bien que $\text{supp}(f)$ est lui-même bien ordonné, et donc isomorphe à un ordinal, son type ! Autrement dit, la condition qui nous intéresse chez le support est simplement que son type soit fini.

Définition 25 (Application ordinale à support fini)

Soient β et α deux ordinaux et $f : \beta \longrightarrow \alpha$.

On dit que f est **à support fini** si et seulement si $\text{type}(\text{supp}(f))$ est un ordinal fini.

On note $\text{sf}(\beta \rightarrow \alpha)$ l'ensemble des applications $\beta \longrightarrow \alpha$ à support fini.

Remarque :

Certains auteurs notent $a^{(\beta)}$ l'ensemble $\text{sf}(\beta \rightarrow \alpha)$.

Exemple :

Essayons de voir à travers deux exemples comment comparer dans les faits deux éléments de $\text{sf}(\beta \rightarrow \alpha)$.

1. Prenons $\beta = \omega = \mathbb{N}$ et $\alpha = 2 = \{0, 1\}$, de sorte qu'une application $f : \beta \rightarrow \alpha$ est simplement une application qui à un entier renvoie 0 ou 1. C'est en soi le principe d'une fonction indicatrice : on renvoie 1 si l'entier est dans la partie, et 0 sinon.

Autrement dit, on peut voir $\mathcal{F}(\omega \rightarrow 2)$ comme l'ensemble des parties de \mathbb{N} , et $\text{sf}(\omega \rightarrow 2)$ comme l'ensemble des parties finies de \mathbb{N} (puisque alors on finit par ne plus valoir que 0, dont on finit par ne plus prendre aucun élément).

Comment alors comparer A et B deux parties finies de \mathbb{N} ? Si les deux sont vides, alors elles sont égales. Si l'une des deux est vide et l'autre non, alors la partie vide est strictement plus petite que la non vide. Sinon A et B sont toutes les deux non vides, on peut donc regarder le dernier élément de chacune, c'est-à-dire $\max(A)$ et $\max(B)$:

- ▶ si $\max(A) < \max(B)$ alors $A \prec B$,
- ▶ si $\max(A) > \max(B)$ alors $A \succ B$,
- ▶ si $\max(A) = \max(B)$ alors on considère A' et B' , qui sont A et B desquelles on a retiré ce maximum, et on réitère le procédé de comparaison sur A' et B' .

On est sûr que l'on va pouvoir s'arrêter car A et B sont finies ! C'est précisément pour ça que la notion de support fini est intéressante : on finira forcément par avoir épousé tout A ou tout B à force de retirer un élément à chaque étape, et donc par pouvoir dire laquelle des deux parties est la plus grande !

2. Plus généralement, si l'on prend deux applications f et g dans $\text{sf}(\beta \rightarrow \alpha)$, comment les comparer ? Considérons $\text{supp}(f)$ et $\text{supp}(g)$. Si les deux sont vides alors $f = g$. Si l'un des deux est vide et l'autre non, alors l'application dont le support est vide est strictement plus petite que l'autre. Sinon par définition les supports sont finis donc ont un maximum :

- ▶ si $\max(\text{supp}(f)) < \max(\text{supp}(g))$ alors $f \prec g$,
- ▶ si $\max(\text{supp}(f)) > \max(\text{supp}(g))$ alors $f \succ g$,
- ▶ si $\max(\text{supp}(f)) = \max(\text{supp}(g))$, notons γ ce maximum :
 - si $f(\gamma) < g(\gamma)$ alors $f \prec g$,
 - si $f(\gamma) > g(\gamma)$ alors $f \succ g$,
 - si $f(\gamma) = g(\gamma)$ alors on recommence le procédé de comparaison entre $f|_{<\gamma}$ et $g|_{<\gamma}$.

Les exemples précédents laissent entrevoir pourquoi nous avons en fait là un ensemble bien ordonné : disposant de X une partie non vide de $\text{sf}(\beta \rightarrow \alpha)$, on va pouvoir suivre la procédure de comparaison pour exhiber le minimum de X . Cette procédure de comparaison est basée sur une récursion : on vide peu à peu des ensembles finis de leur contenu jusqu'à épuisement, et la procédure s'arrête alors.

La proposition qui suit sera démontrée avec moins de rigueur que d'ordinaire : initialement j'ai souhaité la démontrer proprement, mais cela devenait rapidement fastidieux, pour un gain négligeable. Heureusement, celle-ci réussit sans aucun doute le test de rigueur de n'importe quel cursus mathématique, elle manque un peu de rigueur seulement selon le standard que l'on s'est imposé dans le Barbuki

Proposition 72 (Support fini bien ordonné)

Soient α et β deux ordinaux.

Alors $\text{sf}(\beta \rightarrow \alpha)$ muni de l'ordre anti-lexicographique est bien ordonné.

Idée de preuve

1. On s'intéresse aux parties finies de \mathbb{N} , c'est-à-dire comme décrit dans l'exemple ci-dessus au cas $\beta := \omega = \mathbb{N}$ et $\alpha := 2 = \{0, 1\}$.

Considérons X un ensemble non vide de parties finies de \mathbb{N} . On regarde alors l'ensemble $M := \{\max(A) \mid A \in X\}$ des maximums des éléments de X . En effet, cela revient à regarder le dernier élément de chaque $A \in X$: comme avec l'ordre anti-lexicographique la comparaison commence par la fin, c'est par là qu'il faut commencer à regarder. Le minimum de X se trouve alors parmi les $A \in X$ dont le dernier élément est le plus petit, c'est-à-dire les $A \in X$ tels que $\max(A) = \min(M)$.

On considère alors $X_M := \{A \in X \mid \max(A) = \min(M)\}$: le minimum de X se trouve parmi les éléments de X_M . Si X_M est un singleton, alors son unique élément est par définition (de l'ordre anti-lexicographique) le minimum de X . Sinon, on va s'intéresser plus spécifiquement aux éléments de X_M et opérer la comparaison sur eux. Pour cela, on leur retire ce maximum commun, c'est-à-dire que l'on regarde $X' := \{A \setminus \{\min(M)\} \mid A \in X_M\}$, et on recommence alors tout le processus décrit jusqu'ici, avec X' cette fois.

On est sûr que cela va s'arrêter à un moment car tous les éléments de X sont finis : le fait de retirer à chaque fois leur maximum va finir par les épuiser. Pour formaliser cela, on construit cette suite des maximums, qui va être alors une suite décroissante d'ordinaux et donc va finir par stationner d'après la condition de la chaîne descendante.

2. On s'intéresse aux triplets d'ordinaux, c'est-à-dire au cas où $\beta := 3 = \{0, 1, 2\}$ et α est quelconque. Dans ce cas-là, comme β est lui-même fini, cela correspond à tous les triplets possibles. On va cette fois spécifier X afin de fixer les idées.

Prenons $X := \{(45, 8, 1), (\omega^2, 12, 0), (145, 12, 89), (0, 999, 0), (1, 1, 1), (78, 12, 0)\}$.

Chacun de ces triplets a une dernière position non nulle : on note M l'ensemble de toutes les dernières positions non nulles : ici $M = \{1, 2\}$ (on commence à numérotter avec 0 !).

On ne va alors conserver que les triplets dont la dernière composante non nulle est en position $\min(M) = 1$. On considère alors $X_M := \{(\omega^2, 12, 0), (0, 999, 0), (78, 12, 0)\}$.

Par définition de l'ordre anti-lexicographique, le minimum de X se trouve forcément parmi

les éléments de X_M .

On regarde ensuite V l'ensemble des valeurs en cette position $\min(M) = 1$. Cela donne $V := \{12, 999\}$, et on considère alors la plus petite de ces valeurs : $\min(V) = 12$. On ne conserve alors plus que ceux de X_M qui ont cette plus petite valeur 12 en position 1, ce qui donne $X_V := \{(\omega^2, 12, 0), (78, 12, 0)\}$. Par définition de l'ordre anti-lexicographique, le minimum de X se trouve forcément parmi les éléments de X_V .

On considère alors X' l'ensemble des triplets de X_V mais restreints jusqu'à leur position avant 1, ce qui donne simplement $X' := \{\omega^2, 78\}$. On recommence alors la procédure sur X' , pour trouver 78 et donc en revenant à X le minimum est $(78, 12, 0)$.

3. Enfin, décrivons le cas général où α et β sont quelconques. On considère X un ensemble non vide d'applications $\beta \rightarrow \alpha$ à support finis.

On considère alors $M := \{\max(\text{supp}(f)) \mid f \in X\}$: c'est l'ensemble de toutes les dernières positions non nulles de chaque $f \in X$.

On s'intéresse à la plus petite de ces dernières positions $m := \min(M)$.

On ne conserve de X que $X_M := \{f \in X \mid \max(\text{supp}(f)) = m\}$, c'est-à-dire les applications de X dont le support se termine le plus tôt. Par définition de l'ordre anti-lexicographique, le minimum de X se trouve forcément dans X_M .

On regarde alors parmi toutes ces applications la valeur qu'elles prennent en cette dernière position : $V = \{f(m) \mid f \in X_M\}$, puis la plus petite de ces valeurs $v := \min(V)$. On ne conserve de X_M que celles qui prennent la valeur v en la position m , pour considérer $X_V := \{f \in X_M \mid f(m) = v\}$. Le minimum de X se trouve forcément parmi les éléments de X_V , par définition de l'ordre anti-lexicographique. Si X_V est un singleton, alors son unique élément est forcément le minimum de X et donc on s'arrête là. Sinon, on restreint tous les éléments de X_V jusqu'à m (non compris), c'est-à-dire que l'on considère $X' := \{f|_{<m} \mid f \in X_V\}$, et on recommence alors toute la procédure sur X' .

On est sûr que cela va s'arrêter car on épouse à chaque étape peu à peu le support des applications, que l'on sait sont finis. Plus rigoureusement, on construit la suite des minimum des dernières positions m qui va être décroissante, et donc finir par stationner d'après la condition de la chaîne descendante. Comme le fait de restreindre aux positions avant m fait que la prochaine dernière position est strictement plus petite que m , cette suite est strictement décroissante, sauf si l'on a totalement épousé le plus petit des supports. C'est donc juste avant de stationner que la suite des m va nous indiquer qui est le minimum. ■

De même que la proposition précédente n'a pas été démontrée rigoureusement, nous n'allons pas démontrer rigoureusement le théorème suivant, mais simplement en donner une idée de preuve. Redisons-le, cette preuve remplit haut la main les critères de rigueur de l'université : ce n'est que du point de vue de nos standard barbukiens qu'elle est moins rigoureuse.

Théorème 9 (Exponentiation d'ordinaux et supports finis)

Soient α et β deux ordinaux.

On munit $\text{sf}(\beta \rightarrow \alpha)$ de l'ordre anti-lexicographique.

On a alors $\text{type}(\text{sf}(\beta \rightarrow \alpha)) = \alpha^\beta$.

Démonstration

- **Cas où $\alpha = 0$**

Dans ce premier cas, on fixe un ordinal quelconque β .

- ▶ Plaçons-nous dans le cas où $\beta = 0$.

Il existe un unique application $\emptyset \longrightarrow \emptyset$, c'est-à-dire \emptyset lui-même.

Cette application est évidemment à support fini.

On en déduit que $\text{sf}(0 \rightarrow 0) = \text{sf}(\emptyset \rightarrow \emptyset) = \{\emptyset\} = \{0\} = 1 = 0^0$.

- ▶ Plaçons-nous dans le cas où $\beta \neq 0$.

Il n'existe aucune application $\beta \rightarrow \emptyset$, en particulier aucune à support fini.

On en déduit que $\text{sf}(\beta \rightarrow 0) = \text{sf}(\beta \rightarrow \emptyset) = \emptyset = 0 = 0^\beta$.

Dans les deux cas, on a $\text{sf}(\beta \rightarrow 0) = 0^\beta$, donc $\text{sf}(\beta \rightarrow 0) \cong 0^\beta$ par réflexivité de \cong .

- **Cas où $\alpha \neq 0$** .

Ici on ne fixe plus β et on montre la propriété par le principe faible d'induction.

Pour tout ordinal β , posons $P(\beta)$ l'assertion $\text{sf}(\beta \rightarrow \alpha) \cong \alpha^\beta$.

Initialisation

Il existe une unique application $\emptyset \rightarrow \alpha$, c'est-à-dire \emptyset lui-même, et elle est évidemment à support fini, si bien que $\text{sf}(0 \rightarrow \alpha) = \text{sf}(\emptyset \rightarrow \alpha) = \{\emptyset\} = \{0\} = 1 = \alpha^0$.

En particulier on a $\text{sf}(0 \rightarrow \alpha) \cong \alpha^0$ par réflexivité de \cong , et donc $P(0)$.

Hérédité

Soit β un ordinal tel que $P(\beta)$.

Commençons par remarquer que

$$\alpha^{\beta+1} = \alpha^\beta \alpha \text{ par définition de l'exponentiation}$$

$$\begin{aligned} &\cong \alpha \times \alpha^\beta \text{ d'après le théorème 8 page 143} \\ &\cong \alpha \times \text{sf}(\beta \rightarrow \alpha) \text{ d'après } P(\beta) \end{aligned}$$

Ainsi montrer que $\alpha^{\beta+1} \cong \text{sf}(\beta + 1 \rightarrow \alpha)$ revient à montrer que $\alpha \times \text{sf}(\beta \rightarrow \alpha) \cong \text{sf}(\beta + 1 \rightarrow \alpha)$. Pour cela, on introduit l'application suivante :

$$\varphi := \left(\begin{array}{ccc} \text{sf}(\beta + 1 \rightarrow \alpha) & \longrightarrow & \alpha \times \text{sf}(\beta \rightarrow \alpha) \\ f & \longmapsto & (f(\beta), f|_\beta) \end{array} \right)$$

► Montrons que φ est surjective sur $\alpha \times \text{sf}(\beta \rightarrow \alpha)$.

Soit $(\gamma, g) \in \alpha \times \text{sf}(\beta \rightarrow \alpha)$.

En posant $f : \beta + 1 \longrightarrow \alpha$ définie par $f(\beta) = \gamma$ et $f|_\beta = g$, on peut montrer que f est à support fini, et par définition de φ on a $\varphi(f) = (\gamma, g)$.

Le fait que f est à support fini vient du fait que g l'est :

- ou bien $f(\beta) = 0$ auquel cas $\text{supp}(f) = \text{supp}(g)$,
- ou bien $f(\beta) \neq 0$ auquel cas $\text{supp}(f) = \text{supp}(g) \cup \{\beta\}$ qui est fini car $\text{supp}(g)$ l'est.

Ainsi φ est surjective sur $\alpha \times \text{sf}(\beta \rightarrow \alpha)$.

► Montrons que φ est strictement croissante.

On note \prec l'ordre anti-lexicographique strict sur $\text{sf}(\beta + 1 \rightarrow \alpha)$ et $\text{sf}(\beta \rightarrow \alpha)$.

On note \sqsubset l'ordre lexicographique strict sur $\alpha \times \text{sf}(\beta \rightarrow \alpha)$.

Soient f et g dans $\text{sf}(\beta + 1 \rightarrow \alpha)$ telles que $f \prec g$.

Il existe donc $\gamma < \beta + 1$ tel que $\begin{cases} \forall \delta < \beta + 1, (\gamma < \delta \Rightarrow f(\delta) = g(\delta)) \\ f(\gamma) < g(\gamma) \end{cases}$

On a $\gamma < \beta + 1$ donc $\gamma \leq \beta$ d'après la proposition 13 page 33.

Si $\gamma = \beta$ alors $f(\beta) < g(\beta)$ donc $(f(\beta), f|_\beta) \sqsubset (g(\beta), g|_\beta)$.

Plaçons-nous à présent dans le cas où $\gamma < \beta$.

En prenant $\delta := \beta$ dans l'accolade ci-dessus, on obtient $f(\beta) = g(\beta)$.

Toujours dans l'accolade ci-dessus, on obtient en particulier

$$\begin{cases} \forall \delta < \beta, (\gamma < \delta \Rightarrow f|_\beta(\delta) = f(\delta) = g(\delta) = g|_\beta(\delta)) \\ f|_\beta(\gamma) = f(\gamma) < g(\gamma) = g|_\beta(\gamma) \end{cases}$$

si bien que $f|_\beta \prec g|_\beta$, et donc $(f(\beta), f|_\beta) \sqsubset (g(\beta), g|_\beta)$.

Dans les deux cas, on a $(f(\beta), f|_\beta) \sqsubset (g(\beta), g|_\beta)$, c'est-à-dire $\varphi(f) \sqsubset \varphi(g)$.

Ainsi φ est strictement croissante.

Or on a dit lors de la proposition 72 page 166 que $\text{sf}(\beta + 1 \rightarrow \alpha)$ est bien ordonné.

En particulier $\text{sf}(\beta + 1 \rightarrow \alpha)$ est totalement ordonné.

Donc le domaine de φ est totalement ordonné.

Donc φ est croissante et injective, donc φ est croissante et bijective sur $\alpha \times \text{sf}(\beta \rightarrow \alpha)$.

Toujours parce que le domaine de φ est totalement ordonné, on en déduit que φ est un isomorphisme d'ordres.

En particulier $\alpha \times \text{sf}(\beta \rightarrow \alpha) \cong \text{sf}(\beta + 1 \rightarrow \alpha)$, et donc $\alpha^{\beta+1} \cong \text{sf}(\beta + 1 \rightarrow \alpha)$.

On a donc $P(\beta + 1)$.

Ainsi pour tout ordinal β , si $P(\beta)$ alors $P(\beta + 1)$.

Hérité limite

Soit β un ordinal limite non nul tel que $\forall \gamma < \beta, P(\gamma)$.

Autrement dit, pour tout $\gamma < \beta$, $\text{sf}(\gamma \rightarrow \alpha)$ est isomorphe à α^γ .

Donc pour tout $\gamma < \beta$, il y a un isomorphisme $\text{sf}(\gamma \rightarrow \alpha) \longrightarrow \alpha^\gamma$, qui est en fait unique d'après la proposition 26 page 55 : notons-le φ_γ .

► Commençons par remarquer la chose suivante.

Soit $f \in \text{sf}(\beta \rightarrow \alpha)$.

Soit $\gamma < \beta$ un ordinal tel que $\max(\text{supp}(f)) < \gamma$.

Considérons alors $\iota := \begin{pmatrix} \text{sf}(\beta \rightarrow \alpha)_{\prec f} & \longrightarrow & \text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma} \\ g & \longmapsto & g|_\gamma \end{pmatrix}$.

Remarquons que ι est un isomorphisme d'ordres.

En effet, ι est surjective dans $\text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma}$.

Soit $h \in \text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma}$.

On prolonge h à β en complétant par des 0 : l'application g obtenue vérifie par définition $\iota(g) = g|_\gamma = h$. Par définition $h \prec f|_\gamma$ donc il existe $\delta < \gamma$ tel qu'au delà de δ , h et $f|_\gamma$ sont égales et tel que $h(\delta) < f|_\gamma(\delta)$. Mais g au delà de γ vaut 0 par définition, et il en va de même pour f puisque $\max(\text{supp}(f)) < \gamma$. Donc on obtient bien $g \prec f$ si bien que $g \in \text{sf}(\beta \rightarrow \alpha)_{\prec f}$ puisque le support de h étant fini, il en va de même pour g qui n'a pu prendre que des valeurs nulles en plus.

De plus, ι est strictement croissante.

Soient g et g' dans $\text{sf}(\beta \rightarrow \alpha)_{\prec f}$ telles que $g \prec g'$.

Il existe donc $\delta < \beta$ tel que g et g' coïncident au delà de δ et telles que $g(\delta) < g'(\delta)$. Comme $g \prec f$, $g' \prec f$ et $\max(\text{supp}(f)) < \gamma$,

on a aussi $\max(\text{supp}(g)) < \gamma$ et $\max(\text{supp}(g')) < \gamma$ par définition de l'ordre anti-lexicographique. Autrement dit, on a nécessairement $\delta < \gamma$ (sinon on aurait $g(\delta) = 0 = g'(\delta)$), si bien qu'on a $g|_\gamma \prec g'|_\gamma$ par définition de l'ordre anti-lexicographique, et donc $\iota(g) \prec \iota(g')$.

Or $\text{sf}(\beta \rightarrow \alpha)$ est bien ordonné d'après la proposition 72 page 166, donc $\text{sf}(\beta \rightarrow \alpha)_{\prec f}$ l'est aussi, donc en particulier est totalement ordonné d'après la proposition 2 page 10. Ainsi le domaine de ι est totalement ordonné, donc ι est croissante, injective et surjective dans $\text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma}$, et donc toujours car son domaine est totalement ordonné, ι est un isomorphisme d'ordres de $\text{sf}(\beta \rightarrow \alpha)_{\prec f}$ dans $\text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma}$.

Ainsi on a $\text{sf}(\beta \rightarrow \alpha)_{\prec f} \cong \text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma}$.

En particulier on a $\text{type}(\text{sf}(\beta \rightarrow \alpha)_{\prec f}) = \text{type}(\text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma})$.

Or par définition φ_γ est l'isomorphisme d'ordres entre $\text{sf}(\gamma \rightarrow \alpha)$ et son type, à savoir l'ordinal α^γ .

Donc $\varphi_\gamma(f|_\gamma) = \text{type}(\text{sf}(\gamma \rightarrow \alpha)_{\prec f|_\gamma})$ d'après la proposition 28 page 62.

On a donc $\varphi_\gamma(f|_\gamma) = \text{type}(\text{sf}(\beta \rightarrow \alpha)_{\prec f})$ par ce qui précède.

Ainsi on vient de voir que pour tout $\gamma < \beta$ tel que $\max(\text{supp}(f)) < \gamma$, on a $\varphi_\gamma(f|_\gamma) = \text{type}(\text{sf}(\beta \rightarrow \alpha)_{\prec f})$.

Cela veut en particulier dire que pour peu que $\max(\text{supp}(f)) < \gamma$, la valeur de $\varphi_\gamma(f|_\gamma)$ est indépendante de γ .

► On peut désormais construire l'isomorphisme $\text{sf}(\beta \rightarrow \alpha) \longrightarrow \alpha^\beta$.

Posons l'application suivante :

$$\varphi_\beta := \left(\begin{array}{rcl} \text{sf}(\beta \rightarrow \alpha) & \longrightarrow & \alpha^\beta \\ f & \longmapsto & \varphi_\gamma(f|_\gamma) \text{ où } \gamma < \beta \text{ est quelconque tel que } \max(\text{supp}(f)) < \gamma \end{array} \right)$$

Cette application est bien définie :

- Tout d'abord un tel γ existe nécessairement : pour $f \in \text{sf}(\beta \rightarrow \alpha)$, si l'on pose $\delta := \max(\text{supp}(f))$ alors on a par définition $\delta < \beta$, et comme β est limite on a $\delta + 1 < \beta$ d'après la proposition 14 page 37. On peut donc par exemple prendre $\gamma := \delta + 1$. Gardons bien à l'esprit le fait que pour $\gamma < \beta$ tel que $\delta < \gamma$, la valeur de $\varphi_\gamma(f|_\gamma)$ ne dépend pas de γ .
- Pour tout $\gamma < \beta$, on a $\alpha^\gamma < \alpha^\beta$ par stricte croissante de l'exponentiation à

droite, car β est limite non nul, donc $\alpha^\gamma \leq \alpha^\beta$ et donc $\alpha^\gamma \subseteq \alpha^\beta$ par définition de \leq . Comme φ_γ est à valeurs dans α^γ , il est donc en particulier à valeurs dans α^β , et donc l'application φ_β est bien à valeurs dans α^β .

Montrons que φ_β est strictement croissante.

Soient f et g dans $\text{sf}(\beta \rightarrow \alpha)$ telles $f \prec g$.

Considérons $\gamma < \beta$ tel que $\max(\text{supp}(f)) < \gamma$ et $\max(\text{supp}(g)) < \gamma$ (par exemple le max des deux, ajouté de 1).

Alors on a toujours $f|_\gamma \prec g|_\gamma$ (car on ne fait qu'enlever des termes où les deux applications valent 0).

On a donc $\varphi_\gamma(f|_\gamma) < \varphi_\gamma(g|_\gamma)$ par stricte croissance de φ_γ .

Or par définition de φ_β on a $\varphi_\beta(f) = \varphi_\gamma(f|_\gamma)$ et $\varphi_\beta(g) = \varphi_\gamma(g|_\gamma)$.

On a donc $\varphi_\beta(f) < \varphi_\beta(g)$.

Donc φ_β est strictement croissante.

Or son domaine $\text{sf}(\beta \rightarrow \alpha)$ est bien ordonné d'après la proposition 72 page 166.

Donc le domaine de φ_β est totalement ordonné d'après la proposition 2 page 10.

Donc φ_β est croissante et injective.

Montrons que φ_β est surjective dans α^β .

Par définition de φ_β on sait déjà que $\text{im}(\varphi_\beta) \subseteq \alpha^\beta$.

Montrons que $\text{im}(\varphi_\beta) \supseteq \alpha^\beta$.

Soit $\delta \in \alpha^\beta$.

On a donc $\delta < \alpha^\beta$ par définition de $<$.

Or on a $\alpha^\beta = \sup_{\gamma < \beta} \alpha^\gamma$ par définition de l'exponentiation.

Il existe donc $\gamma < \beta$ tel que $\delta < \alpha^\gamma$ et donc $\delta \in \alpha^\gamma$ par définition de $<$.

Or par définition $\varphi_\gamma : \text{sf}(\gamma \rightarrow \alpha) \longrightarrow \alpha^\gamma$ est un isomorphisme d'ordres.

En particulier φ_γ est surjectif dans α^γ .

Il existe donc $g \in \text{sf}(\gamma \rightarrow \alpha)$ tel que $\varphi_\gamma(g) = \delta$.

Posons alors $f : \beta \longrightarrow \alpha$ définie par $f|_\gamma = g$ et que l'on complète à β par des 0. Par définition on a alors $\text{supp}(f) = \text{supp}(g)$.

En particulier comme g est à support fini, f l'est aussi donc $f \in \text{sf}(\beta \rightarrow \alpha)$.

De plus par définition $\max(\text{supp}(g)) < \gamma$ donc $\max(\text{supp}(f)) < \gamma$.

En particulier $\varphi_\beta(f) = \varphi_\gamma(f|_\gamma) = \varphi_\gamma(g) = \delta$.

Donc $\delta \in \text{im}(\varphi_\beta)$.

Ainsi on a $\text{im}(\varphi_\beta) \supseteq \alpha^\beta$ et donc $\text{im}(\varphi_\beta) = \alpha^\beta$.

Ainsi φ_β est surjective dans α^β .

Finalement, $\varphi_\beta : \text{sf}(\beta \rightarrow \alpha) \longrightarrow \alpha^\beta$ est croissante, injective et surjective dans α^β .

Or on a déjà dit que son domaine $\text{sf}(\beta \rightarrow \alpha)$ est totalement ordonné.

Donc $\varphi_\beta : \text{sf}(\beta \rightarrow \alpha) \longrightarrow \alpha^\beta$ est un isomorphisme d'ordres.

En particulier $\text{sf}(\beta \rightarrow \alpha) \cong \alpha^\beta$ et donc $P(\beta)$.

Ainsi pour tout ordinal limite non nul β , si $\forall \gamma < \beta, P(\gamma)$ alors $P(\beta)$.

Finalement, P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal β , on a $P(\beta)$.

Autrement dit pour tout ordinal β , on a $\text{sf}(\beta \rightarrow \alpha) \cong \alpha^\beta$ et donc $\boxed{\text{type}(\text{sf}(\beta \rightarrow \alpha)) = \alpha^\beta}$.

CQFD.

5 Forme normale de Cantor et ε_0

5.1 Logarithme ordinal et forme normale de Cantor

On l'a dit plus tôt, nous allons évoquer l'opération "contraire" de l'exponentiation, à savoir la généralisation de la notion de **logarithme** aux ordinaux. Cela permettra alors de déboucher sur une généralisation de la décomposition d'un entier dans une base donnée (par exemple décomposer un entier en base 10), généralisation qui s'appelle la **forme normale de Cantor**.

Pour l'heure, commençons par revenir un peu sur les assertions fonctionnelles. Nous avons déjà donné du sens à la notion d'assertion fonctionnelle croissante et continue lors de la définition 19 page 102. Intéressons-nous cette fois à la stricte croissance.

Définition 26 (Assertion fonctionnelle strictement croissante)

Soient C et D deux classes d'**ordinaux**, et $F : C \longrightarrow D$ une assertion fonctionnelle. On dit que F est **strictement croissante** si et seulement si pour tout α et β dans C , on a

$$\alpha < \beta \implies F(\alpha) < F(\beta)$$

Premier fait remarquable : une assertion fonctionnelle qui est strictement croissante n'est pas bornée, c'est-à-dire qu'elle finit par dépasser n'importe quel ordinal donné.

Proposition 73 (Stricte croissance et absence de borne)

Soit $F : ON \longrightarrow ON$ une assertion fonctionnelle **strictement croissante**.

1. Pour toute ordinal α , on a $F(\alpha) \geq \alpha$.
 2. Pour tout ordinal β , il existe un ordinal α tel que $\beta < F(\alpha)$.
- On dit que F est **non bornée**.

Démonstration

1. Montrons ce résultat par induction.

Pour tout ordinal α , on pose $P(\alpha)$ l'assertion « $\alpha \leq F(\alpha)$ ».

Initialisation

Par définition F est à valeurs dans les ordinaux donc $F(0)$ est un ordinal.

Or 0 est le plus petit des ordinaux donc $0 \leq F(0)$ et donc $P(0)$.

Hérédité

Soit α un ordinal tel que $P(\alpha)$.

On a $\alpha < \alpha + 1$ donc $F(\alpha) < F(\alpha + 1)$ par stricte croissance de F .

Or on a $\alpha \leq F(\alpha)$ d'après $P(\alpha)$.

On a donc $\alpha < F(\alpha + 1)$ par transitivité.

On a donc $\alpha + 1 \leq F(\alpha + 1)$ d'après la proposition 13 page 33.

Autrement dit on a $P(\alpha + 1)$.

Donc pour tout ordinal α , si $P(\alpha)$ alors $P(\alpha + 1)$.

Hérité de limite

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

Soit β un ordinal tel que $\beta < \alpha$.

On a alors $F(\beta) < F(\alpha)$ par stricte croissance de F .

Or on a aussi $\beta \leq F(\beta)$ d'après $P(\beta)$.

On a donc $\beta < F(\alpha)$ par transitivité.

Ainsi $\forall \beta < \alpha, \beta < F(\alpha)$.

En particulier $\sup_{\beta < \alpha} \beta \leq F(\alpha)$ par minimalité de la borne supérieure.

Or on a $\alpha = \sup_{\beta \in \alpha} \beta = \sup_{\beta < \alpha} \beta$ d'après la proposition 21 page 47.

On a donc $\alpha \leq F(\alpha)$ et donc $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Ainsi P vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal α on a $P(\alpha)$.

Autrement dit pour tout ordinal α , on a $\alpha \leq F(\alpha)$.

2. En particulier pour tout ordinal β , on a $\beta < \beta + 1 \leq F(\beta + 1)$.

Donc F est non bornée.

CQFD.

Le fait que l'on vient de voir va pouvoir nous servir ici : pour α et β deux ordinaux donnés, si $\beta > 1$ alors il y a forcément un moment où β^γ dépassera α , puisque $\gamma \mapsto \beta^\gamma$ est strictement croissante donc n'est pas bornée. On se place juste avant de dépasser α pour trouver la puissance de β juste en dessous de α . Il ne reste alors plus qu'à effectuer la division ordinal de α par cette puissance pour conclure.

Proposition 74 (Logarithme ordinal)

Soient α et β deux ordinaux tels que $\alpha > 0$ et $\beta > 1$.

Il existe des ordinaux λ , δ et σ tels que $\alpha = \beta^\lambda \delta + \sigma$, avec $0 < \delta < \beta$ et $\sigma < \beta^\lambda$.

De plus, de tels ordinaux sont uniques.

 *Démonstration*

Existence

Considérons l'assertion fonctionnelle $F : \begin{pmatrix} ON & \longrightarrow & ON \\ \gamma & \longmapsto & \beta^\gamma \end{pmatrix}$.

Comme $\beta > 1$, F est strictement croissante d'après la proposition 66 page 151.

En particulier F est non bornée d'après la proposition 73 page 174.

Il existe donc un ordinal γ tel que $\alpha < F(\gamma) = \beta^\gamma$.

Considérons la classe $A := \{\varepsilon \in ON \mid \beta^\varepsilon \leq \alpha\}$.

Montrons que A est majorée par γ .

Supposons par l'absurde que γ ne majore pas A .

Il existe donc $\varepsilon \in A$ tel que $\varepsilon \not\leq \gamma$.

Or les ordinaux sont totalement ordonnés donc on a $\gamma < \varepsilon$.

Comme $\beta > 1$, on a donc $\beta^\gamma < \beta^\varepsilon$ par stricte croissance de l'exponentiation à gauche.

Or on a $\alpha < \beta^\gamma$ par définition de γ , donc $\alpha < \beta^\varepsilon$ par transitivité de $<$.

C'est absurde puisque $\varepsilon \in A$.

Ainsi A est majorée par γ .

En particulier A est un ensemble d'après la proposition 12 page 32.

En particulier A admet une borne supérieure d'après la proposition 11 page 29.

Considérons donc $\lambda := \sup(A)$.

On a alors $\beta^\lambda = \beta^{\sup(A)} = \sup_{\varepsilon \in A} \beta^\varepsilon$ par continuité de l'exponentiation à gauche.

Or pour tout $\varepsilon \in A$, on a justement $\beta^\varepsilon \leq \alpha$ par définition de A .

Donc $\sup_{\varepsilon \in A} \beta^\varepsilon \leq \alpha$ par minimalité de la borne supérieure, et donc $\beta^\lambda \leq \alpha$.

Il existe deux ordinaux δ et σ tels que $\boxed{\alpha = \beta^\lambda \delta + \sigma}$ d'après la proposition 63 page 147.

De plus ceux-ci vérifient $\delta \leq \alpha$ et $\boxed{\sigma < \beta^\lambda}$, et sont uniques.

Supposons par l'absurde que $\delta = 0$.

On a alors $\alpha = \beta^\lambda \delta + \sigma = \beta^\lambda \cdot 0 + \sigma = 0 + \sigma = \sigma$.

Or on a dit que $\sigma < \beta^\lambda$ donc $\alpha < \beta^\lambda$.

C'est absurde puisqu'on a justement montré que $\beta^\lambda \leq \alpha$.

On a donc montré par l'absurde que $\boxed{0 < \delta}$.

Supposons par l'absurde que $\beta \leq \delta$.

Il existe deux ordinaux μ et ν tels que $\delta = \beta\mu + \nu$ avec $\mu \leq \delta$ et $\nu < \beta$ d'après la proposition 63 page 147.

On a alors $\alpha = \beta^\lambda \delta + \sigma = \beta^\lambda(\beta\mu + \nu) + \sigma = \beta^\lambda\beta\mu + \beta^\lambda\nu + \sigma = \beta^{\lambda+1}\mu + \beta^\lambda\nu + \sigma$.

Supposons par l'absurde que $\mu = 0$.

On a alors $\delta = \beta\mu + \nu = \beta \cdot 0 + \nu = 0 + \nu = \nu$.

Or on a dit que $\nu < \beta$ donc $\delta < \beta$, ce qui est absurde puisque $\beta \leq \delta$.

Ainsi on a montré par l'absurde que $\mu > 0$, et donc $\mu \geq 1$.

De plus $\beta^\lambda\nu + \sigma$ est un ordinal donc $\beta^\lambda\nu + \sigma \geq 0$.

Autrement dit on a

$$\begin{aligned}\alpha &= \beta^{\lambda+1}\mu + \beta^\lambda\nu + \sigma \\ &\geq \beta^{\lambda+1}\mu + 0 \text{ par croissance de l'addition à gauche} \\ &= \beta^{\lambda+1}\mu \\ &\geq \beta^{\lambda+1} \cdot 1 \text{ par croissance de la multiplication à gauche} \\ &= \beta^{\lambda+1}\end{aligned}$$

Autrement dit on a $\beta^{\lambda+1} \leq \alpha$.

En particulier $\lambda + 1 \in A$: c'est absurde puisque $\lambda = \sup(A)$.

Ainsi on a montré par l'absurde que $\boxed{\beta \leq \delta}$.

Unicité

Soient λ' , δ' et σ' des ordinaux tels que $\alpha = \beta^{\lambda'}\delta' + \sigma'$, avec $0 < \delta' < \beta$ et $\sigma' < \beta^{\lambda'}$.

Comme tout à l'heure on a $\delta' \geq 1$ et $\sigma' \geq 0$, si bien que

$$\alpha = \beta^{\lambda'}\delta' + \sigma' \geq \beta^{\lambda'}\delta' + 0 = \beta^{\lambda'}\delta' \geq \beta^{\lambda'} \cdot 1 = \beta^{\lambda'}$$

Ainsi $\beta^{\lambda'} \leq \alpha$ et donc $\lambda' \in A$ donc $\lambda' \leq \sup(A) = \lambda$.

On a donc $\lambda' = \lambda$ ou $\lambda' < \lambda$.

Supposons par l'absurde que $\lambda' < \lambda$.

En particulier on a $\lambda' + 1 \leq \lambda$ d'après la proposition 13 page 33.

Par définition on a $\delta' < \beta$ donc $\delta' + 1 \leq \beta$ toujours d'après la même proposition.

On a donc

$$\begin{aligned}\beta^{\lambda'}\delta' + \sigma' &< \beta^{\lambda'}\delta' + \beta^{\lambda'} \text{ car par définition } \sigma' < \beta^{\lambda'} \\ &= \beta^{\lambda'}\delta' + \beta^{\lambda'} \cdot 1 = \beta^{\lambda'}(\delta' + 1) \\ &\leq \beta^{\lambda'}\beta \text{ car on a } \delta' + 1 \leq \beta \\ &= \beta^{\lambda'+1} \leq \beta^\lambda \text{ car on a } \lambda' + 1 \leq \lambda \\ &= \beta^\lambda \cdot 1 \leq \beta^\lambda\delta \text{ car } 1 \leq \delta \\ &= \beta^\lambda\delta + 0 \leq \beta^\lambda\delta + \sigma \\ &= \alpha\end{aligned}$$

On a donc $\beta^{\lambda'}\delta' + \sigma' < \alpha$, ce qui est absurde.

Par l'absurde on vient de montrer que $\lambda' = \lambda$.

Ainsi on a $\beta^\lambda\delta + \sigma = \alpha = \beta^\lambda\delta' + \sigma'$.

On a donc $\delta = \delta'$ et $\sigma = \sigma'$ par unicité dans la division ordinaire.

D'où l'unicité voulue.

CQFD.

Abordons à présent la notion de **forme normale de Cantor**. Nous n'allons pas lui donner de démonstration rigoureuse : il s'agit plutôt d'un aperçu de ce qui existe chez les ordinaux pour le lecteur qui serait intéressé pour aller plus loin. Cette discussion est en grande partie inspirée du billet de blog de David Madore, intitulé "*Nombres ordinaux : une (longue) introduction*", daté du 18 septembre 2011.

L'idée est de généraliser la notion de décomposition dans une base donnée que l'on retrouve chez les entiers. Par exemple la base la plus commune est la base 10 : le nombre 734 s'écrit alors $7 \cdot 10^2 + 3 \cdot 10^1 + 4 \cdot 10^0$, et on peut voir que chaque facteur dans cette décomposition (chaque chiffre donc) est un entier strictement inférieur à 10. Plus généralement, un entier a s'écrira dans la base b sous la forme

$$a = d_0 b^{\ell_0} + d_1 b^{\ell_1} + \cdots + d_n b^{\ell_n}$$

avec un nombre n de termes finis. Les d_i sont tous des entiers strictement inférieurs à b et les ℓ_i sont strictement décroissants : $\ell_0 > \ell_1 > \cdots > \ell_n$. Ici la base n'est plus forcément un entier mais un ordinal β (généralement ω). Ainsi, pour un ordinal quelconque α , on pourra écrire

$$\alpha = \beta^{\lambda_0}\delta_0 + \beta^{\lambda_1}\delta_1 + \cdots + \beta^{\lambda_\nu}\delta_\nu$$

avec un nombre ν de termes potentiellement transfinis ! Les δ_i sont tous des ordinaux strictement inférieurs à β , et les λ_i sont strictement décroissants : $\lambda_0 > \lambda_1 > \cdots > \lambda_\nu$. On interdit généralement aux δ_i d'être nuls (sauf pour $\alpha = 0$).

Cette décomposition est alors unique, si bien qu'elle fournit une façon standard de comparer deux ordinaux, la comparaison s'effectuant à nouveau via l'ordre lexicographique. Supposons par exemple que l'on veuille comparer les ordinaux $\alpha = \beta^{\lambda_0}\delta_0 + \beta^{\lambda_1}\delta_1 + \cdots + \beta^{\lambda_\nu}\delta_\nu$ et $\alpha' = \beta^{\kappa_0}\varepsilon_0 + \beta^{\kappa_1}\varepsilon_1 + \cdots + \beta^{\kappa_\mu}\varepsilon_\mu$:

1. On commence par comparer les premiers exposants λ_0 et κ_0 : si $\lambda_0 < \kappa_0$ alors $\alpha < \alpha'$, si $\lambda_0 > \kappa_0$ alors $\alpha > \alpha'$ et dans ces deux cas on s'arrête-là. Sinon, on continue à l'étape suivante.
2. On compare alors les premiers facteurs δ_0 et ε_0 : si $\delta_0 < \varepsilon_0$ alors $\alpha < \alpha'$, si $\delta_0 > \varepsilon_0$ alors $\alpha > \alpha'$ et dans ces deux cas on s'arrête-là. Sinon, on continue à l'étape suivante.
3. On considère alors $\pi := \beta^{\lambda_1}\delta_1 + \cdots + \beta^{\lambda_\nu}\delta_\nu$ et $\pi' := \beta^{\kappa_1}\varepsilon_1 + \cdots + \beta^{\kappa_\mu}\varepsilon_\mu$ puis on recommence l'étape 1 avec π et π' .

L'existence et l'unicité de la décomposition en base β repose simplement sur l'existence et l'unicité du logarithme ordinal. En effet, pour décomposer α dans la base β :

1. On utilise une première fois le logarithme ordinal entre α et β . Il existe des ordinaux λ_0 , δ_0 et α_1 tels que $\alpha = \beta^{\lambda_0}\delta_0 + \alpha_1$. Ces trois ordinaux sont uniques et vérifient en particulier $\alpha_1 < \beta^{\lambda_0}$ et $0 < \delta_0 < \beta$.

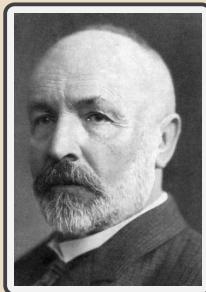
2. On recommence alors le logarithme ordinal entre α_1 et β : il existe trois ordinaux λ_1, δ_1 et α_2 tels que $\alpha_1 = \beta^{\lambda_1} \delta_1 + \alpha_2$. Ces trois ordinaux sont uniques et vérifient en particulier $\alpha_2 < \beta^{\lambda_1}$ et $0 < \delta_1 < \beta$.
3. On a donc $\alpha = \beta^{\lambda_0} \delta_0 + \beta^{\lambda_1} \delta_1 + \alpha_2$. On recommence encore à nouveau avec α_2 .

On remarque alors plusieurs choses :

- Pour chaque i , on a $0 < \delta_i < \beta$, donc les facteurs sont tous non nuls et plus petits que la base, comme demandé.
- Pour chaque i , on a $\alpha_{i+1} < \beta^{\lambda_i}$, et on a vu dans la démonstration du logarithme ordinal que $\beta^{\lambda_{i+1}} \leq \alpha_{i+1}$, si bien que $\beta^{\lambda_{i+1}} < \beta^{\lambda_i}$, et donc $\lambda_{i+1} < \lambda_i$, donc la suite des exposants est bien strictement décroissante, comme demandé.
- Justement, la suite des exposants λ_i est une suite d'ordinaux strictement décroissante : elle finit par s'arrêter au vu de la proposition 34 page 84, ce qui prouve que l'algorithme de décomposition en base β s'arrête à un moment.

Une autre remarque à présent : la forme véritablement normale de Cantor demande en réalité aux exposants λ_i d'être eux-mêmes décomposés sous cette forme : une fois l'algorithme précédent déployé, on recommence celui-ci pour chacun des λ_i . Malheureusement cette nouvelle étape sur les exposants n'est pas garantie de s'arrêter, et nous allons justement voir pourquoi avec l'introduction d'un nouvel ordinal noté ε_0 .

Pour la petite histoire



Georg Cantor (3 mars 1845 – 6 janvier 1918) est un mathématicien allemand, connu pour être le créateur de la théorie des ensembles. Il établit l'importance de la bijection entre les ensembles, définit les ensembles infinis et les ensembles bien ordonnés. Il prouva également que les nombres réels sont « *plus nombreux* » que les entiers naturels. En fait, il démontra même l'existence d'une « *infinité d'infinis* ». C'est aussi à lui que l'on doit la théorie des ordinaux, mais aussi des cardinaux que l'on va voir dans le prochain chapitre. Nous verrons dans quel contexte tout ceci s'est présenté à lui à la toute fin de ce livre.

Cantor a été confronté à la résistance de la part des mathématiciens de son époque, en particulier Kronecker. Poincaré, bien qu'il connût et appréciait les travaux de Cantor, avait de profondes réserves sur son maniement de l'infini en tant que totalité achevée. Dans le but de contrer les détracteurs de Cantor, Hilbert a affirmé : « *Nul ne doit nous exclure du Paradis que Cantor a créé.* »

On peut ainsi dire qu'il est le père des théories abordées dans ces deux premiers ouvrages !

5.2 L'ordinal ε_0 et la classe des points fixes

L'ordinal ε_0 peut être vu comme étant $\omega^{\omega^{\dots}}$ avec ω fois l'exponentiation (une infinité donc!). Le problème dont nous venons de parler vient quand on essaie de décomposer ε_0 dans la base ω :

$$\varepsilon_0 = \omega^{\omega^{\omega^{\dots}}} = \omega^{\omega^{\omega^{\omega^{\dots}}}} = \omega^{\varepsilon_0}$$

Ainsi $\varepsilon_0 = \omega^{\varepsilon_0}$ est lui-même son premier (et dernier) exposant dans la décomposition en base ω . Autrement dit, maintenant qu' ε_0 est décomposé, l'étape suivant est de décomposer tous ses exposants, à savoir à nouveau ε_0 . On se retrouve donc à écrire

$$\varepsilon_0 = \omega^{\varepsilon_0} = \omega^{\omega^{\varepsilon_0}} \text{ puis à l'étape d'après } \varepsilon_0 = \omega^{\omega^{\varepsilon_0}} = \omega^{\omega^{\omega^{\varepsilon_0}}}$$

et cela n'en finit jamais ! On voit donc bien qu'il y a un problème : pour palier celui-ci on considère généralement qu' ε_0 est **un exposant** correctement décomposé, et qu'il n'y a plus rien à faire. On peut donc par exemple dire que les nombres

$$\omega^{\varepsilon_0} \cdot 2 + \omega^3 + 5 \text{ et } \omega^{\varepsilon_0^2} + \omega^2 \cdot 7$$

sont correctement décomposés. Cependant, le problème va de nouveau se présenter avec

$$\varepsilon_1 := \varepsilon_0^{\varepsilon_0^{\varepsilon_0^{\dots}}}$$

exactement pour les mêmes raisons. De la même manière, on va pouvoir définir l'ordinal ε_2 et ainsi de suite. Mais tout d'abord, comment définit les ε_i rigoureusement parlant ?

L'idée est de se dire "on a itéré ω fois des puissances pour définir ε_0 ". De la même manière que l'addition est une itération de successeurs, la multiplication une itération d'additions, et l'exponentiation une itération de multiplication, on pourrait imaginer une opération, la **tétration**, qui serait une itération de l'exponentiation. Par exemple $T(2, 3) = 2^{2^2}$ et plus généralement par récursion on poserait

$$\begin{cases} T(\alpha, 0) := 1 \\ T(\alpha, \beta + 1) := \alpha^{T(\alpha, \beta)} \text{ pour tout ordinal } \beta \\ T(\alpha, \gamma) := \sup_{\delta < \gamma} T(\alpha, \delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

puis il ne reste plus qu'à poser $\varepsilon_0 := T(\omega, \omega)$. Ainsi, ε_0 est la limite de la suite $1, \omega, \omega^\omega, \omega^{\omega^\omega}, \dots$

On peut se contenter à nouveau d'itérer cette construction, encore et encore pour définir ε_1 , ε_2 et ainsi de suite, mais le fait que $\varepsilon_0 = \omega^{\varepsilon_0}$ nous laisse entrevoir qu'il peut être intéressant de creuser du côté des points fixes de l'assertion fonctionnelle $\alpha \mapsto \omega^\alpha$. Pour l'heure, revenons un peu plus longuement sur les assertions fonctionnelles pour justement nous munir de tous les outils nécessaires. Plus précisément, nous allons généraliser et montrer des résultats que nous connaissons bien sur les applications : rien ne devrait nous surprendre, au moins au début.

Retour sur les assertions fonctionnelles

Définition 27 (Assertion fonctionnelle bijective)

Soient F une assertion fonctionnelle et D une classe.

1. On dit que F est **injective** si et seulement si pour tout x et x' dans $\text{dom}(F)$, on a

$$F(x) = F(x') \implies x = x'$$

2. On dit que F est **surjective** sur D si et seulement si $\text{im}(F) = D$.
3. On dit que F est **bijective** de $\text{dom}(F)$ dans D si et seulement si F est injective et surjective sur D .

Exemple :

Pour une classe C donnée, l'assertion fonctionnelle **identité** $\text{id}_C := \begin{pmatrix} C & \longrightarrow & C \\ x & \longmapsto & x \end{pmatrix}$ est évidemment bijective de C dans C .

Définition 28 (Assertion fonctionnelle réciproque)

Soit F une assertion fonctionnelle **injective**.

On appelle **réciproque** de F l'assertion fonctionnelle $F^{-1} : \text{im}(F) \longrightarrow \text{dom}(F)$ définie pour tout $y \in \text{im}(F)$ par $F^{-1}(y)$ est l'unique $x \in \text{dom}(F)$ tel que $y = F(x)$.

Autrement dit pour tout $x \in \text{dom}(F)$ et $y \in \text{im}(F)$, on a l'équivalence

$$y = F(x) \iff F^{-1}(y) = x$$

Exemple :

Pour une classe C donnée, l'assertion fonctionnelle id_C est sa propre réciproque.

Proposition 75 (Propriétés de la réciproque)

Soit F une assertion fonctionnelle **injective**.

1. Pour tout $x \in \text{dom}(F)$, on a $F^{-1}(F(x)) = x$.
2. Pour tout $y \in \text{im}(F)$, on a $F(F^{-1}(y)) = y$.
3. F^{-1} est injective et on a $(F^{-1})^{-1} = F$.

Démonstration

1. Soit $x \in \text{dom}(F)$.

Posons $z := F^{-1}(F(x))$.

Par définition de F^{-1} , on a $F(z) = F(x)$.

Par injectivité de F , on a $z = x$.

On a donc $F^{-1}(F(x)) = x$.

2. Soit $y \in \text{im}(F)$.

Par définition de $\text{im}(F)$, il existe $x \in \text{dom}(F)$ tel que $y = F(x)$.

Par définition de F^{-1} , on a alors $x = F^{-1}(y)$.

Ainsi on a $F(F^{-1}(y)) = F(x) = y$.

3. Montrons que F^{-1} .

Soient y et y' dans $\text{im}(F)$ tels que $F^{-1}(y) = F^{-1}(y')$.

On a donc $F(F^{-1}(y)) = F(F^{-1}(y'))$ et donc $y = y'$ d'après 2.

Donc F^{-1} est injective.

De plus par définition de F^{-1} et de $(F^{-1})^{-1}$, pour tout $x \in \text{dom}(F)$ et $y \in \text{im}(F)$, on a

$$y = (F^{-1})^{-1}(x) \iff F^{-1}(y) = x \iff y = F(x)$$

et donc $(F^{-1})^{-1} = F$.

CQFD.

Ce qui suit pourrait être vu dans le cadre de classes ordonnées, mais cela demanderait de généraliser la notion de relation d'ordres aux classes. Ce ne serait pas spécialement difficile à faire, mais cela ne nous intéresserait pas particulièrement, car les classes ne nous servent ici que dans le cas particulier des classes d'ordinaux. L'avantage est que l'on gagne automatiquement le fait que ce sont des classes totalement ordonnées, si bien que l'on a par exemple le résultat suivant. Rappelons-nous que dans le cas des ensembles ordonnées, l'implication $1 \Rightarrow 2$ n'a pas de raison d'être vraie si l'ensemble de départ n'est pas totalement ordonné.

Proposition 76 (Stricte croissance et injectivité)

Soient C et D deux classes d'ordinaux et $F : C \longrightarrow D$ une assertion fonctionnelle.
Les assertions suivantes sont équivalentes :

1. F est strictement croissante.
2. F est croissante et injective.

 *Démonstration*

$1 \Rightarrow 2$

Supposons que F est strictement croissante.

Soient α et β dans C tels que $\alpha \leq \beta$.

On a donc $\alpha = \beta$ ou $\alpha < \beta$.

Si $\alpha = \beta$ alors $F(\alpha) = F(\beta)$.

Si $\alpha < \beta$ alors $F(\alpha) < F(\beta)$ par stricte croissance de F .

Dans les deux cas on a en particulier $F(\alpha) \leq F(\beta)$.

Donc $[F$ est croissante].

Soient α et β dans C tels que $F(\alpha) = F(\beta)$.

Comme α et β sont des ordinaux, on a $\alpha < \beta$ ou $\alpha = \beta$ ou $\beta < \alpha$.

Supposons par l'absurde que $\alpha \neq \beta$.

On a donc $\alpha < \beta$ ou $\beta < \alpha$.

Si $\alpha < \beta$ alors $F(\alpha) < F(\beta)$ par stricte croissance de F .

Si $\beta < \alpha$ alors $F(\beta) < F(\alpha)$ par stricte croissance de F .

Dans les deux cas on a en particulier $F(\alpha) \neq F(\beta)$.

C'est absurde puisqu'on a justement fait l'hypothèse que $F(\alpha) = F(\beta)$.

Par l'absurde on vient de montrer que $\alpha = \beta$.

Donc $[F$ est injective].

$1 \Leftarrow 2$

Supposons que F est croissante et injective.

Soient α et β dans C tels que $\alpha < \beta$.

En particulier on a $\alpha \leq \beta$ donc $F(\alpha) \leq F(\beta)$ par croissance de F .

On a donc $F(\alpha) < F(\beta)$ ou $F(\alpha) = F(\beta)$.

Supposons par l'absurde que $F(\alpha) = F(\beta)$.

On a alors $\alpha = \beta$ par injectivité de F .

C'est absurde puisqu'on a $\alpha < \beta$ par hypothèse, et $<$ est anti-réfléxive.

Par l'absurde on vient de montrer que l'on a $F(\alpha) \neq F(\beta)$ et donc $F(\alpha) < F(\beta)$.

Donc $[F$ est strictement croissante].

CQFD.

Définition 29 (Isomorphisme entre classes d'ordinaux)

Soient C et D deux classes d'ordinaux, et $F : C \rightarrow D$ une assertion fonctionnelle. On dit que F est un **isomorphisme d'ordres** si et seulement si :

1. F est bijective de C dans D .
2. F est croissante.
3. F^{-1} est croissante.

On dit alors que C et D sont **isomorphes**, et on note $C \cong D$.

Proposition 77 (Propriétés de l'isomorphie d'ordres)

Soient C et D deux classes d'ordinaux.

1. $\text{id}_C : C \rightarrow C$ est un isomorphisme d'ordres.
En particulier on a $C \cong C$: on dit que l'isomorphie d'ordres est **réflexive**.
2. Supposons qu'il existe $F : C \rightarrow D$ un isomorphisme d'ordres.
Alors $F^{-1} : D \rightarrow C$ est un isomorphisme d'ordres.
En particulier si $C \cong D$ alors $D \cong C$.
On dit que l'isomorphie d'ordres est **symétrique**.



Démonstration

1. On a déjà vu lors de précédents exemples que id_C est injective, et est sa propre réciproque. Elle est de plus évidemment surjective dans C .
Elle est aussi évidemment croissante : si $\alpha \leq \beta$ alors par $\text{id}_C(\alpha) = \alpha \leq \beta = \text{id}_C(\beta)$.
Ainsi id_C répond à toutes les conditions d'un isomorphisme d'ordres de C vers C .
2. Supposons qu'il existe $F : C \rightarrow D$ un isomorphisme d'ordres.
Par définition, F est bijective de C vers D .
Toujours par définition, F et F^{-1} sont croissantes.
Par définition de F^{-1} , F^{-1} est bijective de D vers C .
De plus $(F^{-1})^{-1} = F$ d'après la proposition 75 page 181.
Donc F^{-1} et $(F^{-1})^{-1} = F$ sont croissantes.
Donc F^{-1} est un isomorphisme d'ordres de D dans C .
CQFD.

Remarque :

On a aussi la transitivité de \cong , mais nous n'en parlons pas ici car il faudrait d'abord définir la composition d'assertions fonctionnelles. Ce n'est pas compliqué à faire, mais nous n'en avons pas besoin pour la suite et nous ne faisons ici que développer des outils pour les ordinaux, pas développer une théorie sur les assertions fonctionnelles en elles-mêmes. En soi \cong est donc en quelque sorte une relation d'équivalence, en tout cas en un sens généralité.

Encore une fois, le fait d'avoir des classes uniquement d'ordinaux nous offre d'office le fait que l'ordre soit total. En particulier on a la sympathique caractérisation suivante, qui sinon demanderait la totalité de l'ordre pour passer de 3 à 2 et de 2 à 1.

Proposition 78 (Caractérisation d'un isomorphisme)

Soient C et D deux classes d'ordinaux et $F : C \rightarrow D$ une assertion fonctionnelle. Les assertions suivantes sont équivalentes :

1. F est un isomorphisme d'ordres de C vers D .
2. F est croissante et bijective de C vers D .
3. F est strictement croissante et surjective dans D .

Démonstration

On va montrer $1 \iff 2 \iff 3$.

$1 \Rightarrow 2$

Supposons que F est un isomorphisme d'ordres de C vers D .

En particulier par définition $[F \text{ est croissante et bijective de } C \text{ vers } D]$.

$1 \Leftarrow 2$

Supposons que F est croissante et bijective de C vers D .

Il reste à montrer que F^{-1} est croissante.

Soient γ et δ dans $\text{im}(F)$ tels que $\gamma \leq \delta$.

Par définition il existe α et β dans $\text{dom}(F)$ tels que $\gamma = F(\alpha)$ et $\delta = F(\beta)$.

Par définition de F^{-1} on a donc $\alpha = F^{-1}(\gamma)$ et $\beta = F^{-1}(\delta)$.

Comme α et β sont des ordinaux, on a $\alpha \leq \beta$ ou $\beta < \alpha$.

Supposons par l'absurde que $\beta < \alpha$.

Par hypothèse F est croissante et bijective de C vers D .

En particulier F est croissante et injective par définition.

Donc F est strictement croissante d'après la proposition 76 page 182.

On a donc $F(\beta) < F(\alpha)$ par stricte croissance, c'est-à-dire $\delta < \gamma$.

C'est absurde puisque $\gamma \leq \delta$.

Par l'absurde on vient donc de montrer que $\alpha \leq \beta$, c'est-à-dire $F^{-1}(\gamma) \leq F^{-1}(\delta)$.

Ainsi F^{-1} est croissante, et finalement $[F \text{ est un isomorphisme d'ordres de } C \text{ vers } D]$.

$2 \Rightarrow 3$

Supposons que F est croissante et bijective de C vers D .

En particulier F est croissante, injective et surjective dans D par définition.

Donc F est strictement croissante d'après la proposition 76 page 182.

$2 \Leftarrow 3$

Supposons que F est strictement croissante et surjective dans D .

Alors F est croissante et injective d'après la proposition 76 page 182.

Ainsi F est croissante et bijective de C vers D par définition.

CQFD.

Si les propositions précédentes généralisent ce que l'on a vu dans le précédent livre, celle qui suit est une généralisation de la proposition 24 page 51. La preuve est d'ailleurs la même, avec quelques modifications de circonstances liées au fait qu'on a des classes d'ordinaux.

Proposition 79 (Unicité d'un isomorphisme de classes)

Soient C et D deux classes d'ordinaux.

Il y a au plus un seul isomorphisme d'ordres de C vers D .

Démonstration

Supposons qu'il existe un isomorphisme d'ordres $F : C \rightarrow D$.

Soit $G : C \rightarrow D$ un autre isomorphisme d'ordres.

Montrons que $F = G$.

Supposons par l'absurde que $F \neq G$.

Considérons alors la classe $A := \{\gamma \in C \mid F(\gamma) \neq G(\gamma)\}$.

Par hypothèse A est donc une classe non vide de C donc de ON .

Donc A admet un ordinal minimum α d'après la proposition 9 page 24.

Comme $\alpha = \min(A)$, pour tout $\beta \in C$ tel que $\beta < \alpha$ on a $\beta \notin A$ donc $F(\beta) = G(\beta)$.

De plus $\alpha \in A$ donc $F(\alpha) \neq G(\alpha)$.

Or D est une classe d'ordinaux, donc $F(\alpha) < G(\alpha)$ ou $G(\alpha) < F(\alpha)$.

► Plaçons-nous dans le cas où $F(\alpha) < G(\alpha)$.

Soit $\beta \in C$.

Comme C est une classe d'ordinaux, on a $\beta < \alpha$ ou $\alpha \leq \beta$.

Si $\beta < \alpha$ alors $G(\beta) = F(\beta) < F(\alpha)$ par stricte croissance de F .

Si $\alpha \leq \beta$ alors $F(\alpha) < G(\alpha) \leq G(\beta)$ par croissance de G .

Dans les deux cas on a $G(\beta) \neq F(\alpha)$.

Ainsi pour tout $\beta \in C$ on a $G(\beta) \neq F(\alpha)$.

Ainsi $F(\alpha)$ est un élément de D que G n'atteint pas.

C'est absurde puisque G est surjective dans D .

► Plaçons-nous dans le cas où $G(\alpha) < F(\alpha)$.

On montre de la même manière que dans ce cas-là $G(\alpha)$ est un élément de D que F n'atteint pas. C'est absurde puisque F est surjective dans D .

Dans les deux cas on aboutit une absurdité concernant la surjectivité d'une des deux assertions fonctionnelles.

Par l'absurde on vient de montrer que $F = G$, d'où l'unicité.

CQFD.

Un fait assez remarquable et nouveau à présent : quand on prend une classe d'ordinaux, si celle-ci est propre (c'est-à-dire n'est pas associée à un ensemble) alors elle est automatiquement isomorphe à ON tout entier ! C'est assez puissant, mais finalement pas si étonnant : l'isomorphisme consiste simplement à énumérer dans l'ordre tous les éléments de la classe, un pour chaque ordinal de ON .

Proposition 80 (Classes propres de ON et isomorphie avec ON)

Soit C une classe propre de ON .

Alors C et ON sont isomorphes.

Démonstration

Construction de l'assertion fonctionnelle

Construisons un isomorphisme d'ordres $F : ON \longrightarrow C$.

L'idée va être pour F d'énumérer dans l'ordre tous les éléments de C .

Autrement dit, notre objectif est de faire en sorte que $F(\alpha)$ soit le premier des $\gamma \in C$ qui n'a pas déjà été atteint par $F(\beta)$ pour $\beta < \alpha$.

D'après la proposition 9 page 24, toute classe non vide de ON admet un minimum.

Pour tout ordinal α , on va poser $F(\alpha) := \min \left(\{ \gamma \in C \mid \forall \beta < \alpha, F(\beta) < \gamma \} \right)$.

Justification de la définition de F

Pour cela, nous allons utiliser le théorème 6 page 71.

Définissons l'assertion fonctionnelle H pour toute application f tel que $\text{dom}(f)$ est un ordinal et à valeurs dans les ordinaux par

$$H(f) := \min \left(\{ \gamma \in C \mid \forall \beta < \text{dom}(f), f(\beta) < \gamma \} \right)$$

Montrons que H vérifie les conditions du théorème.

Soit α un ordinal et $f : \alpha \longrightarrow ?$ tels que f est H -inductive.

On sait déjà que le domaine de f est un ordinal.

Il suffit donc de montrer que f ne prend que des valeurs ordinales pour conclure que

f est dans le domaine de H .

Soit $\beta < \alpha$.

Par définition f est H -inductive donc $f(\beta) = H(f|_\beta)$.

Or par définition $H(f|_\beta)$ est un minimum d'une classe d'ordinaux.

Donc $H(f|_\beta)$ est un ordinal et donc $f(\beta)$ est un ordinal.

Donc pour tout $\beta < \alpha$, $f(\beta)$ est un ordinal, et donc f est à valeurs dans les ordinaux.

Finalement, f est dans le domaine de H .

Ainsi H vérifie les conditions du théorème.

D'après celui-ci, il existe une unique assertion fonctionnelle F de domaine ON qui est H -inductive. Autrement dit pour tout ordinal α , on a

$$\begin{aligned} F(\alpha) &= H(F|_\alpha) = \min \left(\{ \gamma \in C \mid \forall \beta < \alpha, F|_\alpha(\beta) < \gamma \} \right) \\ &= \min \left(\{ \gamma \in C \mid \forall \beta < \alpha, F(\beta) < \gamma \} \right) \end{aligned}$$

F est un isomorphisme d'ordres

F est définie sur tout ON et à valeurs dans C par définition.

Montrons que F est strictement croissante.

Soient α et α' deux ordinaux tels que $\alpha < \alpha'$.

Considérons $B := \{ \gamma \in C \mid \forall \beta < \alpha', F(\beta) < \gamma \}$.

Par définition $F(\alpha') = \min(B)$ donc en particulier $F(\alpha') \in B$.

Donc $\forall \beta < \alpha', F(\beta) < F(\alpha')$.

En particulier en prenant $\beta := \alpha$, on trouve $F(\alpha) < F(\alpha')$.

Donc F est strictement croissante.

Montrons que F est surjective dans C .

Supposons par l'absurde que F n'est pas surjective dans C .

Il existe alors au moins un $\gamma \in C$ tel que pour tout ordinal, $F(\alpha) \neq \gamma$.

Considérons $B := \{ \gamma \in C \mid \forall \alpha \in ON, F(\alpha) \neq \gamma \}$.

Alors B est une sous-classe non vide de C donc une sous-classe non vide de ON .

B admet donc un ordinal minimum γ_0 d'après la proposition 9 page 24.

Considérons $A := \{ \alpha \in ON \mid F(\alpha) < \gamma_0 \}$.

Par définition $F : ON \longrightarrow C$ donc comme $C \subseteq ON$ on a $F : ON \longrightarrow ON$.

Or on a montré que F est strictement croissante.

Donc F n'est pas bornée : il existe $\varepsilon \in ON$ tel que $\gamma_0 < F(\varepsilon)$.

Soit $\alpha \in A$.

On a alors $F(\alpha) < \gamma_0 < F(\varepsilon)$.

Si $\varepsilon < \alpha$ alors $F(\varepsilon) < F(\alpha)$ par strictement croissance de F .

Par contraposition on a donc $\alpha \leq \varepsilon$.

Donc A est bornée par ε , donc A est un ensemble d'après la proposition 12 page 32.

En particulier A admet une borne supérieure d'après la proposition 11 page 29.

Considérons alors $\alpha_0 := \sup(A)$ et montrons que $F(\alpha_0) = \gamma_0$.

Posons $D := \{\gamma \in C \mid \forall \beta < \alpha_0, F(\beta) < \gamma\}$.

Par définition de F , on a $F(\alpha_0) = \min(D)$.

Montrons que $\gamma_0 \in D$.

En effet, soit $\beta < \alpha_0$.

Comme $\alpha_0 = \sup(A)$, il existe $\alpha \in A$ tel que $\beta \leq \alpha$.

Comme $\alpha \in A$, on a $F(\alpha) < \gamma_0$ par définition de A .

Par croissance de F on a alors $F(\beta) \leq F(\alpha) < \gamma_0$ et donc $F(\beta) < \gamma_0$.

Ainsi $\forall \beta < \alpha_0, F(\beta) < \gamma_0$ donc $\gamma_0 \in D$.

De plus, montrons que γ_0 minore D .

En effet, soit $\gamma \in D$.

Comme ce sont des ordinaux, on a $\gamma_0 \leq \gamma$ ou $\gamma < \gamma_0$.

Supposons par l'absurde que $\gamma < \gamma_0$.

Par définition $\gamma_0 = \min(D)$ donc $\gamma \notin D$.

Par définition de D il existe donc $\alpha \in ON$ tel que $F(\alpha) = \gamma$.

En particulier on a $F(\alpha) < \gamma_0$ donc $\alpha \in A$ par définition de A .

On a donc $\alpha \leq \alpha_0$ car $\alpha_0 = \sup(A)$ par définition.

Ainsi $\alpha \leq \alpha_0$ vérifie $F(\alpha) = \gamma$.

C'est absurde par définition de D puisque $\gamma \in D$ par définition.

Par l'absurde on vient de montrer que $\gamma_0 \leq \gamma$.

Ainsi γ_0 minore D , et comme $\gamma_0 \in D$, on a $\gamma_0 = \min(D)$.

Or on a dit que $F(\alpha_0) = \min(D)$ donc $F(\alpha_0) = \gamma_0$.

C'est absurde car par définition γ_0 n'est pas atteint car $\gamma_0 \in B$.

Par l'absurde on vient de montrer que F est surjective dans C .

Ainsi F est strictement croissante et surjective dans C .

Donc F est un isomorphisme d'ordres de ON dans C .

CQFD.

Fixation d'une assertion fonctionnelle

Étant donné un ensemble E et une application $f : E \rightarrow E$, nous avons vu à la fin du chapitre 1 via la définition 17 page 87 la notion d'itérées de f : intuitivement on prend $x \in E$ et on considère la suite

$$x, f(x), f(f(x)), f(f(f(x))), \dots$$

en itérant f autant de fois que souhaité. Ici on va faire de même pour les assertions fonctionnelles $F : ON \rightarrow ON$. L'avantage d'être chez les ordinaux est de disposer systématiquement d'une borne supérieure et donc il est possible d'itérer un nombre ordinal de fois !

Définition 30 (Itérées d'une assertion fonctionnelle)

Soit $F : ON \rightarrow ON$ une assertion fonctionnelle.

Pour tout ordinal α , on pose

$$\begin{cases} F^0(\alpha) := \alpha \\ F^{\beta+1}(\alpha) := F(F^\beta(\alpha)) \text{ pour tout ordinal } \beta \\ F^\gamma(\alpha) := \sup_{\delta < \gamma} F^\delta(\alpha) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Pour tout ordinal β , on obtient ainsi une assertion fonctionnelle $F^\beta : ON \rightarrow ON$.

Remarque :

Pour justifier cette construction, on peut encore et toujours se reporter à la proposition 36 page 91, en prenant cette fois $\mu_0 := \alpha$ et $G(\xi) := F(\xi)$ pour tout ordinal ξ (on rappelle qu'ici α et F sont fixés à l'avance). On obtient alors une assertion fonctionnelle \mathfrak{F}_α , et pour tout ordinal β on pose $F^\beta(\alpha) := \mathfrak{F}_\alpha(\beta)$.

Imaginons qu'un ordinal γ soit un point fixe de F : on a alors

$$\gamma = F(\gamma) = F(F(\gamma)) = F(F(F(\gamma))) = \dots$$

si bien que γ est un point fixe de toutes les itérées de F . Ainsi les notions de points fixes et d'itérées sont très liées. Un habitué de topologie pourrait même faire la remarque que le théorème du point fixe de Banach-Picard construit un point fixe précisément en itérant une application, et en prenant la limite du procédé. En quelque sorte, itérer un nombre infini de fois une application (ou plus généralement une assertion fonctionnelle) va produire un point fixe ! Fascinant !

Proposition 81 (Point fixe à l'infini)

Soit $F : ON \rightarrow ON$ une assertion fonctionnelle **strictement croissante et continue**.

Pour tout ordinal α , on a :

1. $F^\omega(\alpha)$ est un point fixe de F vérifiant $\alpha \leq F^\omega(\alpha)$.
2. Pour tout γ point fixe de F tel que $\alpha \leq \gamma$, on a $F^\omega(\alpha) \leq \gamma$.

Ainsi $F^\omega(\alpha)$ est le plus petit de ces points fixes de F .

Démonstration

1.

- Commençons par montrer une première égalité.

Pour tout $n < \omega$, on a $n + 1 < \omega$ d'après 15 page 38.

Donc pour tout $n < \omega$, on a $F^{n+1}(\alpha) \leq \sup_{m < \omega} F^m(\alpha)$.

Donc $\sup_{n < \omega} F^{n+1}(\alpha) \leq \sup_{m < \omega} F^m(\alpha)$.

Supposons par l'absurde que $\sup_{n < \omega} F^{n+1}(\alpha) \neq \sup_{m < \omega} F^m(\alpha)$.

On a donc $\sup_{n < \omega} F^{n+1}(\alpha) < \sup_{m < \omega} F^m(\alpha)$ par ce qui précède.

Il existe donc $m < \omega$ tel que $\sup_{n < \omega} F^{n+1}(\alpha) < F^m(\alpha)$.

En particulier $F(F^m(\alpha)) = F^{m+1}(\alpha) \leq \sup_{n < \omega} F^{n+1}(\alpha) < F^m(\alpha)$.

Or F est strictement croissante.

On a donc $F^m(\alpha) \leq F(F^m(\alpha))$ d'après la proposition 73 page 174.

C'est absurde puisqu'on vient justement de dire que $F(F^m(\alpha)) < F^m(\alpha)$.

Par l'absurde on vient de montrer que $\sup_{n < \omega} F^{n+1}(\alpha) = \sup_{m < \omega} F^m(\alpha)$.

Or $F^\omega(\alpha) = \sup_{m < \omega} F^m(\alpha)$ par définition de F^ω , donc $\boxed{\sup_{n < \omega} F^{n+1}(\alpha) = F^\omega(\alpha)}$.

Notons (\star) cette égalité.

- On peut en conclure que $F^\omega(\alpha)$ est un point fixe de F .

On a les égalités suivantes :

$$\begin{aligned} F(F^\omega(\alpha)) &= F\left(\sup_{n < \omega} F^n(\alpha)\right) \text{ par définition de } F^\omega \\ &= \sup_{n < \omega} F(F^n(\alpha)) \text{ par continuité de } F \\ &= \sup_{n < \omega} F^{n+1}(\alpha) \text{ par définition de } F^{n+1} \\ &= F^\omega(\alpha) \text{ d'après } (\star) \end{aligned}$$

Ainsi on a $F(F^\omega(\alpha)) = F^\omega(\alpha)$ donc $\boxed{F^\omega(\alpha) \text{ est un point fixe de } F}$.

- De plus $\alpha = F^0(\alpha) \leq \sup_{n < \omega} F^n(\alpha) = F^\omega(\alpha)$, donc $\boxed{\alpha \leq F^\omega(\alpha)}$.

2. Montrons que $F^\omega(\alpha)$ est le plus petit de tels points fixes.

Soit γ un point fixe de F tel que $\alpha \leq \gamma$.

Pour tout entier naturel n , posons $P(n)$ l'assertion « $F^n(\alpha) \leq \gamma$ ».

Initialisation

Par définition de F^0 , on a $F^0(\alpha) = \alpha \leq \gamma$ donc $P(0)$.

Héritéité

Soit n un entier naturel tel que $P(n)$, c'est-à-dire $F^n(\alpha) \leq \gamma$.

F est strictement croissante donc F est croissante d'après la prop. 76 p. 182.

On a donc $F(F^n(\alpha)) \leq F(\gamma)$ par croissance.

Or $F^{n+1}(\alpha) = F(F^n(\alpha))$ par définition de F^{n+1} .

De plus $F(\gamma) = \gamma$ car γ est un point fixe de F .

On a donc $F^{n+1}(\alpha) \leq \gamma$, c'est-à-dire $P(n+1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n+1)$.

Ainsi P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$, c'est-à-dire $F^n(\alpha) \leq \gamma$.

En particulier on a $F^\omega(\alpha) = \sup_{n < \omega} F^n(\alpha) \leq \gamma$.

Ainsi $F^\omega(\alpha)$ est plus petit que tout point fixe γ de F tel que $\alpha \leq \gamma$.

CQFD.

Ainsi si l'on prend une assertion fonctionnelle strictement croissante et continue, alors elle admet des points fixes. Mieux, ceux-ci forment une classe propre, comme le montre le théorème suivant. En particulier en les énumérant on obtient un isomorphisme de ON vers cette classe d'après la proposition 80 page 187. De plus, c'est le seul isomorphisme d'après la proposition 79 page 186. On peut donc lui donner un petit nom, et c'est lui qui va nous permettre de définir les ε_α .

Théorème 10 (Fixation d'une assertion fonctionnelle)

Soit $F : ON \longrightarrow ON$ une assertion fonctionnelle.

Notons $\text{fix}(F)$ la classe des points fixes de F .

Si F est strictement croissante et continue alors $\text{fix}(F)$ est une classe propre de ON .

On appelle **fixation** de F l'unique isomorphisme d'ordres $ON \longrightarrow \text{fix}(F)$.

On le note F° .

*Démonstration*

Supposons que F est strictement croissante et continue.

Pour prouver que $\text{fix}(F)$ est une classe propre, nous allons montré qu'elle n'est pas bornée : comme tout ensemble d'ordinaux doit admettre une borne supérieure (donc être majorée), $\text{fix}(F)$ n'est pas un ensemble donc une classe propre. Autrement dit nous allons montrer

que pour tout ordinal α , il existe $\beta \in \text{fix}(F)$ tel que $\alpha < \beta$.

Soit α un ordinal.

Considérons $\beta := F^\omega(\alpha + 1)$.

D'après la proposition 81 page 190, on a $\beta \in \text{fix}(F)$, et $\alpha + 1 \leq \beta$.

Or on a $\alpha < \alpha + 1$ d'après la proposition 13 page 33.

On a donc $\alpha < \beta$ par transitivité.

Donc $\text{fix}(F)$ n'est pas bornée et est donc une classe propre.

CQFD.

Remarque :

Dans la littérature, le nom que l'on retrouve le plus est celui de **dérivée** et non fixation, que l'on note alors F' plutôt que F° . L'auteur a préféré éviter de conserver cela pour ne pas donner de faux espoirs aux lecteurs férus d'analyse !

Exemple :

Nous avons donc à notre disposition une autre façon d'envisager ε_0 , puis ε_1 et ainsi de suite ! En effet, comme annoncé dans l'introduction, il suffit de considérer l'assertion fonctionnelle $F := \begin{pmatrix} ON & \longrightarrow & ON \\ \gamma & \longmapsto & \omega^\gamma \end{pmatrix}$. Comme on l'avait vu, ε_0 est le premier point fixe de F , ε_1 le deuxième, et ainsi de suite.

D'après la proposition 66 page 151, F est strictement croissante. D'après la proposition 68 page 155, F est continue. Autrement dit on peut considérer sa fixation F° qui énumère tous les points fixes de F , et donc pour tout ordinal α on peut simplement poser $\varepsilon_\alpha := F^\circ(\alpha)$.

On peut donc grâce au concept de fixation commencer à aborder des ordinaux démesurés. Cela ne fait pourtant que commencer. Pour cela, remarquons avec la proposition qui suit que la fixation admet elle-même une fixation !

Proposition 82 (Fixation de la fixation)

Soit $F : ON \longrightarrow ON$ une assertion fonctionnelle **strictement croissante et continue**. Alors sa fixation F° est elle-même strictement croissante et continue.

Démonstration

- Supposons que F est croissante et continue.

Elle admet donc une fixation F° d'après le théorème 10 page 192.

Par définition F° est l'isomorphisme d'ordres $ON \longrightarrow \text{fix}(F)$.

En particulier F° est strictement croissante d'après la proposition 78 page 185.

- Montrons que $\text{fix}(F)$ est stable par passage à la borne supérieure.

Soit X un ensemble d'éléments de $\text{fix}(F)$.

Autrement dit pour tout $\xi \in X$, on a $F(\xi) = \xi$.

Considérons $\sigma := \sup(X)$.

Par continuité de F , on a $F(\sigma) = F\left(\sup_{\xi \in X} \xi\right) = \sup_{\xi \in X} F(\xi) = \sup_{\xi \in X} \xi = \sigma$.

Donc σ est un point fixe de F et donc $\sigma \in \text{fix}(F)$.

Ainsi pour tout ensemble $X \subseteq \text{fix}(F)$, on a $\sup(X) \in \text{fix}(F)$.

- Montrons que F° est continue.

Supposons par l'absurde que F° n'est pas continue.

Utilisons la proposition 42 page 103.

Il existe donc γ un ordinal limite non nul tel que $F^\circ(\gamma) \neq \sup_{\delta < \gamma} F^\circ(\delta)$.

Or par stricte croissance de F° , pour tout $\delta < \gamma$ on a $F^\circ(\delta) < F^\circ(\gamma)$.

On a donc $\sup_{\delta < \gamma} F^\circ(\delta) \leq F^\circ(\gamma)$ et donc $\sup_{\delta < \gamma} F^\circ(\delta) < F^\circ(\gamma)$ par ce qui précède.

Posons alors $\alpha := \sup_{\delta < \gamma} F^\circ(\delta)$: on vient donc de voir que $\alpha < F^\circ(\gamma)$.

F° est à valeurs dans $\text{fix}(F)$ donc pour tout $\delta < \gamma$, $F^\circ(\delta) \in \text{fix}(F)$.

Donc $\alpha \in \text{fix}(F)$ par la stabilité de la borne supérieure que l'on a montrée plus tôt.

Par définition F° est surjective dans $\text{fix}(F)$.

Il existe donc β un ordinal tel que $F^\circ(\beta) = \alpha$.

Or on a dit que $\alpha < F^\circ(\gamma)$, si bien que $F^\circ(\beta) < F^\circ(\gamma)$.

On a donc $\beta < \gamma$ car F° est un isomorphisme d'ordres.

Mais γ est limite non nul donc on a $\beta + 1 < \gamma$ d'après la proposition 14 page 37.

On a donc $\alpha = F^\circ(\beta) < F^\circ(\beta + 1) \leq \sup_{\delta < \gamma} F^\circ(\delta) = \alpha$.

Ainsi $\alpha < \alpha$, ce qui est absurde.

Par l'absurde, on vient de montrer que F° est continue.

CQFD.

Exemple :

Reprendons $F := \begin{pmatrix} ON & \longrightarrow & ON \\ \gamma & \longmapsto & \omega^\gamma \end{pmatrix}$ dont nous avons parlé juste au-dessus.

On l'a dit, on a alors $F^\circ = \begin{pmatrix} ON & \longrightarrow & ON \\ \gamma & \longmapsto & \varepsilon_\gamma \end{pmatrix}$. On vient de le voir, F° admet elle-même une fixation : on peut donc considérer $F^{\circ\circ}$. On pose alors $\zeta_0 := F^{\circ\circ}(0)$. Mais que représente ζ_0 au juste ? C'est le premier point fixe de $\gamma \longmapsto \varepsilon_\gamma$. Autrement dit on a $\zeta_0 = \varepsilon_{\zeta_0}$.

On se rend donc compte que l'on a $\zeta_0 = \varepsilon_{\varepsilon_{\varepsilon_{\varepsilon_{\varepsilon_{\dots}}}}}$!

Avant de conclure ce chapitre, qui avouons-le était un peu fastidieux, essayons d'obtenir une expression un peu plus explicite de F° en fonction de F . Celle-ci passe par les itérées de F , comme on l'a pressenti avec la proposition 81 page 190, et se fait par récursion. Prenons bien garde à ne pas confondre F^β pour un ordinal β , qui signifie l'itération β fois de F , et F° qui signifie sa fixation.

Proposition 83 (Expression explicite de la fixation)

Soit $F : ON \rightarrow ON$ une assertion fonctionnelle **strictement croissante et continue**.
On a alors :

$$\begin{cases} F^\circ(0) = F^\omega(0) \\ F^\circ(\alpha + 1) = F^\omega(F^\circ(\alpha) + 1) \text{ pour tout ordinal } \alpha \\ F^\circ(\gamma) = \sup_{\delta < \gamma} F^\circ(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Démonstration

D'après la proposition 36 page 91, en prenant $\mu_0 := F^\omega(0)$ et $G(\xi) := F^\omega(\xi + 1)$ pour tout ordinal ξ , il existe une unique assertion fonctionnelle $\mathfrak{F} : ON \rightarrow ON$ telle que

$$\begin{cases} \mathfrak{F}(0) = F^\omega(0) \\ \mathfrak{F}(\alpha + 1) = F^\omega(\mathfrak{F}(\alpha) + 1) \text{ pour tout ordinal } \alpha \\ \mathfrak{F}(\gamma) = \sup_{\delta < \gamma} \mathfrak{F}(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Montrons que $F^\circ = \mathfrak{F}$ par induction transfinie.

Pour tout ordinal α , on pose $P(\alpha)$ l'assertion « $F^\circ(\alpha) = \mathfrak{F}(\alpha)$ ».

Initialisation

Par définition, F° est à valeurs dans les points fixes de F .

En particulier $F^\circ(0)$ est un point fixe de F .

De plus comme 0 est le plus petit des ordinaux, on a nécessairement $0 \leq F^\circ(0)$.

Ainsi $F^\circ(0)$ est un point fixe γ de F tel que $0 \leq \gamma$.

On a donc $F^\omega(0) \leq F^\circ(0)$ par minimalité de $F^\omega(0)$, d'après la proposition 81 page 190.

Mais $F^\omega(0)$ est un point fixe de F d'après cette même proposition, donc $F^\omega(0) \in \text{fix}(F)$.

Or par définition de F° , on a $F^\circ(0) = \min(\text{fix}(F))$ et donc $F^\circ(0) \leq F^\omega(0)$.

On en conclut que $F^\omega(0) = F^\circ(0)$ par antisymétrie de \leq .

Autrement dit on a $F^\circ(0) = \mathfrak{F}(0)$, c'est-à-dire $P(0)$.

Héritéité

Soit α un ordinal tel que $P(\alpha)$, c'est-à-dire $F^\circ(\alpha) = \mathfrak{F}(\alpha)$.

En particulier on a $\mathfrak{F}(\alpha + 1) = F^\omega(F^\circ(\alpha) + 1)$ par définition de \mathfrak{F} .

On a $\alpha < \alpha + 1$ d'après la proposition 13 page 33.

Donc $F^\circ(\alpha) < F^\circ(\alpha + 1)$ par stricte croissance de F° .

On a donc $F^\circ(\alpha) + 1 \leq F^\circ(\alpha + 1)$ d'après la proposition 13 page 33.

Or par définition $\text{im}(F^\circ) \subseteq \text{fix}(F)$ donc $F^\circ(\alpha + 1) \in \text{fix}(F)$.

Ainsi $F^\circ(\alpha + 1)$ est un point fixe γ de F tel que $F^\circ(\alpha) + 1 \leq \gamma$.

On a donc $F^\omega(F^\circ(\alpha) + 1) \leq F^\circ(\alpha + 1)$ d'après la proposition 81 page 190.

Autrement dit on a $\mathfrak{F}(\alpha + 1) \leq F^\circ(\alpha + 1)$.

On applique à nouveau la proposition 81 page 190.

D'après celle-ci $F^\omega(F^\circ(\alpha) + 1)$ est un point fixe γ de F tel que $F^\circ(\alpha) + 1 \leq \gamma$.

Ainsi $F^\omega(F^\circ(\alpha) + 1) \in \text{fix}(F)$ et $F^\circ(\alpha) + 1 \leq F^\omega(F^\circ(\alpha) + 1)$.

Il existe donc $\beta \in ON$ tel que $F^\omega(F^\circ(\alpha) + 1) = F^\circ(\beta)$ par surjectivité de F° .

De plus $F^\circ(\alpha) < F^\omega(F^\circ(\alpha) + 1)$ d'après la proposition 13 page 33.

Ainsi $F^\circ(\alpha) < F^\circ(\beta)$ donc $\alpha < \beta$ car F° est un isomorphisme d'ordres.

On a donc $\alpha + 1 \leq \beta$ d'après la proposition 13 page 33.

On a donc $F^\circ(\alpha + 1) \leq F^\circ(\beta)$ par croissance de F .

Or on a dit que $F^\circ(\beta) = F^\omega(F^\circ(\alpha) + 1) = \mathfrak{F}(\alpha + 1)$.

On a donc $F^\circ(\alpha + 1) \leq \mathfrak{F}(\alpha + 1)$.

On en conclut que $F^\circ(\alpha + 1) = \mathfrak{F}(\alpha + 1)$ par antisymétrie de \leq , c'est-à-dire $P(\alpha + 1)$.

Ainsi pour tout ordinal α , si l'on a $P(\alpha)$ alors on a $P(\alpha + 1)$.

Héritéité limite

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, P(\beta)$.

On a alors

$$\begin{aligned}\mathfrak{F}(\alpha) &= \sup_{\beta < \alpha} \mathfrak{F}(\beta) \text{ par définition de } \mathfrak{F} \\ &= \sup_{\beta < \alpha} F^\circ(\beta) \text{ car } \forall \beta < \alpha, P(\beta) \\ &= F^\circ\left(\sup_{\beta < \alpha} \beta\right) \text{ par continuité de } F^\circ \\ &= F^\circ\left(\sup_{\beta \in \alpha} \beta\right) \text{ par définition de } <\end{aligned}$$

$$= F^\circ(\sup(\alpha)) = F^\circ(\alpha) \text{ d'après la prop. 21 p. 47}$$

On a donc $\mathfrak{F}(\alpha) = F^\circ(\alpha)$, c'est-à-dire $P(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, P(\beta)$ alors $P(\alpha)$.

Finalement P vérifie les trois conditions du principe faible d'induction transfinie.

Donc pour tout ordinal α , on a $P(\alpha)$, c'est-à-dire $\mathfrak{F}(\alpha) = F^\circ(\alpha)$.

CQFD.

Ainsi on vient de se donner la possibilité de désigner des ordinaux démesurément grands, on fait face à l'immensité du monde des ordinaux. On s'en doute au vu du chemin parcouru, il est toujours possible d'aller plus loin, mais ce dont on ne se rend pas forcément bien compte à ce stade, c'est que dans le monde merveilleux des ordinaux, on n'a fait qu'effleurer la surface.

Le prochain chapitre va nous montrer des ordinaux immensément plus grands que tous ceux que l'on a pu aborder jusqu'à présent, de sorte que même ζ_0 n'est rien comparaison de ce qui nous attend. Il est temps d'aborder enfin les **cardinaux**, notion qui sera sans doute une des plus utiles de tout ce livre pour la suite de nos aventures !

Chapitre 3

Cardinaux



Note de l'auteur

Terminons ce livre avec sans doute l'application la plus utile de la théorie des ordinaux : la notion de cardinal. Intuitivement, il s'agit de compter le nombre d'éléments d'un ensemble, et ce nombre est alors appelé **cardinal de l'ensemble**. Nous l'avons vu, les nombres entiers naturels sont des cas particuliers d'ordinaux : il n'est donc pas si étonnant que les ordinaux permettent aussi de définir la notion de cardinal d'un ensemble.

Ce chapitre va commencer par revenir plus en détails sur la notion d'injection et de bijections, qui sont au cœur comme nous le verrons de l'idée de *compter le nombre d'éléments d'un ensemble*. Cela permettra alors de distinguer certains ordinaux jouant un rôle particulier : on les appellera **nombres cardinaux**. Étant donné un ensemble E , si jamais il existe un bon ordre sur E , alors nous verrons comme associer à E un unique cardinal, appelé naturellement cardinal de E .

Grâce à l'**axiome du choix** et deux grands théorèmes équivalents, nous verrons comment faire pour que tout ensemble admette un cardinal. Nous aurons alors l'occasion d'effectuer des opérations sur les cardinaux. On finira enfin par définir la notion d'ensembles finis et infinis, et d'ensembles dénombrables et indénombrables.

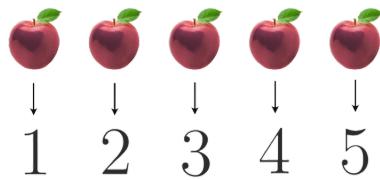
Sommaire

1	Équipotence et subpotence	200
1.1	Équipotence et subpotence	200
1.2	Théorème de Cantor	209
1.3	Équipotence et opérations	215
2	Nombres cardinaux	230
2.1	Les cardinaux	230
2.2	Le cardinal d'un ensemble	236
3	Les grands théorèmes	242
3.1	Choix, Zorn et Zermelo	242
3.2	Théorème et cardinal de Hartogs	256
4	Opérations sur les cardinaux	265
5	Ensembles finis et ensembles dénombrables	284
5.1	Ensembles finis	284
5.2	Ensembles dénombrables	302

1 Équipotence et subpotence

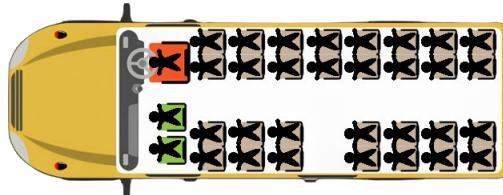
1.1 Équipotence et subpotence

Qu'est-ce que compter le nombre d'éléments d'un ensemble ? Quand nous avons appris à compter, nous avons commencé par apprendre une *comptine*, c'est-à-dire la liste des noms des premiers entiers naturels. Pour alors compter le nombre de pommes devant nous, il suffit de pointer tour à tour chacune des pommes en récitant à chaque fois un élément de plus de la comptine. Si par exemple on a dit « *un deux trois quatre cinq* » en ayant pointé une et une seule fois chaque pomme, on sait qu'il y a 5 pommes devant nous. Ainsi, on a mis en **bijection** l'ensemble des pommes avec l'ensemble $\{1, 2, 3, 4, 5\}$.



On a réalisé une bijection : il y a donc 5 pommes.

Prenons un autre exemple : une façon de savoir qu'il y autant de passagers d'un bus que de sièges dans celui-ci consiste à faire s'asseoir chaque passager sur un siège et de constater qu'aucun passager n'est resté debout ni qu'aucun siège n'est vide. Ainsi, on sait qu'il y a le même nombre de passagers que de sièges sans pour autant n'avoir eu besoin de compter ni les passagers ni les sièges.



*Il y a autant de personnes dans le bus que de places assises.
Pas besoin de compter pour le voir !*

Encore une fois, on a ici réalisé une bijection entre l'ensemble des passagers et l'ensemble des sièges. On comprend donc que la notion de quantité est fortement liée à la notion de bijection, d'où la définition suivante.

Définition 31 (Équipotence)

Soient E et F deux ensembles.

On dit que E et F sont **équipotents** si et seulement s'il existe une bijection de E vers F .

On note alors $E \approx F$.

Ainsi au vu de ce que l'on a dit plus haut, dire que deux ensembles sont équipotents, c'est dire qu'ils ont le même nombre d'éléments. C'est d'une certaine manière une relation d'équivalence,

en tout cas en un sens généralisé aux classes et non au ensembles, puisqu'on a l'a dit l'ensemble de tous les ensembles n'existe pas.

Proposition 84 (Propriétés de l'équipotence)

Soient E , F et G trois ensembles.

1. On a $E \approx E$. On dit que \approx est **réflexive**.
2. Si $E \approx F$ alors $F \approx E$. On dit que \approx est **symétrique**
3. Si $E \approx F$ et $F \approx G$ alors $E \approx G$. On dit que \approx est **transitive**.



Démonstration

1. On sait que $\text{id}_E : E \longrightarrow E$ est une bijection.

On a donc $[E \approx E]$.

2. Supposons que $E \approx F$.

Il existe alors $f : E \longrightarrow F$ une bijection.

Alors $f^{-1} : F \longrightarrow E$ est une bijection.

Donc $[F \approx E]$.

3. Supposons que $E \approx F$ et $F \approx G$.

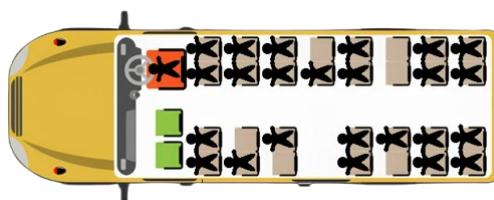
Il existe alors $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux bijections.

Alors $g \circ f : E \longrightarrow G$ est une bijection.

Donc $[E \approx G]$.

CQFD

Dans l'exemple où l'on demande aux passagers d'un bus de s'asseoir sur les sièges, que conclure dans le cas où il reste des places vides ? Cela veut dire qu'il y a moins de passagers que de sièges, et donc que l'ensemble des passagers est plus petit que l'ensemble des sièges. Cette fois-ci, on a donc affaire à la notion d'**injection**.



*Il y a moins de personnes dans le bus que de places assises.
Pas besoin de compter pour le voir !*

Définition 32 (Subpotence)

Soient E et F deux ensembles.

On dit que E est **subpotent** à F si et seulement s'il existe une injection $E \rightarrow F$.

On note alors $E \preccurlyeq F$.

Remarque :

On a vu dans le précédent livre que c'est équivalent à l'existence d'une **surjection** $F \rightarrow E$. Cela nécessite cependant l'**axiome du choix**. Cela revient donc à dire que F a plus d'éléments que E .

De la même manière que l'équipotence est une relation d'équivalence généralisée aux classes, la subpotence peut être vu comme une relation d'ordre généralisée au classes. En réalité pas tout à fait, car on n'a malheureusement pas l'antisymétrie : en effet, ce n'est pas parce que deux ensembles ont le même nombre d'éléments qu'ils sont égaux.

Proposition 85 (Propriétés de la subpotence)

Soient E , F et G trois ensembles.

1. On a $E \preccurlyeq E$: on dit que \preccurlyeq est **réflexive**.
2. Si $E \preccurlyeq F$ et $F \preccurlyeq G$ alors $E \preccurlyeq G$: on dit que \preccurlyeq est **transitive**.

Démonstration

1. L'application $\text{id}_E : E \rightarrow E$ est injective donc $[E \preccurlyeq E]$.

2. Supposons que $E \preccurlyeq F$ et $F \preccurlyeq G$.

Il existe donc deux injections $f : E \rightarrow F$ et $g : F \rightarrow G$.

Alors $g \circ f : E \rightarrow G$ est une injection et donc $[E \preccurlyeq G]$.

CQFD.

La propriété qui suit est en quelque sorte une généralisation de la réflexivité. Si deux ensembles sont équipotents, alors ils sont subpotents l'un par rapport à l'autre.

Proposition 86 (Équipotence implique double subpotence)

Soient E et F deux ensembles.

Si $E \approx F$ alors ($E \preccurlyeq F$ et $F \preccurlyeq E$).

Démonstration

Supposons que $E \approx F$.

Il existe donc $f : E \rightarrow F$ une bijection.

En particulier f est injection et donc $[E \preccurlyeq F]$.

De plus $f^{-1} : F \rightarrow E$ est aussi une bijection donc une injection, et donc $[F \preccurlyeq E]$.

CQFD.

On l'a dit, nous n'avons pas d'antisymétrie de la subpotence.

Autrement dit, on n'a pas l'implication $(E \preccurlyeq F \preccurlyeq E) \Rightarrow E = F$.

Ok très bien, mais peut-on au moins dire que l'on a $(E \preccurlyeq F \preccurlyeq E) \Rightarrow E \approx F$?

La réponse est oui, mais pour pouvoir le montrer, nous allons tout d'abord nous intéresser à une version plus faible, c'est-à-dire $(E \preccurlyeq F \subseteq E) \Rightarrow E \approx F$.

C'est l'objet de la proposition qui suit.

Proposition 87 (Subpotence et inclusion)

Soient E et F deux ensembles.

1. Si $F \subseteq E$ alors $F \preccurlyeq E$.
2. Si $F \subseteq E \preccurlyeq F$ alors $E \approx F$.

Démonstration

1. Supposons que $F \subseteq E$.

On a $\text{id}_F : F \rightarrow F$ donc $\text{id}_E : F \rightarrow E$.

Or id_F est injective donc $[F \preccurlyeq E]$.

2.

- Supposons que $F \subseteq E \preccurlyeq F$.

Comme $E \preccurlyeq F$, il existe une injection $u : E \rightarrow F$.

On construit la suite $(G_n)_{n \in \mathbb{N}}$ par récurrence de la manière suivante :

$$\begin{cases} G_0 := E \setminus F \\ G_{n+1} := u^{-1}(G_n) \text{ pour tout } n \in \mathbb{N} \end{cases}$$

On pose alors $G := \bigcup_{n \in \mathbb{N}} G_n$.

Remarquons que pour tout $x \in E$, comme u est à valeurs dans F on a $u(x) \in F$.

De plus pour tout $x \in E$, si $x \notin G$ alors en particulier $x \notin G_0 = E \setminus F$ donc $x \in F$.

On peut donc définir l'application

$$v := \begin{pmatrix} E & \longrightarrow & F \\ x & \longmapsto & \begin{cases} u(x) & \text{si } x \in G \\ x & \text{si } x \notin G \end{cases} \end{pmatrix}$$

- Montrons que v est injective.

Soient x et x' tels que $v(x) = v(x')$.

► Plaçons-nous dans le cas où $x \in G$ et $x' \in G$.

On a donc $u(x) = v(x) = v(x') = u(x')$ donc $u(x) = u(x')$.

Or u est injective par définition donc $x = x'$.

► Plaçons-nous dans le cas où $x \notin G$ et $x' \notin G$.

On a donc $x = v(x) = v(x') = x'$ et donc $x = x'$.

► Plaçons-nous dans le cas où $x \in G$ et $x' \notin G$.

On a donc $u(x) = v(x) = v(x') = x'$ donc $u(x) = x'$.

Or $x \in G = \bigcup_{n \in \mathbb{N}} G_n$ donc il existe $n \in \mathbb{N}$ tel que $x \in G_n$.

On a donc $x' = u(x) \in u^{\rightarrow}(G_n) = G_{n+1} \subseteq G$.

Ainsi $x' \in G$, ce qui est absurde puisque par hypothèse $x' \notin G$.

Ce cas est donc impossible.

► Pour exactement la même raison, le cas $x \notin G$ et $x' \in G$ est impossible.

Ainsi dans les deux cas possibles, on a $x = x'$.

Donc v est injective.

- Montrons que v est surjective dans F .

Par définition de v on sait déjà que $\text{im}(v) \subseteq F$.

Soit $y \in F$.

► Plaçons-nous dans le cas où $y \in G$.

On a $G = \bigcup_{n \in \mathbb{N}} G_n$ donc il existe $n \in \mathbb{N}$ tel que $y \in G_n$.

Si $n = 0$ alors $y \in G_0 = E \setminus F$ donc $y \notin F$, ce qui est absurde.

On est donc forcément dans le cas où $n > 0$.

Il existe donc $m \in \mathbb{N}$ tel que $n = m + 1$ et donc $y \in G_{m+1} = u^{\rightarrow}(G_m)$.

Il existe donc $x \in G_m$ tel que $y = u(x)$.

Comme $x \in G_m$, on a $x \in G$ donc $v(x) = u(x) = y$, et donc $y \in \text{im}(v)$.

► Plaçons-nous dans le cas où $y \notin C$.

On a alors $v(y) = y$ donc $y \in \text{im}(v)$.

Dans les deux cas on a $y \in \text{im}(v)$.

Ainsi $\text{im}(v) \supseteq F$ et donc $\text{im}(v) = F$.

Ainsi v est surjective dans F .

Finalement v est injective et surjective dans F .

Donc $v : E \rightarrow F$ est une bijection, et donc $E \approx F$.

CQFD.

Voilà, nous sommes désormais en mesure de montrer l'implication $(E \preccurlyeq F \preccurlyeq E) \Rightarrow E \approx F$. Elle porte le nom de théorème de Cantor-(Schröder-)Bernstein.

Théorème 11 (de Cantor-Schröder-Bernstein)

Soient E et F deux ensembles.

Si $(E \preccurlyeq F \text{ et } F \preccurlyeq E)$, alors $E \approx F$.

On se propose de donner deux démonstrations de ce résultat.



Démonstration

Première démonstration

C'est celle qui fait intervenir la proposition 87 page 203.

Supposons que $E \preccurlyeq F$ et $F \preccurlyeq E$.

Il existe donc $f : E \rightarrow F$ et $g : F \rightarrow E$ deux injections.

Alors $g \circ f : E \rightarrow E$ est injective, et $\text{im}(g \circ f) \subseteq \text{im}(g)$.

Posons ensuite $G := \text{im}(g)$, de sorte que $\text{im}(g \circ f) \subseteq G$ et donc $g \circ f : E \rightarrow G$.

Comme $g \circ f$ est injective, on a donc $E \preccurlyeq G$.

Or g est à valeurs dans E donc $G = \text{im}(g) \subseteq E$.

Ainsi on a $G \subseteq E \preccurlyeq G$ donc $E \approx G$ d'après la proposition 87 page 203.

Il existe donc $v : E \rightarrow G$ une bijection.

De plus $g : F \rightarrow E$ est injective et $\text{im}(g) = G$ donc $g : F \rightarrow G$ est bijective.

Donc $g^{-1} : G \rightarrow F$ est une bijection.

Alors $g^{-1} \circ v : E \rightarrow F$ est une bijection, et donc $E \approx F$.

CQFD.

Démonstration

Deuxième démonstration

C'est celle qui fait intervenir le **théorème de Knaster-Tarski** abordé dans le précédent livre.

Supposons que $E \preccurlyeq F$ et $F \preccurlyeq E$.

Il existe donc $f : E \longrightarrow F$ et $g : F \longrightarrow E$ deux injections.

$$\text{Posons } \varphi := \begin{pmatrix} \mathcal{P}(E) & \longrightarrow & \mathcal{P}(E) \\ G & \longmapsto & E \setminus g^\rightarrow(F \setminus f^\rightarrow(G)) \end{pmatrix}$$

En munissant $\mathcal{P}(E)$ de l'inclusion, montrons que φ est croissante.

Soient G et G' deux parties de E telles que $G \subseteq G'$.

On a alors $f^\rightarrow(G) \subseteq f^\rightarrow(G')$ par croissance l'image directe.

On a donc $F \setminus f^\rightarrow(G) \supseteq F \setminus f^\rightarrow(G')$ par décroissance de la différence.

Donc $g^\rightarrow(F \setminus f^\rightarrow(G)) \supseteq g^\rightarrow(F \setminus f^\rightarrow(G'))$ par croissance de l'image directe.

Donc $E \setminus g^\rightarrow(F \setminus f^\rightarrow(G)) \subseteq E \setminus g^\rightarrow(F \setminus f^\rightarrow(G'))$ par décroissance de la différence.

Autrement dit on a $\varphi(G) \subseteq \varphi(G')$.

Ainsi φ est croissante.

D'après le théorème de Knaster-Tarski, φ admet un point fixe M , c'est-à-dire $\varphi(M) = M$.

On a donc $E \setminus g^\rightarrow(F \setminus f^\rightarrow(M)) = M$ et donc $E \setminus M = g^\rightarrow(F \setminus f^\rightarrow(M))$.

Autrement dit, pour tout $x \in E$, si $x \notin M$ alors $x \in g^\rightarrow(F \setminus f^\rightarrow(M))$ et donc $x \in \text{im}(g)$.

On a dit $g : F \longrightarrow E$ est injective par définition donc $g : F \longrightarrow \text{im}(g)$ est bijective.

On peut donc considérer sa réciproque $g^{-1} : \text{im}(g) \longrightarrow F$.

On peut donc poser

$$h := \begin{pmatrix} E & \longrightarrow & F \\ x & \longmapsto & \begin{cases} f(x) & \text{si } x \in M \\ g^{-1}(x) & \text{si } x \notin M \end{cases} \end{pmatrix}$$

- Montrons que h est injective.

Soit x et x' tels que $h(x) = h(x')$.

- Plaçons-nous dans le cas où $x \in M$ et $x' \in M$.

On a alors $f(x) = h(x) = h(x') = f(x')$ donc $f(x) = f(x')$.

Or f est injective par définition donc $x = x'$.

- Plaçons-nous dans le cas où $x \notin M$ et $x' \notin M$.

On a alors $g^{-1}(x) = h(x) = h(x') = g^{-1}(x')$ donc $g^{-1}(x) = g^{-1}(x')$.

Or g^{-1} est injective donc $x = x'$.

► Plaçons-nous dans le cas où $x \in M$ et $x' \notin M$.

On a donc $f(x) = h(x) = h(x') = g^{-1}(x')$ et donc $f(x) = g^{-1}(x')$.

Comme $x \in M$ on a donc $f(x) \in f^{\rightarrow}(M)$ et donc $g^{-1}(x') \in f^{\rightarrow}(M)$.

Or $x' \notin M$ et $E \setminus M = g^{\rightarrow}(F \setminus f^{\rightarrow}(M))$ donc $x' \in g^{\rightarrow}(F \setminus f^{\rightarrow}(M))$.

On a donc $g^{-1}(x') \in F \setminus f^{\rightarrow}(M)$ et donc $g^{-1}(x') \notin f^{\rightarrow}(M)$.

C'est impossible puisqu'on a justement dit que $g^{-1}(x') \in f^{\rightarrow}(M)$.

Ce cas est donc impossible.

► Le cas où $x \notin M$ et $x' \in M$ est impossible pour la même raison.

Dans les deux cas possibles on a donc $x = x'$.

Donc h est injective.

• Montrons que h est surjective dans F .

Par définition de h on sait déjà que $\text{im}(h) \subseteq F$.

Soit $y \in F$.

► Plaçons-nous dans le cas où $y \in f^{\rightarrow}(M)$.

Il existe donc $x \in M$ tel que $y = f(x)$.

Mais on a $h(x) = f(x)$ donc $h(x) = y$ et donc $y \in \text{im}(h)$.

► Plaçons-nous dans le cas où $y \notin f^{\rightarrow}(M)$.

Ainsi on a $y \in F \setminus f^{\rightarrow}(M)$.

Considérons $x := g(y)$, de sorte que $x \in g^{\rightarrow}(F \setminus f^{\rightarrow}(M))$.

Or on a dit que $M = E \setminus g^{\rightarrow}(F \setminus f^{\rightarrow}(M))$ donc $x \notin M$.

On a donc $h(x) = g^{-1}(x) = g^{-1}(g(y)) = y$ et donc $y \in \text{im}(h)$.

Dans les deux cas on a donc $y \in \text{im}(h)$.

Ainsi $\text{im}(h) \supseteq F$ et donc $\text{im}(h) = F$.

Ainsi h est surjective dans F .

Finalement h est injective et surjective dans F .

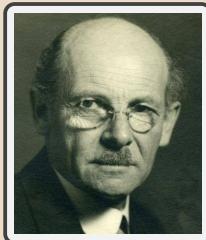
Donc $h : E \longrightarrow F$ est bijective, et donc $\boxed{E \approx F}$.

CQFD.

Remarque :

Ce théorème est souvent simplement nommé théorème de Cantor-Bernstein.

Pour la petite histoire

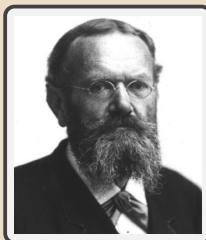


Felix Bernstein (24 février 1878 – 3 décembre 1956) est un mathématicien allemand.

Il a été l'élève de Cantor puis a soutenu sa thèse sur la théorie des ensembles sous la direction de Hilbert. Ses centres d'intérêt passent des ensembles aux probabilités et statistiques. À Gottingen en 1918 il fonde un institut de statistiques, où il traitera notamment de biostatistiques et mathématiques sur les assurances. Il fait des découvertes sur les transmissions génétiques des groupes sanguins.

Cantor énonce sans démonstration en 1887 le théorème de Cantor-Schröder-Bernstein. Felix Bernstein en produit une démonstration à l'âge de 18 ans en 1896, qui sera publiée deux ans plus tard par Borel, soit la même année que Schröder.

Pour la petite histoire



Ernst Schröder (25 novembre 1841 – 16 juin 1902) est un mathématicien allemand.

Son travail porte sur la logique et l'algèbre de Boole. C'est un personnage majeur de l'histoire de la logique mathématique, car il fit une synthèse des œuvres de Boole, De Morgan, MacColl, et particulièrement Sanders Peirce, et poursuivit leurs travaux. Il est connu en particulier pour son œuvre monumentale, les *Vorlesungen über die Algebra der Logik* (leçons sur l'algèbre de la logique), qui a aidé au développement de la logique mathématique en tant que discipline autonome au cours du 20^{ème} siècle.

En 1898 il propose une démonstration du théorème de Cantor-Schröder-Bernstein mais qui malheureusement comporte une erreur qui ne sera décelée par lui-même que 3 ans plus tard. Il reconnaîtra donc la paternité de la démonstration à Bernstein.

1.2 Théorème de Cantor

Il est évidemment possible qu'un ensemble ait strictement moins d'éléments qu'un autre. On parle alors de **stricte subpotence**.

Définition 33 (Stricte subpotence)

Soient A et B deux ensembles.

On dit que A est **strictement subpotent** à B si et seulement si $A \preccurlyeq B$ et $A \not\approx B$.

On note alors $A \prec B$.

Ainsi, la strict subpotence est un peu une version généralisée de l'ordre strict associé à une relation d'ordre. En particulier on retrouve la même idée de transitivité croisée. Passons en revue tous les cas possibles, car bien souvent ces situations se présenteront à nous.

Proposition 88 (Transitivité croisée de la subpotence)

Soient E , F et G trois ensembles.

1. Supposons avoir l'une des assertions suivantes :

- (a) $E \preccurlyeq F \approx G$
- (b) $E \approx F \preccurlyeq G$

alors $E \preccurlyeq G$.

2. Supposons avoir l'une des assertions suivantes :

- (a) $E \preccurlyeq F \prec G$
- (b) $E \approx F \prec G$
- (c) $E \prec F \preccurlyeq G$
- (d) $E \prec F \approx G$

alors on a $E \prec G$.

Démonstration

1. Supposons avoir $E \preccurlyeq F \approx G$ ou $E \approx F \preccurlyeq G$.

On a en particulier $E \preccurlyeq F \preccurlyeq G$ d'après la proposition 86 page 202.

On a donc $\boxed{E \preccurlyeq G}$ par transitivité de \preccurlyeq .

2. (a) Supposons avoir $E \preccurlyeq F \prec G$.

Comme $F \prec G$, on a $F \preccurlyeq G$ et $F \not\approx G$ par définition de \prec .

Ainsi on a $E \preccurlyeq F \preccurlyeq G$ et donc $E \preccurlyeq G$ par transitivité de \preccurlyeq .

Supposons par l'absurde avoir $E \approx G$.

On a en particulier $G \preccurlyeq E$ d'après la proposition 86 page 202.

On a donc $G \preccurlyeq E \preccurlyeq F$ donc $G \preccurlyeq F$ par transitivité de \preccurlyeq .

Comme $F \preccurlyeq G$, on a $F \approx G$ d'après le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on a justement dit que $F \not\approx G$.

Par l'absurde, on vient de montrer que $E \not\approx G$.

Comme on a $E \preccurlyeq G$, on en conclut que $[E \prec G]$ par définition de \prec .

(b) Supposons avoir $E \approx F \prec G$.

On a en particulier $E \preccurlyeq F \prec G$ d'après la proposition 86 page 202.

On a donc $[E \prec G]$ d'après (a).

(c) Supposons avoir $E \prec F \preccurlyeq G$.

Comme $E \prec F$, on a $E \preccurlyeq F$ et $E \not\approx F$ par définition de \prec .

Ainsi on a $E \preccurlyeq F \preccurlyeq G$ et donc $E \preccurlyeq G$ par transitivité de \preccurlyeq .

Supposons par l'absurde avoir $E \approx G$.

On a en particulier $G \preccurlyeq E$ d'après la proposition 86 page 202.

On a donc $F \preccurlyeq G \preccurlyeq E$ donc $F \preccurlyeq E$ par transitivité de \preccurlyeq .

Comme $E \preccurlyeq F$, on a $E \approx F$ d'après le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on a justement dit que $E \not\approx F$.

Par l'absurde, on vient de montrer que $E \not\approx G$.

Comme on a $E \preccurlyeq G$, on en conclut que $[E \prec G]$ par définition de \prec .

(d) Supposons avoir $E \prec F \approx G$.

On a en particulier $E \prec F \preccurlyeq G$ d'après la proposition 86 page 202.

On a donc $[E \prec G]$ d'après (c).

CQFD.

On l'a dit, la strict subpotence se comporte comme un ordre strict vis à vis de la subpotence. C'est d'autant plus vrai que l'on a les deux propriétés suivantes.

Proposition 89 (Strict subpotence et ordre strict)

Soient E , F et G trois ensembles.

1. On a $E \not\prec E$. La strict subpotence est **antiréflexive**.
2. Si $E \prec F \prec G$ alors $E \prec G$. La strict subpotence est **transitive**.



Démonstration

1. On a $E \approx E$ par réflexivité de \approx .

En particulier on a $\text{non}(E \not\approx E)$.

En particulier on a $\text{non}(E \preccurlyeq E \text{ et } E \not\approx E)$.

On a donc $E \not\prec E$ par définition de \prec .

2. Supposons avoir $E \prec F \prec G$.

On a en particulier $E \preccurlyeq F \prec G$ par définition de \prec .

On a donc $E \prec G$ d'après la proposition 88 page 209.

La proposition qui suit est très importante : étant donné une application, il existe toujours un ensemble qui n'est pas dans l'image de celle-ci. Son principe n'est pas sans rappeler celui du paradoxe de Russell, avec l'ensemble des ensembles qui ne s'appartiennent pas, ou le barbier qui ne rase que ceux qui ne se rasent pas.

Proposition 90 (Argument diagonal de Cantor)

Soit f une application.

Considérons $D := \{x \in \text{dom}(f) \mid x \notin f(x)\}$.

Alors $D \notin \text{im}(f)$.

Démonstration

Supposons par l'absurde que $D \in \text{im}(f)$.

Il existe donc $x \in \text{dom}(f)$ tel que $D = f(x)$.

► Plaçons-nous dans le cas où $x \in D$.

Alors par définition de D on a $x \notin f(x)$.

Mais par définition de x on a $f(x) = D$, si bien que $x \notin D$.

C'est absurde puisqu'on est justement dans le cas où $x \in D$.

► Plaçons-nous dans le cas où $x \notin D$.

Alors par définition de D on a $x \in f(x)$.

Mais par définition de x on a $f(x) = D$, si bien que $x \in D$.

C'est absurde puisqu'on est justement dans le cas où $x \notin D$.

Dans les deux cas on aboutit à une absurdité.

Par l'absurde on vient de montrer que $D \notin \text{im}(f)$.

CQFD.

On l'appelle **argument diagonal de Cantor** parce qu'il est possible de visualiser l'argument dans le cas où $f : E \longrightarrow \mathcal{P}(E)$, pour montrer que f ne peut pas être surjective dans $\mathcal{P}(E)$, ce qui est l'idée sous-jacente du théorème de Cantor qui va suivre. En effet, f va prendre chaque élément de E et lui associer une partie de E . Prenons par exemple le cas où $E = \{a, b, c, d\}$, et donnons un exemple d'une fonction $f : E \longrightarrow \mathcal{P}(E)$. Chaque ligne représente l'image d'un élément de E , et chaque colonne indique si un élément est dans la partie (la case est alors noire), ou non (la case est laissée blanche).

	$a \in$	$b \in$	$c \in$	$d \in$
$f(a)$	■	□	■	■
$f(b)$	□	□	■	■
$f(c)$	□	■	■	□
$f(d)$	■	□	□	□

Ainsi en lisant la première ligne, on peut voir que $a \in f(a)$, que $b \notin f(a)$, que $c \in f(a)$ et $d \in f(a)$, et donc $f(a) = \{a, c, d\}$. De même, on peut voir en lisant la deuxième ligne que $f(b) = \{c, d\}$, en lisant la troisième ligne que $f(c) = \{b, c\}$ et en lisant la quatrième ligne que $f(d) = \{a\}$.

L'argument diagonal consiste alors à former la partie donnée par la diagonale de ce tableau (celle qui part d'en haut à gauche pour aller en bas à droite), c'est-à-dire ■□■□. Comment interpréter cette diagonale ? Par exemple, le fait que sa première case soit noire indique que $a \in f(a)$, le fait que sa deuxième case soit blanche indique que $b \notin f(b)$, et de même $c \in f(c)$ et $d \notin f(d)$. Autrement dit, la diagonale nous donne la partie $\{a, c\}$, qui sont les éléments x de E vérifiant $x \in f(x)$. Il suffit alors de considérer son complémentaire, c'est-à-dire $\{b, d\}$, qui s'obtient en inversant les couleurs des cases, c'est-à-dire □■□■. Par définition, il s'agit justement des éléments x de E tels que $x \notin f(x)$, qui est bien la partie D que nous avons considérée dans l'énoncé.

Cette partie D ne peut pas être dans l'image de f . En effet, cela reviendrait à dire que D serait l'une des lignes du tableau, mais justement elle ne peut pas être la première ligne puisqu'elles ont leurs premières cases différentes, elle ne peut être la deuxième ligne car elles ont leurs deuxièmes cases différentes, et ainsi de suite elle ne peut ni être la troisième ni la quatrième ligne. D a été construite justement pour différer de chaque ligne du tableau, donc n'est pas dans l'image de f .

On peut donc se servir de cet argument pour démontrer le théorème suivant, qui est plus important qu'il n'y paraît au premier abord.

Théorème 12 (de Cantor)

Soit E un ensemble.

On a alors $E \prec \mathcal{P}(E)$.

Démonstration

- Montrons que $E \prec \mathcal{P}(E)$.

Pour cela, considérons $f := \begin{pmatrix} E & \longrightarrow & \mathcal{P}(E) \\ x & \longmapsto & \{x\} \end{pmatrix}$.

Soient x et x' dans E tels que $f(x) = f(x')$.

On a donc $\{x\} = \{x'\}$ et donc $x = x'$.

Ainsi $f : E \longrightarrow \mathcal{P}(E)$ est injective et donc $E \prec \mathcal{P}(E)$.

- Montrons que $E \not\approx \mathcal{P}(E)$.

Supposons par l'absurde que $E \approx \mathcal{P}(E)$.

Il existe donc une bijection $g : E \longrightarrow \mathcal{P}(E)$.

En particulier g est surjective dans $\mathcal{P}(E)$ donc $\text{im}(g) = \mathcal{P}(E)$.

Considérons alors $D := \{x \in E \mid x \notin g(x)\}$.

Ainsi $D \in \mathcal{P}(E) = \text{im}(g)$.

C'est absurde car d'après l'argument diagonal de Cantor, $D \notin \text{im}(g)$.

Par l'absurde, on vient de montrer que $E \not\approx \mathcal{P}(E)$.

Finalement on a $E \preccurlyeq \mathcal{P}(E)$ et $E \not\approx \mathcal{P}(E)$ donc $E \prec \mathcal{P}(E)$.

CQFD.

Le théorème de Cantor est évident dans le cas fini, mais sa conclusion est étonnante dans le cas infini : en effet, cela veut dire que \mathbb{N} a strictement moins d'éléments que $\mathcal{P}(\mathbb{N})$! Il y a donc différentes tailles d'infinis ! C'était déjà annoncé dans la partie sur les ordinaux (par exemple $\omega < \omega + 1$), mais cela provenait simplement de la façon d'ordonner les éléments, de disposer les bâtons devant soi. Ici ce qui est étonnant, c'est que cela se moque de la façon de présenter les éléments en quantité infini : il y a dans $\mathcal{P}(\mathbb{N})$ beaucoup plus d'éléments que dans \mathbb{N} , et ce de manière intrinsèque.

Dans l'illustration que nous avons donnée de l'argument de la diagonale de Cantor, nous avons représenté chaque partie de E comme une ligne ayant plusieurs cases, chaque case pouvant être ou bien noire, ou bien blanche. Cela revient donc à associer une partie de E à une application $E \longrightarrow \{0, 1\}$: la valeur 0 représentant les cases blanches, et la valeur 1 les cases noires. On tombe sur le principe des indicatrices, dont voici la définition.

Définition 34 (Indicatrice d'une partie)

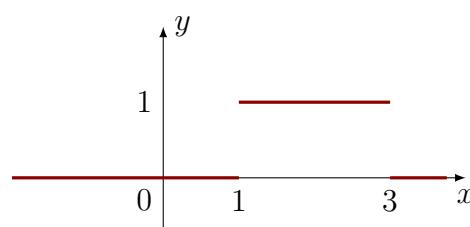
Soient E un ensemble et F une partie de E .

On appelle **indicatrice** de F (au sein de E) l'application

$$\mathbf{1}_F := \begin{pmatrix} E & \longrightarrow & \{0, 1\} \\ x & \longmapsto & \begin{cases} 1 & \text{si } x \in F \\ 0 & \text{si } x \notin F \end{cases} \end{pmatrix}$$

Exemple :

Dans le cas où $E = \mathbb{R}$, on peut parfois représenter graphiquement une indicatrice, par exemple ci-dessous l'indicatrice du segment $[1, 3]$.



Nous l'avons dit, avec l'illustration du tableau de tout à l'heure, chaque partie peut être associée à une ligne de cases noires ou blanches, et donc d'indicatrices. Mais cela fonctionne aussi dans l'autre sens, de sorte qu'on peut en fait réaliser une bijection entre les parties de E et les indicatrices de ces parties, c'est-à-dire l'ensemble des applications $E \rightarrow \{0, 1\}$. Rappelons que par définition $2 = S(1) = 1 \cup \{1\} = \{0\} \cup \{1\} = \{0, 1\}$.

Proposition 91 (Applications à valeurs dans 2 et parties)

Soit E un ensemble.

Alors $\mathcal{F}(E \rightarrow 2) \approx \mathcal{P}(E)$.

Démonstration

Pour rappel, $2 = \{0, 1\}$.

Montrons qu'il existe une bijection $\mathcal{P}(E) \rightarrow \mathcal{F}(E \rightarrow \{0, 1\})$.

Considérons $\varphi := \begin{pmatrix} \mathcal{P}(E) & \longrightarrow & \mathcal{F}(E \rightarrow \{0, 1\}) \\ F & \longmapsto & \mathbb{1}_F \end{pmatrix}$.

- Montrons que φ est injective.

Soient F et F' des parties de E telles que $\varphi(F) = \varphi(F')$.

On a donc $\mathbb{1}_F = \mathbb{1}_{F'}$.

Donc pour tout $x \in E$, $\mathbb{1}_F(x) = \mathbb{1}_{F'}(x)$.

Soit $z \in F$.

On a alors $\mathbb{1}_{F'}(z) = \mathbb{1}_F(z) = 1$ donc $z \in F'$.

Ainsi $F \subseteq F'$.

Par le même argument on montre que $F \supseteq F'$ et donc $F = F'$.

Ainsi φ est injective.

- Montrons que φ est surjective dans $\mathcal{F}(E \rightarrow \{0, 1\})$.

Par définition de φ on sait déjà que $\text{im}(\varphi) \subseteq \mathcal{F}(E \rightarrow \{0, 1\})$.

Soit $f : E \rightarrow \{0, 1\}$.

Considérons $F := \{x \in E \mid f(x) = 1\}$.

Montrons que $\mathbb{1}_F = f$.

Par définition $\text{dom}(\mathbb{1}_F) = E = \text{dom}(f)$.

Soit $x \in E$.

On a l'équivalence $\mathbb{1}_F(x) = 1 \iff x \in F \iff f(x) = 1$.

De même $\mathbb{1}_F(x) = 0 \iff x \notin F \iff f(x) = 0$.

Comme $\mathbb{1}_F$ et f sont à valeurs dans $\{0, 1\}$, on a donc $\mathbb{1}_F(x) = f(x)$.

Donc $\forall x \in E$, $\mathbb{1}_F(x) = f(x)$ et donc $\mathbb{1}_F = f$.

Autrement dit $f = \varphi(F)$ et donc $f \in \text{im}(\varphi)$.

Ainsi $\text{im}(\varphi) \supseteq \mathcal{F}(E \rightarrow \{0, 1\})$ et donc $\text{im}(\varphi) = \mathcal{F}(E \rightarrow \{0, 1\})$.

Autrement dit φ est surjective dans $\mathcal{F}(E \rightarrow \{0, 1\})$.

Finalement φ est injective et surjective dans $\mathcal{F}(E \rightarrow \{0, 1\})$.

Donc $\varphi : \mathcal{P}(E) \longrightarrow \mathcal{F}(E \rightarrow \{0, 1\})$ est bijective, et donc $\boxed{\mathcal{F}(E \rightarrow 2) \approx \mathcal{P}(E)}$.

CQFD.

1.3 Équipotence et opérations

Intéressons-nous à présent au comportement de l'équipotence et de la subpotence vis à vis des opérations ensemblistes. Commençons par le produit cartésien.

Proposition 92 (Produit cartésien, équipotence et subpotence)

Soient A, B, E et F quatre ensembles.

1. Si B est **non vide** alors $A \preccurlyeq A \times B$.
2. Si $A \preccurlyeq E$ et $B \preccurlyeq F$ alors $A \times B \preccurlyeq E \times F$.
3. Si $A \approx E$ et $B \approx F$ alors $A \times B \approx E \times F$.

Démonstration

1. Supposons que B est non vide.

Il existe donc $b_0 \in B$.

Considérons $f := \begin{pmatrix} A & \longrightarrow & A \times B \\ a & \longmapsto & (a, b_0) \end{pmatrix}$.

Montrons que f est injective.

Soient a et a' tels que $f(a) = f(a')$.

On a donc $(a, b_0) = (a', b_0)$ donc $a = a'$.

Donc $f : A \longrightarrow A \times B$ est injective et donc $\boxed{A \preccurlyeq A \times B}$.

2. Supposons que $A \preccurlyeq E$ et $B \preccurlyeq F$.

Il existe donc $f : A \longrightarrow E$ et $g : B \longrightarrow F$ deux injections.

Posons alors $\varphi := \begin{pmatrix} A \times B & \longrightarrow & E \times F \\ (a, b) & \longmapsto & (f(a), g(b)) \end{pmatrix}$.

Montrons que φ est injective.

Soient (a, b) et (a', b') dans $A \times B$ tels que $\varphi(a, b) = \varphi(a', b')$.

On a donc $(f(a), g(b)) = (f(a'), g(b'))$.

On a donc $f(a) = f(a')$ et $g(b) = g(b')$.

Or f et g sont injectives par définition.

Donc $a = a'$ et $b = b'$, et donc $(a, b) = (a', b')$.

Donc $\varphi : A \times B \longrightarrow E \times F$ est injective et donc $[A \times B \preccurlyeq E \times F]$.

3. Supposons que $A \approx E$ et $B \approx F$.

On a donc $A \preccurlyeq E$ et $E \preccurlyeq A$ et $B \preccurlyeq F$ et $F \preccurlyeq B$ d'après la proposition 86 page 202.

D'après 2, on a $A \times B \preccurlyeq E \times F$ et $E \times F \preccurlyeq A \times B$.

On a donc $[A \times B \approx E \times F]$ d'après le théorème de Cantor-Schröder-Bernstein.

CQFD.

Quand on fait l'union de deux ensembles A et B , si jamais A et B ont des éléments en communs (donc ne sont pas disjoints), ceux-ci ne se retrouveront qu'en un seul exemplaire dans $A \cup B$ puisqu'il est impossible d'avoir plusieurs fois le même élément dans un ensemble. C'est pour cette raison que nous avons introduit l'union disjointe $A \amalg B$, qui s'assure d'avoir deux exemplaires de chaque élément en commun de A et de B . On retrouve donc naturellement le résultat suivant.

Proposition 93 (Équipotence entre union et union disjointe)

Soient A et B deux ensembles.

1. On a $A \cup B \preccurlyeq A \amalg B$.
2. Si A et B sont disjoints alors $A \cup B \approx A \amalg B$.

Démonstration

1. Remarquons que pour $x \in A \cup B$:

- si $x \in A$ alors par définition $(0, x) \in A \amalg B$,
- et si $x \notin A$ alors $x \in B$ donc $(1, x) \in A \amalg B$.

On peut donc considérer l'application $f := \begin{cases} A \cup B & \longrightarrow A \amalg B \\ x & \longmapsto \begin{cases} (0, x) & \text{si } x \in A \\ (1, x) & \text{si } x \notin A \end{cases} \end{cases}$.

Montrons que f est injective.

Soient x et x' dans $A \cup B$ tels que $f(x) = f(x')$.

► Plaçons-nous dans le cas où $x \in A$ et $x' \in A$.

On a alors $(0, x) = f(x) = f(x') = (0, x')$ donc $(0, x) = (0, x')$.

On a donc $x = x'$.

► Plaçons-nous dans le cas où $x \notin A$ et $x' \notin A$.

On a alors $(1, x) = f(x) = f(x') = (1, x')$ donc $(1, x) = (1, x')$.

On a donc $x = x'$.

► Plaçons-nous dans le cas où $x \in A$ et $x \notin A$.

On a alors $(0, x) = f(x) = f(x') = (1, x')$ donc $(0, x) = (1, x')$.

On a donc $0 = 1$, ce qui est absurde, donc ce cas est impossible.

► Le cas où $x \notin A$ et $x' \in A$ est impossible pour la même raison.

Donc dans les deux cas possibles on a $x = x'$.

Donc f est injective.

Or $f : A \cup B \longrightarrow A \amalg B$ donc $[A \cup B \preccurlyeq A \amalg B]$.

2. Supposons que A et B sont disjoints.

Montrons que f est surjective dans $A \amalg B$.

Par définition de f on sait déjà que $\text{im}(f) \subseteq A \amalg B$.

Soit $(i, x) \in A \amalg B$.

► Plaçons-nous dans le cas où $i = 0$.

On a alors $x \in A$ par définition de $A \amalg B$.

On a donc $f(x) = (0, x) = (i, x)$ donc $(i, x) \in \text{im}(f)$.

► Plaçons-nous dans le cas où $i = 1$.

On a alors $x \in B$ par définition de $A \amalg B$.

Comme A et B sont **disjoints**, on a $x \notin A$.

On a donc $f(x) = (1, x) = (i, x)$ et donc $(i, x) \in \text{im}(f)$.

Dans les deux cas on a $(i, x) \in \text{im}(f)$.

Ainsi $\text{im}(f) \supseteq A \amalg B$ et donc $\text{im}(f) = A \amalg B$.

Ainsi f est surjective dans $A \amalg B$.

Finalement f injective et surjective dans $A \amalg B$.

Donc $f : A \cup B \longrightarrow A \amalg B$ est bijective, et donc $[A \cup B \approx A \amalg B]$.

CQFD.

On retrouve la même propriété pour l'union disjointe que pour le produit cartésien que l'on a vu plus tôt.

Proposition 94 (Équipotence et union disjointe)

Soient A, B, E et F quatre ensemble.

1. Si $A \preccurlyeq E$ et $B \preccurlyeq F$ alors $A \amalg B \preccurlyeq E \amalg F$.

2. Si $A \approx E$ et $B \approx F$ alors $A \amalg B \approx E \amalg F$.

 *Démonstration*

1. Supposons que $A \preccurlyeq E$ et $B \preccurlyeq F$.

Il existe donc $f : A \longrightarrow E$ et $g : B \longrightarrow F$ deux injections.

Posons alors $\varphi := \begin{cases} A \amalg B & \longrightarrow E \amalg F \\ (0, a) & \longmapsto (0, f(a)) \quad \text{pour tout } a \in A \\ (1, b) & \longmapsto (1, g(b)) \quad \text{pour tout } b \in B \end{cases}$.

Montrons que φ est injective.

Soient (i, x) et (j, y) dans $A \amalg B$ tels que $\varphi(i, x) = \varphi(j, y)$.

► Plaçons-nous dans le cas où $i = 0 = j$.

Alors $x \in A$ et $y \in A$.

On a donc $(0, f(x)) = \varphi(0, x) = \varphi(i, x) = \varphi(j, y) = \varphi(0, y) = (0, f(y))$.

Ainsi $(0, f(x)) = (0, f(y))$ donc $f(x) = f(y)$.

Mais f est injective par définition, donc $x = y$.

Comme $i = 0 = j$, on donc $(i, x) = (j, y)$.

► Plaçons-nous dans le cas où $i = 1 = j$.

Alors $x \in B$ et $y \in B$.

On a donc $(1, g(x)) = \varphi(1, x) = \varphi(i, x) = \varphi(j, y) = \varphi(1, y) = (1, g(y))$.

Ainsi $(1, g(x)) = (1, g(y))$ donc $g(x) = g(y)$.

Mais g est injective par définition, donc $x = y$.

Comme $i = 1 = j$, on donc $(i, x) = (j, y)$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

Alors $x \in A$ et $y \in B$.

On a donc $(0, f(x)) = \varphi(0, x) = \varphi(i, x) = \varphi(j, y) = \varphi(1, y) = (1, g(y))$.

Ainsi $(0, f(x)) = (1, g(y))$, ce qui est impossible puisqu'alors $0 = 1$.

► Le cas où $i = 1$ et $j = 0$ est impossible pour la même raison.

Ainsi dans les deux cas possibles, on a $(i, x) = (j, y)$.

Ainsi $\varphi : A \amalg B \longrightarrow E \amalg F$ est injective, et donc $[A \amalg B \preccurlyeq E \amalg F]$.

2. Supposons que $A \approx E$ et $B \approx F$.

On a alors $A \preccurlyeq E$ et $E \preccurlyeq A$ d'après la proposition 86 page 202.

De même on a $B \preccurlyeq F$ et $F \preccurlyeq B$.

On a donc $A \amalg B \preccurlyeq E \amalg F$ et $E \amalg F \preccurlyeq A \amalg B$ d'après 1.

On a donc $[A \amalg B \approx E \amalg F]$ d'après le théorème de Cantor-Schröder-Bernstein.

CQFD.

Vis à vis de l'équipotence (plutôt que de l'égalité), l'union disjointe et le produit cartésien

entretiennent le même rapport que l'addition et la multiplication, comme l'indique la proposition suivante. Les propriétés 1 et 3 rappellent la commutativité, les propriétés 2 et 4 rappellent l'associativité, et la propriété 5 rappelle la distributivité.

Proposition 95 (Union disjointe, produit cartésien, équipotence)

Soient A , B et C trois ensembles.

On a alors :

1. $A \amalg B \approx B \amalg A$.
2. $(A \amalg B) \amalg C \approx A \amalg (B \amalg C)$.
3. $A \times B \approx B \times A$.
4. $(A \times B) \times C \approx A \times (B \times C)$.
5. $A \times (B \amalg C) \approx (A \times B) \amalg (A \times C)$.

Démonstration

1.

Considérons $\tau : \{0, 1\} \longrightarrow \{0, 1\}$ définie par $\tau(0) := 1$ et $\tau(1) := 0$.

Remarquons que $\tau(\tau(0)) = \tau(1) = 0$ et $\tau(\tau(1)) = \tau(0) = 1$.

Considérons $f := \begin{pmatrix} A \amalg B & \longrightarrow & B \amalg A \\ (i, x) & \longmapsto & (\tau(i), x) \end{pmatrix}$ et $g := \begin{pmatrix} B \amalg A & \longrightarrow & A \amalg B \\ (i, x) & \longmapsto & (\tau(i), x) \end{pmatrix}$.

Montrons que f et g sont réciproques l'une de l'autre.

Soit $(i, x) \in A \amalg B$.

On a $(g \circ f)(i, x) = g(f(i, x)) = g(\tau(i), x) = (\tau(\tau(i)), x) = (i, x) = \text{id}_{A \amalg B}(i, x)$.

Ainsi $g \circ f = \text{id}_{A \amalg B}$, et on montre de même que $f \circ g = \text{id}_{B \amalg A}$.

En particulier $f : A \amalg B \longrightarrow B \amalg A$ est bijective, et donc $[A \amalg B \approx B \amalg A]$.

2. Commençons par observer à quoi ressemblent les éléments de $(A \amalg B) \amalg C$ et $A \amalg (B \amalg C)$.

Soit $x \in (A \amalg B) \amalg C$.

Trois cas se présentent à nous, par définition de l'union disjointe :

- ou bien il existe $a \in A$ tel que $x = (0, (0, a))$
- ou bien il existe $b \in B$ tel que $x = (0, (1, b))$
- ou bien il existe $c \in C$ tel que $x = (1, c)$

De même, soit $y \in A \amalg (B \amalg C)$.

Trois cas se présentent à nous, par définition de l'union disjointe :

- ou bien il existe $a \in A$ tel que $y = (0, a)$
- ou bien il existe $b \in B$ tel que $y = (1, (0, b))$

► ou bien il existe $c \in C$ tel que $y = (1, (1, c))$

$$\text{Considérons } f := \begin{pmatrix} (A \amalg B) \amalg C & \longrightarrow & A \amalg (B \amalg C) \\ (0, (0, a)) & \longmapsto & (0, a) & \text{pour } a \in A \\ (0, (1, b)) & \longmapsto & (1, (0, b)) & \text{pour } b \in B \\ (1, c) & \longmapsto & (1, (1, c)) & \text{pour } c \in C \end{pmatrix}.$$

$$\text{Considérons aussi } g := \begin{pmatrix} A \amalg (B \amalg C) & \longrightarrow & (A \amalg B) \amalg C \\ (0, a) & \longmapsto & (0, (0, a)) & \text{pour } a \in A \\ (1, (0, b)) & \longmapsto & (0, (1, b)) & \text{pour } b \in B \\ (1, (1, c)) & \longmapsto & (1, c) & \text{pour } c \in C \end{pmatrix}.$$

Montrons que f et g sont réciproques l'une de l'autre.

Soit $x \in (A \amalg B) \amalg C$.

► Plaçons-nous dans le cas où il existe $a \in A$ tel que $x = (0, (0, a))$.

Alors $(g \circ f)(x) = g(f(x)) = g(f(0, (0, a))) = g(0, a) = (0, (0, a)) = x$.

► Plaçons-nous dans le cas où il existe $b \in B$ tel que $x = (0, (1, b))$.

Alors $(g \circ f)(x) = g(f(x)) = g(f(0, (1, b))) = g(1, (0, b)) = (0, (1, b)) = x$.

► Plaçons-nous dans le cas où il existe $c \in C$ tel que $x = (1, c)$.

Alors $(g \circ f)(x) = g(f(x)) = g(f(1, c)) = g(1, (1, c)) = (1, c) = x$.

Dans tous les cas on a donc $(g \circ f)(x) = x = \text{id}_{(A \amalg B) \amalg C}(x)$.

On a donc $g \circ f = \text{id}_{(A \amalg B) \amalg C}$.

Soit $y \in A \amalg (B \amalg C)$.

► Plaçons-nous dans le cas où il existe $a \in A$ tel que $y = (0, a)$.

Alors $(f \circ g)(y) = f(g(y)) = f(g(0, a)) = f(0, (0, a)) = (0, a) = y$.

► Plaçons-nous dans le cas où il existe $b \in B$ tel que $y = (1, (0, b))$.

Alors $(f \circ g)(y) = f(g(y)) = f(g(1, (0, b))) = f(0, (1, b)) = (1, (0, b)) = y$.

► Plaçons-nous dans le cas où il existe $c \in C$ tel que $y = (1, (1, c))$.

Alors $(f \circ g)(y) = f(g(y)) = f(g(1, (1, c))) = f(1, c) = (1, (1, c)) = y$.

Dans tous les cas on a $(f \circ g)(y) = \text{id}_{A \amalg (B \amalg C)}(y)$.

On a donc $f \circ g = \text{id}_{A \amalg (B \amalg C)}$.

Finalement f et g sont réciproques l'une de l'autre.

En particulier $f : (A \amalg B) \amalg C \longrightarrow A \amalg (B \amalg C)$ est bijective, et donc $(A \amalg B) \amalg C \approx A \amalg (B \amalg C)$.

3. Commençons par remarquer la chose suivante.

Soient x et y deux ensembles.

On a alors $(x, y) \in A \times B \iff (x \in A \text{ et } y \in B) \iff (y, x) \in B \times A$.

Posons alors $f := \begin{pmatrix} A \times B & \longrightarrow & B \times A \\ (x, y) & \longmapsto & (y, x) \end{pmatrix}$ et $g := \begin{pmatrix} B \times A & \longrightarrow & A \times B \\ (y, x) & \longmapsto & (x, y) \end{pmatrix}$.

Montrons que f et g sont réciproques l'une de l'autre.

Soit $(x, y) \in A \times B$.

On a alors $(g \circ f)(x, y) = g(f(x, y)) = g(y, x) = (x, y) = \text{id}_{A \times B}(x, y)$.

Ainsi $g \circ f = \text{id}_{A \times B}$.

Soit $(y, x) \in B \times A$.

On a alors $(f \circ g)(y, x) = f(g(y, x)) = f(x, y) = (y, x) = \text{id}_{B \times A}(y, x)$.

Ainsi $f \circ g = \text{id}_{B \times A}$.

Finalement f et g sont réciproques l'une de l'autre.

En particulier $f : A \times B \longrightarrow B \times A$ est bijective, et donc $[A \times B \approx B \times A]$.

4. Commençons par observer à quoi ressemblent leurs éléments.

Pour tout $x \in (A \times B) \times C$, il existe $a \in A$, $b \in B$ et $c \in C$ tel que $x = ((a, b), c)$.

Pour tout $y \in A \times (B \times C)$, il existe $a \in A$, $b \in B$ et $c \in C$ tel que $y = (a, (b, c))$.

Considérons alors $f := \begin{pmatrix} (A \times B) \times C & \longrightarrow & A \times (B \times C) \\ ((a, b), c) & \longmapsto & (a, (b, c)) \end{pmatrix}$.

Considérons aussi $g := \begin{pmatrix} A \times (B \times C) & \longrightarrow & (A \times B) \times C \\ (a, (b, c)) & \longmapsto & ((a, b), c) \end{pmatrix}$.

Montrons que f et g sont réciproques l'une de l'autre.

Soit $x \in (A \times B) \times C$: il existe $a \in A$, $b \in B$ et $c \in C$ tels que $x = ((a, b), c)$.

Alors $g(f(x)) = g(f((a, b), c)) = g(a, (b, c)) = ((a, b), c) = x$.

Ainsi $(g \circ f)(x) = \text{id}_{(A \times B) \times C}(x)$.

On a donc $g \circ f = \text{id}_{(A \times B) \times C}$.

Soit $y \in A \times (B \times C)$: il existe $a \in A$, $b \in B$ et $c \in C$ tels que $y = (a, (b, c))$.

Alors $f(g(y)) = f(g(a, (b, c))) = f((a, b), c) = (a, (b, c)) = y$.

Ainsi $(f \circ g)(y) = \text{id}_{A \times (B \times C)}(y)$.

On a donc $f \circ g = \text{id}_{A \times (B \times C)}$.

Finalement f et g sont réciproques l'une de l'autre.

En particulier $f : (A \times B) \times C \longrightarrow A \times (B \times C)$ est une bijection.

On a donc $[(A \times B) \times C \approx A \times (B \times C)]$.

5. Commençons par observer à quoi ressemblent leurs éléments.

Soit $x \in A \times (B \amalg C)$.

Il existe alors $a \in A$ et $z \in B \amalg C$ tel que $x = (a, z)$.

Il existe donc $i \in \{0, 1\}$ et $u \in B \cup C$ tel que $z = (i, u)$.

Si $i = 0$ alors $u \in B$ et si $i = 1$ alors $u \in C$.

Finalement $x = (a, z) = (a, (i, u))$.

Soit $y \in (A \times B) \amalg (A \times C)$.

Il existe alors $j \in \{0, 1\}$ et un ensemble v tel que $y = (j, v)$.

Si $j = 0$ alors $v \in A \times B$ donc il existe $s \in A$ et $t \in B$ tel que $v = (s, t)$.

Si $j = 1$ alors $v \in A \times C$ donc il existe $s \in A$ et $t \in C$ tel que $v = (s, t)$.

Ainsi $y = (j, (s, t))$, avec $s \in A$, si $j = 0$ alors $t \in B$ et si $j = 1$ alors $t \in C$.

Remarquons en particulier que $(j, t) \in B \amalg C$.

$$\text{Considérons alors } f := \begin{cases} A \times (B \amalg C) & \longrightarrow (A \times B) \amalg (A \times C) \\ (a, (i, u)) & \longmapsto (i, (a, u)) \end{cases}.$$

$$\text{Considérons aussi } g := \begin{cases} (A \times B) \amalg (A \times C) & \longrightarrow A \times (B \amalg C) \\ (j, (s, t)) & \longmapsto (s, (j, t)) \end{cases}.$$

Montrons que f et g sont réciproques l'une de l'autre.

Soit $x \in A \times (B \amalg C)$.

Il existe alors $a \in A$, $i \in \{0, 1\}$ et $u \in B \cup C$ tel que $x = (a, (i, u))$.

Alors $g(f(x)) = g(f(a, (i, u))) = g(i, (a, u)) = (a, (i, u)) = x$.

On a donc $(g \circ f)(x) = \text{id}_{A \times (B \amalg C)}(x)$.

Ainsi $g \circ f = \text{id}_{A \times (B \amalg C)}$.

Soit $y \in (A \times B) \amalg (A \times C)$.

Il existe donc $s \in A$ et $(j, t) \in B \amalg C$ tel que $y = (j, (s, t))$.

Alors $f(g(y)) = f(g(j, (s, t))) = f(s, (j, t)) = (j, (s, t)) = y$.

On a donc $(f \circ g)(y) = \text{id}_{(A \times B) \amalg (A \times C)}(y)$.

Ainsi $f \circ g = \text{id}_{(A \times B) \amalg (A \times C)}(y)$.

Finalement f et g sont réciproques l'une de l'autre.

En particulier $f : A \times (B \amalg C) \longrightarrow (A \times B) \amalg (A \times C)$ est bijective.

On a donc $\boxed{A \times (B \amalg C) \approx (A \times B) \amalg (A \times C)}$.

CQFD.

Vis à vis de la subpotence et de l'équipotence, le vide joue un rôle particulier : être plus grand qu'un non vide fait de nous un non vide, être plus petit que le vide fait de nous le vide, et la seule façon d'être aussi grand que le vide, c'est d'être soi-même vide.

Proposition 96 (Vide, subpotence et équipotence)

Soient E et F deux ensembles.

1. On suppose que $E \preccurlyeq F$.
 - (a) Si E est non vide alors F est non vide.
 - (b) Si F est vide alors E est vide.
2. On suppose que $E \approx F$.
Alors E est vide si et seulement si F est vide.



Démonstration

1. Supposons que $E \preccurlyeq F$.

Il existe donc une injection $f : E \longrightarrow F$.

- (a) Supposons que E est non vide.

Il existe donc $x_0 \in E$.

Alors $f(x_0) \in F$ et donc F est non vide.

- (b) C'est simplement la contraposée de (a).

2. Supposons que $E \approx F$.

On a alors $E \preccurlyeq F$ et $F \preccurlyeq E$ d'après la proposition 86 page 202.

Comme $E \preccurlyeq F$, si F est vide alors E est vide d'après 1.(b).

Comme $F \preccurlyeq E$, si E est vide alors F est vide d'après 1.(b).

On a donc l'équivalence E est vide si et seulement si F est vide.

CQFD.

On observe des liens similaires avec les ensembles d'applications.

Proposition 97 (Ensembles d'applications et équipotence)

Soient A , B , E et F quatre ensembles.

1. Si $E \preccurlyeq F$ alors $\mathcal{F}(B \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)$.
2. Supposons que parmi E ou A , au moins l'un des deux est **non vide**.
 - (a) Si $A \preccurlyeq B$ alors $\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow E)$.
 - (b) En particulier si $A \preccurlyeq B$ et $E \preccurlyeq F$ alors $\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)$.
3. Si $A \approx B$ et $E \approx F$ alors $\mathcal{F}(A \rightarrow E) \approx \mathcal{F}(B \rightarrow F)$.

 *Démonstration*

1. Supposons que $E \preccurlyeq F$.

Par hypothèse on sait que $E \preccurlyeq F$ donc il existe $\phi : E \rightarrow F$ une injection.

Pour tout $f : B \rightarrow E$, on a alors $\phi \circ f : B \rightarrow F$.

On peut donc poser $\Phi := \begin{pmatrix} \mathcal{F}(B \rightarrow E) & \longrightarrow & \mathcal{F}(B \rightarrow F) \\ f & \longmapsto & \phi \circ f \end{pmatrix}$.

Montrons que Φ est injective.

Soient f et g deux applications $B \rightarrow E$ telles que $\Phi(f) = \Phi(g)$.

On sait que $\phi : E \rightarrow F$ est injective donc $\phi : E \rightarrow \text{im}(\phi)$ est bijective.

Donc $\phi^{-1} : \text{im}(\phi) \rightarrow E$ est bijective et vérifie $\phi^{-1} \circ \phi = \text{id}_E$.

On a donc

$$\begin{aligned} f &= \text{id}_E \circ f = (\phi^{-1} \circ \phi) \circ f = \phi^{-1} \circ (\phi \circ f) = \phi^{-1} \circ \Phi(f) \\ &= \phi^{-1} \circ \Phi(g) = \phi^{-1} \circ (\phi \circ g) = (\phi^{-1} \circ \phi) \circ g = \text{id}_E \circ g \\ &= g \end{aligned}$$

et donc $f = g$.

Ainsi $\Phi : \mathcal{F}(B \rightarrow E) \rightarrow \mathcal{F}(B \rightarrow F)$ est injective, et donc $\boxed{\mathcal{F}(B \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)}$.

2. On suppose ici que parmi E et A , au moins l'un des deux est non vide.

On a donc $E \neq \emptyset$ ou ($E = \emptyset$ et $A \neq \emptyset$).

On suppose aussi que $A \preccurlyeq B$: il existe $\varphi : A \rightarrow B$ une injection.

(a) Montrons que $\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow E)$.

• Plaçons-nous dans le cas où $E = \emptyset$ et $A \neq \emptyset$.

Alors il n'existe aucune application $A \rightarrow E$ donc $\mathcal{F}(A \rightarrow E) = \emptyset$.

Mais $A \preccurlyeq B$ et $A \neq \emptyset$, donc $B \neq \emptyset$ d'après la proposition 96 page 223.

On a donc pour la même raison $\mathcal{F}(B \rightarrow E) = \emptyset$.

Ainsi $\mathcal{F}(A \rightarrow E) = \mathcal{F}(B \rightarrow E)$, donc $\boxed{\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow E)}$ par réflexivité de \preccurlyeq .

• Plaçons-nous dans le cas où $E \neq \emptyset$.

On a dit que $\varphi : A \rightarrow B$ est une injection, donc $\varphi : A \rightarrow \text{im}(\varphi)$ est une bijection.

Donc $\varphi^{-1} : \text{im}(\varphi) \rightarrow A$ est une bijection.

Or E est **non vide** par hypothèse, donc il existe $y_0 \in E$.

Soit $f : A \rightarrow E$.

Construisons une application $B \rightarrow E$.

Autrement dit, on veut à $b \in B$ associer un élément de E .

Pour cela deux possibilités : ou bien $b \in \text{im}(\varphi)$, ou bien $b \notin \text{im}(\varphi)$.

Si $b \in \text{im}(\varphi)$ alors $\varphi^{-1}(b) \in A$ et donc on peut considérer $f(\varphi^{-1}(b)) \in E$.

Si $b \notin \text{im}(\varphi)$, on peut simplement associer y_0 .

Autrement dit, on peut considérer l'application

$$\psi_f := \begin{pmatrix} B & \longrightarrow & E \\ b & \longmapsto & \begin{cases} f(\varphi^{-1}(b)) & \text{si } b \in \text{im}(\varphi) \\ y_0 & \text{si } b \notin \text{im}(\varphi) \end{cases} \end{pmatrix}$$

$$\text{On peut alors poser } \psi := \begin{pmatrix} \mathcal{F}(A \rightarrow E) & \longrightarrow & \mathcal{F}(B \rightarrow E) \\ f & \longmapsto & \psi_f \end{pmatrix}.$$

Montrons que ψ est injective.

Soient f et g deux applications $A \rightarrow E$ telles que $\psi(f) = \psi(g)$.

Donc pour tout $b \in B$, on a $\psi_f(b) = \psi_g(b)$.

En particulier pour tout $b \in \text{im}(\varphi)$, on a $\psi_f(b) = \psi_g(b)$.

Autrement dit pour tout $b \in \text{im}(\varphi)$, $f(\varphi^{-1}(b)) = g(\varphi^{-1}(b))$.

Soit $a \in A$.

Posons $b := \varphi(a)$ de sorte que $a = \varphi^{-1}(b)$.

On a alors $f(a) = f(\varphi^{-1}(b)) = g(\varphi^{-1}(b)) = g(a)$.

Donc pour tout $a \in A$, on a $f(a) = g(a)$ et donc $f = g$.

Donc $\psi : \mathcal{F}(A \rightarrow E) \rightarrow \mathcal{F}(B \rightarrow E)$, et donc $\boxed{\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow E)}$.

(b) On suppose de plus que $E \preccurlyeq F$.

D'après 1, on a alors $\mathcal{F}(B \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)$.

Or on vient de montrer que $\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow E)$.

On a donc $\boxed{\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)}$ par transitivité de \preccurlyeq .

3. Ici on ne suppose pas particulier que parmi E ou A , au moins l'un des deux est non vide.

Quatre cas s'offrent alors à nous :

- $E = \emptyset$ et $A = \emptyset$.
- $E = \emptyset$ et $A \neq \emptyset$.
- $E \neq \emptyset$ et $A = \emptyset$.
- $E \neq \emptyset$ et $A \neq \emptyset$.

On suppose de plus que $A \approx B$ et $E \approx F$.

- Plaçons-nous dans le cas où $E = \emptyset$ et $A = \emptyset$.

Comme $E \approx F$ et $A \approx B$, on a $F = \emptyset$ et $B = \emptyset$ d'après la proposition 96 page 223.

On a donc $\mathcal{F}(A \rightarrow E) = \mathcal{F}(\emptyset \rightarrow \emptyset) = \mathcal{F}(B \rightarrow F)$.

En particulier $\boxed{\mathcal{F}(A \rightarrow E) \approx \mathcal{F}(B \rightarrow F)}$ par réflexivité de \approx .

- Plaçons-nous dans le cas où $E = \emptyset$ et $A \neq \emptyset$.

Comme $E \approx F$ et $A \approx B$, on a $F = \emptyset$ et $B \neq \emptyset$ d'après la proposition 96 page 223.

Il n'existe aucune application $A \rightarrow \emptyset$ et $B \rightarrow \emptyset$.

On a donc $\mathcal{F}(A \rightarrow E) = \mathcal{F}(A \rightarrow \emptyset) = \emptyset = \mathcal{F}(B \rightarrow \emptyset) = \mathcal{F}(B \rightarrow F)$.

En particulier $\boxed{\mathcal{F}(A \rightarrow E) \approx \mathcal{F}(B \rightarrow F)}$ par réflexivité de \approx .

- Plaçons-nous dans le cas où $E \neq \emptyset$ et $A = \emptyset$.

Comme $E \approx F$ et $A \approx B$, on a $F \neq \emptyset$ et $B = \emptyset$ d'après la proposition 96 page 223.

Il existe une unique application $\emptyset \rightarrow E$ et $\emptyset \rightarrow F$, c'est \emptyset lui-même.

On a donc $\mathcal{F}(A \rightarrow E) = \mathcal{F}(\emptyset \rightarrow E) = \{\emptyset\} = \mathcal{F}(\emptyset \rightarrow F) = \mathcal{F}(B \rightarrow F)$.

En particulier $\boxed{\mathcal{F}(A \rightarrow E) \approx \mathcal{F}(B \rightarrow F)}$ par réflexivité de \approx .

- Plaçons-nous dans le cas où $E \neq \emptyset$ et $A \neq \emptyset$.

Comme $E \approx F$ et $A \approx B$, on a $F \neq \emptyset$ et $B \neq \emptyset$ d'après la proposition 96 page 223.

On a $E \approx F$ donc $E \preccurlyeq F$ et $F \preccurlyeq E$ d'après la proposition 86 page 202.

On a $A \approx B$ donc $A \preccurlyeq B$ et $B \preccurlyeq A$ d'après la proposition 86 page 202.

Ainsi parmi E et A , au moins l'un des deux est non vide (les deux le sont).

Or $A \preccurlyeq B$ et $E \preccurlyeq F$ donc $\mathcal{F}(A \rightarrow E) \preccurlyeq \mathcal{F}(B \rightarrow F)$ d'après 2.(b).

De même, parmi F et B , au moins l'un des deux est non vide (les deux le sont).

Or $B \preccurlyeq E$ et $F \preccurlyeq A$ donc $\mathcal{F}(B \rightarrow F) \preccurlyeq \mathcal{F}(A \rightarrow E)$ d'après 2.(b).

On a donc $\boxed{\mathcal{F}(A \rightarrow E) \approx \mathcal{F}(B \rightarrow F)}$ d'après le théorème de Cantor-Schröder-Bernstein.

CQFD.

La proposition qui suit est analogue aux propriétés de l'exponentiation. En effet, si l'on se rappelle que $\mathcal{F}(B \rightarrow A)$ se note parfois A^B alors on va avoir

$$(A^B)^C \approx A^{B \times C} \text{ et } A^B \times A^C \approx A^{B \amalg C}$$

le produit cartésien remplaçant la multiplication et l'union disjointe l'addition.

Proposition 98 (Ensembles d'applications, union et produit)

Soient A, B et C trois ensembles.

1. On a $\mathcal{F}(C \rightarrow \mathcal{F}(B \rightarrow A)) \approx \mathcal{F}((C \times B) \rightarrow A)$.
2. On a $\mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) \approx \mathcal{F}(B \amalg C \rightarrow A)$.
3. Si B et C sont disjoints alors $\mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) \approx \mathcal{F}(B \cup C \rightarrow A)$.

 *Démonstration*

1.

Soit $f : C \rightarrow \mathcal{F}(B \rightarrow A)$.

Ainsi pour tout $c \in C$, $f(c) : B \rightarrow A$.

Ainsi pour tout $c \in C$ et tout $b \in B$, $f(c)(b) \in A$.

Posons alors $\varphi_f := \begin{pmatrix} C \times B & \longrightarrow & A \\ (c, b) & \longmapsto & f(c)(b) \end{pmatrix}$.

Considérons alors l'application

$$\varphi := \begin{pmatrix} \mathcal{F}(C \rightarrow \mathcal{F}(B \rightarrow A)) & \longrightarrow & \mathcal{F}((C \times B) \rightarrow A) \\ f & \longmapsto & \varphi_f \end{pmatrix}$$

Montrons que φ est injective.

Soient f et g deux applications $C \rightarrow \mathcal{F}(B \rightarrow A)$ telles que $\varphi(f) = \varphi(g)$.

Ainsi pour tout $(c, b) \in C \times B$, on a $\varphi_f(c, b) = \varphi_g(c, b)$.

Donc pour tout $c \in C$ et $b \in B$ on a $f(c)(b) = g(c)(b)$.

Donc pour tout $c \in C$, on a $f(c) = g(c)$, et donc $f = g$.

Donc φ est injective.

Montrons que φ est surjective dans $\mathcal{F}((C \times B) \rightarrow A)$.

Par définition de φ on sait déjà que $\text{im}(\varphi) \subseteq \mathcal{F}((C \times B) \rightarrow A)$.

Soit $g : (C \times B) \rightarrow A$.

Pour tout $c \in C$, posons $g_c := \begin{pmatrix} B & \longrightarrow & A \\ b & \longmapsto & g(c, b) \end{pmatrix}$ et $f := \begin{pmatrix} C & \longrightarrow & \mathcal{F}(B \rightarrow A) \\ c & \longmapsto & g_c \end{pmatrix}$.

Montrons que $\varphi(f) = g$.

Soit $(c, b) \in C \times B$.

On a $\varphi(f)(c, b) = \varphi_f(c, b) = f(c)(b) = g_c(b) = g(c, b)$.

Ainsi pour tout (c, b) on a $\varphi(f)(c, b) = g(c, b)$ donc $\varphi(f) = g$ et donc $g \in \text{im}(\varphi)$.

Ainsi $\text{im}(\varphi) \subseteq \mathcal{F}((C \times B) \rightarrow A)$ et donc $\text{im}(\varphi) = \mathcal{F}((C \times B) \rightarrow A)$.

Ainsi φ est surjective dans $\mathcal{F}((C \times B) \rightarrow A)$.

Finalement φ est injective et surjective dans $\mathcal{F}((C \times B) \rightarrow A)$.

Donc $\varphi : \mathcal{F}(C \rightarrow \mathcal{F}(B \rightarrow A)) \rightarrow \mathcal{F}((C \times B) \rightarrow A)$ est bijective.

On a donc $\boxed{\mathcal{F}(C \rightarrow \mathcal{F}(B \rightarrow A)) \approx \mathcal{F}((C \times B) \rightarrow A)}$.

2. Commençons par remarquer ceci.

Soit $(f, g) \in \mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A)$.

Ainsi $f : B \rightarrow A$ et $g : C \rightarrow A$.

Soit $(i, x) \in B \amalg C$.

Si $i = 0$ alors $x \in B$ et donc $f(x) \in A$.

Si $i = 1$ alors $x \in C$ et donc $g(x) \in A$.

Posons alors $\varphi_{f,g} := \begin{pmatrix} B \amalg C & \longrightarrow & A \\ (i, x) & \longmapsto & \begin{cases} f(x) & \text{si } i = 0 \\ g(x) & \text{si } i = 1 \end{cases} \end{pmatrix}$.

Considérons alors $\varphi := \begin{pmatrix} \mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) & \longrightarrow & \mathcal{F}(B \amalg C \rightarrow A) \\ (f, g) & \longmapsto & \varphi_{f,g} \end{pmatrix}$.

Montrons que φ est injective.

Soient (f, g) et (f', g') dans $\mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A)$ tels que $\varphi(f, g) = \varphi(f', g')$.

Ainsi pour tout $(i, x) \in B \amalg C$ on a $\varphi_{f,g}(i, x) = \varphi_{f',g'}(i, x)$.

Montrons que $f = f'$.

Soit $x \in B$.

On a alors $f(x) = \varphi_{f,g}(0, x) = \varphi_{f',g'}(0, x) = f'(x)$.

Donc pour tout $x \in B$ on a $f(x) = f'(x)$ et donc $f = f'$.

Montrons que $g = g'$.

Soit $x \in C$.

On a alors $g(x) = \varphi_{f,g}(1, x) = \varphi_{f',g'}(1, x) = g(x)$.

Donc pour tout $x \in C$, on a $g(x) = g'(x)$ et donc $g = g'$.

Ainsi $f = f'$ et $g = g'$ donc $(f, g) = (f', g')$.

Donc φ est injective.

Montrons que φ est surjective dans $\mathcal{F}(B \amalg C \rightarrow A)$.

Par définition de φ on sait déjà que $\text{im}(\varphi) \subseteq \mathcal{F}(B \amalg C \rightarrow A)$.

Soit $h : B \amalg C \rightarrow A$.

$$\text{Considérons } f := \begin{pmatrix} B & \longrightarrow & A \\ b & \longmapsto & h(0, b) \end{pmatrix} \text{ et } g := \begin{pmatrix} C & \longrightarrow & A \\ x & \longmapsto & h(1, c) \end{pmatrix}.$$

Montrons que $\varphi(f, g) = h$.

Soit $(i, x) \in B \amalg C$.

Si $i = 0$ alors $\varphi(f, g)(i, x) = \varphi_{f,g}(i, x) = f(x) = h(i, x)$.

Si $i = 1$ alors $\varphi(f, g)(i, x) = \varphi_{f,g}(i, x) = g(x) = h(i, x)$.

Dans les deux cas on a $\varphi(f, g)(i, x) = h(i, x)$.

Ainsi pour tout $(i, x) \in B \amalg C$ on a $\varphi(f, g)(i, x) = h(i, x)$.

On a donc $\varphi(f, g) = h$ et donc $h \in \text{im}(\varphi)$.

Ainsi $\text{im}(\varphi) \supseteq \mathcal{F}(B \amalg C \rightarrow A)$ et donc $\text{im}(\varphi) = \mathcal{F}(B \amalg C \rightarrow A)$.

Donc φ est surjective dans $\mathcal{F}(B \amalg C \rightarrow A)$.

Finalement φ est injective et surjective dans $\mathcal{F}(B \amalg C \rightarrow A)$.

Donc $\varphi : \mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) \rightarrow \mathcal{F}(B \amalg C \rightarrow A)$ est bijective.

On a donc $\boxed{\mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) \approx \mathcal{F}(B \amalg C \rightarrow A)}$.

3. Supposons que B et C sont disjoints.

On a alors $B \amalg C \approx B \cup C$ d'après la proposition 93 page 216.

De plus on sait que $A \approx A$ par réflexivité de \approx .

On a donc $\mathcal{F}(B \amalg C \rightarrow A) \approx \mathcal{F}(B \cup C \rightarrow A)$ d'après la proposition ?? page ??.

Finalement, on a $\boxed{\mathcal{F}(B \rightarrow A) \times \mathcal{F}(C \rightarrow A) \approx \mathcal{F}(B \cup C \rightarrow A)}$ par transitivité de \approx .

CQFD.

2 Nombres cardinaux

2.1 Les cardinaux

Maintenant que nous en savons un peu plus sur l'équipotence et la subpotence, observons quelques propriétés qu'entretiennent les ordinaux vis à vis de celles-ci.

Proposition 99 (Ordinaux, équipotence et subpotence)

Soient E un ensemble, et α, β et γ trois ordinaux.

1. Si $\alpha \leq \beta$ alors $\alpha \preccurlyeq \beta$.
2. Si $\alpha \leq \beta \leq \gamma$ et $\alpha \approx \gamma$ alors $\alpha \approx \beta \approx \gamma$.
3. Si $E \preccurlyeq \alpha$ alors il existe un ordinal $\delta \leq \alpha$ tel que $E \approx \delta$.



Démonstration

1. Supposons que $\alpha \leq \beta$.

Par définition de \leq on a donc $\alpha \subseteq \beta$.

On a donc $\boxed{\alpha \preccurlyeq \beta}$ d'après la proposition 87 page 203.

2. Supposons que $\alpha \leq \beta \leq \gamma$ et $\alpha \approx \gamma$.

D'après 1, on a alors $\alpha \preccurlyeq \beta \preccurlyeq \gamma$.

Mais $\alpha \approx \gamma$ donc $\gamma \preccurlyeq \alpha$ d'après la proposition 86 page 202.

Ainsi on a $\gamma \preccurlyeq \alpha \preccurlyeq \beta \preccurlyeq \gamma$.

On a donc $\boxed{\alpha \approx \beta \approx \gamma}$ d'après le théorème de Cantor-Schröder-Benrstein.

3. Supposons que $E \preccurlyeq \alpha$.

Il existe donc $f : E \longrightarrow \alpha$ une application injective.

Donc $f : E \longrightarrow \text{im}(f)$ est bijective.

Or f est à valeurs dans l'ordinal α donc $\text{im}(f)$ est une partie de α .

En particulier $\text{im}(f)$ est un ensemble d'ordinaux.

Donc $\text{im}(f)$ est totalement ordonné d'après le théorème 1 page 21.

Posons alors $\delta := \text{type}(\text{im}(f), \leq)$.

Comme $\text{im}(f) \subseteq \alpha$ on a $\text{type}(\text{im}(f)) \leq \text{type}(\alpha)$ d'après la proposition 27 page 60.

Or α est un ordinal, et donc on a $\boxed{\delta \leq \alpha}$

Par définition du type, $\text{im}(f)$ est isomorphe à δ .

Il existe donc $\varphi : \text{im}(f) \longrightarrow \delta$ un isomorphisme d'ordres.

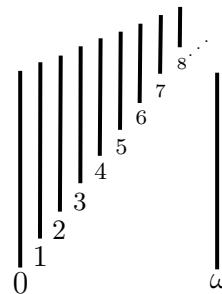
En particulier $\varphi : \text{im}(f) \longrightarrow \delta$ est une bijection, si bien que $\text{im}(f) \approx \delta$.

Or on a dit que $f : E \longrightarrow \text{im}(f)$ est une bijection, si bien que $E \approx \text{im}(f)$.

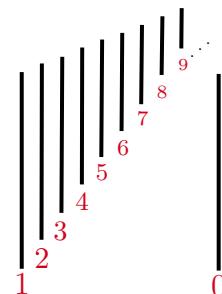
Finalement on a $E \approx \delta$ par transitivité de \approx .
CQFD.

Il est temps de définir ce que l'on appelle les cardinaux. On l'a vu lors de la proposition 84 page 201, l'équipotence est une relation d'équivalence. En cela, elle possède des classes d'équivalence : les éléments d'une même classe sont tous ceux qui sont équipotents deux à deux, donc tous ceux qui ont intuitivement le même nombre d'éléments. L'idée est de dire que parmi tous ces éléments se trouvent des ordinaux. Cependant, il n'y a pas de raisons qu'il n'y ait qu'un seul ordinal dans toute la classe, c'est-à-dire qu'il est possible que deux ordinaux différents soient pourtant équipotents.

Prenons l'exemple de $\omega + 1$, dont voici la représentation habituelle avec des bâtons.



Comme la notion de bijection se moque de l'ordre (contrairement à la notion d'isomorphisme d'ordre), il est tout à fait possible de renuméroter chacun des bâtons, de façon à mettre en bijection $\omega + 1$ avec ω , comme ceci.



Ainsi, ω et $\omega + 1$ ont beau être différents, ils sont équipotents. Pour s'en sortir, on va simplement se souvenir que parmi toutes les ordinaux de la classe, il y en a forcément un qui est plus petit que tous les autres : c'est lui que nous appelons cardinal de la classe. Autrement dit, un cardinal est un ordinal qui n'est pas équipotent à un ordinal plus petit, c'est le tout premier de sa classe.

Définition 35 (Cardinaux)

Soit α un ordinal.
 On dit que α est un **cardinal** si et seulement si $\forall \beta < \alpha, \beta \prec \alpha$.

La proposition qui suit est une simple reformulation de la définition mais qui s'avère très souvent pratique pour des démonstrations par l'absurde, puisqu'elle exhibe l'existence d'un élément.

Proposition 100 (Ne pas être un cardinal)

Soit α un ordinal.

Les assertions suivantes sont équivalentes :

1. α n'est pas un cardinal.
2. Il existe un ordinal β tel que $\beta < \alpha$ et $\beta \approx \alpha$.

Démonstration

$1 \Rightarrow 2$

Supposons que α n'est pas un cardinal.

Il existe donc un ordinal β tel que $\beta < \alpha$ et $\beta \not\approx \alpha$.

On a $\beta < \alpha$ donc en particulier $\beta \leq \alpha$.

On a donc $\beta \preccurlyeq \alpha$ d'après la proposition 99 page 230.

Ainsi on a $\beta \preccurlyeq \alpha$ et $\beta \not\approx \alpha$.

On a donc $\beta \preccurlyeq \alpha$ et non($\beta \preccurlyeq \alpha$ et $\beta \not\approx \alpha$).

On en conclut donc que $\beta \approx \alpha$.

$1 \Leftarrow 2$

Supposons qu'il existe un ordinal β tel que $\beta < \alpha$ et $\beta \approx \alpha$.

En particulier $\beta < \alpha$ et non($\beta \preccurlyeq \alpha$ et $\beta \not\approx \alpha$).

Donc $\beta < \alpha$ et $\beta \not\approx \alpha$.

Donc α n'est pas un cardinal.

CQFD.

Ainsi avec cette notion de cardinal, on est assuré qu'il n'y en a qu'un seul par classe d'équivalence. C'est ce que précise la proposition suivante.

Proposition 101 (Égalité entre deux cardinaux)

Soient κ et λ deux cardinaux.

On a l'équivalence $\kappa = \lambda \iff \kappa \approx \lambda$.

Démonstration

\Rightarrow

Supposons que $\kappa = \lambda$.

Alors $\kappa \approx \lambda$ par réflexivité de \approx .

\Leftarrow

Supposons que $\kappa \approx \lambda$.

Supposons par l'absurde que $\kappa \neq \lambda$.

Les cardinaux sont des ordinaux, et tous les ordinaux sont comparables.

On a donc $\kappa < \lambda$ ou $\lambda < \kappa$.

► Plaçons-nous dans le cas où $\kappa < \lambda$.

On a donc $\kappa \prec \lambda$ par λ est un cardinal.

En particulier $\kappa \not\approx \lambda$ par définition de \prec .

C'est absurde puisqu'on a justement supposé que $\kappa \approx \lambda$.

► Le cas où $\lambda < \kappa$ est absurde pour la même raison.

Dans les deux cas on about à une absurdité.

Par l'absurde, on vient de prouver que $\boxed{\kappa = \lambda}$.

CQFD.

Le théorème qui suit nous dit plusieurs choses, notamment le lien qu'entretiennent ω et les entiers naturels avec la notion de cardinal. Premièrement, les cardinaux à partir de ω sont tous limites : en effet pour la même raison que $\omega \approx \omega + 1$, un ordinal après ω est toujours équivalent à son successeur, ce qui empêche ce successeur d'être un cardinal.

Le théorème nous dit aussi ce que l'on sait depuis l'enfance : les entiers naturels sont aussi bien des ordinaux que des cardinaux, c'est-à-dire que l'on peut s'en servir pour donner des positions (le premier, le deuxième, etc) mais aussi pour compter.

Théorème 13 (Propriétés des cardinaux)

1. Pour tout cardinal α , si $\omega \leq \alpha$ alors α est limite.
2. Tout entier naturel est un cardinal.
3. Pour tout ensemble A de cardinaux, $\sup(A)$ est un cardinal.
4. ω est un cardinal.



Démonstration

1. Soit α un ordinal tel que $\omega \leq \alpha$.

Supposons que α n'est pas limite.

Il existe donc un ordinal β tel que $\alpha = \beta + 1$.

Comme $\omega \leq \alpha$ et ω est limite, on a $\omega \neq \alpha$ donc $\omega < \alpha$.

Ainsi $\omega < \beta + 1$ donc $\omega \leq \beta$ d'après la proposition 13 page 33.

Ainsi on a $\omega \subseteq \beta \subseteq \alpha$ par définition de \leq .

On peut donc définir l'application $\alpha \longrightarrow \beta$ suivante :

$$f := \begin{cases} \alpha \longrightarrow \beta \\ \gamma \longmapsto \begin{cases} \gamma + 1 & \text{si } \gamma < \omega \\ \gamma & \text{si } \omega \leq \gamma < \beta \\ 0 & \text{si } \gamma = \beta \end{cases} \end{cases}$$

Montrons que f est injective.

Soient γ et γ' dans α tels que $f(\gamma) = f(\gamma')$.

- Plaçons-nous dans le cas où $\gamma < \omega$ et $\gamma' < \omega$.
Alors $\gamma + 1 = f(\gamma) = f(\gamma') = \gamma' + 1$ et donc $\gamma = \gamma'$.
- Plaçons-nous dans le cas où $\omega \leq \gamma < \beta$ et $\omega \leq \gamma' < \beta$.
On a alors $\gamma = f(\gamma) = f(\gamma') = \gamma'$ et donc $\gamma = \gamma'$.
- Le cas où $\gamma = \beta = \gamma'$ donne directement $\gamma = \gamma'$.
- Les autres cas sont impossibles : l'image d'un entier naturel donne un entier naturel non nul donc différent d'un γ tel que $\omega \leq \gamma < \beta$ et différent de l'image nulle de β .

Dans les trois cas possibles, on a donc $\gamma = \gamma'$.

Donc $f : \alpha \longrightarrow \beta$ est injective, et donc $\alpha \preccurlyeq \beta$.

Comme $\beta \subseteq \alpha$, on a $\beta \preccurlyeq \alpha$ d'après la proposition 87 page 203.

Finalement on a $\alpha \approx \beta$ d'après le théorème de Cantor-Schröder-Bernstein.

En particulier on n'a pas $\beta \prec \alpha$.

Comme $\beta < \alpha$, on en conclut que α n'est pas un cardinal.

Donc si α n'est pas limite alors α n'est pas un cardinal.

Par contraposition, si α est un cardinal alors $\boxed{\alpha \text{ est limite}}$.

2. Démontrons-le par induction.

Pour tout entier naturel n , on pose $P(n)$ l'assertion « n est un cardinal ».

Initialisation

Pour tout ordinal β , l'implication $\beta < 0 \Rightarrow \beta \prec 0$ est vraie car sa prémissse est fausse.

Autrement dit on a $\forall \beta < 0, \beta \prec 0$ donc 0 est un cardinal et donc $P(0)$.

Hérité

Soit n un entier naturel tel que $P(n)$.

Ainsi n est un cardinal.

Supposons par l'absurde que $n + 1$ n'est pas un cardinal.

Il existe donc un ordinal $\beta < n + 1$ tel que $\beta \approx n + 1$ d'après la prop. 100 p. 232.

Il existe donc une bijection $f : \beta \longrightarrow n + 1$.

Il est impossible que f soit surjective dans $n + 1$ si $\beta = 0 = \emptyset$.

On a donc $\beta > 0$.

n est un entier naturel donc $n + 1$ aussi d'après la proposition 15 page 38.

Or $\beta < n + 1$, donc β est un entier naturel d'après la proposition 15 page 38.

Étant non nul, il existe un entier naturel m tel que $\beta = m + 1$.

Ainsi $f : m + 1 \longrightarrow n + 1$ est une bijection.

► Plaçons-nous dans le cas où $f(m) = n$.

Alors $f|_m$ est injective comme restriction d'une application injective.

Donc $\text{im}(f|_m) = \text{im}(f) \setminus \{f(m)\} = (n + 1) \setminus \{n\} = n$.

Donc $f|_m : m \longrightarrow n$ est une bijection.

► Plaçons-nous dans le cas où $f(m) < n$.

Considérons alors $i < m$ l'antécédent de n par f .

Dans ce cas-là, considérons $g : m + 1 \longrightarrow n + 1$ définie par

$$\begin{cases} g(j) = f(j) \text{ pour tout } j \neq i \text{ et } j \neq m \\ g(i) = f(m) \\ g(m) = f(i) = n \end{cases}$$

Alors $g : m + 1 \longrightarrow n + 1$ est une bijection car f l'est.

De plus elle vérifie $g(m) = n$ donc on est de nouveau dans le cas précédent.

Autrement dit $g|_m : m \longrightarrow n$ est une bijection.

Dans les deux cas on a construit une bijection $m \longrightarrow n$.

Ainsi $m \approx n$.

Mais on a dit que $m + 1 < n + 1$ donc $m < n$.

C'est en contradiction avec le fait que n est un cardinal.

Par l'absurde on vient de montrer que $n + 1$ est un cardinal, c'est-à-dire $P(n + 1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n + 1)$.

Ainsi P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout $n \in \mathbb{N}$, on a $P(n)$.

Autrement dit pour tout n entier naturel, n est un cardinal.

3. Soit A un ensemble de cardinaux.

Supposons par l'absurde que $\sup(A)$ n'est pas un cardinal.

Il existe donc un ordinal $\beta < \sup(A)$ et $\beta \approx \sup(A)$ d'après la prop. 100 p. 232.

Mais $\beta < \sup(A)$ donc il existe $\alpha \in A$ tel que $\beta < \alpha$.

Comme $\alpha \in A$, on a $\alpha \leq \sup(A)$.

Ainsi $\beta < \alpha \leq \sup(A)$ et $\beta \approx \sup(A)$ donc $\beta \approx \alpha$ d'après la proposition 99 page 230.

Ainsi on a $\beta < \alpha$ et $\beta \approx \alpha$ donc α n'est pas un cardinal d'après la prop. 100 p. 232.

C'est absurde puisque $\alpha \in A$ et A est un ensemble de cardinaux.

Par l'absurde, on vient de montrer que $\boxed{\sup(A) \text{ est un cardinal}}$.

4. D'après 2, ω est un ensemble de cardinaux.

Donc d'après 3, $\sup(\omega)$ est un cardinal.

Mais ω est limite donc $\sup(\omega) = \omega$ d'après la proposition 21 page 47.

Donc $\boxed{\omega \text{ est un cardinal}}$.

CQFD.

2.2 Le cardinal d'un ensemble

Nous l'avons dit, les nombres cardinaux sont là pour représenter les classes d'équipotence. On a prétendu que cela venait du fait que toute classe d'équipotence admettait au moins un ordinal en son sein, et c'est le plus petit d'entre ces ordinaux qui est le cardinal de la classe. Mais comment s'assurer qu'il y a au moins un ordinal dans la classe ?

Pour l'instant nous ne pouvons l'affirmer. Nous pouvons cependant nous intéresser aux ensembles munissables d'un bon ordre, car nous verrons qu'eux sont toujours équipotents à un ordinal, et donc à un cardinal.

Définition 36 (Ensemble bien ordonnable)

Soit E un ensemble.

On dit que E est **bien ordonnable** si et seulement s'il existe un bon ordre sur E .

La proposition qui suit explique l'intérêt porté aux ensembles bien ordonnables : pour eux nous sommes assurés d'être équipotents avec un ordinal, et donc admettre un cardinal. Cela n'est pas étonnant : les ordinaux ont été définis justement pour qu'ils soient isomorphes (et donc en particulier équipotents) aux ensembles bien ordonnés !

Proposition 102 (Bonne ordonnabilité et ordinal équivalent)

Soit E un ensemble.

Les assertions suivantes sont équivalentes :

1. E est bien ordonnable.
2. Il existe un ordinal α tel que $E \approx \alpha$.



Démonstration

$1 \Rightarrow 2$

Supposons que E est bien ordonnable.

Il existe donc \leq un bon ordre sur E .

Posons alors $\alpha := \text{type}(E, \leq)$.

Par définition du type, (E, \leq) et α sont isomorphes.

Il existe donc $f : E \rightarrow \alpha$ un isomorphisme d'ordres.

En particulier $f : E \rightarrow \alpha$ est une bijection et donc $[E \approx \alpha]$.

$1 \Leftarrow 2$

Supposons qu'il existe un ordinal α tel que $E \approx \alpha$.

Il existe donc $f : E \rightarrow \alpha$ une bijection.

Pour tout x et y dans E , on pose alors $x \leq y \iff f(x) \leq f(y)$.

On a vu dans le premier livre :

- que \leq est une relation d'ordre sur E .
- que $f : E \rightarrow \alpha$ est un isomorphisme d'ordres.

En particulier $f^{-1} : \alpha \rightarrow E$ est un isomorphisme d'ordres.

Or α est un ordinal donc est bien ordonné.

Donc (E, \leq) est bien ordonné d'après la proposition 22 page 48.

Donc $[E \text{ est bien ordonnable}]$.

CQFD.

Nous y sommes : pour un ensemble bien ordonnable, on peut définir son cardinal.

Proposition 103 (Cardinal d'un ensemble bien ordonnable)

Soit E un ensemble **bien ordonnable**.

Il existe un unique cardinal équivalent à E .

On l'appelle **cardinal** de E et on le note $\text{card}(E)$.

Démonstration

Existence

La classe $C := \{\alpha \in ON \mid E \approx \alpha\}$ est non vide d'après la proposition 102 page 237.

Elle admet donc un ordinal minimum α d'après la proposition 9 page 24.

En particulier $E \approx \alpha$ puisque $\alpha \in C$.

Montrons que α est un cardinal.

Supposons par l'absurde que α n'est pas un cardinal.

Il existe donc un ordinal $\beta < \alpha$ tel que $\beta \approx \alpha$ d'après la proposition 100 page 232.

On a dit que $E \approx \alpha$ donc $E \approx \beta$ par transitivité.

En particulier $\beta \in C$.

Mais on a aussi $\beta < \alpha$ et $\alpha = \min(C)$, d'où l'absurdité.

Par l'absurde, on vient de montrer que α est un cardinal.

Unicité

Soit α' un cardinal tel que $E \approx \alpha'$.

En particulier on a $\alpha \approx \alpha'$ par transitivité.

On a donc $\alpha = \alpha'$ d'après la proposition 101 page 232, d'où l'unicité.

CQFD.

Remarque :

Certains auteurs le notent aussi $|E|$ ou $\#E$.

On a vu dans la démonstration que si β est un ordinal tel que $E \approx \beta$ alors $\text{card}(E) \leq \beta$.

Malheureusement au point où nous en sommes, rien ne nous garantit que tout ensemble est bien ordonnable et donc rien ne nous garantit que tout ensemble admet un cardinal. Voici cependant une proposition qui nous permet de dire que si on est plus petit qu'un ensemble bien ordonnable, alors on est aussi bien ordonnable.

Proposition 104 (Bonne ordonnabilité et subpotence)

Soient E et F deux ensembles. On suppose que E est **bien ordonnable**.

1. Si $F \preceq E$ alors F est bien ordonnable et $\text{card}(F) \leq \text{card}(E)$.
2. En particulier si $F \subseteq E$ alors F est bien ordonnable et $\text{card}(F) \leq \text{card}(E)$.

Démonstration

1. Supposons que $F \preceq E$.

Par définition du cardinal, on a $E \approx \text{card}(E)$.

On a donc $F \preceq \text{card}(E)$ d'après la proposition 88 page 209.

Or $\text{card}(E)$ est un ordinal.

Il existe donc δ un ordinal tel que $F \approx \delta \leq \text{card}(E)$ d'après la proposition 99 page 230.

En particulier F est bien ordonnable d'après la proposition 102 page 237.

On a dit que $F \preccurlyeq \text{card}(E)$ et que $\text{card}(E)$ est un ordinal.

On a donc en particulier $\text{card}(F) \leq \text{card}(E)$ par définition du cardinal.

2. Supposons que $F \subseteq E$.

On a alors $F \preccurlyeq E$ d'après la proposition 87 page 203.

On peut donc conclure car on s'est ramené à 1.

CQFD.

La proposition qui suit est en elle-même évidente : on l'encadre surtout parce qu'elle revient souvent.

Proposition 105 (Type et équipotence)

Soit (E, \leq) un ensemble **bien ordonné**.

Alors $E \approx \text{type}(E, \leq)$.

En particulier si $\text{type}(E, \leq)$ est un cardinal, alors $\text{card}(E) = \text{type}(E, \leq)$.

Démonstration

Par définition du type, il existe $f : (E, \leq) \longrightarrow \text{type}(E, \leq)$ un isomorphisme d'ordres.

En particulier $f : E \longrightarrow \text{type}(E, \leq)$ est une bijection, et donc $E \approx \text{type}(E, \leq)$.

CQFD.

Étant donnés deux ensembles bien ordonnables, le cardinal de chacun d'eux se comporte bien vis à vis de l'équipotence et de la subpotence.

Proposition 106 (Cardinal, équipotence et (strict) subpotence)

Soient E et F deux ensembles **bien ordonnables**.

1. $E \preccurlyeq F \iff \text{card}(E) \leq \text{card}(F)$
2. $E \approx F \iff \text{card}(E) = \text{card}(F)$
3. $E \prec F \iff \text{card}(E) < \text{card}(F)$

 *Démonstration*

1. \Rightarrow

Supposons que $E \preccurlyeq F$.

Alors $\text{card}(E) \leq \text{card}(F)$ d'après la proposition 104 page 238.

\Leftarrow

Supposons que $\text{card}(E) \leq \text{card}(F)$.

Alors $\text{id}_{\text{card}(E)} : \text{card}(E) \longrightarrow \text{card}(F)$ est injective.

Par définition du cardinal, il existe $\varphi : E \longrightarrow \text{card}(E)$ et $\psi : \text{card}(F) \longrightarrow F$ deux bijections. Alors $\psi \circ \text{id}_{\text{card}(E)} \circ \varphi : E \longrightarrow F$ est injective.

Donc $[E \preccurlyeq F]$.

2. On a les équivalences suivantes :

$$\begin{aligned} E \approx F &\iff E \preccurlyeq F \text{ et } F \preccurlyeq E \text{ d'après le théorème de Cantor-Schröder-Bernstein} \\ &\iff \text{card}(E) \leq \text{card}(F) \text{ et } \text{card}(F) \leq \text{card}(E) \text{ d'après 1} \\ &\iff \text{card}(E) = \text{card}(F) \end{aligned}$$

Et donc finalement $[E \approx F \iff \text{card}(E) = \text{card}(F)]$.

3. On a les équivalences suivantes :

$$\begin{aligned} E \prec F &\iff E \preccurlyeq F \text{ et } E \not\approx F \text{ par définition} \\ &\iff \text{card}(E) \leq \text{card}(F) \text{ et } \text{card}(E) \neq \text{card}(F) \text{ d'après 1 et 2} \\ &\iff \text{card}(E) < \text{card}(F) \end{aligned}$$

Et donc finalement $[E \prec F \iff \text{card}(E) < \text{card}(F)]$.

CQFD.

On a déjà plus ou moins constaté le résultat suivant, mais faisons-en une proposition pour pouvoir l'utiliser systématiquement.

Proposition 107 (Cardinal plus petit et subpotence)

Soient E un ensemble **bien ordonnable** et κ un cardinal.

On a l'équivalence $\text{card}(E) \leq \kappa \iff E \preccurlyeq \kappa$.

 *Démonstration*



Supposons que $\text{card}(E) \leq \kappa$.

On a donc $\text{card}(E) \preccurlyeq \kappa$ d'après la proposition 99 page 230.

Or par définition du cardinal, on a $E \approx \text{card}(E)$.

On a donc $E \preccurlyeq \kappa$ d'après la proposition 88 page 209.



Supposons que $E \preccurlyeq \kappa$.

On a donc $\text{card}(E) \leq \text{card}(\kappa)$ d'après la proposition 106 page 239.

Or κ est un cardinal donc $\kappa = \text{card}(\kappa)$.

On a donc $\text{card}(E) \leq \kappa$.

CQFD.

Mais alors, existe-t-il des ensembles sans cardinal ? Pour dire qu'un ensemble admette un cardinal, il faut et il suffit que cet ensemble soit bien ordonnable, d'après la proposition 102 page 237. Existe-t-il donc des ensembles qui ne sont pas bien ordonnables ? La réponse est heureusement non : tout ensemble est bien ordonnable ! Enfin, c'est vrai ... mais à condition d'admettre l'**axiome du choix** ! C'est l'objet de la section suivante, qui va détailler tout cela.

3 Les grands théorèmes

3.1 Choix, Zorn et Zermelo

Dans cette section, nous allons utiliser l'axiome du choix pour démontrer deux théorèmes très importants des mathématiques : le lemme de Zorn et le théorème de Zermelo. Nous avons déjà démontré l'équivalence entre le lemme de Zorn et l'axiome du choix dans le précédent livre, mais sa démonstration y était très fastidieuse et pas très intuitive. Maintenant que l'on dispose des ordinaux, sa démonstration va prendre bien moins de pages, et on pourra essayer d'en dégager une intuition.

Nous allons pour cela à nouveau avoir besoin des classes. Dans le précédent livre, nous avons montré que pour deux ensembles E et F , s'il existe une injection $E \rightarrow F$ alors il existe une surjection $F \rightarrow E$. Si l'on accepte l'axiome du choix (ce qui est notre cas dans les Barbuki), alors c'est une équivalence. Dans le cas des classes, nous n'allons pas généraliser l'axiome du choix et donc n'allons retrouver que l'implication directe, que voici.

Proposition 108 (Injection et surjection avec des classes)

Soient C et D deux classes, avec C **non vide**.

Supposons qu'il existe une assertion fonctionnelle injective $C \rightarrow D$.

Alors il existe une assertion fonctionnelle surjective $D \rightarrow C$.



Démonstration

Supposons qu'il existe une assertion fonctionnelle injective $F : C \rightarrow D$.

On peut alors considérer sa réciproque $F^{-1} : \text{im}(F) \rightarrow C$.

Par définition C est non vide : il existe $x_0 \in C$.

Posons alors $G := \begin{cases} D & \rightarrow C \\ y & \longmapsto \begin{cases} F^{-1}(y) & \text{si } y \in \text{im}(F) \\ x_0 & \text{si } y \notin \text{im}(F) \end{cases} \end{cases}$.

Montrons que G est surjective dans C .

Par définition de G , on sait déjà que $\text{im}(G) \subseteq C$.

Soit $x \in C$.

Posons $y := F(x)$, de sorte que $y \in \text{im}(F)$.

Par définition de G , on a alors $G(y) = F^{-1}(y) = F^{-1}(F(x)) = x$.

On a donc $x \in \text{im}(G)$.

Ainsi $\text{im}(G) \supseteq C$ et donc $\text{im}(G) = C$.

Ainsi $G : D \rightarrow C$ est surjective dans C .

CQFD.

Le résultat qui suit ne devrait pas nous étonner : une classe propre ne pouvant pas être associée à un ensemble, cela veut dire qu'elle est trop grosse pour un ensemble. C'est en particulier le cas de la classe propre ON : il est impossible de l'injecter dans un ensemble.

Proposition 109 (Pas d'injection d' ON dans un ensemble)

Soient E un ensemble et C une classe **propre**.

Il n'existe pas d'assertion fonctionnelle injective $C \rightarrow E$.

Autrement dit on a $C \not\leq E$.



Démonstration

Supposons par l'absurde qu'il existe $F : C \rightarrow E$ une assertion fonctionnelle injective.

Il existe donc $G : E \rightarrow C$ une assertion fonctionnelle surjective d'après la prop. 108 p. 242.

G étant surjective dans C , on a $G^{-1}(E) = \text{im}(G) = C$.

Mais E est un ensemble donc $G^{-1}(E)$ est un ensemble d'après l'axiome de remplacement.

Ainsi C est un ensemble, ce qui est absurde puisque justement C est une classe propre.

CQFD.

Revenons désormais sur le lemme de Zorn, et démontrons-le à l'aide des ordinaux. Rappelons les définitions suivantes, pour que tout soit clair : pour un ensemble ordonné (E, \leq) ,

- un **élément maximal** m de E est un élément tel que $\forall x \in E, (m \leq x \Rightarrow m = x)$, c'est-à-dire que m est le seul élément de E à être plus grand que m .
- une **chaîne** C de E est une partie de E qui est totalement ordonnée, c'est-à-dire que tous les éléments de C sont comparables deux à deux.
- on dit que E est **inductif** si et seulement si toutes ses chaînes sont majorées.

L'idée du lemme de Zorn est alors de dire qu'un ensemble inductif admet forcément un élément maximal, c'est-à-dire un élément au delà duquel il n'est pas possible d'aller. Sa preuve repose sur l'idée de construire une chaîne par récursion (c'est là qu'interviennent les ordinaux) : comme la classe des ordinaux ON est vraiment plus grande que E (proposition 109 page 243), nécessairement on aura épousé à un moment tout ce que E pouvait nous offrir pour prolonger la chaîne, et donc il ne sera plus possible de continuer. On s'est alors arrêté sur un élément maximal de E .

Théorème 14 (Lemme de Zorn)

Soit E un ensemble ordonné.

Si E est non vide et inductif alors E admet un élément maximal.

 *Démonstration*

Notons \leq l'ordre sur E et \lhd l'ordre strict associé.

- Supposons que E est inductif.

En particulier E est non vide.

D'après l'**axiome du choix**, il existe une application $f : \mathcal{P}(E) \setminus \{\emptyset\} \rightarrow E$ telle que pour toute partie non vide A de E , on a $f(A) \in A$.

E est un ensemble et la classe des ensembles U est propre d'après le paradoxe de Russell.

Autrement dit il existe y un ensemble tel que $y \notin E$.

On prolonge alors f en une application $g : \mathcal{P}(E) \rightarrow E \cup \{y\}$ de la façon suivante :

$$g := \begin{cases} \mathcal{P}(E) & \longrightarrow E \cup \{y\} \\ A & \longmapsto \begin{cases} f(A) & \text{si } A \neq \emptyset \\ y & \text{si } A = \emptyset \end{cases} \end{cases}$$

Définissons une assertion fonctionnelle $G : ON \rightarrow E \cup \{y\}$ par récursion de la façon suivante : pour tout ordinal α , on pose $G(\alpha) := g(\{x \in E \mid \forall \beta < \alpha, G(\beta) \lhd x\})$.

D'après la proposition 109 page 243, G n'est pas injective.

Pour tout ordinal α , posons $A_\alpha := \{x \in E \mid \forall \beta < \alpha, G(\beta) \lhd x\}$.

Ainsi pour tout ordinal α on a $G(\alpha) = g(A_\alpha)$.

- Montrons que pour tout ordinal α , si $A_\alpha \neq \emptyset$ alors $\forall \beta < \alpha, G(\beta) \lhd G(\alpha)$.

Soit un ordinal α tel que $A_\alpha \neq \emptyset$.

On a alors $G(\alpha) = g(A_\alpha) = f(A_\alpha) \in A_\alpha$.

Autrement dit $G(\alpha) \in \{x \in E \mid \forall \beta < \alpha, G(\beta) \lhd x\}$.

Donc $\forall \beta < \alpha, G(\beta) \lhd G(\alpha)$.

Ainsi pour tout ordinal α , si $A_\alpha \neq \emptyset$ alors $\forall \beta < \alpha, G(\beta) \lhd G(\alpha)$.

Notons (\star_1) ce fait.

- Montrons que $y \in \text{im}(G)$.

Supposons par l'absurde que $y \notin \text{im}(G)$.

Par définition de G on a $\text{im}(G) \subseteq E \cup \{y\}$ donc $\text{im}(G) \subseteq E$.

Donc pour tout ordinal α , $G(\alpha) \in E$ et donc $g(A_\alpha) \in E$.

Or par définition de g , pour tout $A \in \mathcal{P}(E)$, $g(A) \in E \iff A \neq \emptyset$.

Donc pour tout ordinal α , $A_\alpha \neq \emptyset$ donc $\forall \beta < \alpha, G(\beta) \lhd G(\alpha)$ d'après (\star_1) .

Montrons que G est injective.

Soient α et α' des ordinaux tels que $G(\alpha) = G(\alpha')$.

Comme l'ordre est total chez les ordinaux, on a $\alpha = \alpha'$ ou $\alpha < \alpha'$ ou $\alpha' < \alpha$.

Si $\alpha < \alpha'$ alors $G(\alpha) \triangleleft G(\alpha')$ d'après ce qui précède.

Si $\alpha' < \alpha$ alors $G(\alpha') \triangleleft G(\alpha)$ d'après ce qui précède.

Dans ces deux cas on a $G(\alpha) \neq G(\alpha')$ par antiréflexivité de \triangleleft .

Autrement dit le seul cas possible est $\alpha = \alpha'$.

Ainsi $G : ON \longrightarrow E$ est injective, ce qui est absurde d'après la proposition 109 page 243.

On vient de montrer par l'absurde que $y \in \text{im}(G)$.

Il existe donc un ordinal α tel que $y = G(\alpha) = g(Y_\alpha)$ et donc $Y_\alpha = \emptyset$.

- Soit α_0 le plus petit ordinal tel que $Y_{\alpha_0} = \emptyset$.

Soit $\gamma < \alpha_0$.

Par minimalité de α_0 on a $Y_\gamma \neq \emptyset$.

On a donc $\forall \beta < \gamma, G(\beta) \triangleleft G(\gamma)$ d'après (\star_1) .

Ainsi $\forall \gamma < \alpha_0, \forall \beta < \gamma, G(\beta) \triangleleft G(\gamma)$.

Notons (\star_2) ce fait.

- Considérons alors $X := \{G(\beta) \mid \beta < \alpha_0\}$.

Montrons que X est une chaîne de E .

Soient z et z' dans X .

Il existe alors $\gamma < \alpha_0$ et $\gamma' < \alpha_0$ tels que $z = G(\gamma)$ et $z' = G(\gamma')$.

Comme les ordinaux sont totalement ordonnés on a $\gamma < \gamma'$ ou $\gamma' < \gamma$ ou $\gamma = \gamma'$.

► Plaçons-nous dans le cas où $\gamma < \gamma'$.

Comme $\gamma' < \alpha_0$, on a $\forall \beta < \gamma', G(\beta) \triangleleft G(\gamma')$ d'après (\star_2) .

En particulier on a $G(\gamma) \triangleleft G(\gamma')$.

► On montre de même que si $\gamma' < \gamma$ alors $G(\gamma') \triangleleft G(\gamma)$.

► Enfin si $\gamma = \gamma'$ on a évidemment $G(\gamma) = G(\gamma')$.

On a donc $(\gamma) \triangleleft G(\gamma')$ ou $G(\gamma') \triangleleft G(\gamma)$ ou $G(\gamma) = G(\gamma')$.

Donc $G(\gamma)$ et $G(\gamma')$ sont comparables, et donc z et z' sont comparables.

Ainsi tous les éléments de X sont comparables, donc X est une chaîne de E .

Or E est **inductif** par hypothèse, donc X admet un majorant m .

- Montrons que m est un élément maximal de E .

Supposons par l'absurde que m n'est pas un élément maximal de E .

Il existe donc $m' \in E$ tel que $m \triangleleft m'$ d'après la proposition 1 page 8.

Or m majore X donc m' majore strictement X .

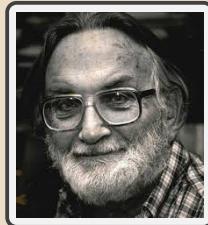
Autrement dit $\forall \beta < \alpha_0, G(\beta) < m'$.

Donc $m' \in Y_{\alpha_0}$, ce qui est absurde puisque $Y_{\alpha_0} = \emptyset$ par définition de α_0 .

Par l'absurde on vient de montrer que m est un élément maximal de E .

CQFD.

Pour la petite histoire



Max Zorn (6 juin 1906 – 9 mars 1993) est un mathématicien américain né en Allemagne. Ses travaux portent sur l'algèbre, la théorie des groupes et l'analyse numérique. Il est surtout connu pour le lemme de Zorn que nous venons de voir. Il a notamment été l'élève d'Artin.

Le lemme de Zorn a été énoncé par de nombreux mathématiciens dès le début du 20^{ème} siècle dans le cadre de la *Crise des fondements* : démontré pour la première fois par Kuratowski en 1922, il a été retrouvé de façon indépendante par Zorn en 1935. Si c'est finalement Zorn qui lui a donné son nom, c'est parce qu'il a été le premier à utiliser ce résultat pour simplifier de nombreuses preuves d'algèbre déjà existantes. Max Zorn lui-même n'en revendiquait pas la paternité : « *Ce n'est pas un lemme, et il n'est pas de moi* ».

Pour la petite histoire



Kazimierz Kuratowski (2 février 1896 – 18 juin 1980) est un mathématicien polonais.

Les travaux de Kuratowski portent sur la théorie des fonctions de variable réelle, la topologie – en particulier, en compagnie d'Hausdorff, la définition d'une topologie en termes de *voisinages*, les *espaces métriques* et la notion de *compacité* –, la théorie des ensembles et la théorie des graphes, laquelle lui est redevable d'un important théorème

portant désormais son nom. Avec Alfred Tarski et Wacław Sierpiński, il développe la théorie des espaces polonais (nommés ainsi en hommage au groupe de mathématiciens polonais à l'origine de cette théorie).

Dans le domaine de la théorie des ensembles, on doit à Kuratowski, dans son article *Une méthode d'élimination des nombres transfinis des raisonnements mathématiques*, publié en 1922, une première version du lemme de Zorn.

Le lemme de Zorn est en fait équivalent à l'axiome du choix. Autrement dit, dressons la liste de tous les axiomes que l'on a considérés jusque là, et retirons-en l'axiome du choix. À la place, mettons-y le lemme de Zorn. De ce nouveau système d'axiome, il est alors possible de déduire l'axiome du choix, qui en devient donc un théorème. On peut en retrouver une démonstration directe dans le précédent livre. On va donc ici le faire d'une manière détournée, avec le théorème qui va suivre.

Comme promis, tout ensemble est bien ordonnable et donc admet un cardinal. C'est ce que l'on appelle le **théorème de Zermelo**.

Théorème 15 (de Zermelo)

Tout ensemble est bien ordonnable.

On se propose deux démonstrations :

1. En utilisant le lemme de Zorn.
2. En utilisant directement l'axiome du choix.

Les deux démonstrations reposent sur l'idée qu'étant donné un ensemble E , une partie munie d'un bon ordre peut être vu comme la donnée d'un ordinal α et d'une injection $\alpha \longrightarrow E$. On peut donc raisonner sur les ordinaux plutôt que les parties de E .

Démonstration

Première démonstration : en utilisant le lemme de Zorn.

L'idée est de considérer tous les ordinaux qui s'injectent dans E : à chaque fois cela fournit un bon ordre sur les parties de E . Grâce au lemme de Zorn, cela va permettre de trouver un ordinal maximal s'injectant dans E , et par maximalité il sera en fait en bijection avec E , si bien que E est bien ordonnable au vu de la proposition 102 page 237.

Soit E un ensemble.

Montrons que E est bien ordonnable.

Posons $\mathcal{E} := \{(A, \leq) \mid A \subseteq E \text{ et } \leq \text{ est un bon ordre sur } A\}$.

Notre objectif est de montrer qu'il existe un bon ordre \leq sur E , et donc que $(E, \leq) \in \mathcal{E}$.

Comme indiqué en amont de la démonstration, on va plutôt raisonner sur les ordinaux et

les injections de ces ordinaux dans E .

Pour cela, posons la classe $\mathcal{F} := \{(\alpha, f) \mid \alpha \in ON \text{ et } f : \alpha \longrightarrow E \text{ est injective}\}$.

Nous allons appliquer le lemme de Zorn à \mathcal{F} .

- Montrons que \mathcal{F} est un ensemble.

Considérons pour cela l'assertion fonctionnelle

$$\varphi := \left(\begin{array}{l} \mathcal{E} \longrightarrow \mathcal{F} \\ (A, \leq) \longmapsto (\alpha, f) \text{ avec } \left\{ \begin{array}{l} \alpha = \text{type}(A, \leq) \\ \text{et } f : \alpha \longrightarrow (A, \leq) \text{ l'isomorphisme associé} \end{array} \right. \end{array} \right)$$

D'après l'axiome de remplacement, comme \mathcal{E} est un ensemble, $\text{im}(\varphi)$ aussi.

Montrons que $\text{im}(\varphi) = \mathcal{F}$, c'est-à-dire que φ est surjective dans \mathcal{F} .

Par définition de φ , on sait déjà que $\text{im}(\varphi) \subseteq \mathcal{F}$.

Soit $(\alpha, f) \in \mathcal{F}$.

Par définition de \mathcal{F} , α est un ordinal et $f : \alpha \longrightarrow E$ une injection.

Considérons $A := \text{im}(f)$, de sorte que $f : \alpha \longrightarrow A$ est bijective.

Autrement dit, $f^{-1} : A \longrightarrow \alpha$ est bijective.

On pose alors pour tout a et b dans A , $a \trianglelefteq b \iff f^{-1}(a) \leq f^{-1}(b)$.

On a vu dans le précédent livre qu'alors \trianglelefteq est une relation d'ordre sur A et que $f^{-1} : (A, \trianglelefteq) \longrightarrow \alpha$ est un isomorphisme d'ordres.

En particulier $f : \alpha \longrightarrow (A, \trianglelefteq)$ est un isomorphisme d'ordres.

Comme α est bien ordonné, (A, \trianglelefteq) l'est aussi d'après la proposition 22 page 48.

En particulier $\alpha = \text{type}(A, \trianglelefteq)$ et $f : \alpha \longrightarrow (A, \trianglelefteq)$ est l'isomorphisme associé.

Autrement dit $(\alpha, f) = \varphi(A, \trianglelefteq)$ et donc $(\alpha, f) \in \text{im}(\varphi)$.

Ainsi $\text{im}(\varphi) \supseteq \mathcal{F}$ et donc $\text{im}(\varphi) = \mathcal{F}$.

En particulier comme annoncé, \mathcal{F} est un ensemble.

- Construisons une relation d'ordre sur \mathcal{F} .

Pour tout (α, f) et (β, g) dans \mathcal{F} , on pose $(\alpha, f) \sqsubseteq (\beta, g) \iff (\alpha \leq \beta \text{ et } f = g|_{\alpha})$.

Montrons que \sqsubseteq est une relation d'ordre sur \mathcal{F} .

Réflexivité

Soit (α, f) dans \mathcal{F} .

Par réflexivité de \leq , on a $\alpha \leq \alpha$.

On a $f : \alpha \longrightarrow E$ donc en particulier $f|_{\alpha} = f$.

On a donc bien $(\alpha, f) \sqsubseteq (\alpha, f)$.

Donc \sqsubseteq est réflexive sur \mathcal{F} .

Antisymétrie

Soient (α, f) et (β, g) dans \mathcal{F} tels que $(\alpha, f) \sqsubseteq (\beta, g)$ et $(\beta, g) \sqsubseteq (\alpha, f)$.

On a donc $\alpha \leq \beta$ et $\beta \leq \alpha$ et $g|_\alpha = f$ et $f|_\beta = g$.

On a donc $\alpha = \beta$ par antisymétrie de \leq .

De même, pour une application, le fait d'être une restriction revient à être inclus.

Autrement dit $f \subseteq g$ et $g \subseteq f$, si bien que $f = g$ par antisymétrie de l'inclusion.

On a donc $(\alpha, f) = (\beta, g)$.

Donc \sqsubseteq est antisymétrique.

Transitivité

Soient (α, f) , (β, g) et (γ, h) dans \mathcal{F} tels que $(\alpha, f) \sqsubseteq (\beta, g)$ et $(\beta, g) \sqsubseteq (\gamma, h)$.

On a en particulier $\alpha \leq \beta \leq \gamma$ donc $\alpha \leq \gamma$ par transitivité de \leq .

De plus on a $g|_\alpha = f$ et $h|_\beta = g$ donc $h|_\alpha = (h|_\beta)|_\alpha = g|_\alpha = f$.

On a donc $(\alpha, f) \sqsubseteq (\gamma, h)$.

Donc \sqsubseteq est transitive.

Finalement \sqsubseteq est réflexive sur \mathcal{F} , antisymétrique et transitive.

Donc \sqsubseteq est une relation d'ordre sur \mathcal{F} .

- Montrons que \mathcal{F} muni de \sqsubseteq est inductif.

Soit \mathcal{C} une chaîne de \mathcal{F} .

Considérons alors $\mathcal{A} := \{\alpha \mid \exists f, (\alpha, f) \in \mathcal{C}\}$ et $\mathcal{B} := \{f \mid \exists \alpha, (\alpha, f) \in \mathcal{C}\}$.

Posons alors $\alpha^* := \bigcup \mathcal{A}$ et $f^* := \bigcup \mathcal{B}$.

\mathcal{A} est un ensemble d'ordinaux donc $\alpha^* = \sup(\mathcal{A})$ est un ordinal.

On a montré dans le précédent livre que f^* est une application si et seulement si les éléments de \mathcal{B} se recollent deux à deux.

Soient f et g dans \mathcal{B} .

En posant $\alpha := \text{dom}(f)$ et $\beta := \text{dom}(g)$, on a $(\alpha, f) \in \mathcal{C}$ et $(\beta, g) \in \mathcal{C}$.

Par définition \mathcal{C} est une chaîne donc tous ses éléments sont comparables.

On a donc $(\alpha, f) \sqsubseteq (\beta, g)$ ou $(\beta, g) \sqsubseteq (\alpha, f)$.

- ▶ Plaçons-nous dans le cas où $(\alpha, f) \sqsubseteq (\beta, g)$.

On a alors $\alpha \leq \beta$ et $f = g|_\alpha$.

En particulier $\text{dom}(f) \cap \text{dom}(g) = \alpha \cap \beta = \alpha$.

Donc pour tout $\gamma \in \text{dom}(f) \cap \text{dom}(g)$, $f(\gamma) = g|_\alpha(\gamma) = g(\gamma)$.

- ▶ Plaçons-nous dans le cas où $(\beta, g) \sqsubseteq (\alpha, f)$.

On montre de la même manière que $\forall \gamma \in \text{dom}(f) \cap \text{dom}(g), f(\gamma) = g(\gamma)$.

Dans les deux cas, f et g se recollent.

Donc tous les éléments de \mathcal{B} se recollent deux à deux.

Donc $f^* = \bigcup \mathcal{B}$ est une application d'après le premier livre.

Montrons que $\text{dom}(f^*) = \alpha^*$.

Pour tout x , on a les équivalences

$$\begin{aligned} x \in \text{dom}(f^*) &\iff \exists y, y = f^*(x) \text{ par définition du domaine} \\ &\iff \exists y, (x, y) \in f^* \\ &\iff \exists y, \exists f \in \mathcal{B}, (x, y) \in f \text{ car } f^* = \bigcup \mathcal{B} \\ &\iff \exists y, \exists f \in \mathcal{B}, y = f(x) \\ &\iff \exists f \in \mathcal{B}, \exists y, y = f(x) \\ &\iff \exists f \in \mathcal{B}, x \in \text{dom}(f) \text{ par définition du domaine} \\ &\iff \exists \alpha \in \mathcal{A}, x \in \alpha \text{ par définition de } \mathcal{A} \text{ et } \mathcal{B} \\ &\iff x \in \alpha^* \text{ car } \alpha^* = \bigcup \mathcal{A} \end{aligned}$$

On a donc $\text{dom}(f^*) = \alpha^*$.

Montrons que f^* est injective.

Soient $\gamma < \alpha^*$ et $\gamma' < \alpha^*$ tels que $f^*(\gamma) = f^*(\gamma')$.

Posons y cette image commune.

On a $y = f^*(\gamma)$ donc il existe $f \in \mathcal{B}$ tel que $y = f(\gamma)$.

De même $y = f^*(\gamma')$ donc il existe $f' \in \mathcal{B}$ tel que $y = f'(\gamma')$.

Posons $\alpha := \text{dom}(f)$ et $\alpha' := \text{dom}(f')$.

Par définition de \mathcal{B} on a $(\alpha, f) \in \mathcal{C}$ et $(\alpha', f') \in \mathcal{C}$.

Or par définition \mathcal{C} est une chaîne donc est totalement ordonné.

On a donc $(\alpha, f) \sqsubseteq (\alpha', f')$ ou $(\alpha', f') \sqsubseteq (\alpha, f)$.

► Plaçons-nous dans le cas où $(\alpha, f) \sqsubseteq (\alpha', f')$.

On a alors $\alpha \leq \alpha'$ et $f'_{|\alpha} = f$.

En particulier $f'(\gamma) = f'_{|\alpha}(\gamma) = f(\gamma) = y = f'(\gamma')$.

Or par définition $f' : \alpha \longrightarrow E$ est injective.

Donc $\gamma = \gamma'$.

► Plaçons-nous dans le cas $(\alpha', f') \sqsubseteq (\alpha, f)$.

On montre de la même manière que $\gamma = \gamma'$.

Dans les deux cas on a $\gamma = \gamma'$.

Donc f^* est injective.

Montrons que $f^* : \alpha^* \rightarrow E$, c'est-à-dire que $\text{im}(f^*) \subseteq E$.

Soit $y \in \text{im}(f^*)$.

Il existe donc $\gamma < \alpha^*$ tel que $y = f^*(\gamma)$.

Comme $f^* = \bigcup \mathcal{B}$, il existe $f \in \mathcal{B}$ tel que $y = f(\gamma)$.

Or par définition $f : \text{dom}(f) \rightarrow E$ donc $\text{im}(f) \subseteq E$ et donc $y \in E$.

Ainsi $\text{im}(f^*) \subseteq E$, et donc $f^* : \alpha^* \rightarrow E$.

Comme f^* est injective et α^* un ordinal, on a $(\alpha^*, f^*) \in \mathcal{C}$.

Montrons que (α^*, f^*) majore \mathcal{C} .

Soit $(\alpha, f) \in \mathcal{C}$.

On a en particulier $\text{dom}(f) = \alpha$ et $f : \alpha \rightarrow E$ injective.

Alors $\alpha \in \mathcal{A}$ et $f \in \mathcal{B}$ par définition de ces deux ensembles.

On a alors $\alpha \subseteq \bigcup \mathcal{A} = \alpha^*$ donc $\alpha \leq \alpha^*$ par définition de \leq .

De même $f \subseteq \bigcup \mathcal{B} = f^*$ donc en particulier $f|_{\alpha} = f$.

On a donc $(\alpha, f) \sqsubseteq (\alpha^*, f^*)$.

Ainsi (α^*, f^*) majore \mathcal{C} .

Ainsi toute chaîne de $(\mathcal{F}, \sqsubseteq)$ est majorée, donc $(\mathcal{F}, \sqsubseteq)$ est inductif.

D'après **le lemme de Zorn**, $(\mathcal{F}, \sqsubseteq)$ admet un élément maximal (α^*, f^*) .

Par définition $f^* : \alpha^* \rightarrow E$ est injective, et on a aussi $\text{im}(f^*) \subseteq E$.

Montrons que f^* est surjective dans E , c'est-à-dire que $\text{im}(f^*) = E$.

Supposons par l'absurde que $\text{im}(f^*) \subsetneq E$.

Il existe donc $y \in E \setminus \text{im}(f^*)$.

Posons alors $g := \begin{cases} \alpha^* + 1 & \longrightarrow E \\ \gamma & \longmapsto \begin{cases} f^*(\gamma) & \text{si } \gamma < \alpha^* \\ y & \text{si } \gamma = \alpha^* \end{cases} \end{cases}$.

Montrons que g est injective.

Soient $\gamma < \alpha^* + 1$ et $\gamma' < \alpha^* + 1$ tels que $g(\gamma) = g(\gamma')$.

► Plaçons-nous dans le cas où $\gamma < \alpha^*$ et $\gamma' < \alpha^*$.

On a donc $f^*(\gamma) = g(\gamma) = g(\gamma') = f^*(\gamma')$.

Or f^* est injective donc $\gamma = \gamma'$.

► Le cas où $\gamma = \alpha^* = \gamma'$ conduit immédiatement à $\gamma = \gamma'$.

► Plaçons-nous dans le cas où $\gamma < \alpha^*$ et $\gamma' = \alpha^*$.

On a alors $f(\gamma) = g(\gamma) = g(\gamma') = y$.

En particulier $y \in \text{im}(f)$, ce qui est impossible par définition de y .

► Le cas où $\gamma = \alpha^*$ et $\gamma' < \alpha^*$ est impossible pour la même raison.

Dans les deux cas possibles, on a donc $\gamma = \gamma'$.

Ainsi $g : \alpha^* + 1 \longrightarrow E$ est injective, et donc $(\alpha^* + 1, g) \in \mathcal{C}$.

Or $\alpha^* < \alpha^* + 1$ d'après la proposition 13 page 33.

On a aussi $g|_{\alpha^*} = f^*$ par définition de g .

On a donc $(\alpha^*, f^*) \sqsubset (\alpha^* + 1, g)$.

C'est absurde puisque (α^*, f^*) est un élément maximal de \mathcal{C} .

Par l'absurde on vient de montrer que $f^* : \alpha^* \longrightarrow E$ est surjective dans E .

Ainsi $f^* : \alpha^* \longrightarrow E$ est une bijection.

En particulier $(f^*)^{-1} : E \longrightarrow \alpha^*$ est une bijection.

• Construisons \trianglelefteq un ordre sur E .

Pour tout x et y dans E , posons $x \trianglelefteq y \iff (f^*)^{-1}(x) \leq (f^*)^{-1}(y)$.

D'après le précédent livre, \trianglelefteq est un ordre sur E et $(f^*)^{-1} : (E, \trianglelefteq) \longrightarrow \alpha^*$ est un isomorphisme d'ordres. En particulier $f^* : \alpha^* \longrightarrow (E, \trianglelefteq)$ est un isomorphisme d'ordres.

Or α^* est un ordinal donc est bien ordonné.

Donc (E, \trianglelefteq) est bien ordonné d'après la proposition 22 page 48.

En particulier E est bien ordonnable.

CQFD.

On peut ne pas passer par le lemme de Zorn et directement utiliser l'axiome du choix, mais le lecteur avisé pourrait bien remarquer que cette démonstration et celle du lemme de Zorn se ressemblent beaucoup : on est en fait en train de redémontrer Zorn dans le cas particulier du théorème de Zermelo.

Démonstration

Deuxième démonstration : en utilisant l'axiome du choix.

Soit E un ensemble : montrons qu'il est bien ordonnable.

Si E est vide, on le sait déjà que car $0 = \emptyset$ est un ordinal.

On suppose à présent que E est non vide.

D'après l'**axiome du choix**, il existe une application $f : \mathcal{P}(E) \setminus \{\emptyset\} \longrightarrow E$ telle que pour tout $A \in \mathcal{P}(E) \setminus \{\emptyset\}$, on a $f(A) \in A$.

E est un ensemble et la classe des ensembles U est propre d'après le paradoxe de Russell.

Autrement dit il existe un ensemble y tel que $y \notin E$.

On prolonge alors f en une application $g : \mathcal{P}(E) \rightarrow E \cup \{y\}$ de la façon suivante :

$$g := \begin{cases} \mathcal{P}(E) & \longrightarrow E \cup \{y\} \\ A & \longmapsto \begin{cases} f(A) & \text{si } A \neq \emptyset \\ y & \text{si } A = \emptyset \end{cases} \end{cases}$$

Définissons une assertion fonctionnelle $G : ON \rightarrow E \cup \{y\}$ par récursion de la façon suivante : pour tout ordinal α , on pose $G(\alpha) := g(\{x \in E \mid \forall \beta < \alpha, G(\beta) \neq x\})$.

D'après la proposition 109 page 243, G n'est pas injective.

Pour tout ordinal α , posons $A_\alpha := \{x \in E \mid \forall \beta < \alpha, G(\beta) \neq x\}$.

Ainsi pour tout ordinal α , on a $G(\alpha) = g(A_\alpha)$.

- Pour tout ordinal α , si $A_\alpha \neq \emptyset$ alors $\forall \beta < \alpha, G(\beta) \neq G(\alpha)$.

En effet, soit un ordinal α tel que $A_\alpha \neq \emptyset$.

On a alors $G(\alpha) = g(A_\alpha) = f(A_\alpha) \in A_\alpha$.

Autrement dit $G(\alpha) \in \{x \in E \mid \forall \beta < \alpha, G(\beta) \neq x\}$.

Donc $\forall \beta < \alpha, G(\beta) \neq G(\alpha)$.

Ainsi pour tout ordinal α , si $A_\alpha \neq \emptyset$ alors $\forall \beta < \alpha, G(\beta) \neq G(\alpha)$.

Notons (\star_1) ce fait.

- Montrons que $y \in \text{im}(G)$.

Supposons par l'absurde que $y \notin \text{im}(G)$.

Donc pour tout ordinal α , on a $G(\alpha) \in E$ et donc $g(A_\alpha) \in E$.

Or par définition de g , pour tout $A \subseteq E$ on a $g(A) \in E \iff A \neq \emptyset$.

Donc pour tout ordinal α , on a $A_\alpha \neq \emptyset$.

Donc pour tout ordinal $\alpha, \forall \beta < \alpha, G(\beta) \neq G(\alpha)$ d'après (\star_1) .

Montrons que G est injective.

Soient α et α' deux ordinaux tels que $G(\alpha) = G(\alpha')$.

Comme l'ordre est total chez les ordinaux, on a $\alpha = \alpha'$ ou $\alpha < \alpha'$ ou $\alpha' < \alpha$.

Si $\alpha < \alpha'$ alors $G(\alpha) \neq G(\alpha')$ d'après ce qui précède.

Si $\alpha' < \alpha$ alors $G(\alpha') \neq G(\alpha)$ d'après ce qui précède.

Ainsi le seul cas possible est $\alpha = \alpha'$.

Donc G est injective.

C'est absurde puisqu'on a justement dit que G n'est pas injective.

Par l'absurde on vient de montrer que $y \in \text{im}(G)$.

Il existe donc un ordinal α tel que $y = G(\alpha) = g(A_\alpha)$ et donc $A_\alpha = \emptyset$.

- Soit α_0 le plus petit ordinal tel que $A_\alpha = \emptyset$.

Soit $\gamma < \alpha_0$.

Par minimalité de α_0 , on a $A_\gamma \neq \emptyset$.

On a donc $\forall \beta < \gamma, G(\beta) \neq G(\gamma)$ d'après (\star_1) .

Ainsi $\forall \gamma < \alpha_0, \forall \beta < \gamma, G(\beta) \neq G(\gamma)$.

Notons (\star_2) ce fait.

- Considérons alors $h := G|_{\alpha_0}$, de sorte que $h : \alpha_0 \longrightarrow E \cup \{y\}$.

Pour tout $\gamma < \alpha_0$, on a $A_\gamma \neq \emptyset$ donc $h(\gamma) = G(\gamma) = g(A_\gamma) = f(A_\gamma) \in E$.

On a donc $h : \alpha_0 \longrightarrow E$.

Montrons que h est injective.

Soient γ et γ' tels que $h(\gamma) = h(\gamma')$.

Comme l'ordre est total chez les ordinaux, on a $\gamma = \gamma'$ ou $\gamma < \gamma'$ ou $\gamma' < \gamma$.

Si $\gamma < \gamma'$ alors $h(\gamma) = G(\gamma) \neq G(\gamma') = h(\gamma')$ d'après (\star_2) .

Si $\gamma' < \gamma$ alors $h(\gamma') = G(\gamma') \neq G(\gamma) = h(\gamma)$ d'après (\star_2) .

Dans ces deux cas on a $h(\gamma) \neq h(\gamma')$ et donc le seul cas possible est $\gamma = \gamma'$.

Donc h est injective.

- Montrons que h est surjective dans E .

Par définition de α_0 , on a $A_{\alpha_0} = \emptyset$.

Autrement dit $\{x \in E \mid \forall \beta < \alpha_0, G(\beta) \neq x\} = \emptyset$.

Donc pour tout $x \in E$, il existe $\beta < \alpha_0$ tel que $G(\beta) = x$.

Comme $h = G|_{\alpha_0}$, pour tout $x \in E$, il existe $\beta < \alpha_0$ tel que $h(\beta) = x$.

Donc h est surjective dans E .

- Ainsi $h : \alpha_0 \longrightarrow E$ est bijective.

En particulier $h^{-1} : E \longrightarrow \alpha_0$ est bijective.

Pour tout x et x' dans E , posons $x \trianglelefteq x' \iff h^{-1}(x) \leq h^{-1}(x')$.

On a vu dans le précédent livre que \trianglelefteq est une relation d'ordre sur E .

On a aussi vu que $h^{-1} : (E, \trianglelefteq) \longrightarrow \alpha_0$ est un isomorphisme d'ordres.

En particulier $h : \alpha_0 \longrightarrow (E, \trianglelefteq)$ est un isomorphisme d'ordres.

Or α_0 est un ordinal donc est bien ordonné.

Donc (E, \trianglelefteq) est bien ordonné d'après la proposition 22 page 48.

En particulier E est bien ordonnable.

CQFD.

Le théorème de Zermelo est équivalent à l'axiome du choix. Autrement dit, dressons la liste de tous les axiomes que l'on a considérés jusque là, et retirons-en l'axiome du choix. À la place, mettons-y le théorème de Zermelo. De ce nouveau système d'axiome, il est alors possible de

déduire l'axiome du choix, qui en devient donc un théorème.

Idée de preuve

Supposons le théorème de Zermelo, c'est-à-dire que tout ensemble est bien ordonnable.

Soit E un ensemble.

D'après le **théorème de Zermelo**, E est bien ordonnable.

Soit \trianglelefteq un bon ordre sur E .

En particulier, toute partie non vide de E admet un minimum pour \trianglelefteq .

Posons donc $f := \begin{pmatrix} \mathcal{P}(E) \setminus \{\emptyset\} & \longrightarrow & E \\ A & \longmapsto & \min(A) \end{pmatrix}$.

Alors pour tout $A \in \mathcal{P}(E) \setminus \{\emptyset\}$, $f(A) = \min(A) \in A$.

CQFD.

Pour la petite histoire



Ernst Zermelo (27 juillet 1871 – 21 mai 1953) est un mathématicien allemand. Il s'est principalement intéressé aux fondations des mathématiques et à la philosophie. Il a donné des développements importants à la théorie des ensembles et est un des précurseurs de la théorie des jeux. Sa thèse de doctorat en 1894 porte sur le calcul des variations : il aura travaillé une partie de sa vie sur la mécanique.

En 1883, Cantor avait affirmé sans démonstration que tout ensemble est bien ordonnable. Zermelo envoie la démonstration de son théorème à Hilbert en 1904, et utilise justement l'axiome du choix pour cela. Il est d'ailleurs le premier à formuler explicitement l'axiome du choix.

Zermelo a formulé en 1908 la plupart des axiomes utilisés aujourd'hui en théorie des ensembles, notamment ceux décrits dans le précédent livre (et l'axiome de l'infini décrit dans le premier chapitre). Plus tard, Fraenkel et Skolem compléteront ses travaux pour donner naissance à l'axiomatique de Zermelo-Fraenkel, abrégée ZF. En ajoutant à celle-ci l'axiome du choix, on obtient l'axiomatique de ZFC.

Nous avons établi que l'axiome du choix implique le lemme de Zorn qui implique le théorème de Zermelo, qui lui-même implique l'axiome du choix. Ces trois énoncés sont donc équivalents.

Le mathématicien Jerry Bona disait en plaisantant à ce propos : « *L'axiome du choix est trivialement vrai, le théorème de Zermelo est trivialement faux, et que pouvons-nous dire du lemme de Zorn ?* ».

Le fait que tout ensemble est ordonnable nous assure que tout ensemble est équivalent à un ordinal d'après la proposition 102 page 237. Or nous le savons depuis le premier chapitre (théorème 1 page 21) : tous les ordinaux sont comparables, si bien que tous les ensembles sont comparables.

Théorème 16 (Tous les ensembles sont comparables)

Soient E et F deux ensembles.

On a alors $E \preccurlyeq F$ ou $F \preccurlyeq E$.

Démonstration

D'après le **théorème de Zermelo**, E et F sont bien ordonnables.

Ils admettent donc tous deux un cardinal d'après la proposition 103 page 237.

Les cardinaux sont des ordinaux donc sont comparables.

On a donc $\text{card}(E) \leq \text{card}(F)$ ou $\text{card}(F) \leq \text{card}(E)$.

On a donc $\text{card}(E) \preccurlyeq \text{card}(F)$ ou $\text{card}(F) \preccurlyeq \text{card}(E)$ d'après la proposition 99 page 230.

► Plaçons-nous dans le cas où $\text{card}(E) \preccurlyeq \text{card}(F)$.

Par définition du cardinal on a $E \approx \text{card}(E)$ et $F \approx \text{card}(F)$.

Ainsi on a $E \approx \text{card}(E) \preccurlyeq \text{card}(F) \approx F$.

On a donc $E \preccurlyeq F$ d'après la proposition 88 page 209.

► Plaçons-nous dans le cas où $\text{card}(F) \preccurlyeq \text{card}(E)$.

On montre de la même manière qu'alors $F \preccurlyeq E$.

Ainsi on a donc $[E \preccurlyeq F \text{ ou } F \preccurlyeq E]$.

CQFD.

Ainsi nous avons utilisé le théorème de Zermelo pour montrer ce résultat, et donc l'axiome du choix. Le fait que tous les ensembles sont comparables est en fait équivalent à l'axiome du choix : voici à nouveau un résultat que l'on aurait pu prendre comme axiome à la place de l'axiome du choix. Cependant pour voir qu'il implique l'axiome du choix, nous avons besoin d'un résultat dû à Hartogs, que nous allons voir à présent.

3.2 Théorème et cardinal de Hartogs

Le théorème d'Hartogs qui suit nous dit que pour un ensemble E donné, il existe un cardinal κ tel que κ ne s'injecte pas dans E , c'est-à-dire $\kappa \not\preccurlyeq E$. Ce résultat, d'après le théorème 16 page 256 (donc d'après l'axiome du choix), revient à dire que $E \prec \kappa$. Mieux, d'après le théorème de Cantor on a $E \prec \mathcal{P}(E)$, et avec le théorème de Zermelo (donc avec l'axiome du choix),

$\mathcal{P}(E)$ admet un cardinal κ , qui vérifie donc $E \prec \kappa$, ce qui rend la démonstration du théorème immédiate.

Cependant, on introduit ici ce théorème pour démontrer que le théorème 16 page 256 implique l'axiome du choix, donc on ne peut pas se servir de celui-ci.

Théorème 17 (de Hartogs)

Soit E un ensemble.

Il existe un cardinal κ tel qu'il n'existe pas d'injection $\kappa \rightarrow E$.

Démonstration

- Construisons κ .

Posons $\kappa := \{\text{type}(A, \leq) \mid A \subseteq E \text{ et } \leq \text{ est un bon ordre sur } A\}$.

Montrons que κ est un ordinal.

Par définition, κ est un ensemble d'ordinaux.

En particulier (κ, \in) est strictement bien ordonné d'après le théorème 1 page 21.

Montrons que κ est transitif.

Soit $\alpha \in \kappa$.

Par définition il existe $A \subseteq E$ et \leq un bon ordre sur A tel que $\alpha := \text{type}(A, \leq)$.

Soit $f : \alpha \rightarrow A$ l'isomorphisme d'ordres associé.

Soit $\beta \in \alpha$.

Comme f est injective, $f|_{\beta}$ est injective.

Posons alors $B := \text{im}(f|_{\beta})$, de sorte que $f|_{\beta} : \beta \rightarrow B$ est bijective.

Comme f est croissante, $f|_{\beta}$ est croissante.

$\text{dom}(f|_{\beta}) = \beta$ est un ordinal donc est bien ordonné donc totalement ordonné.

On en conclut que $f|_{\beta} : \beta \rightarrow B$ est un isomorphisme d'ordres.

En particulier $\beta = \text{type}(B, \leq)$.

Or $B \subseteq A \subseteq E$, et donc $\beta \in \kappa$.

Donc $\forall \beta \in \alpha, \beta \in \kappa$ et donc $\alpha \subseteq \kappa$.

Donc $\forall \alpha \in \kappa, \alpha \subseteq \kappa$, si bien que κ est inductif.

On en conclut donc que κ est un ordinal.

- Montrons que κ ne s'injecte pas dans E .

Supposons par l'absurde qu'il existe une injection $g : \kappa \rightarrow E$.

Posons $A := \text{im}(g)$, de sorte que $g : \kappa \rightarrow A$ est une bijection.

En particulier $g^{-1} : A \rightarrow \kappa$ est une bijection.

Pour tout a et a' dans A , on pose $a \trianglelefteq a' \iff g^{-1}(a) \leq g^{-1}(a')$.

On a vu dans le premier livre qu'alors \trianglelefteq est une relation d'ordre sur A .

On a aussi vu qu'alors $g^{-1} : A \longrightarrow \kappa$ est un isomorphisme d'ordres.

Donc $g : \kappa \longrightarrow A$ est un isomorphisme d'ordres.

Or κ est un ordinal donc est bien ordonné.

Donc (A, \leq) est bien ordonné d'après la proposition 22 page 48.

En particulier $\kappa = \text{type}(A, \leq)$.

Comme $A \subseteq E$, on en déduit que $\kappa \in \kappa$ par définition de κ .

C'est absurde car \in est antiréflexive sur les ordinaux.

Par l'absurde, on vient de montrer que $\boxed{\kappa \text{ ne s'injecte pas dans } E}$.

- Montrons que κ est un cardinal.

Soit α un ordinal tel que $\alpha < \kappa$.

On a donc $\alpha \in \kappa$ par définition de $<$.

Il existe donc $A \subseteq E$ et \leq un bon ordre sur A tel que $\alpha = \text{type}(A, \leq)$.

Soit $f : \alpha \longrightarrow A$ l'isomorphisme associé, de sorte que $f : \alpha \longrightarrow E$ est une injection.

Supposons par l'absurde que $\alpha \approx \kappa$.

On a en particulier $\kappa \preccurlyeq \alpha$ d'après la proposition 86 page 202.

Il existe donc une injection $g : \kappa \longrightarrow \alpha$.

Alors $g \circ f : \kappa \longrightarrow E$ est une injection.

C'est absurde puisqu'on a montré que κ ne s'injecte pas dans E .

Par l'absurde, on vient de montrer que $\alpha \not\approx \kappa$.

Or $\alpha < \kappa$ donc $\alpha \leq \kappa$ et donc $\alpha \preccurlyeq \kappa$ d'après la proposition 99 page 230.

Ainsi $\alpha \preccurlyeq \kappa$ et $\alpha \not\approx \kappa$ donc $\alpha \prec \kappa$.

Ainsi $\forall \alpha < \kappa, \alpha \prec \kappa$, et donc $\boxed{\kappa \text{ est un cardinal}}$.

CQFD.

Pour la petite histoire



Friedrich Moritz Hartogs, dit Fritz Hartogs (20 mai 1874 – 18 août 1943), est un mathématicien allemand connu pour ses importantes contributions à la théorie des fonctions de plusieurs variables complexes. C'est à lui que l'on doit le théorème qui précède, qu'il a démontré en 1915. Notons bien que cette démonstration ne nécessite pas l'axiome du choix !

On peut désormais montrer que le fait que tous les ensembles sont comparables implique l'axiome du choix, plus précisément le théorème de Zermelo.

Idee de preuve

Supposons que tous les ensembles soient comparables.

Soit E un ensemble.

D'après le théorème d'Hartogs, il existe un cardinal κ tel qui ne s'injecte pas dans E .

Or par hypothèse E et κ sont comparables.

La seule possibilité est donc que $E \preccurlyeq \kappa$.

Il existe donc $\alpha \leq \kappa$ tel que $E \approx \alpha$ d'après la proposition 99 page 230.

En particulier E est bien ordonnable d'après la proposition 102 page 237.

CQFD.

Revenons dans notre cadre usuel : l'axiome du choix et toutes ses dérivées sont admis. D'après le théorème de Hartogs, pour un ensemble E il existe au moins un cardinal κ tel que $E \prec \kappa$. Le plus petit de tels cardinaux est alors appelé **cardinal de Hartogs** de E . Son existence est encore une fois assurée par le fait qu'une classe d'ordinaux admet un plus petit élément. C'est ainsi le cardinal qui suit directement celui de E , son successeur en somme (successeur au sens des cardinaux, pas des ordinaux !).

Définition 37 (Cardinal de Hartogs)

Soit E un ensemble.

On appelle **cardinal de Hartogs** de E le plus petit cardinal κ tel que $E \prec \kappa$.

On le note $\aleph(E)$.

Exemple :

Pour n un entier naturel, on a $\aleph(n) = n + 1$. En effet, $n + 1$ est aussi un entier naturel donc est un cardinal d'après le théorème 13 page 233. C'est aussi le plus petit ordinal strictement plus grand que n d'après la proposition 13 page 33.

Proposition 110 (Cardinal de Hartogs d'un ordinal)

Soit α un ordinal.

Alors $\alpha < \aleph(\alpha)$.

Démonstration

Comme tous les ordinaux sont comparables, on a $\alpha < \aleph(\alpha)$ ou $\aleph(\alpha) \leq \alpha$.

Supposons par l'absurde que $\aleph(\alpha) \leq \alpha$.

On alors $\aleph(\alpha) \preccurlyeq \alpha$ d'après la proposition 99 page 230.

Or par définition du cardinal de Hartogs, on a $\alpha \prec \aleph(\alpha)$.

On a donc $\alpha \preccurlyeq \aleph(\alpha)$ et $\alpha \not\approx \aleph(\alpha)$.

En particulier on a $\aleph(\alpha) \preccurlyeq \alpha \preccurlyeq \aleph(\alpha)$.

On a donc $\alpha \approx \aleph(\alpha)$ d'après le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on vient justement de dire que $\alpha \not\approx \aleph(\alpha)$.

Par l'absurde, on vient de montrer que $\alpha < \aleph(\alpha)$.

CQFD.

Définition 38 (Fonction de Hartogs)

On définit l'assertion fonctionnelle $\aleph : ON \longrightarrow ON$ suivante par récursion :

$$\left\{ \begin{array}{l} \aleph_0 := \omega \\ \aleph_{\alpha+1} := \aleph(\aleph_\alpha) \text{ pour tout ordinal } \alpha \\ \aleph_\gamma := \sup_{\delta < \gamma} \aleph_\delta \text{ pour tout ordinal limite non nul } \gamma \end{array} \right.$$

On pose aussi $\omega_\alpha := \aleph_\alpha$ pour tout ordinal α , les deux notations étant tout autant employées.

Remarque :

1. Ainsi $\omega_0 = \omega$ et ω_1 est le cardinal qui vient juste après ω . Nous verrons plus tard que ω_1 joue un rôle particulier avec l'ensemble des nombres réels \mathbb{R} , et plus généralement avec tous les ensembles dit indénombrables.
2. Pour justifier rigoureusement que cette définition a du sens, on utilise encore et toujours la proposition 36 page 91, avec $\mu_0 := \omega$ et $G(\xi) := \aleph(\xi)$ pour tout ordinal ξ .

Le fait que tout ensemble admette un cardinal strictement plus grand (c'est le théorème d'Hartogs) implique qu'il n'est pas possible de considérer l'ensemble de tous les cardinaux : la classe des cardinaux est donc une classe propre. En particulier la classe C_∞ des cardinaux **infinis** est propre. D'après les propositions 79 page 186 et 80 page 187, il existe donc un unique isomorphisme $ON \longrightarrow C_\infty$. Il s'avère que c'est justement \aleph .

Théorème 18 (La fonction d'Hartogs est un isomorphisme)

Soit C_∞ la classe des cardinaux **infinis**.

Alors \aleph est l'unique isomorphisme d'ordres de ON vers C_∞ .

De plus \aleph est continue.

Démonstration

- Montrons que \aleph est strictement croissante.

Fixons α un ordinal.

Pour tout ordinal β , posons $P(\beta)$ l'assertion « $\alpha < \beta \Rightarrow \aleph_\alpha < \aleph_\beta$ ».

Initialisation

La prémissse $\alpha < 0$ étant fausse, l'implication $\alpha < 0 \Rightarrow \aleph_\alpha < \aleph_0$ est vraie.

Autrement dit, on a $P(0)$.

Héritéité

Soit β un ordinal tel que $P(\beta)$.

Par définition, on a $\aleph_{\beta+1} = \aleph(\aleph_\beta)$.

Or on a $\aleph_\beta < \aleph(\aleph_\beta)$ d'après la proposition 110 page 259, donc $\aleph_\beta < \aleph_{\beta+1}$.

Supposons que $\alpha < \beta + 1$.

On a donc $\alpha \leq \beta$ d'après la proposition 13 page 33.

On a donc $\alpha < \beta$ ou $\alpha = \beta$.

► Plaçons-nous dans le cas où $\alpha < \beta$.

On a alors $\aleph_\alpha < \aleph_\beta$ d'après $P(\beta)$.

Or on a dit que $\aleph_\beta < \aleph_{\beta+1}$.

On a donc $\aleph_\alpha < \aleph_{\beta+1}$ par transitivité de $<$.

► Plaçons-nous dans le cas où $\alpha = \beta$.

On a alors $\aleph_\alpha = \aleph_\beta$.

Or on a dit que $\aleph_\beta < \aleph_{\beta+1}$.

On a donc $\aleph_\alpha < \aleph_{\beta+1}$.

Dans les deux cas, on a $\aleph_\alpha < \aleph_{\beta+1}$.

Ainsi, si $\alpha < \beta + 1$ alors $\aleph_\alpha < \aleph_{\beta+1}$, et donc $P(\beta + 1)$.

Ainsi pour tout ordinal β , si $P(\beta)$ alors $P(\beta + 1)$.

Héritéité limite

Soit γ un ordinal limite non nul tel que $\forall \delta < \gamma, P(\delta)$.

Supposons que $\alpha < \gamma$.

On a $\gamma = \sup(\gamma) = \sup_{\delta \in \gamma} \delta = \sup_{\delta < \gamma} \delta$ d'après la proposition 21 page 47.

Il existe donc $\delta < \gamma$ tel que $\delta \not\leq \alpha$ par minimalité de la borne supérieure.

Comme les ordinaux sont totalement ordonnés, on a donc $\alpha < \delta$.

On en déduit que $\aleph_\alpha < \aleph_\delta$ d'après $P(\delta)$.

Or par définition $\aleph_\gamma = \sup_{\varepsilon < \gamma} \aleph_\varepsilon$.

Puisque $\delta < \gamma$, on a $\aleph_\delta \leq \aleph_\gamma$ car la borne supérieure est un majorant.

On a donc $\aleph_\alpha < \aleph_\gamma$ par transitivité.

Ainsi, si $\alpha < \gamma$ alors $\aleph_\alpha < \aleph_\gamma$, et donc $P(\gamma)$.

Donc pour tout ordinal limite non nul, si $\forall \delta < \gamma, P(\delta)$ alors $P(\gamma)$.

Ainsi \aleph vérifie les trois conditions du principe faible d'induction.

Donc pour tout ordinal β , on a $P(\beta)$.

Autrement dit pour tout ordinal β , si $\alpha < \beta$ alors $\aleph_\alpha < \aleph_\beta$.

Ainsi \aleph est strictement croissante.

- Montrons que \aleph est continue.

Comme \aleph est strictement croissante, \aleph est croissante d'après la proposition 76 page 182.

Par définition de \aleph , pour tout ordinal limite non nul γ , on a $\aleph_\gamma = \sup_{\delta < \gamma} \aleph_\delta$.

Donc \aleph est continue d'après la proposition 42 page 103.

Autrement dit, pour tout ensemble d'ordinaux non vide X , on a $\aleph_{\sup(X)} = \sup(\aleph_X)$.

- Montrons que \aleph est à valeurs dans C_∞ la classe des cardinaux infinis.

Soit α un ordinal.

On a $0 \leq \alpha$ donc $\aleph_0 \leq \aleph_\alpha$ par croissance de \aleph .

Or par définition $\aleph_0 = \omega$, donc $\omega \leq \aleph_\alpha$.

Or chez les ordinaux, « être fini » veut dire « être strictement inférieur à ω ».

Donc \aleph_α est infini, et est évidemment un cardinal par définition de \aleph .

Ainsi pour tout ordinal α , \aleph_α est un cardinal infini donc $\aleph_\alpha \in C_\infty$.

- Montrons que \aleph est surjective dans C_∞ .

Soit κ un cardinal infini.

Considérons alors la classe $A := \{\alpha \in ON \mid \aleph_\alpha \leq \kappa\}$.

On a dit que \aleph est strictement croissante.

En particulier \aleph est non bornée d'après la proposition 73 page 174.

Il existe donc $\beta \in ON$ tel que $\kappa < \aleph_\beta$.

Supposons par l'absurde que β ne majore pas A .

Il existe donc $\alpha \in A$ tel que $\alpha \not\leq \beta$.

Comme les ordinaux sont totalement ordonnés, on a donc $\beta < \alpha$.

Alors $\aleph_\beta < \aleph_\alpha$ par stricte croissance de \aleph .

En particulier $\kappa < \aleph_\alpha$ par transitivité de $<$.

C'est absurde puisque $\alpha \in A$.

Par l'absurde, on vient de montrer que β majore A .

Ainsi A est majorée donc A est un ensemble d'après la proposition 12 page 32.

En particulier $\aleph_A := \aleph^\frown(A)$ est un ensemble d'après l'axiome de remplacement.

On en déduit que A et \aleph_A admettent une borne supérieure car ensembles d'ordinaux.

Posons alors $\alpha^* := \sup(A)$.

Par définition κ est un cardinal infini donc $\omega \leq \kappa$.

Or $\aleph_0 = \omega$ donc $\aleph_0 \leq \kappa$ et donc $0 \in A$ par définition de A .

Ainsi A est non vide, donc $\sup(\aleph_A) = \aleph_{\sup(A)}$ par continuité de \aleph .

On a donc $\sup(\aleph_A) = \aleph_{\alpha^*}$.

Or $\forall \alpha \in A, \aleph_\alpha \leq \kappa$ par définition de A .

Donc $\sup(\aleph_A) \leq \kappa$ par minimalité de la borne supérieure, et donc $\aleph_{\alpha^*} \leq \kappa$.

On a $\alpha^* < \alpha^* + 1$ d'après la proposition 13 page 33.

Alors $\alpha^* + 1 \notin A$ car la borne supérieure est un majorant.

On a donc $\kappa < \aleph_{\alpha^* + 1}$ par définition de A .

Ainsi on a $\aleph_{\alpha^*} \leq \kappa < \aleph_{\alpha^* + 1}$, c'est-à-dire $\aleph_{\alpha^*} \leq \kappa < \aleph(\aleph_{\alpha^*})$ par définition de \aleph .

On a donc $\aleph_{\alpha^*} = \kappa$ par minimalité de $\aleph(\aleph_{\alpha^*})$.

En particulier $\kappa \in \text{im}(\aleph)$.

Ainsi \aleph est surjective dans C_∞ .

Ainsi $\aleph : ON \longrightarrow C_\infty$ est strictement croissante et surjective dans C_∞ .

Donc \aleph est un isomorphisme d'ordres de ON dans C_∞ d'après la proposition 78 page 185.

CQFD.

Remarque :

En particulier on peut appliquer le théorème 10 page 192 : celui-ci nous dit que la classe $\text{fix}(\aleph)$ des points fixes de \aleph est propre. Elle est en particulier non vide ! Ceci signifie qu'il existe des ordinaux ξ tels que $\aleph_\xi = \xi$.

Considérons un ensemble **totalement** ordonné (E, \leq) et \lhd sont ordre strict associé. On a vu dans le précédent livre qu'alors nécessairement pour tout x et y dans E , on a

$$x \leq y \text{ ou } y \lhd x$$

ainsi que l'équivalence qui en découle

$$x \not\leq y \iff y \lhd x$$

Or on a vu grâce à l'axiome du choix que tous les ensembles sont comparables vis à vis de la subpotence. Nous avons donc naturellement le même phénomène avec la subpotence.

Proposition 111 (Subpotence strict et comparabilité totale)

Soient E et F deux ensembles.

1. On a $E \preccurlyeq F$ ou $F \prec E$.
2. On a l'équivalence $E \not\preccurlyeq F \iff F \prec E$.



Démonstration

1. D'après le théorème 16 page 256, on a $E \preccurlyeq F$ ou $F \preccurlyeq E$.

Supposons que $E \not\preccurlyeq F$.

On a donc $F \preccurlyeq E$ par ce qui précède.

Or on a $E \approx F \implies E \preccurlyeq F$ d'après la proposition 86 page 202.

On a donc $E \not\preccurlyeq F \implies E \not\approx F$ par contraposition.

On a donc $E \not\approx F$ par modus ponens.

Comme $F \preccurlyeq E$, on a donc $F \prec E$ par définition de \prec .

On a donc l'implication $E \not\preccurlyeq F \Rightarrow F \prec E$.

Autrement dit on a $(\text{non}(E \not\preccurlyeq F) \text{ ou } F \prec E)$, c'est-à-dire $(E \preccurlyeq F \text{ ou } F \prec E)$.

2. On a montré juste avant que l'on a l'implication $E \not\preccurlyeq F \Rightarrow F \prec E$.



Supposons que l'on a $F \prec E$.

On a donc $F \preccurlyeq E$ et $F \not\approx E$ par définition de \prec .

Supposons par l'absurde que l'on a $E \preccurlyeq F$.

On a donc $E \approx F$ d'après le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on a justement dit que $E \not\approx F$.

Par l'absurde, on vient de montrer que $E \not\preccurlyeq F$.

On a donc l'implication $F \prec E \Rightarrow E \not\preccurlyeq F$.

Finalement, on a l'équivalence $E \not\preccurlyeq F \iff F \prec E$.

CQFD.

4 Opérations sur les cardinaux

À la manière des ordinaux, il est possible d'effectuer les trois opérations sur les cardinaux : l'addition, la multiplication et l'exponentiation. Attention, les cardinaux étant des cas particuliers d'ordinaux, il est déjà possible de faire ces opérations.

L'idée est ici de donner d'autres définitions à ces opérations spécifiquement dans le cadre des cardinaux : cela ne donnera pas nécessairement le même résultat. Par exemple on a vu que ω est un cardinal : $\omega + \omega$ en tant qu'addition ordinaire nous l'avons vu donne $\omega \cdot 2$, alors que nous le verrons $\omega + \omega$ en tant qu'addition cardinale va juste donner ω .

Pourquoi définir de nouvelles opérations ? Rappelons-nous que l'idée derrière la théorie des cardinaux est de représenter les classes d'équipotence, et donc c'est à cet égard que ces nouvelles opérations seront définies. Dans toute la suite, pour α et β deux ordinaux, on notera :

- ▶ $\overset{\mathcal{O}}{\alpha + \beta}$ pour signifier l'addition ordinale de α par β .
- ▶ $\overset{\mathcal{O}}{\alpha \cdot \beta}$ pour signifier la multiplication ordinale de α par β .
- ▶ $\alpha^{\mathcal{O}\beta}$ pour signifier l'exponentiation ordinale de α par β .

Tentons de justifier sommairement les définitions suivantes, avec le cas d'ensembles finis. On ne fait pas un raisonnement rigoureux, on développe simplement notre intuition. Prenons $E := \{a, b\}$ et $F := \{a, c, d\}$. Ils ont respectivement 2 et 3 éléments.

- ▶ Comment former naturellement un ensemble ayant $2 + 3 = 5$ éléments ? Leur réunion $E \cup F = \{a, b, c, d\}$ n'en a que 4. Il faut en fait considérer leur union disjointe $E \amalg F$. En effet, on a $E \amalg F = \{(0, a), (0, b), (1, a), (1, c), (1, d)\}$, qui a bien 5 élément, car on note avec 0 ou 1 la provenance de chacun d'entre eux. Pour définir la somme de deux cardinaux, on va donc utiliser leur union disjointe.
- ▶ Comment former naturellement un ensemble ayant $2 \cdot 3 = 6$ éléments ? Leur produit cartésien est $E \times F = \{(a, a), (a, c), (a, d), (b, a), (b, c), (b, d)\}$, qui a bien 6 éléments. Pour définir le produit de deux cardinaux, on va donc utiliser leur produit cartésien.
- ▶ Comment former naturellement un ensemble ayant $2^3 = 8$ éléments ?
L'ensemble $\mathcal{F}(F \rightarrow E)$ des applications de F dans E y répond : en effet, pour $a \in F$ en tant qu'antécédent, on a 2 choix possibles d'images dans E , et de même pour c et d . Il y a donc $2 \cdot 2 \cdot 2$ choix d'applications possibles. Pour définir l'exponentiation de deux cardinaux, on va donc utiliser les ensembles d'applications.

Définition 39 (Opérations sur les cardinaux)

Soient κ et λ deux cardinaux.

On pose

1. $\kappa + \lambda := \text{card}(\kappa \amalg \lambda)$
2. $\kappa \cdot \lambda := \text{card}(\kappa \times \lambda)$ où \times est le produit cartésien.
3. $\kappa^\lambda = \text{card}(\mathcal{F}(\lambda \rightarrow \kappa))$

Nous l'avons dit, les opérations cardinales que nous venons de définir n'ont a priori pas de raison de coïncider avec les opérations ordinaires abordées au chapitre 2. Heureusement, dans le cas des entiers naturels, tout va bien !

Proposition 112 (Coïncidence chez les entiers naturels)

Soient n et m deux entiers naturels.

1. On a $n + m = n \overset{\sigma}{+} m$
2. On a $n \cdot m = n \overset{\sigma}{\cdot} m$
3. On a $n^m = n^{\sigma m}$

Autrement dit, les opérations cardinales et ordinaires coïncident chez les entiers naturels.

Démonstration

1. D'après le théorème 7 page 120, on a $n \overset{\sigma}{+} m = \text{type}(n \amalg m)$.

Or $n \overset{\sigma}{+} m$ est un entier naturel d'après la proposition 39 page 96.

Donc $n \overset{\sigma}{+} m$ est un cardinal d'après le théorème 13 page 233.

Donc $\text{type}(n \amalg m)$ est un cardinal.

Or $n \amalg m$ et $\text{type}(n \amalg m)$ sont équivalents d'après la proposition 105 page 239.

Donc $\text{card}(n \amalg m) = \text{type}(n \amalg m)$ et donc $\text{card}(n \overset{\sigma}{+} m) = n \overset{\sigma}{+} m$.

Or par définition $n + m = \text{card}(n \amalg m)$, si bien que $n + m = n \overset{\sigma}{+} m$.

2. D'après le théorème 8 page 143, on a $n \overset{\sigma}{\cdot} m = \text{type}(n \times m)$.

Or $n \overset{\sigma}{\cdot} m$ est un entier naturel d'après la proposition 54 page 129 .

Donc $n \overset{\sigma}{\cdot} m$ est un cardinal d'après le théorème 13 page 233.

Donc $\text{type}(n \times m)$ est un cardinal.

Or $n \times m$ et $\text{type}(n \times m)$ sont équivalents d'après la proposition 105 page 239.

Donc $\text{card}(n \times m) = \text{type}(n \times m)$ et donc $\text{card}(n \overset{\sigma}{\cdot} m) = n \overset{\sigma}{\cdot} m$.

Or par définition $n \cdot m = \text{card}(n \times m)$, si bien que $n \cdot m = n \overset{\sigma}{\cdot} m$.

3. D'après le théorème 9 page 168, on a $n^{\sigma m} = \text{type}(\text{sf}(m \rightarrow n))$.

Montrons que $\text{sf}(m \rightarrow n) = \mathcal{F}(m \rightarrow n)$.

Par définition, on sait déjà que $\text{sf}(m \rightarrow n) \subseteq \mathcal{F}(m \rightarrow n)$.

Soit $f : m \rightarrow n$.

Posons $S := \text{supp}(f)$ et montrons que S est de type fini.

On a $S \subseteq m$ donc $\text{type}(S) \leq \text{type}(m)$ d'après la proposition 27 page 60.

Or m est un ordinal donc $\text{type}(m) = m$ et donc $\text{type}(S) \leq m$.

Or m est un entier naturel.

Donc $\text{type}(S)$ est un entier naturel d'après la proposition 15 page 38.

Ainsi $\text{type}(S)$ est fini donc f est à support fini, donc $f \in \text{sf}(m \rightarrow n)$.

On a donc $\text{sf}(m \rightarrow n) \supseteq \mathcal{F}(m \rightarrow n)$ et donc $\text{sf}(m \rightarrow n) = \mathcal{F}(m \rightarrow n)$.

En particulier on a $n^{\mathcal{O}m} = \text{type}(\mathcal{F}(m \rightarrow n))$.

Or $n^{\mathcal{O}m}$ est un entier naturel d'après la proposition 65 page 150.

Donc $n^{\mathcal{O}m}$ est un cardinal d'après le théorème 13 page 233.

Donc $\text{type}(\mathcal{F}(m \rightarrow n))$ est un cardinal.

Or $\mathcal{F}(m \rightarrow n)$ et $\text{type}(\mathcal{F}(m \rightarrow n))$ sont équivalents d'après la proposition 105 page 239.

Donc $\text{card}(\mathcal{F}(m \rightarrow n)) = \text{type}(\mathcal{F}(m \rightarrow n))$ et donc $\text{card}(\mathcal{F}(m \rightarrow n)) = n^{\mathcal{O}m}$.

Or par définition $n^m = \text{card}(\mathcal{F}(m \rightarrow n))$, si bien que $n^m = n^{\mathcal{O}m}$.

CQFD.

Comme nous l'avons dit, dès qu'un des deux cardinaux est infinis, il n'y a plus de raison que ça coïncide. Prenons l'exemple de l'exponentiation. Rappelons-nous que pour tout entier naturel n , on a $2^{\mathcal{O}n} \in \mathbb{N}$ d'après la proposition 65 page 150. Autrement dit

$$\forall n < \omega, 2^{\mathcal{O}n} < \omega$$

Par minimalité de la borne supérieure, on a donc

$$\sup_{n < \omega} 2^{\mathcal{O}n} \leq \omega$$

Or ω est un ordinal limite non nul, donc par définition de l'exponentiation ordinaire on a

$$2^{\mathcal{O}\omega} = \sup_{n < \omega} 2^{\mathcal{O}n}$$

ce qui permet de dire que

$$2^{\mathcal{O}\omega} \leq \omega$$

On pourrait montrer qu'on a exactement $2^{\mathcal{O}\omega} = \omega$, mais l'inégalité nous suffit déjà pour conclure : en effet on a en particulier $2^{\mathcal{O}\omega} \preccurlyeq \omega$ d'après la proposition 99 page 230. En se rappelant qu'on a aussi introduit les notations $\omega = \mathbb{N} = \aleph_0$, on a donc $2^{\mathcal{O}\aleph_0} \preccurlyeq \mathbb{N}$. D'après le théorème de Cantor, on a $\mathbb{N} \prec \mathcal{P}(\mathbb{N})$ et donc $2^{\mathcal{O}\aleph_0} \prec \mathcal{P}(\mathbb{N})$. Pourtant, avec l'exponentiation cardinale, on a le résultat suivant.

Proposition 113 (Cardinal des parties d'un ensemble)

1. Pour tout ensemble E , on a $\text{card}(\mathcal{P}(E)) = 2^{\text{card}(E)}$.
2. En particulier $\text{card}(\mathcal{P}(\mathbb{N})) = 2^{\aleph_0}$.

Démonstration

1. Posons $\kappa := \text{card}(E)$.

Par définition de l'exponentiation cardinale, on a $2^\kappa = \text{card}(\mathcal{F}(\kappa \rightarrow 2))$.

En particulier on a $2^\kappa \approx \mathcal{F}(\kappa \rightarrow 2)$ par définition du cardinal.

Mais $\kappa = \text{card}(E)$ donc $\kappa \approx E$ par définition du cardinal.

On a aussi $2 \approx 2$ par réflexivité de \approx .

On a donc $\mathcal{F}(\kappa \rightarrow 2) \approx \mathcal{F}(E \rightarrow 2)$ d'après la proposition 97 page 223.

On a donc $2^\kappa \approx \mathcal{F}(E \rightarrow 2)$ par transitivité de \approx .

Enfin, on a $\mathcal{F}(E \rightarrow 2) \approx \mathcal{P}(E)$ d'après la proposition 91 page 214.

On a donc $2^\kappa \approx \mathcal{P}(E)$ par transitivité de \approx .

Or 2^κ est un cardinal par définition, donc $2^\kappa = \text{card}(\mathcal{P}(E))$ par définition du cardinal.

Enfin comme $\kappa = \text{card}(E)$, on a donc $\boxed{\text{card}(\mathcal{P}(E)) = 2^{\text{card}(E)}}$.

2. D'après 1, on a $\text{card}(\mathcal{P}(\mathbb{N})) = 2^{\text{card}(\mathbb{N})}$.

Or \mathbb{N} est un cardinal d'après le théorème 13 page 233.

On a donc $\text{card}(\mathbb{N}) = \mathbb{N}$, et comme on l'a aussi noté \aleph_0 , on a bien $\boxed{\text{card}(\mathcal{P}(\mathbb{N})) = 2^{\aleph_0}}$.

CQFD.

Ainsi $2^{\aleph_0} \approx \mathcal{P}(\mathbb{N})$, et donc d'après ce que l'on a dit plus tôt, $2^{O\aleph_0} \prec 2^{\aleph_0}$.

Pour montrer que cette différence est la règle plutôt que l'exception, observons à présent un théorème très important pour toute la suite : multiplier un cardinal infini par lui-même redonne le cardinal de départ.

Théorème 19 (Multiplication par lui-même d'un cardinal infini)

Soit κ un cardinal **infini**.

Alors $\kappa \cdot \kappa = \kappa$.

Démonstration

- Par définition de la multiplication cardinale, on a $\kappa \cdot \kappa = \text{card}(\kappa \times \kappa)$.

Il nous faut donc montrer que $\kappa \times \kappa \approx \kappa$.

Nous allons à cet effet munir $\kappa \times \kappa$ d'un ordre un peu particulier.

- Pour cela, comme d'ordinaire munissons κ de \leq l'ordre ordinal.

Munissons alors $\kappa \times \kappa$ de l'ordre lexicographique associé \trianglelefteq .

Alors $(\kappa \times \kappa, \trianglelefteq)$ est bien ordonné d'après la proposition 4 page 12.

En ayant munit κ de \leq et $\kappa \times \kappa$ de \trianglelefteq , on munit $\kappa \times (\kappa \times \kappa)$ de l'ordre lexicographique associé \sqsubseteq , c'est en quelque sorte un super ordre lexicographique.

Comme \leq et \trianglelefteq sont de bons ordres, \sqsubseteq est un bon ordre d'après la proposition 4 page 12.

L'idée va être de s'en servir pour construire un deuxième ordre \trianglelefteq sur $\kappa \times \kappa$, qui lui va nous être utile.

Posons alors $g := \begin{pmatrix} \kappa \times \kappa & \longrightarrow & \kappa \times (\kappa \times \kappa) \\ (\alpha, \beta) & \longmapsto & (\max(\alpha, \beta), (\alpha, \beta)) \end{pmatrix}$.

Montrons que g est injective.

Soient (α, β) et (α', β') dans $\kappa \times \kappa$ tels que $g(\alpha, \beta) = g(\alpha', \beta')$.

On a alors $(\max(\alpha, \beta), (\alpha, \beta)) = (\max(\alpha', \beta'), (\alpha', \beta'))$.

En particulier $(\alpha, \beta) = (\alpha', \beta')$.

Donc $g : \kappa \times \kappa \longrightarrow \kappa \times (\kappa \times \kappa)$ est injective.

Posons $A := \text{im}(g)$, de sorte que $g : \kappa \times \kappa \longrightarrow A$ est bijective.

On pose alors pour tout (α, β) et (α', β') dans $\kappa \times \kappa$:

$$(\alpha, \beta) \trianglelefteq (\alpha', \beta') \iff g(\alpha, \beta) \sqsubseteq g(\alpha', \beta')$$

D'après le premier livre, \trianglelefteq est une relation d'ordre sur $\kappa \times \kappa$.

De plus, $g : (\kappa \times \kappa, \trianglelefteq) \longrightarrow (A, \sqsubseteq)$ est un isomorphisme d'ordres.

Donc $g^{-1} : (A, \sqsubseteq) \longrightarrow (\kappa \times \kappa, \trianglelefteq)$ est un isomorphisme d'ordres.

Or (A, \sqsubseteq) est bien ordonné d'après la proposition 3 page 11.

Donc $(\kappa \times \kappa, \trianglelefteq)$ est bien ordonné d'après la proposition 22 page 48.

Notons que \trianglelefteq induit un ordre sur tous les $B \times B$ avec $B \subseteq \kappa$, donc en particulier sur les $\alpha \times \alpha$ avec α un ordinal tel que $\alpha \leq \kappa$.

- Montrons que $\text{type}(\kappa \times \kappa, \trianglelefteq) = \kappa$.

Supposons par l'absurde que $\text{type}(\kappa \times \kappa, \trianglelefteq) \neq \kappa$.

Soit κ_0 le plus petit cardinal infini tel que $\text{type}(\kappa_0 \times \kappa_0, \trianglelefteq) \neq \kappa_0$.

Soit α un ordinal tel que $\alpha < \kappa_0$.

► Plaçons-nous dans le cas où α est fini.

Alors $\alpha \cdot \alpha = \alpha^{\omega} \cdot \alpha$ d'après la proposition 112 page 266.

Or $\alpha^{\omega} \cdot \alpha$ est fini d'après la proposition 54 page 129.

Donc $\alpha \cdot \alpha$ est fini, mais κ_0 est infini.

On a donc $\alpha \cdot \alpha < \kappa_0$, et on a $\alpha \cdot \alpha = \text{card}(\alpha \times \alpha)$ par définition.

On a donc $\text{card}(\alpha \times \alpha) < \kappa_0$.

► Plaçons-nous dans le cas où α est infini.

Posons alors $\lambda := \text{card}(\alpha)$.

Par définition du cardinal, λ est un cardinal et $\lambda \approx \alpha$.

Par définition d'être un cardinal, on a donc $\lambda \leq \alpha$, et donc $\lambda \leq \alpha < \kappa_0$.

Par minimalité de κ_0 , on a donc $\text{type}(\lambda \times \lambda, \preceq) = \lambda$.

Donc $\lambda \times \lambda$ et λ sont équipotents d'après la proposition 105 page 239.

Comme λ est un cardinal, on a $\text{card}(\lambda \times \lambda) = \lambda$.

Comme $\lambda < \kappa_0$, on a donc $\text{card}(\lambda \times \lambda) < \kappa_0$.

Or $\lambda = \text{card}(\alpha)$, donc $\lambda \approx \alpha$.

Donc $\lambda \times \lambda \approx \alpha \times \alpha$ d'après la proposition 92 page 215.

Donc $\text{card}(\alpha \times \alpha) = \text{card}(\lambda \times \lambda)$ d'après la prop. 106 p. 239.

On a donc $\text{card}(\alpha \times \alpha) < \kappa_0$.

Dans tous les cas, on a $\text{card}(\alpha \times \alpha) < \kappa_0$.

On a donc $\text{card}(\alpha \times \alpha) \prec \kappa_0$ car κ_0 est un cardinal.

Ainsi pour tout ordinal $\alpha < \kappa_0$, on a $\text{card}(\alpha \times \alpha) \prec \kappa_0$.

Notons (*) cette assertion.

Posons $\delta := \text{type}(\kappa_0 \times \kappa_0, \preceq)$.

Comme les ordinaux sont totalement ordonnés, on a ($\delta = \kappa_0$ ou $\delta < \kappa_0$ ou $\kappa_0 < \delta$).

Par définition de κ_0 on a $\delta \neq \kappa_0$ donc ($\delta < \kappa_0$ ou $\kappa_0 < \delta$).

► Plaçons-nous dans le cas où $\delta < \kappa_0$.

Comme κ_0 est un cardinal, on a $\delta \prec \kappa_0$.

Or $\kappa_0 \preccurlyeq \kappa_0 \times \kappa_0$ d'après la proposition 92 page 215.

De plus $\kappa_0 \times \kappa_0 \approx \delta$ d'après la proposition 105 page 239.

Ainsi $\kappa_0 \preccurlyeq \kappa_0 \times \kappa_0 \approx \delta$ donc $\kappa_0 \preccurlyeq \delta$ d'après la proposition 88 page 209.

Ainsi $\kappa_0 \preccurlyeq \delta \prec \kappa_0$ donc $\kappa_0 \prec \kappa_0$ d'après la même proposition.

En particulier $\kappa_0 \not\approx \kappa_0$, ce qui est absurde par réflexivité de \approx .

► Plaçons-nous dans le cas où $\kappa_0 < \delta$.

Il existe $f : \delta \longrightarrow (\kappa_0 \times \kappa_0, \preceq)$ un isomorphisme par définition du type.

Alors $\kappa_0 \in \delta = \text{dom}(f)$ donc on peut considérer $f(\kappa_0) \in \kappa_0 \times \kappa_0$.

Posons $(\xi, \zeta) = f(\kappa_0)$, si bien que $\xi \in \kappa_0$ et $\zeta \in \kappa_0$.

Ainsi $\xi < \kappa_0$ et $\zeta < \kappa_0$ donc $\max(\xi, \zeta) < \kappa_0$.

Posons alors $\alpha := \max(\xi, \zeta) + 1$.

Or κ_0 est un cardinal **infini** donc κ_0 est limite d'après le théorème 13 page 233.

On a donc $\alpha < \kappa_0$ d'après la proposition 14 page 37.

On a donc $\text{card}(\alpha \times \alpha) \prec \kappa_0$ d'après (*).

Notons (**) ce fait.

Montrons que $f^\rightarrow(\kappa_0) \subseteq \alpha \times \alpha$.

Soit $y \in f^\rightarrow(\kappa_0)$.

Il existe donc $\gamma < \kappa_0$ tel que $y = f(\gamma)$.

f est un isomorphisme d'ordres donc est croissant, et donc $f(\gamma) \sqsubseteq f(\kappa_0)$.

Or f est à valeurs dans $\kappa_0 \times \kappa_0$.

Il existe donc $\mu < \kappa_0$ et $\nu < \kappa_0$ tels que $f(\gamma) = (\mu, \nu)$.

Comme $f(\gamma) \sqsubseteq f(\kappa_0)$, on a donc $(\mu, \nu) \sqsubseteq (\xi, \zeta)$.

Par définition de \sqsubseteq on a donc $(\max(\mu, \nu), (\mu, \nu)) \sqsubseteq (\max(\xi, \zeta), (\xi, \zeta))$.

En particulier on a $\max(\mu, \nu) \leq \max(\xi, \zeta)$ par définition de \sqsubseteq .

On a donc $\max(\mu, \nu) < \max(\xi, \zeta) + 1 = \alpha$ donc $\mu < \alpha$ et $\nu < \alpha$.

Ainsi $\mu \in \alpha$ et $\nu \in \alpha$ donc $y = f(\gamma) = (\mu, \nu) \in \alpha \times \alpha$.

On a donc $f^\rightarrow(\kappa_0) \subseteq \alpha \times \alpha$.

On a donc $f^\rightarrow(\kappa_0) \preccurlyeq \alpha \times \alpha$ d'après la proposition 87 page 203.

Or f est un isomorphisme d'ordres donc en particulier f est injectif.

Donc $f|_{\kappa_0}$ est injectif, donc $f|_{\kappa_0} : \kappa_0 \longrightarrow f^\rightarrow(\kappa_0)$ est bijectif.

Ainsi $\kappa_0 \approx f^\rightarrow(\kappa_0)$ et donc $\kappa_0 \preccurlyeq \alpha \times \alpha$ d'après la proposition 88 page 209.

C'est absurde puisque d'après (**) on a justement $\alpha \times \alpha \prec \kappa_0$.

Ainsi dans les deux cas on aboutit à une absurdité.

Par l'absurde, on vient de montrer que $\text{type}(\kappa \times \kappa, \sqsubseteq) = \kappa$.

On a donc $\kappa \times \kappa \approx \kappa$ d'après la proposition 105 page 239.

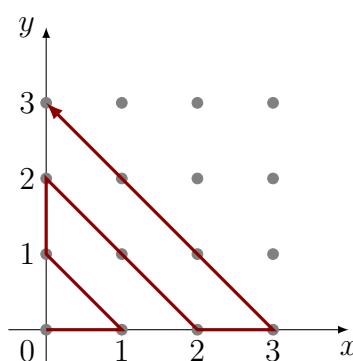
On a donc $\text{card}(\kappa \times \kappa) = \text{card}(\kappa)$ d'après la proposition 106 page 239.

Or κ est un cardinal et $\kappa \cdot \kappa = \text{card}(\kappa \times \kappa)$ par définition.

On a donc $\boxed{\kappa \cdot \kappa = \kappa}$.

CQFD.

Ainsi pour κ un cardinal infini, on a $\kappa \times \kappa \approx \kappa$. On peut en développer une intuition dans le cas où $\kappa = \mathbb{N}$ avec le dessin suivant.



Ainsi on peut développer l'intuition qu'en continuant de la tracer indéfiniment en zigzag, la flèche rouge va parcourir une et une seule fois chaque point gris, et il suffit alors d'associer à chaque entier naturel n le $n^{\text{ème}}$ point gris parcouru, ce qui fournira une bijection entre \mathbb{N} et $\mathbb{N} \times \mathbb{N}$.

Nous avons vu que si κ et λ deux cardinaux finis, alors les opérations cardinales coïncident avec les opérations ordinaires. Cependant, si au moins l'un des deux est infini, alors l'addition et la multiplication cardinales perdent de leur intérêt. En effet, on a les résultats suivants.

Proposition 114 (Addition et multiplication de cardinaux infinis)

Soient κ et λ deux cardinaux, dont au moins l'un des deux est **infini**.

On a alors :

1. $\kappa + \lambda = \max(\kappa, \lambda)$
2. Si κ et λ sont **non nuls** alors $\kappa \cdot \lambda = \max(\kappa, \lambda)$.



Démonstration

Plaçons-nous dans le cas où $\kappa \leq \lambda$.

Comme au moins l'un des deux est infinis, λ est infini.

On a donc les équivalences et égalités suivantes :

$$\begin{aligned} \lambda \times \lambda &\approx \text{card}(\lambda \times \lambda) \text{ par définition du cardinal} \\ &= \lambda \cdot \lambda \text{ par définition de la multiplication cardinale} \\ &= \lambda \text{ d'après le théorème 19 page 268} \end{aligned}$$

et donc $\lambda \times \lambda \approx \lambda$.

Notons (*) ce fait.

1. On a $\lambda \subseteq \kappa \cup \lambda$.

On a donc $\lambda \underset{87 \text{ p. 203}}{\preccurlyeq} \kappa \cup \lambda \underset{93 \text{ p. 216}}{\preccurlyeq} \kappa \amalg \lambda$ et donc $\lambda \preccurlyeq \kappa \amalg \lambda$ par transitivité de \preccurlyeq .

Montrons que $\kappa \amalg \lambda \preccurlyeq \lambda \times \lambda$.

Soit (i, α) dans $\kappa \amalg \lambda$.

Si $i = 0$ alors $\alpha < \kappa$ et comme $\kappa \leq \lambda$, on a $\alpha < \lambda$.

Si $i = 1$ alors $\alpha < \lambda$.

Dans les deux cas on a donc $\alpha < \lambda$.

Or λ est un cardinal **infini** donc est limite d'après le théorème 13 page 233.

On a donc $\alpha + 1 < \lambda$, c'est-à-dire $\alpha + 1 \in \lambda$.

De même λ est infini donc $0 < \lambda$, c'est-à-dire $0 \in \lambda$.

On a donc $(\alpha + 1, 0) \in \lambda \times \lambda$ et $(0, \alpha + 1) \in \lambda \times \lambda$.

On peut donc considérer $f := \begin{cases} \kappa \amalg \lambda & \longrightarrow \lambda \times \lambda \\ (i, \alpha) & \longmapsto \begin{cases} (\alpha + 1, 0) & \text{si } i = 0 \\ (0, \alpha + 1) & \text{si } i = 1 \end{cases} \end{cases}$.

Montrons que f est injective.

Soient (i, α) et (j, β) dans $\kappa \amalg \lambda$ tels que $f(i, \alpha) = f(j, \beta)$.

► Plaçons-nous dans le cas où $i = 0 = j$.

On a alors $(\alpha + 1, 0) = f(i, \alpha) = f(j, \beta) = (\beta + 1, 0)$.

En particulier $\alpha + 1 = \beta + 1$ donc $\alpha = \beta$ d'après la proposition 13 page 33.

Comme $i = j$ on a donc $(i, \alpha) = (j, \beta)$.

► Plaçons-nous dans le cas où $i = 1 = j$.

On a alors $(0, \alpha + 1) = f(i, \alpha) = f(j, \beta) = (0, \beta + 1)$.

En particulier $\alpha + 1 = \beta + 1$ donc $\alpha = \beta$ d'après la proposition 13 page 33.

Comme $i = j$ on a donc $(i, \alpha) = (j, \beta)$.

► Plaçons-nous dans le cas où $i = 0$ et $j = 1$.

On a alors $(\alpha + 1, 0) = f(i, \alpha) = f(j, \beta) = (0, \beta + 1)$.

En particulier $\alpha + 1 = 0$, donc $\alpha < 0$, ce qui est impossible.

► Le cas où $i = 1$ et $j = 0$ est impossible pour la même raison.

Dans les deux cas possibles, on a $(i, \alpha) = (j, \beta)$.

Donc $f : \kappa \amalg \lambda \longrightarrow \lambda \times \lambda$ est injective et donc $\kappa \amalg \lambda \preccurlyeq \lambda \times \lambda$.

Avec (\star) on a donc montré $\lambda \preccurlyeq \kappa \amalg \lambda \preccurlyeq \lambda \times \lambda \approx \lambda$.

On a donc $\lambda \approx \kappa \amalg \lambda$ par le théorème de Cantor-Schröder-Bernstein.

Donc $\text{card}(\lambda) = \text{card}(\kappa \amalg \lambda)$ d'après la proposition 106 page 239.

Or λ est un cardinal et $\kappa + \lambda = \text{card}(\kappa \amalg \lambda)$ par définition.

On a donc $\lambda = \kappa + \lambda$, et comme $\kappa \leq \lambda$, on a $\boxed{\kappa + \lambda = \max(\kappa, \lambda)}$.

2. On est toujours dans le cas où $\kappa \leq \lambda$.

On a alors $\kappa \preccurlyeq \lambda$ d'après la proposition 87 page 203.

De même $\lambda \preccurlyeq \lambda$ par réflexivité de \preccurlyeq .

On a donc $\kappa \times \lambda \preccurlyeq \lambda \times \lambda$ d'après la proposition 92 page 215.

De plus avec (\star) on a vu que $\lambda \times \lambda \approx \lambda$.

On a donc $\kappa \times \lambda \preccurlyeq \lambda$ d'après la proposition 88 page 209.

Mais on a aussi $\lambda \preccurlyeq \kappa \times \lambda$ d'après la proposition 92 page 215.

On a donc $\lambda \approx \kappa \times \lambda$ d'après le théorème de Cantor-Schröder-Bernstein.

On a donc $\text{card}(\lambda) = \text{card}(\kappa \times \lambda)$ d'après la proposition 106 page 239.

Or λ est un cardinal et $\text{card}(\kappa \times \lambda) = \kappa \cdot \lambda$ par définition.

On a donc $\kappa \cdot \lambda = \lambda$ et comme $\kappa \leq \lambda$, on a donc $\boxed{\kappa \cdot \lambda = \max(\kappa, \lambda)}$.

Le cas où $\lambda \leq \kappa$ se montre exactement de la même manière.

CQFD.

Ainsi les opérations d'addition et de multiplication cardinales manquent d'intérêt dans le cas où au moins l'un des deux cardinaux est infini, puisqu'il s'agit simplement du maximum des deux. Remarquons tout de même que l'on a cette propriété de croissance, que l'on retrouvait aussi chez les ordinaux.

Proposition 115 (Croissance des opérations sur les cardinaux)

Soient κ, λ, σ et θ quatre cardinaux.

On suppose que $\kappa \leq \sigma$ et $\lambda \leq \theta$.

1. On a $\kappa + \lambda \leq \sigma + \theta$.
2. On a $\kappa \cdot \lambda \leq \sigma \cdot \theta$.
3. Supposons que parmi κ et λ , au moins un des deux est **non nul**.
On a alors $\kappa^\lambda \leq \sigma^\theta$.

Démonstration

1. On a $\kappa \leq \sigma$ et $\lambda \leq \theta$.

On a donc $\kappa \preccurlyeq \sigma$ et $\lambda \preccurlyeq \theta$ d'après la proposition 99 page 230.

On a donc $\kappa \amalg \lambda \preccurlyeq \sigma \amalg \theta$ d'après la proposition 94 page 217.

On a donc $\text{card}(\kappa \amalg \lambda) \leq \text{card}(\sigma \amalg \theta)$ d'après la proposition 106 page 239.

On a donc $\boxed{\kappa + \lambda \leq \sigma + \theta}$ par définition de l'addition cardinale.

2. On a $\kappa \leq \sigma$ et $\lambda \leq \theta$.

On a donc $\kappa \preccurlyeq \sigma$ et $\lambda \preccurlyeq \theta$ d'après la proposition 99 page 230.

On a donc $\kappa \times \lambda \preccurlyeq \sigma \times \theta$ d'après la proposition 92 page 215.

On a donc $\text{card}(\kappa \amalg \lambda) \leq \text{card}(\sigma \amalg \theta)$ d'après la proposition 106 page 239.

On a donc $\boxed{\kappa \cdot \lambda \leq \sigma \cdot \theta}$ par définition de la multiplication cardinale.

3. On suppose ici que parmi κ et λ , au moins un des deux est non nul.

On a $\kappa \leq \sigma$ et $\lambda \leq \theta$.

On a donc $\kappa \preccurlyeq \sigma$ et $\lambda \preccurlyeq \theta$ d'après la proposition 99 page 230.

On a donc $\mathcal{F}(\lambda \rightarrow \kappa) \preccurlyeq \mathcal{F}(\theta \rightarrow \sigma)$ d'après la proposition 97 page 223.

On a donc $\text{card}(\mathcal{F}(\lambda \rightarrow \kappa)) \leq \text{card}(\mathcal{F}(\theta \rightarrow \sigma))$ d'après la proposition 106 page 239.

On a donc $\boxed{\kappa^\lambda \leq \sigma^\theta}$ par définition de l'exponentiation cardinale.

CQFD.

On a vu lors du chapitre 2 que les opérations ordinaires ne sont pas commutatives. On a cependant vu peu après que dans le cas des entiers naturels, l'addition et la multiplication sont bel et bien commutatives. En ce qui concerne l'addition et la multiplication cardinale, il y a cette fois commutativité ! On retrouve de plus les propriétés usuelles de ces opérations.

Proposition 116 (Propriétés des opérations cardinales)

Soient κ , λ et θ trois cardinaux.

Addition

1. On a $\kappa + \lambda = \lambda + \kappa$.
On dit que l'addition cardinale est **commutative**.
2. On a $(\kappa + \lambda) + \theta = \kappa + (\lambda + \theta)$.
On dit que l'addition cardinale est **associative**.

Multiplication

3. On a $\kappa \cdot \lambda = \lambda \cdot \kappa$.
On dit que la multiplication cardinale est **commutative**.
4. On a $(\kappa \cdot \lambda) \cdot \theta = \kappa \cdot (\lambda \cdot \theta)$.
On dit que la multiplication cardinale est **associative**.

Addition, multiplication et exponentiation

5. On a $\kappa \cdot (\lambda + \theta) = \kappa \cdot \lambda + \kappa \cdot \theta$.
On dit que la multiplication cardinale est **distributive** sur l'addition cardinale.
6. On a $\kappa^{\lambda \cdot \theta} = (\kappa^\lambda)^\theta$.
7. On a $\kappa^{\lambda+\theta} = \kappa^\lambda \cdot \kappa^\theta$.



Démonstration

1. On a $\kappa \amalg \lambda \approx \lambda \amalg \kappa$ d'après la proposition 95 page 219.

On a donc $\text{card}(\kappa \amalg \lambda) = \text{card}(\lambda \amalg \kappa)$ d'après la proposition 106 page 239.

On a donc $\boxed{\kappa + \lambda = \lambda + \kappa}$ par définition de l'addition cardinale.

2. Par définition du cardinal, on a $\text{card}(\kappa \amalg \lambda) \approx \kappa \amalg \lambda$.

On a aussi $\theta \approx \theta$ par réflexivité de \approx .

On a donc $\text{card}(\kappa \amalg \lambda) \amalg \theta \approx (\kappa \amalg \lambda) \amalg \theta$ d'après la proposition 94 page 217.

Notons $(*)_1$ cette assertion.

De même, on a $\text{card}(\lambda \amalg \theta) \approx \lambda \amalg \theta$ par définition du cardinal.

On a aussi $\kappa \approx \kappa$ par réflexivité de \approx .

On a donc $\kappa \amalg \text{card}(\lambda \amalg \theta) \approx \kappa \amalg (\lambda \amalg \theta)$ 94 page 217.

Notons (\star_2) cette assertion.

On a les égalités et équipotences suivantes :

$$\begin{aligned}
 (\kappa + \lambda) + \theta &= \text{card}((\kappa + \lambda) \amalg \theta) \text{ par définition de l'addition cardinale} \\
 &\approx (\kappa + \lambda) \amalg \theta \text{ par définition du cardinal} \\
 &\approx \text{card}(\kappa \amalg \lambda) \amalg \theta \text{ par définition de l'addition cardinale} \\
 &\approx (\kappa \amalg \lambda) \amalg \theta \text{ d'après } (\star_1) \\
 &\approx \kappa \amalg (\lambda \amalg \theta) \text{ d'après la prop. 95 p. 219} \\
 &\approx \kappa \amalg \text{card}(\lambda \amalg \theta) \text{ d'après } (\star_2) \\
 &\approx \kappa \amalg (\lambda + \theta) \text{ par définition de l'addition cardinale} \\
 &\approx \text{card}(\kappa \amalg (\lambda + \theta)) \text{ par définition du cardinal} \\
 &= \kappa + (\lambda + \theta) \text{ par définition de l'addition cardinale}
 \end{aligned}$$

Ainsi on a $(\kappa + \lambda) + \theta \approx \kappa + (\lambda + \theta)$.

Or ce sont tous deux des cardinaux, donc $\boxed{(\kappa + \lambda) + \theta = \kappa + (\lambda + \theta)}$.

3. On a $\kappa \times \lambda \approx \lambda \times \kappa$ d'après la proposition 95 page 219.

On a donc $\text{card}(\kappa \times \lambda) = \text{card}(\lambda \times \kappa)$ d'après la proposition 106 page 239.

On a donc $\boxed{\kappa \cdot \lambda = \lambda \cdot \kappa}$ par définition de la multiplication cardinale.

4. Par définition du cardinal, on a $\text{card}(\kappa \times \lambda) \approx \kappa \times \lambda$.

On a aussi $\theta \approx \theta$ par réflexivité de \approx .

On a donc $\text{card}(\kappa \times \lambda) \times \theta \approx (\kappa \times \lambda) \times \theta$ d'après la proposition 92 page 215

Notons (\star_3) cette assertion.

De même, on a $\text{card}(\lambda \times \theta) \approx \lambda \times \theta$ par définition du cardinal.

On a aussi $\kappa \approx \kappa$ par réflexivité de \approx .

On a donc $\kappa \times \text{card}(\lambda \times \theta) \approx \kappa \times (\lambda \times \theta)$ 92 page 215.

Notons (\star_4) cette assertion.

On a les égalités et équipotences suivantes :

$$\begin{aligned}
 (\kappa \cdot \lambda) \cdot \theta &= \text{card}((\kappa \cdot \lambda) \times \theta) \text{ par définition de la multiplication cardinale} \\
 &\approx (\kappa \cdot \lambda) \times \theta \text{ par définition du cardinal} \\
 &\approx \text{card}(\kappa \times \lambda) \times \theta \text{ par définition de la multiplication cardinale}
 \end{aligned}$$

$$\begin{aligned}
 &\approx (\kappa \times \lambda) \times \theta \text{ d'après } (\star_3) \\
 &\approx \kappa \times (\lambda \times \theta) \text{ d'après la prop. 95 p. 219} \\
 &\approx \kappa \times \text{card}(\lambda \times \theta) \text{ d'après } (\star_4) \\
 &\approx \kappa \times (\lambda \cdot \theta) \text{ par définition de la multiplication cardinale} \\
 &\approx \text{card}(\kappa \times (\lambda \cdot \theta)) \text{ par définition du cardinal} \\
 &= \kappa \cdot (\lambda \cdot \theta) \text{ par définition de la multiplication cardinale}
 \end{aligned}$$

Ainsi on a $(\kappa \cdot \lambda) \cdot \theta \approx \kappa \cdot (\lambda \cdot \theta)$.

Or ce sont tous deux des cardinaux, donc $(\kappa \cdot \lambda) \cdot \theta = \kappa \cdot (\lambda \cdot \theta)$.

5. On a $\kappa \cdot (\lambda + \theta) = \text{card}(\kappa \times (\lambda + \theta))$ par définition de la multiplication cardinale.

On a donc $\kappa \cdot (\lambda + \theta) \approx \kappa \times (\lambda + \theta)$ par définition du cardinal d'un ensemble.

Notons (\star_5) cette assertion.

On a $\lambda + \theta = \text{card}(\lambda \amalg \theta)$ par définition de l'addition cardinale.

On a donc $\lambda + \theta \approx \lambda \amalg \theta$ par définition du cardinal d'un ensemble.

Or $\kappa \approx \kappa$ par réflexivité de \approx .

On a donc $\kappa \times (\lambda + \theta) \approx \kappa \times (\lambda \amalg \theta)$ d'après la proposition 92 page 215.

Notons (\star_6) cette assertion.

On a $\kappa \cdot \lambda = \text{card}(\kappa \times \lambda)$ et $\kappa \cdot \theta = \text{card}(\kappa \times \theta)$ par définition de la multiplication cardinale.

On a donc $\kappa \cdot \lambda \approx \kappa \times \lambda$ et $\kappa \cdot \theta \approx \kappa \times \theta$ par définition du cardinal d'un ensemble.

On a donc $(\kappa \cdot \lambda) \amalg (\kappa \cdot \theta) \approx (\kappa \times \lambda) \amalg (\kappa \times \theta)$ d'après la proposition 94 page 217.

Notons (\star_7) cette assertion.

Enfin $(\kappa \cdot \lambda) + (\kappa \cdot \theta) = \text{card}((\kappa \cdot \lambda) \amalg (\kappa \cdot \theta))$ par définition de l'addition cardinale.

On a donc $(\kappa \cdot \lambda) + (\kappa \cdot \theta) \approx (\kappa \cdot \lambda) \amalg (\kappa \cdot \theta)$ par définition du cardinal d'un ensemble.

Notons (\star_8) cette assertion.

En combinant ces quatre assertions, on obtient donc

$$\begin{aligned}
 \kappa \cdot (\lambda + \theta) &\approx \kappa \times (\lambda + \theta) \text{ d'après } (\star_5) \\
 &\approx \kappa \times (\lambda \amalg \theta) \text{ d'après } (\star_6) \\
 &\approx (\kappa \times \lambda) \amalg (\kappa \times \theta) \text{ d'après la prop. 95 p. 219} \\
 &\approx (\kappa \cdot \lambda) \amalg (\kappa \cdot \theta) \text{ d'après } (\star_7) \\
 &\approx (\kappa \cdot \lambda) + (\kappa \cdot \theta) \text{ d'après } (\star_8)
 \end{aligned}$$

On a donc $\kappa \cdot (\lambda + \theta) \approx (\kappa \cdot \lambda) + (\kappa \cdot \theta)$ par transitivité de \approx .

Comme ce sont tous deux des cardinaux, on a donc $\boxed{\kappa \cdot (\lambda + \theta) = (\kappa \cdot \lambda) + (\kappa \cdot \theta)}$.

6. On a $(\kappa^\lambda)^\theta = \text{card}(\mathcal{F}(\theta \rightarrow \kappa^\lambda))$ par définition de l'exponentiation cardinale.

On a donc $(\kappa^\lambda)^\theta \approx \mathcal{F}(\theta \rightarrow \kappa^\lambda)$ par définition du cardinal d'un ensemble.

Notons (\star_9) cette assertion.

On a $\kappa^\lambda = \text{card}(\mathcal{F}(\lambda \rightarrow \kappa))$ par définition de l'exponentiation cardinale.

On a donc $\kappa^\lambda \approx \mathcal{F}(\lambda \rightarrow \kappa)$ par définition du cardinal d'un ensemble.

De plus $\theta \approx \theta$ par réflexivité de \approx .

On a donc $\mathcal{F}(\theta \rightarrow \kappa^\lambda) \approx \mathcal{F}(\theta \rightarrow \mathcal{F}(\lambda \rightarrow \kappa))$ d'après la proposition 97 page 223.

Notons (\star_{10}) cette assertion.

On a $\theta \cdot \lambda = \text{card}(\theta \times \lambda)$ par définition de la multiplication cardinale.

On a donc $\theta \cdot \lambda \approx \theta \times \lambda$ par définition du cardinal d'un ensemble.

De plus $\kappa \approx \kappa$ par réflexivité de \approx .

On a donc $\mathcal{F}((\theta \cdot \lambda) \rightarrow \kappa) \approx \mathcal{F}((\theta \times \lambda) \rightarrow \kappa)$ d'après la proposition 97 page 223.

Notons (\star_{11}) cette assertion.

On a $\kappa^{\theta \cdot \lambda} = \text{card}(\mathcal{F}((\theta \cdot \lambda) \rightarrow \kappa))$ par définition de l'exponentiation cardinale.

On a donc $\kappa^{\theta \cdot \lambda} \approx \mathcal{F}((\theta \cdot \lambda) \rightarrow \kappa)$ par définition du cardinal d'un ensemble.

Notons (\star_{12}) cette assertion.

En combinant ces quatre assertions, on a donc

$$\begin{aligned} (\kappa^\lambda)^\theta &\approx \mathcal{F}(\theta \rightarrow \kappa^\lambda) \text{ d'après } (\star_9) \\ &\approx \mathcal{F}(\theta \rightarrow \mathcal{F}(\lambda \rightarrow \kappa)) \text{ d'après } (\star_{10}) \\ &\approx \mathcal{F}((\theta \times \lambda) \rightarrow \kappa) \text{ d'après la prop. 98 p. 227} \\ &\approx \mathcal{F}((\theta \cdot \lambda) \rightarrow \kappa) \text{ d'après } (\star_{11}) \\ &\approx \kappa^{\theta \cdot \lambda} \text{ d'après } (\star_{12}) \\ &= \kappa^{\lambda \cdot \theta} \text{ par commutativité de la multiplication cardinale} \end{aligned}$$

On a donc $(\kappa^\lambda)^\theta \approx \kappa^{\lambda \cdot \theta}$ par transitivité de \approx .

Comme ce sont tous deux des cardinaux, on a $\boxed{(\kappa^\lambda)^\theta = \kappa^{\lambda \cdot \theta}}$.

7. On a $\kappa^{\lambda+\theta} = \text{card}(\mathcal{F}((\lambda + \theta) \rightarrow \kappa))$ par définition de l'exponentiation cardinale.

On a donc $\kappa^{\lambda+\theta} \approx \mathcal{F}((\lambda + \theta) \rightarrow \kappa)$ par définition du cardinal d'un ensemble.

Notons (\star_{13}) cette assertion.

On a $\lambda + \theta = \text{card}(\lambda \amalg \theta)$ par définition de l'addition cardinale.

On a donc $\lambda + \theta \approx \lambda \amalg \theta$ par définition du cardinal d'un ensemble.

De plus $\kappa \approx \kappa$ par réflexivité de \approx .

On a donc $\mathcal{F}((\lambda + \theta) \rightarrow \kappa) \approx \mathcal{F}((\lambda \amalg \theta) \rightarrow \kappa)$ d'après la proposition 97 page 223.

Notons (\star_{14}).

On a $\kappa^\lambda = \text{card}(\mathcal{F}(\lambda \rightarrow \kappa))$ par définition de l'exponentiation cardinale.

On a donc $\kappa^\lambda \approx \mathcal{F}(\lambda \rightarrow \kappa)$ par définition du cardinal d'un ensemble.

On montre de même que $\kappa^\theta \approx \mathcal{F}(\theta \rightarrow \kappa)$.

On a donc $\kappa^\lambda \times \kappa^\theta \approx \mathcal{F}(\lambda \rightarrow \kappa) \times \mathcal{F}(\theta \rightarrow \kappa)$ d'après la proposition 92 page 215.

Notons (\star_{15}) cette assertion.

On a $\kappa^\lambda \cdot \kappa^\theta = \text{card}(\kappa^\lambda \times \kappa^\theta)$ par définition de la multiplication cardinale.

On a donc $\kappa^\lambda \cdot \kappa^\theta \approx \kappa^\lambda \times \kappa^\theta$ par définition du cardinal d'un ensemble.

Notons (\star_{16}) cette assertion.

En combinant ces quatre assertions, on obtient donc :

$$\begin{aligned}\kappa^{\lambda+\theta} &\approx \mathcal{F}((\lambda + \theta) \rightarrow \kappa) \text{ d'après } (\star_{13}) \\ &\approx \mathcal{F}((\lambda \amalg \theta) \rightarrow \kappa) \text{ d'après } (\star_{14}) \\ &\approx \mathcal{F}(\lambda \rightarrow \kappa) \times \mathcal{F}(\theta \rightarrow \kappa) \text{ d'après la prop. 98 p. 227} \\ &\approx \kappa^\lambda \times \kappa^\theta \text{ d'après } (\star_{15}) \\ &\approx \kappa^\lambda \cdot \kappa^\theta \text{ d'après } (\star_{16})\end{aligned}$$

Ainsi $\kappa^{\lambda+\theta} \approx \kappa^\lambda \cdot \kappa^\theta$ par transitivité de \approx .

Comme ce sont tous deux des cardinaux, on a donc $\boxed{\kappa^{\lambda+\theta} = \kappa^\lambda \cdot \kappa^\theta}$.

CQFD.

On l'a vu lors de la proposition 113 page 267, on a $\text{card}(\mathcal{P}(\mathbb{N})) = 2^{\aleph_0}$. On peut cependant remarquer qu'à la place de 2, on aurait en fait pu prendre 3 ou bien 4 et conserver le même cardinal : plus généralement, on a la proposition suivante.

Proposition 117 (Puissance de 2 et cardinal infini)

Soient κ et λ deux cardinaux.

Supposons que λ est **infini** et que l'on a $2 \leq \kappa \leq 2^\lambda$.

On a alors $\kappa^\lambda = 2^\lambda$.

Démonstration

On a $2 \leq \kappa \leq 2^\lambda$ donc $2^\lambda \leq \kappa^\lambda \leq (2^\lambda)^\lambda$ par croissance de l'exponentiation cardinale.

Or on a

$$\begin{aligned} (2^\lambda)^\lambda &= 2^{\lambda \cdot \lambda} \text{ d'après la prop. 116 p. 275} \\ &= 2^\lambda \text{ d'après le théorème 19 page 268, car } \lambda \text{ est infini} \end{aligned}$$

On a donc $(2^\lambda)^\lambda = 2^\lambda$, si bien que $2^\lambda \leq \kappa^\lambda \leq 2^\lambda$ par ce qui précède.

On en déduit $\boxed{\kappa^\lambda = 2^\lambda}$ par antisymétrie de \leq .

CQFD.

Une propriété importante à présent. Un fait bien connu que l'on démontrera plus tard est que l'union dénombrable d'une famille dénombrable est dénombrable. C'est grâce à ce théorème plus général qu'on le montrera : c'est vrai pour tout cardinal infini.

Théorème 20 (Cardinaux infinis et stabilité par union)

Soit κ un cardinal **infini**.

Soit \mathcal{A} un ensemble tel que :

1. $\text{card}(\mathcal{A}) \leq \kappa$.
2. $\forall X \in \mathcal{A}, \text{card}(X) \leq \kappa$.

Alors $\text{card}(\bigcup \mathcal{A}) \leq \kappa$.

Démonstration

- Si $\mathcal{A} = \emptyset$ alors $\bigcup \mathcal{A} = \bigcup \emptyset = \emptyset = 0$ donc $\boxed{\text{card}(\bigcup \mathcal{A}) = \text{card}(0) = 0 \leq \kappa}$.

- On considère désormais que \mathcal{A} est non vide.

Par hypothèse on a $\text{card}(\mathcal{A}) \leq \kappa$ donc $\mathcal{A} \preccurlyeq \kappa$ d'après la proposition 107 page 240.

Il existe donc une injection $f : \mathcal{A} \longrightarrow \kappa$.

Soit $X \in \mathcal{A}$.

Par hypothèse, on a $\text{card}(X) \leq \kappa$ donc $X \preccurlyeq \kappa$ d'après la proposition 107 page 240.

Il existe donc au moins une injection $X \longrightarrow \kappa$.

Ainsi l'ensemble $S_X := \{g : X \rightarrow \kappa \mid g \text{ est une injection}\}$ est non vide.

Posons alors $\mathcal{S} := \{S_X \mid X \in \mathcal{A}\}$.

Par ce qui précède, pour tout $X \in \mathcal{A}, S_X \neq \emptyset$ donc $\emptyset \notin \mathcal{S}$.

D'après l'**axiome du choix**, il existe une fonction de choix $\varphi := \mathcal{S} \longrightarrow ?$.

Autrement dit pour tout $S \in \mathcal{S}, \varphi(S) \in S$ donc pour tout $X \in \mathcal{A}, \varphi(S_X) \in S_X$.

Dit encore autrement, pour tout $X \in \mathcal{A}$, $g_X := \varphi(S_X) : X \longrightarrow \kappa$ est une injection.

- Considérons $\coprod \mathcal{A} := \{(X, x) \mid X \in \mathcal{A} \text{ et } x \in X\}$.

Posons alors $\psi := \begin{pmatrix} \coprod \mathcal{A} & \longrightarrow & \bigcup \mathcal{A} \\ (X, x) & \longmapsto & x \end{pmatrix}$.

Montrons que $\psi : \coprod \mathcal{A} \longrightarrow \bigcup \mathcal{A}$ est une surjection.

Par définition $\text{im}(\psi) \subseteq \bigcup \mathcal{A}$.

Soit $x \in \bigcup \mathcal{A}$.

Par définition de la réunion, il existe $X \in \mathcal{A}$ tel que $x \in X$.

Alors $(X, x) \in \coprod \mathcal{A}$ et $x = \psi(X, x)$ donc $x \in \text{im}(\psi)$.

On a donc $\text{im}(\psi) \supseteq \bigcup \mathcal{A}$ et donc $\text{im}(\psi) = \bigcup \mathcal{A}$.

Ainsi $\psi : \coprod \mathcal{A} \longrightarrow \bigcup \mathcal{A}$ est une surjection et donc $\boxed{\bigcup \mathcal{A} \preccurlyeq \coprod \mathcal{A}}$.

- Rappelons que l'on a $f : \mathcal{A} \longrightarrow \kappa$ donc $\forall X \in \mathcal{A}, f(X) \in \kappa$.

Rappelons aussi que pour tout $X \in \mathcal{A}$, on a $g_X : X \longrightarrow \kappa$ donc $\forall x \in X, g_X(x) \in \kappa$.

On peut donc poser $h := \begin{pmatrix} \coprod \mathcal{A} & \longrightarrow & \kappa \times \kappa \\ (X, x) & \longmapsto & (f(X), g_X(x)) \end{pmatrix}$.

Montrons que h est injective.

Soient (X, x) et (Y, y) dans $\coprod \mathcal{A}$ tels que $h(X, x) = h(Y, y)$.

On a donc $(f(X), g_X(x)) = (f(Y), g_Y(y))$.

En particulier $f(X) = f(Y)$, et donc $X = Y$ car f est injective.

De plus $g_X(x) = g_Y(y)$ donc par ce qui précède $y \in X$ et $g_X(x) = g_X(y)$.

On a donc $x = y$ car g_X est injective.

Ainsi $X = Y$ et $x = y$ donc $(X, x) = (Y, y)$.

Ainsi $h : \coprod \mathcal{A} \longrightarrow \kappa \times \kappa$ est injective, et donc $\boxed{\coprod \mathcal{A} \preccurlyeq \kappa \times \kappa}$.

- Enfin $\kappa \cdot \kappa = \text{card}(\kappa \times \kappa)$ par définition de la multiplication cardinale.

Or $\kappa \cdot \kappa = \kappa$ car κ est **infini**, d'après le théorème 19 page 268.

On a donc $\kappa = \text{card}(\kappa \times \kappa)$ et donc $\boxed{\kappa \times \kappa \approx \kappa}$ par définition du cardinal.

Finalement, on a $\bigcup \mathcal{A} \preccurlyeq \coprod \mathcal{A} \preccurlyeq \kappa \times \kappa \approx \kappa$.

On a donc $\boxed{\bigcup \mathcal{A} \preccurlyeq \kappa}$ d'après la proposition 88 page 209.

CQFD.

On l'a dit juste avant, l'usage habituel est celui d'une union de famille. On retrouve donc plutôt cette formulation-là.

Proposition 118 (Cardinaux infinis, union de famille)

Soit κ un cardinal **infini**.

Soit I un ensemble tel que $\text{card}(I) \leq \kappa$.

Soit $(E_i)_{i \in I}$ une famille telle que $\forall i \in I, \text{card}(E_i) \leq \kappa$.

Alors $\text{card}\left(\bigcup_{i \in I} E_i\right) \leq \kappa$.

 *Démonstration*

Posons $\mathcal{A} := \{E_i \mid i \in I\}$, de sorte que $\bigcup \mathcal{A} = \bigcup_{i \in I} E_i$.

On a $\mathcal{A} \preccurlyeq I$ d'après le précédent livre.

On a donc $\text{card}(\mathcal{A}) \leq \text{card}(I)$ d'après la proposition 106 page 239.

Or $\text{card}(I) \leq \kappa$ par définition, donc $\text{card}(\mathcal{A}) \leq \kappa$ par transitivité de \leq .

De plus pour tout $X \in \mathcal{A}$, il existe $i \in I$ tel que $X = E_i$.

Donc pour tout $X \in \mathcal{A}$, $\text{card}(X) \leq \kappa$ par définition.

Alors $\text{card}(\bigcup \mathcal{A}) \leq \kappa$ d'après le théorème 20 page 280.

Autrement dit, $\boxed{\text{card}\left(\bigcup_{i \in I} E_i\right) \leq \kappa}$.

CQFD.

Observons enfin comment se comporte le cardinal vis à vis des opérations ensemblistes. Ici rien ne devrait nous étonner : on a précisément défini les opérations cardinales pour avoir ces égalités.

Proposition 119 (Cardinal et opérations ensemblistes)

Soient E et F deux ensembles.

1. On a $\text{card}(E \amalg F) = \text{card}(E) + \text{card}(F)$.
2. On a $\text{card}(E \cup F) \leq \text{card}(E) + \text{card}(F)$.
3. Si E et F sont disjoints alors $\text{card}(E \cup F) = \text{card}(E) + \text{card}(F)$.
4. On a $\text{card}(E \times F) = \text{card}(E) \cdot \text{card}(F)$.
5. On a $\text{card}(\mathcal{F}(E \rightarrow F)) = \text{card}(F)^{\text{card}(E)}$.

 *Démonstration*

1. Par définition du cardinal d'un ensemble, on a $E \approx \text{card}(E)$ et $F \approx \text{card}(F)$.

On a donc $E \amalg F \approx \text{card}(E) \amalg \text{card}(F)$ d'après la proposition 94 page 217.

On a donc $\text{card}(E \amalg F) = \text{card}(\text{card}(E) \amalg \text{card}(F))$ d'après la proposition 106 page 239.

Or $\text{card}(E) + \text{card}(F) = \text{card}(\text{card}(E) \amalg \text{card}(F))$ par définition de l'addition cardinale.

On a donc $\boxed{\text{card}(E \amalg F) = \text{card}(E) + \text{card}(F)}.$

2. On a $E \cup F \preccurlyeq E \amalg F$ d'après la proposition 93 page 216.

On a donc $\text{card}(E \cup F) \leq \text{card}(E \amalg F)$ d'après la proposition 106 page 239.

On a donc $\boxed{\text{card}(E \cup F) \leq \text{card}(E) + \text{card}(F)}$ d'après 1.

3. Supposons que E et F sont disjoints.

On a alors $E \cup F \approx E \amalg F$ d'après la proposition 93 page 216.

On a donc $\text{card}(E \cup F) = \text{card}(E \amalg F)$ d'après la proposition 106 page 239.

On a donc $\boxed{\text{card}(E \cup F) = \text{card}(E) + \text{card}(F)}$ d'après 1.

4. Posons $\kappa := \text{card}(E)$ et $\lambda := \text{card}(F)$.

Par définition du cardinal d'un ensemble, on a $E \approx \kappa$ et $F \approx \lambda$.

On a donc $E \times F \approx \kappa \times \lambda$ d'après la proposition 92 page 215.

On a donc $\text{card}(E \times F) = \text{card}(\kappa \times \lambda)$ d'après la proposition 106 page 239.

Or $\kappa \cdot \lambda = \text{card}(\kappa \times \lambda)$ par définition de la multiplication cardinale.

On a donc $\text{card}(E \times F) = \kappa \cdot \lambda$, c'est-à-dire $\boxed{\text{card}(E \times F) = \text{card}(E) \cdot \text{card}(F)}.$

5. Posons $\kappa := \text{card}(E)$ et $\lambda := \text{card}(F)$.

Par définition du cardinal d'un ensemble, on a $E \approx \kappa$ et $F \approx \lambda$.

On a donc $\mathcal{F}(E \rightarrow F) \approx \mathcal{F}(\kappa \rightarrow \lambda)$ d'après la proposition 97 page 223.

On a donc $\text{card}(\mathcal{F}(E \rightarrow F)) = \text{card}(\mathcal{F}(\kappa \rightarrow \lambda))$ d'après la proposition 106 page 239.

Or $\lambda^\kappa = \text{card}(\mathcal{F}(\kappa \rightarrow \lambda))$ par définition de l'exponentiation cardinale.

On a donc $\lambda^\kappa = \text{card}(\mathcal{F}(E \rightarrow F))$, c'est-à-dire $\boxed{\text{card}(\mathcal{F}(E \rightarrow F)) = \text{card}(F)^{\text{card}(E)}}.$

CQFD.

5 Ensembles finis et ensembles dénombrables

5.1 Ensembles finis

Il est enfin temps de définir la notion d'ensemble fini. Chez les ordinaux, on a déjà donné du sens à cela : c'est la même chose qu'être un entier naturel. Au fond, il nous suffit donc de simplement déclarer qu'être fini, c'est avoir son cardinal fini, c'est-à-dire être équivalent à un entier naturel.

Comment le justifier intuitivement ? On peut par exemple l'envisager sous l'angle d'une machine qui piocherait dans l'ensemble et en retirerait les éléments, à la vitesse d'un élément par seconde : quitte à attendre extrêmement longtemps, il existera un moment où l'ensemble aura été vidé. Le nombre de seconde écoulée jusqu'à épuisement est alors le cardinal de l'ensemble, puisque l'on a mis en bijection (par le fait de piocher) les secondes écoulées avec les éléments de l'ensemble. Dans le cas d'un ensemble infini, la machine n'aura au contraire jamais épuisé totalement l'ensemble.

Définition 40 (Ensembles finis et ensembles infinis)

Soit E un ensemble.

On dit que E est **fini** si et seulement s'il existe un entier naturel n tel que $E \approx n$.

Dans le cas contraire, on dit que E est **infini**.

Remarque :

D'après le théorème 13 page 233, tout entier naturel est un cardinal.

En particulier si E est fini avec $E \approx n$, alors $\text{card}(E) = n$.

Exemple :

1. $\emptyset = 0 \approx 0$ donc \emptyset est fini et $\text{card}(\emptyset) = 0$.
2. Pour tout ensemble x , $\{x\} \approx \{0\} = 1$ donc $\{x\}$ est fini et $\text{card}(\{x\}) = 1$.
3. Pour tout ensembles x et y avec $x \neq y$, on a $\{x, y\} \approx \{0, 1\} = 2$ donc $\{x, y\}$ est fini et $\text{card}(\{x, y\}) = 2$.
4. Pour tout entier naturel n , on a $n \approx n$ donc n est fini et $\text{card}(n) = n$.
5. \mathbb{N} est infini : en effet $\mathbb{N} = \omega$ est lui-même un cardinal d'après le théorème 13 page 233. Donc pour tout entier naturel n , comme $n \in \mathbb{N}$ on a $n < \omega$ donc $\omega \not\approx n$.
On symbolise souvent le cardinal de \mathbb{N} par la notation de Hartogs \aleph_0 que l'on a introduite plus tôt. Ainsi $\text{card}(\mathbb{N}) = \aleph_0$, mais c'est aussi \mathbb{N} lui-même, ω ou encore ω_0 , toutes ces notations désignent le même ensemble, elles ont simplement des rôles différents.
6. Plus généralement quand on aura défini ces ensembles de nombres, nous verrons que $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et \mathbb{C} sont infinis car \mathbb{N} s'injecte dedans.

Justement en parlant du fait que \mathbb{N} s'injecte dans d'autres ensembles, voici une caractérisation avec \mathbb{N} de la finitude et de l'infinitude. Ainsi \mathbb{N} représente en quelque sorte la frontière entre le fini et l'infini : c'est le premier cardinal infini.

Proposition 120 (N et la frontière du fini)

Soit E un ensemble.

Alors E est fini si et seulement si $E \prec \mathbb{N}$.

De manière équivalente, E est infini si et seulement si $\mathbb{N} \preccurlyeq E$.



Démonstration

\Rightarrow Supposons que E est fini.

Par définition, il existe n un entier naturel tel que $E \approx n$.

Par définition de \mathbb{N} , on a $n \in \mathbb{N}$ donc $n < \mathbb{N}$ par définition de $<$.

Or n et \mathbb{N} sont des cardinaux d'après le théorème 13 page 233.

En particulier $\text{card}(E) = n$ et $\text{card}(\mathbb{N}) = \mathbb{N}$, si bien que $\text{card}(E) < \text{card}(\mathbb{N})$.

On a donc $[E \prec \mathbb{N}]$ d'après la proposition 106 page 239.

\Leftarrow Supposons que $E \prec \mathbb{N}$.

On a alors $\text{card}(E) < \text{card}(\mathbb{N})$ d'après la proposition 106 page 239.

Or \mathbb{N} est un cardinal d'après le théorème 13 page 233.

On a donc $\text{card}(\mathbb{N}) = \mathbb{N}$, si bien que $\text{card}(E) < \mathbb{N}$.

Posons $\kappa := \text{card}(E)$, de sorte que $\kappa < \mathbb{N}$ (et $E \approx \kappa$).

On a donc $\kappa \in \mathbb{N}$ par définition de $<$, donc κ est un entier naturel par définition de \mathbb{N} .

Comme $E \approx \kappa$, il s'en suit que $[E \text{ est fini}]$.

CQFD.

On en déduit une propriété de stabilité de la finitude et de l'infinitude : être plus petit qu'un ensemble fini fait de nous un ensemble fini, et inversement, être plus grand qu'un ensemble infini fait de nous un ensemble infini.

Proposition 121 (Plus petit qu'un fini et plus gros qu'un infini)

Soient E et F deux ensembles.

Supposons que $E \subseteq F$ ou que $E \preccurlyeq F$.

1. Si F est fini alors E est fini.
2. Si E est infini alors F est infini.



Démonstration

• Plaçons-nous dans le cas où $E \preccurlyeq F$.

1.

Supposons que F est fini.

On a alors $F \prec \mathbb{N}$ d'après la proposition 120 page 285.

On a donc $F \preccurlyeq \mathbb{N}$ et $F \not\approx \mathbb{N}$ par définition de \prec .

Comme $E \preccurlyeq F$, on a $E \preccurlyeq \mathbb{N}$ par transitivité de \preccurlyeq .

Supposons par l'absurde que $E \approx \mathbb{N}$.

On a alors $\mathbb{N} \preccurlyeq E$ d'après la proposition 86 page 202.

Comme $E \preccurlyeq F$, on a $\mathbb{N} \preccurlyeq F$ par transitivité de \preccurlyeq ;

Ainsi $F \preccurlyeq \mathbb{N} \preccurlyeq F$ donc $F \approx \mathbb{N}$ par le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on a justement dit que $F \not\approx \mathbb{N}$.

Par l'absurde, on vient de montrer que $E \not\approx \mathbb{N}$.

On a donc $E \prec \mathbb{N}$ et donc E est fini d'après la proposition 120 page 285.

Ainsi si F est fini alors E est fini.

2. C'est le contraposée de 1.

- Plaçons-nous dans le cas où $E \subseteq F$.

On a alors $E \preccurlyeq F$ d'après la proposition 87 page 203.

On est donc ramené au premier cas.

CQFD.

Le fait d'être fini est stable par certaines opérations, notamment pour l'union et le produit cartésien.

Proposition 122 (Opérations ensemblistes d'ensembles finis)

Soient E et F deux ensembles **finis**.

Alors $E \amalg F$, $E \cup F$, $E \times F$ et $\mathcal{F}(E \rightarrow F)$ sont finis.

Démonstration

Posons $m := \text{card}(E)$ et $n := \text{card}(F)$, qui sont par définition des entiers naturels.

- On a $\text{card}(E \amalg F) = m + n$ d'après la proposition 119 page 282.

Or $m + n = \overset{\mathcal{O}}{m} + \overset{\mathcal{O}}{n}$ d'après la proposition 112 page 266.

De plus $\overset{\mathcal{O}}{m} + \overset{\mathcal{O}}{n}$ est un entier naturel d'après la proposition 39 page 96.

Donc $m + n$ est un entier naturel, donc $\text{card}(E \amalg F)$ est un entier naturel.

Ainsi donc $E \amalg F$ est fini.

- On a $E \cup F \preccurlyeq E \amalg F$ d'après la proposition 93 page 216.

Or on vient de montrer que $E \amalg F$ est fini.

Donc $\boxed{E \cup F \text{ est fini}}$ d'après la proposition 121 page 285.

- On a $\text{card}(E \times F) = m \cdot n$ d'après la proposition 119 page 282.

Or $m \cdot n = m^{\mathcal{O}} n$ d'après la proposition 112 page 266.

De plus $m^{\mathcal{O}} n$ est un entier naturel d'après la proposition 54 page 129.

Donc $m \cdot n$ est un entier naturel, donc $\text{card}(E \times F)$ est un entier naturel.

Ainsi donc $\boxed{E \times F \text{ est fini}}$.

- On a $\text{card}(\mathcal{F}(E \rightarrow F)) = n^m$ d'après la proposition 119 page 282.

Or $n^m = n^{\mathcal{O}m}$ d'après la proposition 112 page 266.

De plus $n^{\mathcal{O}m}$ est un entier naturel d'après la proposition 65 page 150.

Donc n^m est un entier naturel, donc $\text{card}(\mathcal{F}(E \rightarrow F))$ est un entier naturel.

Ainsi donc $\boxed{\mathcal{F}(E \rightarrow F) \text{ est fini}}$.

CQFD.

Une union finie d'ensembles finis est finie. À première vue, il semblerait que ce soit une application directe de la proposition 118 page 282. Mais avec quel cardinal infini ? ω ? Le soucis, c'est qu'on obtient alors une inégalité large, ce qui veut dire qu'on ne peut pas conclure à la finitude. C'est pour ça que la démonstration qui suit ne s'en sert pas.

Proposition 123 (Union finie d'ensembles finis)

Soient I un ensemble **fini** et $(E_i)_{i \in I}$ une famille d'ensembles **finis**.

Alors $\bigcup_{i \in I} E_i$ est fini.

Démonstration

On fait une démonstration par induction sur le cardinal de I .

Pour tout entier naturel n , on note $P(n)$ l'assertion suivante :

« Pour tout I tel que $\text{card}(I) = n$ et toute famille $(E_i)_{i \in I}$ d'ensembles finis, $\bigcup_{i \in I} E_i$ est fini. »

Initialisation

Soit I un ensemble tel que $\text{card}(I) = 0$ et $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal, on a $I \approx \emptyset$ donc $I = \emptyset$ d'après la proposition 96 page 223.

Alors $\bigcup_{i \in I} E_i = \bigcup_{i \in \emptyset} E_i = \emptyset$ donc est fini.

Ainsi on a $P(0)$.

Hérité

Soit n un entier naturel tel que $P(n)$.

Soit I un ensemble tel que $\text{card}(I) = n + 1$.

Soit $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal, on a $I \approx n + 1$.

Soit $f : n + 1 \rightarrow I$ une bijection.

Considérons alors $J := f^{-1}(n)$.

Alors $f|_n : n \rightarrow J$ est injective car f l'est.

De plus $f|_n : n \rightarrow J$ est aussi surjective dans J par définition de J .

Donc $f|_n : n \rightarrow J$ est une bijection, donc $J \approx n$.

Ainsi $\text{card}(J) = n$ et $(E_i)_{i \in J}$ est une famille d'ensembles finis par définition.

Donc $\bigcup_{i \in J} E_i$ est fini d'après $P(n)$.

Or $f : n + 1 \rightarrow I$ est une bijection donc $f^{-1}(n + 1) = \text{im}(f) = I$.

Or $f^{-1}(n + 1) = f^{-1}(n \cup \{n\}) = f^{-1}(n) \cup f^{-1}(\{n\}) = J \cup \{f(n)\}$.

Donc $I = J \cup \{f(n)\}$: posons $a := f(n)$, de sorte que $I = J \cup \{a\}$.

On a alors $\bigcup_{i \in I} E_i = \bigcup_{i \in J \cup \{a\}} E_i = \left(\bigcup_{i \in J} E_i \right) \cup E_a$.

Or on a dit que $\bigcup_{i \in J} E_i$ est fini, et E_a est fini par définition.

Donc $\bigcup_{i \in I} E_i$ est fini d'après la proposition 122 page 286.

Ainsi on a $P(n + 1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n + 1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$.

CQFD.

Pour E et F deux ensembles, on a défini l'union disjointe $E \amalg F$ par $(\{0\} \times E) \cup (\{1\} \times F)$, de sorte à noter sur chaque éléments de E et de F leur provenance avec 0 pour E et 1 pour F .

Comment généraliser cette idée pour définir l'union disjointe d'une famille quelconque $(E_i)_{i \in I}$? Il faut marquer la provenance de chacun des éléments : on peut le faire en indiquant le i dont ils sont issus. Ainsi, les éléments de E_i seront identifiés aux couples (i, x) avec $x \in E_i$.

Définition 41 (Union disjointe généralisée)

Soient I un ensemble et $(E_i)_{i \in I}$ une famille.
On appelle **union disjointe** de $(E_i)_{i \in I}$ l'ensemble

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

Remarque :

$$\coprod_{i \in \{0,1\}} E_i = \{(0, x) \mid x \in E_0\} \cup \{(1, x) \mid x \in E_1\} = (\{0\} \times E_0) \cup (\{1\} \times E_1) = E_0 \amalg E_1.$$

Proposition 124 (Union disjointe finie d'ensembles finis)

Soient I un ensemble **fini** et $(E_i)_{i \in I}$ une famille d'ensembles **finis**.

Alors $\coprod_{i \in I} E_i$ est fini.



Démonstration

On fait une démonstration par induction sur le cardinal de I .

Pour tout entier naturel n , on note $P(n)$ l'assertion suivante :

« Pour tout I tel que $\text{card}(I) = n$ et toute famille $(E_i)_{i \in I}$ d'ensembles finis, $\coprod_{i \in I} E_i$ est fini ».

Initialisation

Soient I un ensemble tel que $\text{card}(I) = 0$ et $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal on a $I \approx 0$.

On a donc $I = \emptyset$ d'après la proposition 96 page 223.

Donc $\coprod_{i \in I} E_i = \coprod_{i \in \emptyset} E_i = \{(i, x) \mid i \in \emptyset \text{ et } x \in E_i\} = \emptyset$.

Donc $\coprod_{i \in I} E_i$ est fini.

Ainsi on a $P(0)$.

Hérédité

Soit n un entier naturel tel que $P(n)$.

Soit I un ensemble tel que $\text{card}(I) = n + 1$.

Soit $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal, on a $I \approx n + 1$.

Il existe donc $f : n + 1 \longrightarrow I$ une bijection.

Autrement dit on a $(E_i)_{i \in I} = (E_{f(k)})_{k < n+1}$.

On a donc

$$\begin{aligned}
 \coprod_{i \in I} E_i &= \coprod_{k < n+1} E_{f(k)} \\
 &= \{(k, x) \mid k < n+1 \text{ et } x \in E_{f(k)}\} \\
 &= \{(k, x) \mid k < n \text{ et } x \in E_{f(k)}\} \cup \{(n, x) \mid x \in E_{f(n)}\} \\
 &= \left(\coprod_{k < n} E_{f(k)} \right) \cup (\{n\} \times E_{f(n)})
 \end{aligned}$$

Ainsi $\coprod_{i \in I} E_i = \left(\coprod_{k < n} E_{f(k)} \right) \cup (\{n\} \times E_{f(n)})$.

Or $\{n\}$ et $E_{f(n)}$ sont finis.

Donc $\{n\} \times E_{f(n)}$ est fini d'après la proposition 122 page 286.

De plus $\coprod_{k < n} E_{f(k)}$ est fini d'après $P(n)$.

Donc $\left(\coprod_{k < n} E_{f(k)} \right) \cup (\{n\} \times E_{f(n)})$ est fini d'après la proposition 122 page 286.

Donc $\coprod_{i \in I} E_i$ est fini.

On a donc $P(n+1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n+1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$.

CQFD.

Pour prouver que l'union (ou l'union disjointe) finie d'ensembles finis est finie, on a utilisé la proposition 122 page 286 et une récurrence : on passe de n à $n+1$ en rajoutant un ensemble supplémentaire. On va réappliquer la même stratégie dans le cas du produit cartésien, mais pour cela on a besoin de la proposition suivante, qui explique comment justement rajouter un ensemble supplémentaire.

Proposition 125 (Produit cartésien généralisé et binaire)

Soient I un ensemble **non vide** et $(E_i)_{i \in I}$ une famille.

Soient $i_0 \in I$ et $J := I \setminus \{i_0\}$.

On a alors $\prod_{i \in I} E_i \approx \left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

 *Démonstration*

Commençons par remarquer la chose suivante.

Soit $f \in \prod_{i \in I} E_i$.

Alors $f : I \longrightarrow ?$ et $\forall i \in I, f(i) \in E_i$.

Ainsi $\forall j \in J, f(j) \in E_j$ et $f(i_0) \in E_{i_0}$.

On a donc $\forall j \in J, f|_J(j) \in E_j$, et évidemment $f|_J : J \longrightarrow ?$.

On a donc $f|_J \in \prod_{j \in J} E_j$.

On peut donc poser $\varphi := \begin{pmatrix} \prod_{i \in I} E_i & \longrightarrow & \left(\prod_{j \in J} E_j \right) \times E_{i_0} \\ f & \longmapsto & (f|_J, f(i_0)) \end{pmatrix}$.

- Montrons que φ est injective.

Soient f et g dans $\prod_{i \in I} E_i$ tels que $\varphi(f) = \varphi(g)$.

On a donc $(f|_J, f(i_0)) = (g|_J, g(i_0))$ donc $f|_J = g|_J$ et $f(i_0) = g(i_0)$.

On a donc $\forall j \in J, f(j) = g(j)$ et $f(i_0) = g(i_0)$.

Or $I = J \cup \{i_0\}$ donc $\forall i \in I, f(i) = g(i)$, et donc $f = g$.

Ainsi φ est injective.

- Montrons que φ est surjective dans $\left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

Par définition de φ , on sait déjà que $\text{im}(\varphi) \subseteq \left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

Soit $(g, y) \in \left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

Ainsi on a $g : J \longrightarrow ?$ ainsi que $\forall j \in J, g(j) \in E_j$, et $y \in E_{i_0}$.

Posons alors $f := \begin{pmatrix} I & \longrightarrow & ? \\ i & \longmapsto & \begin{cases} g(i) & \text{si } i \in J \\ y & \text{si } i = i_0 \end{cases} \end{pmatrix}$.

Alors pour tout $i \in I$:

- ▶ ou bien $i \in J$ et $f(i) = g(i) \in E_i$,
- ▶ ou bien $i = i_0$ et $f(i) = y \in E_i$.

Ainsi $\forall i \in I, f(i) \in E_i$, et donc $f \in \prod_{i \in I} E_i$.

De plus, $f|_J = g$ par définition, et $f(i_0) = y \in E_{i_0}$.

On a donc $\varphi(f) = (f|_J, f(i_0)) = (g, y)$ et donc $(g, y) \in \text{im}(\varphi)$.

Ainsi $\text{im}(\varphi) \supseteq \left(\prod_{j \in J} E_j \right) \times E_{i_0}$, et donc $\text{im}(\varphi) = \left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

Ainsi φ est surjective dans $\left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

Finalement, $\varphi : \prod_{i \in I} E_i \longrightarrow \left(\prod_{j \in J} E_j \right) \times E_{i_0}$ est bijective.

En particulier $\prod_{i \in I} E_i \approx \left(\prod_{j \in J} E_j \right) \times E_{i_0}$.

CQFD.

Proposition 126 (Produit cartésien fini d'ensembles finis)

Soient I un ensemble **fini** et $(E_i)_{i \in I}$ une famille d'ensembles **finis**.

Alors $\prod_{i \in I} E_i$ est fini.

Démonstration

On fait une démonstration par induction sur le cardinal de I .

Pour tout entier naturel n , on note $P(n)$ l'assertion suivante :

« Pour tout I tel que $\text{card}(I) = n$ et toute famille $(E_i)_{i \in I}$ d'ensembles finis, $\prod_{i \in I} E_i$ est fini ».

Initialisation

Soient I un ensemble tel que $\text{card}(I) = 0$ et $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal on a $I \approx 0$.

On a donc $I = \emptyset$ d'après la proposition 96 page 223.

On a donc $\bigcup_{i \in I} E_i = \bigcup_{i \in \emptyset} E_i = \emptyset$.

Donc $\mathcal{F}\left(I \rightarrow \bigcup_{i \in I} E_i\right) = \mathcal{F}(\emptyset \rightarrow \emptyset) = \{0\}$ qui est fini.

Or $\prod_{i \in I} E_i \subseteq \mathcal{F}\left(I \rightarrow \bigcup_{i \in I} E_i\right)$, donc $\prod_{i \in I} E_i$ est fini d'après la proposition 121 page 285.

Ainsi on a $P(0)$.

Hérédité

Soit n un entier naturel tel que $P(n)$.

Soit I un ensemble tel que $\text{card}(I) = n + 1$.

Soit $(E_i)_{i \in I}$ une famille d'ensembles finis.

Par définition du cardinal, on a $I \approx n + 1$.

Il existe donc $f : n + 1 \longrightarrow I$ une bijection.

Autrement dit on a $(E_i)_{i \in I} = (E_{f(k)})_{k < n+1}$.

Donc $\prod_{i \in I} E_i = \prod_{k < n+1} E_{f(k)} \approx \left(\prod_{k < n} E_{f(k)} \right) \times E_{f(n)}$ d'après la prop. 125 p. 290.

Or $\prod_{k < n} E_{f(k)}$ est fini d'après $P(n)$, et $E_{f(n)}$ est fini par définition.

Donc $\left(\prod_{k < n} E_{f(k)} \right) \times E_{f(n)}$ est fini d'après la proposition 122 page 286.

Donc $\prod_{i \in I} E_i$ est fini.

On a donc $P(n + 1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n + 1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$.

CQFD.

La proposition qui suit donne une caractérisation intéressante des ensembles infinis. Être infini, c'est donc avoir autant d'éléments qu'une de ses parties propres, c'est-à-dire une partie qui n'est pas l'ensemble tout entier. Cela jure avec notre intuition de la vie quotidienne : retirer un objet d'un sac fait que le sac contient moins d'éléments. Oui, mais dans la vie de tous les jours, les sacs sont finis.

Proposition 127 (Ensemble infini et partie propre)

Soit E un ensemble.

Les assertions suivantes sont équivalentes :

1. E est infini.
2. Il existe A une partie **propre** de E telle que $\text{card}(A) = \text{card}(E)$.

 *Démonstration*



Supposons que E est infini.

On a alors $\mathbb{N} \preccurlyeq E$ d'après la proposition 120 page 285.

Il existe donc $f : \mathbb{N} \longrightarrow E$ injective.

Posons $A := E \setminus \{f(0)\}$, de sorte que A est une partie propre de E .

Considérons alors $g := \begin{cases} E & \longrightarrow A \\ x & \longmapsto \begin{cases} x & \text{si } x \notin f^{-1}(\mathbb{N}) \\ f(n+1) & \text{si } x = f(n) \text{ avec } n \in \mathbb{N} \end{cases} \end{cases}$.

Montrons que g est injective.

Soient x et x' dans E tels que $g(x) = g(x')$.

► Plaçons-nous dans le cas où $x \notin f^{\rightarrow}(\mathbb{N})$ et $x' \notin f^{\rightarrow}(\mathbb{N})$.

On a alors $x = g(x) = g(x') = x'$ donc $x = x'$.

► Plaçons-nous dans le cas où $x \in f^{\rightarrow}(\mathbb{N})$ et $x' \in f^{\rightarrow}(\mathbb{N})$.

Il existe donc $n \in \mathbb{N}$ et $n' \in \mathbb{N}$ tels que $x = f(n)$ et $x' = f(n')$.

Alors $f(n+1) = g(x) = g(x') = f(n'+1)$ donc $f(n+1) = f(n'+1)$.

On a donc $n+1 = n'+1$ par injectivité de f .

On a donc $n = n'$ d'après la proposition 13 page 33.

On a donc $x = f(n) = f(n') = x'$ et donc $x = x'$.

► Plaçons-nous dans le cas où $x \in f^{\rightarrow}(\mathbb{N})$ et $x' \notin f^{\rightarrow}(\mathbb{N})$.

Il existe donc $n \in \mathbb{N}$ tel que $x = f(n)$.

Alors $f(n+1) = g(x) = g(x') = x'$ donc $x' = f(n+1)$ donc $x' \in f^{\rightarrow}(\mathbb{N})$.

C'est absurde puisqu'on est justement dans le cas où $x' \notin f^{\rightarrow}(\mathbb{N})$.

► Le cas où $x \notin f^{\rightarrow}(\mathbb{N})$ et $x' \in f^{\rightarrow}(\mathbb{N})$ est impossible pour la même raison.

Ainsi dans les deux cas possibles, on a $x = x'$.

Ainsi $g : E \longrightarrow A$ est injective, et donc $E \preccurlyeq A$.

Or $A \subseteq E$ donc $A \preccurlyeq E$ d'après la proposition 87 page 203.

On a donc $A \approx E$ d'après le théorème de Cantor-Schröder-Bernstein.

En particulier $\boxed{\text{card}(A) = \text{card}(E)}$ d'après la proposition 106 page 239.



Supposons qu'il existe A une partie propre de E tel que $\text{card}(A) = \text{card}(E)$.

On a alors $A \approx E$ d'après la proposition 106 page 239.

Il existe donc $g : E \longrightarrow A$ une bijection.

- Pour tout $n \in \mathbb{N}$, on pose $E_n := (g^n)^{\rightarrow}(E \setminus A)$.

Ainsi on a $E_0 = (g^0)^{\rightarrow}(E \setminus A) = \text{id}_E^{\rightarrow}(E \setminus A) = E \setminus A$.

On a donc pour tout $n \in \mathbb{N}$, $E_n = (g^n)^{\rightarrow}(E \setminus A) = (g^n)^{\rightarrow}(E_0)$.

Pour tout $n \in \mathbb{N}$, on a aussi

$$\begin{aligned} E_{n+1} &= (g^{n+1})^{\rightarrow}(E_0) \\ &= (g^n \circ g)^{\rightarrow}(E_0) \text{ par définition des itérées de } g \\ &= (g \circ g^n)^{\rightarrow}(E_0) \text{ car les itérées d'une application commutent} \\ &= g^{\rightarrow}((g^n)^{\rightarrow}(E_0)) = g^{\rightarrow}(E_n) \end{aligned}$$

et donc $E_{n+1} = g^{\rightarrow}(E_n)$.

- Montrons que E_0 est disjoint de tout E_n pour $n \in \mathbb{N}$ tel que $n \geq 1$.

Soit $n \in \mathbb{N}$ tel que $n \geq 1$.

Il existe donc $m \in \mathbb{N}$ tel que $n = m + 1$.

On a vu que $E_0 = E \setminus A$ donc E_0 et A sont disjoints.

Or par définition de g on a $\text{im}(g) = A$ donc E_0 et $\text{im}(g)$ sont disjoints.

Or $E_n = E_{m+1} = g^\rightarrow(E_m) \subseteq \text{im}(g)$ donc E_0 et E_n sont disjoints.

Ainsi E_0 est disjoint de tout E_n pour $n \in \mathbb{N}$ tel que $n \geq 1$.

Notons cela (\star_1) .

- Montrons que les termes de $(E_n)_{n \in \mathbb{N}}$ sont disjoints deux à deux.

Soient n et m deux entiers naturels tels que $m < n$.

Il existe s un entier naturel non nul tel que $n = m + s$ d'après la prop. 51 p. 124.

Notons que comme g est injective, g^m l'est aussi d'après la proposition 35 page 87.

On a donc

$$\begin{aligned} E_m \cap E_n &= E_m \cap E_{m+s} \\ &= (g^m)^\rightarrow(E_0) \cap (g^{m+s})^\rightarrow(E_0) \\ &= (g^m)^\rightarrow(E_0) \cap (g^m \circ g^s)^\rightarrow(E_0) \text{ d'après la prop. 46 p. 109} \\ &= (g^m)^\rightarrow(E_0) \cap (g^m)^\rightarrow((g^s)^\rightarrow(E_0)) \\ &= (g^m)^\rightarrow(E_0) \cap (g^m)^\rightarrow(E_s) \\ &= (g^m)^\rightarrow(E_0 \cap E_s) \text{ car } g^m \text{ est injective} \end{aligned}$$

Ainsi $E_m \cap E_n = (g^m)^\rightarrow(E_0 \cap E_s)$.

Or E_0 et E_s sont disjoints d'après (\star_1) .

Donc $E_m \cap E_n = (g^m)^\rightarrow(\emptyset) = \emptyset$ donc E_m et E_n sont disjoints.

Ainsi les termes de $(E_n)_{n \in \mathbb{N}}$ sont disjoints deux à deux.

Notons cela (\star_2) .

- Par définition A est une partie propre de E .

En particulier $E \setminus A = E_0$ est non vide : il existe $x_0 \in E_0$.

On pose alors $f := \begin{pmatrix} \mathbb{N} & \longrightarrow & E \\ n & \longmapsto & g^n(x_0) \end{pmatrix}$.

Pour tout $n \in \mathbb{N}$, comme $x_0 \in E_0$, on a $f(n) = g^n(x_0) \in (g^n)^\rightarrow(E_0) = E_n$.

Montrons que f est injective.

Soient n et m deux entiers naturels tels que $n \neq m$.

Alors E_n et E_m sont disjoints d'après (\star_2) .

Or $f(n) \in E_n$ et $f(m) \in E_m$ d'après ce qui précède.

Mais $f(n) = f(m) \Rightarrow f(n) \in E_n \cap E_m \Rightarrow E_n \cap E_m \neq \emptyset$.

Donc par contraposition $E_n \cap E_m = \emptyset \Rightarrow f(n) \neq f(m)$.

On a donc $f(n) \neq f(m)$ par modus ponens.

Donc $f : \mathbb{N} \longrightarrow E$ est injective, donc $\mathbb{N} \preccurlyeq E$.

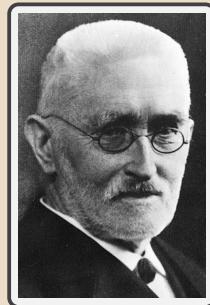
Ainsi E est infini d'après la proposition 120 page 285.

CQFD.

Remarque :

C'est Dedekind qui a proposé cette définition pour la notion d'infini. On parle donc parfois d'ensemble infini **au sens de Dedekind**.

Pour la petite histoire



Richard Dedekind (6 octobre 1831 – 12 février 1916) est un mathématicien allemand. À Göttingen, il suivra notamment les cours de Gauss, Dirichlet, Stern et Weber, et rencontrera Riemann, dont il deviendra assistant en 1851.

Son œuvre mathématique est immense. En 1872, à la même époque que Méray, Cantor et Weierstrass, il propose une construction des nombres réels, à l'aide de la méthode des coupures. Il est le fondateur avec Kummer et Kronecker de la théorie des nombres algébriques. En 1879, il définit la notion d'idéaux en l'honneur des nombres idéaux de Kummer.

De sa correspondance avec Cantor naîtra la notion d'infini que nous venons d'exposer, et il définira d'ailleurs la notion de chaîne (partie totalement ordonnée) d'un ensemble ordonné.

On dit toujours qu'une application est injective si et seulement si un élément de l'ensemble d'arrivée a au plus un antécédent, qu'elle est surjective si et seulement si un élément de l'ensemble d'arrivée a au moins un antécédent, et qu'elle est bijective si et seulement si un élément de l'ensemble d'arrivée a exactement un antécédent. On peut désormais expliciter rigoureusement ces affirmations avec la proposition suivante.

Proposition 128 (Injectivité, surjectivité et cardinaux)

Soient E et F deux ensembles, et $f : E \rightarrow F$.

1. f est injective si et seulement si $\forall y \in F, \text{card}(f^{-1}(\{y\})) \leq 1$.
2. f est surjective dans F si et seulement si $\forall y \in F, \text{card}(f^{-1}(\{y\})) \geq 1$.
3. f est bijective dans F si et seulement si $\forall y \in F, \text{card}(f^{-1}(\{y\})) = 1$.

Démonstration

1.

\Rightarrow Supposons que f est injective.

Soit $y \in F$.

► Plaçons-nous dans le cas où $y \notin \text{im}(f)$.

Alors $f^{-1}(\{y\}) = \emptyset$ donc $\text{card}(f^{-1}(\{y\})) = \text{card}(\emptyset) = 0 \leq 1$.

► Plaçons-nous dans le cas où $y \in \text{im}(f)$.

Il existe donc $x \in E$ tel que $y = f(x)$.

Montrons que $f^{-1}(\{y\}) = \{x\}$.

Par définition de x on sait déjà que $f^{-1}(\{y\}) \supseteq \{x\}$.

Soit $x' \in f^{-1}(\{y\})$.

On a donc $f(x') = y$ donc $f(x') = f(x)$.

Or f est injective par définition, donc $x = x'$ et donc $x' \in \{x\}$.

On a donc $f^{-1}(\{y\}) \subseteq \{x\}$ et donc $f^{-1}(\{y\}) = \{x\}$.

On a donc $\text{card}(f^{-1}(\{y\})) = \text{card}(\{x\}) = 1$.

Dans les deux cas on a $\text{card}(f^{-1}(\{y\})) \leq 1$.

Ainsi $\boxed{\forall y \in F, \text{card}(f^{-1}(\{y\})) \leq 1}$.

\Leftarrow Supposons que $\forall y \in F, \text{card}(f^{-1}(\{y\})) \leq 1$.

Supposons par l'absurde que f n'est pas injective.

Il existe donc x et x' tel que $x \neq x'$ et $f(x) = f(x')$.

Posons $y := f(x)$.

On a donc $\{x, x'\} \subseteq f^{-1}(\{y\})$.

On a donc $\text{card}(\{x, x'\}) \leq \text{card}(f^{-1}(\{y\}))$ d'après la proposition 104 page 238.

Ainsi $2 \leq \text{card}(f^{-1}(\{y\}))$.

C'est absurde puisque par hypothèse on a $\text{card}(f^{-1}(\{y\})) \leq 1$.

Par l'absurde on vient de montrer que $\boxed{f \text{ est injective}}$.

2.

\Rightarrow Supposons que f est surjective dans F .

Soit $y \in F$.

Comme f est surjective dans F , il existe $x \in E$ tel que $y = f(x)$.

On a donc $x \in f^{-1}(\{y\})$ donc $\{x\} \subseteq f^{-1}(\{y\})$.

On a donc $\text{card}(\{x\}) \leq \text{card}(f^{-1}(\{y\}))$ d'après la proposition 104 page 238.

Ainsi $1 \leq \text{card}(f^{-1}(\{y\}))$.

Ainsi $\boxed{\forall y \in F, \text{card}(f^{-1}(\{y\})) \geq 1}$.

\Leftarrow Supposons que $\forall y \in F, \text{card}(f^{-1}(\{y\})) \geq 1$.

Montrons que $\text{im}(f) = F$.

Par définition de f on sait déjà que $\text{im}(f) \subseteq F$.

Supposons par l'absurde que $\text{im}(f) \neq F$, donc que $\text{im}(f) \subsetneq F$.

Il existe donc $y \in F$ tel que $y \notin \text{im}(f)$.

Donc pour tout $x \in E$, on a $f(x) \neq y$ donc $x \notin f^{-1}(\{y\})$.

Or $f^{-1}(\{y\}) \subseteq E$ par définition.

On a donc $f^{-1}(\{y\}) = \emptyset$, donc $\text{card}(f^{-1}(\{y\})) = \text{card}(\emptyset) = 0$.

C'est absurde puisque par hypothèse on a $\text{card}(f^{-1}(\{y\})) \geq 1$.

Par l'absurde on vient de montrer que $\text{im}(f) = F$.

Ainsi $\boxed{f \text{ est surjective dans } F}$.

3.

En se servant de 1 et de 2, on a les équivalences suivantes :

$$\begin{aligned} & f \text{ est bijective dans } F \\ \iff & f \text{ est injective et surjective dans } F \\ \iff & \forall y \in F, \text{card}(f^{-1}(\{y\})) \leq 1 \text{ et } \forall y \in F, \text{card}(f^{-1}(\{y\})) \geq 1 \\ \iff & \forall y \in F, \text{card}(f^{-1}(\{y\})) = 1 \end{aligned}$$

D'où l'équivalence recherchée.

CQFD.

En algèbre linéaire, on se souvient souvent du fait qu'entre deux espaces vectoriels de même dimension, il y a équivalence pour une application linéaire entre être injective, être surjective et être bijective. On a un phénomène similaire dans le cas où deux ensembles sont finis et de même cardinal.

Théorème 21 (Injection, surjection et ensembles finis)

Soient E et F deux ensembles **finis** tels que $\text{card}(E) = \text{card}(F)$.

Soit $f : E \rightarrow F$.

Les assertions suivantes sont équivalentes :

1. f est une bijection de E vers F .
2. f est une injection de E dans F .
3. f est une surjection dans E sur F .



Démonstration

Nous allons montrer $1 \Leftrightarrow 2$ et $1 \Leftrightarrow 3$.

1 \Rightarrow 2 et 3

Supposons que f est une bijection de E vers F .

En particulier f est une injection de E dans F et f est une surjection dans E sur F .

2 \Rightarrow 1

Supposons que f est une injection de E dans F .

Alors $f : E \rightarrow \text{im}(f)$ est une bijection donc $E \approx \text{im}(f)$.

Or $\text{card}(E) = \text{card}(F)$ donc $E \approx F$ d'après la proposition 106 page 239.

On a donc $\text{im}(f) \approx F$ par transitivité de \approx .

En particulier $\text{card}(\text{im}(f)) = \text{card}(F)$ d'après la proposition 106 page 239.

Or F est **fini**, et $\text{im}(f)$ est une partie de F par définition de f .

Donc $\text{im}(f)$ n'est pas une partie propre de F d'après la proposition 127 page 293.

Autrement dit $\text{im}(f) = F$, et donc f est une surjection dans E sur F .

En particulier f est une bijection de E vers F .

3 \Rightarrow 1

Supposons que f est une surjection dans E sur F .

Soit $y \in F$.

Comme f est surjective dans F , il existe au moins un $x \in E$ tel que $y = f(x)$.

Autrement dit, il existe au moins un $x \in E$ tel que $x \in f^{-1}(\{y\})$.

Ainsi $f^{-1}(\{y\}) \neq \emptyset$ donc $\text{card}(f^{-1}(\{y\})) \geq 1$ d'après la proposition 96 page 223.

Supposons par l'absurde que $\text{card}(f^{-1}(\{y\})) \geq 2$.

Posons $p := \text{card}(f^{-1}(\{y\}))$: on a donc $p \geq 2$.

Il existe donc $q \in \mathbb{N}$ tel que $p = q + 2$ d'après la proposition 51 page 124.

Considérons alors $E' := E \setminus f^{-1}(\{y\})$ et $F' := F \setminus \{y\}$.

Ainsi $E = E' \cup f^{-1}(\{y\})$ et E' et $f^{-1}(\{y\})$ sont disjoints.

On a donc $\text{card}(E) = \text{card}(E') + \text{card}(f^{-1}(\{y\}))$ d'après la prop. 119 p. 282.

Ainsi $\text{card}(E) = \text{card}(E') + p = \text{card}(E') + q + 2 = \text{card}(E') + q + 1 + 1$.

De même $F = F' \cup \{y\}$ et F' et $\{y\}$ sont disjoints.

Donc $\text{card}(F) = \text{card}(F') + \text{card}(\{y\})$.

Ainsi $\text{card}(F) = \text{card}(F') + 1$ d'après la prop. 119 p. 282.

Or par définition on a $\text{card}(E) = \text{card}(F)$.

On a donc $\text{card}(E') + q + 1 + 1 = \text{card}(F') + 1$.

On a donc $\text{card}(E') + q + 1 = \text{card}(F')$ d'après la proposition 13 page 33.

On a donc $\text{card}(E') + q < \text{card}(F')$ d'après cette même proposition.

Donc $\text{card}(E') = \text{card}(E') + 0 \leq \text{card}(E') + q < \text{card}(F')$.

Ainsi $\text{card}(E') < \text{card}(F')$ et donc $E' \prec F'$ d'après la proposition 106 page 239.

Ainsi $E' \preccurlyeq F'$ et $E' \not\approx F'$.

Posons $f' := f|_{E'}$.

Montrons que $\text{im}(f') = F'$.

\subseteq

Soit $z \in \text{im}(f')$.

En particulier $z \in F$ car $\text{im}(f') \subseteq \text{im}(f) \subseteq F$.

De plus il existe $x \in E'$ tel que $z = f'(x)$.

Comme $x \in E' = E \setminus f^{-1}(\{y\})$, on a $x \notin f^{-1}(\{y\})$.

Autrement dit $f(x) \notin \{y\}$, c'est-à-dire $z = f(x) \neq y$.

Ainsi $z \notin \{y\}$ donc $z \in F \setminus \{y\} = F'$.

On a donc $\text{im}(f') \subseteq F'$, et donc $f' : E' \longrightarrow F'$.

\supseteq

Soit $z \in F'$.

Comme $F' = F \setminus \{y\}$, on a $z \in F$ et $z \neq y$.

Or f est surjective dans F par hypothèse.

Il existe donc $x \in E$ tel que $z = f(x)$.

Or $f(x) \neq y$ donc $x \notin f^{-1}(\{y\})$.

On a donc $x \in E \setminus f^{-1}(\{y\}) = E'$.

Ainsi $z = f(x) = f|_{E'}(x) = f'(x)$ donc $z \in \text{im}(f')$.

On a donc $\text{im}(f') \supseteq F'$ et donc $\text{im}(f') = F'$.

Ainsi $f' : E' \longrightarrow F'$ est surjective donc $F' \preccurlyeq E'$.

Or on a dit que $E' \preccurlyeq F'$.

On a donc $E' \approx F'$ d'après le théorème de Cantor-Schröder-Bernstein.

C'est absurde puisqu'on a justement dit que $E' \not\approx F'$.

Par l'absurde, on vient de montrer que $\text{card}(f^{-1}(\{y\})) = 1$.

Ainsi pour tout $y \in F$, on a $\text{card}(f^{-1}(\{y\})) = 1$.

Donc f est bijective de E vers F d'après la proposition 128 page 297.

CQFD.

5.2 Ensembles dénombrables

Un ensemble dénombrable, c'est un ensemble que l'on peut dénombrer. Que veut alors dire "*dénombrer un ensemble*" ? Il s'agit d'être capable de dresser la liste de ses éléments, et ici "*liste*" est à comprendre au sens de "*suite*". Autrement dit, il existe une suite parcourant tous les éléments de cet ensemble. Dit encore autrement, \mathbb{N} se surjecte dans l'ensemble E en question, c'est-à-dire $E \preccurlyeq \mathbb{N}$.

Notons cependant que dans la pratique mathématique, deux usages différents du terme "*dénombrable*" sont utilisés. Le premier est celui que nous venons d'évoquer : $E \preccurlyeq \mathbb{N}$. Mais bien souvent, on va plutôt utiliser le terme "*dénombrable*" pour désigner le fait d'être carrément en bijection avec \mathbb{N} . Pour cette raison, on va dans cette définition distinguer les deux usages par les qualificatifs de "*au plus dénombrable*" et "*infini dénombrable*".

Enfin, notons que le terme "*indénombrable*" s'oppose quant à lui à "*au plus dénombrable*".

Définition 42 (Ensembles dénombrables)

Soit E un ensemble.

1. On dit que E est **au plus dénombrable** si et seulement si $E \preccurlyeq \mathbb{N}$.
2. On dit que E est **infini dénombrable** si et seulement si $E \approx \mathbb{N}$.
3. On dit que E est **indénombrable** si et seulement si $\mathbb{N} \prec E$.

Remarque :

On emploie souvent simplement le terme *dénombrable*. Malheureusement, en fonction du contexte, cela peut tout aussi bien désigner *au plus dénombrable* que *infini dénombrable*. Dans les Barbuki, on s'efforcera de toujours préciser le sens.

Exemple :

1. Les ensembles finis sont au plus dénombrables.
2. Les ensembles infinis dénombrables sont au plus dénombrables.
3. \mathbb{N} est dénombrable, et nous verrons dans les prochains livres que \mathbb{Z} et \mathbb{Q} sont infinis dénombrables.
4. D'après le théorème de Cantor, $\mathcal{P}(\mathbb{N})$ est indénombrable.
5. Nous verrons dans les prochains livres que \mathbb{R} et \mathbb{C} sont indénombrables.

De la même manière qu'être plus petit qu'un fini fait de nous un fini, et qu'être plus grand qu'un infini fait de nous un infini, on a le même genre de lien avec la dénombrabilité.

Proposition 129 (Plus petit qu'un au plus dénombrable)

Soient E et F deux ensembles tels que $E \subseteq F$ ou $E \preccurlyeq F$.

1. Si E est indénombrable alors F est indénombrable.
2. Si F est au plus dénombrable alors E est au plus dénombrable.

 *Démonstration*

Si $E \subseteq F$ alors $E \preccurlyeq F$ d'après la proposition 87 page 203.

On va donc montrer le cas général $E \preccurlyeq F$.

1. Supposons que E est indénombrable.

On a alors $\mathbb{N} \prec E \preccurlyeq F$ donc $\mathbb{N} \prec F$ d'après la proposition 88 page 209.

Ainsi F est indénombrable.

2. Supposons que F est au plus dénombrable.

On a alors $E \preccurlyeq F \preccurlyeq \mathbb{N}$ donc $E \preccurlyeq \mathbb{N}$ par transitivité de \preccurlyeq .

Ainsi E est au plus dénombrable.

CQFD.

Rappelons-nous que nous avons défini $\omega_0 = \aleph_0$ comme étant le cardinal de \mathbb{N} , c'est-à-dire le plus petit cardinal infini. À partir de là, on a défini $\omega_1 = \aleph_1$ comme étant $\aleph(\aleph_0)$, c'est-à-dire le cardinal venant juste après $\omega_0 = \aleph_0$. Cela tombe bien, être indénombrable, c'est être strictement plus grand qu'un ensemble infini dénombrable. On retrouve donc naturellement la proposition suivante.

Proposition 130 (aleph1 et la frontière de l'indénombrable)

Soit E un ensemble.

Alors E est au plus dénombrable si et seulement si $E \prec \aleph_1$.

De manière équivalente, E est indénombrable si et seulement si $\aleph_1 \preccurlyeq E$.

 *Démonstration*



Supposons que E est au plus dénombrable.

On a donc $E \preccurlyeq \mathbb{N}$ par définition.

De plus on a $\mathbb{N} = \aleph_0$ par définition, donc $\mathbb{N} \approx \aleph_0$ par réflexivité de \approx .

On a aussi $\aleph_0 \prec \aleph(\aleph_0)$ par définition du cardinal d'Hartogs.

Enfin on a $\aleph(\aleph_0) = \aleph_1$ par définition, donc $\aleph(\aleph_0) \approx \aleph_1$ par réflexivité de \approx .

Ainsi on a $E \preccurlyeq \mathbb{N} \approx \aleph_0 \prec \aleph(\aleph_0) \approx \aleph_1$.

On a donc $E \prec \aleph_1$ d'après la proposition 88 page 209.



Supposons que $E \prec \aleph_1$.

On a donc $\text{card}(E) < \text{card}(\aleph_1)$ d'après la proposition 106 page 239.

Or \aleph_1 est un cardinal par définition, donc $\text{card}(\aleph_1) = \aleph_1$.

Ainsi on a $\text{card}(E) < \aleph_1$. Or $\aleph_1 = \aleph(\aleph_0)$ par définition.

Donc \aleph_1 est le plus petit cardinal strictement supérieur à \aleph_0 .

Donc $\text{card}(E) \leq \aleph_0$, et donc $E \preccurlyeq \aleph_0$ d'après la proposition 107 page 240.

Comme $\aleph_0 = \mathbb{N}$ par définition, on a $E \preccurlyeq \mathbb{N}$.

Autrement dit E est au plus dénombrable.

CQFD.

On a vu lors du théorème de Cantor que $\mathbb{N} \prec \mathcal{P}(\mathbb{N})$, ce qui avec la proposition 113 page 267 montre que $\aleph_0 < 2^{\aleph_0}$. D'après ce qui précède, on a donc $\aleph_1 \leq 2^{\aleph_0}$. A-t-on égalité $\aleph_1 = 2^{\aleph_0}$? On ne peut pas répondre à cette question dans ZFC : en 1938, Gödel a montré que l'on ne pouvait pas réfuter cette affirmation dans ZFC, et en 1968, Cohen a montré que l'on ne pouvait pas prouver cette affirmation dans ZFC. Elle est donc indépendante de ZFC. Considérer cette affirmation vraie, c'est faire ce que l'on appelle l'**hypothèse du continu**.

Plus généralement, faire l'hypothèse que pour tout ordinal α , on a $\aleph_{\alpha+1} = 2^{\aleph_\alpha}$, c'est faire l'**hypothèse généralisée du continu**. Cet énoncé, que l'on peut reformuler en « *Pour tout ensembles X et Y , si X est infini et si $X \preccurlyeq Y \prec \mathcal{P}(X)$, alors $X \approx Y$* », implique l'axiome du choix !

Tout comme la notion de finitude est stable par les opérations ensemblistes élémentaires (proposition 122 page 286), il en va de même pour la dénombrabilité. Notez l'absence des ensembles d'applications, nous reviendrons dessus juste après.

Proposition 131 (Opérations ensemblistes et dénombrabilité)

Soient E et F deux ensembles.

1. Si E et F sont infinis dénombrables alors $E \cup F$, $E \amalg F$ et $E \times F$ sont infinis dénombrables.
2. Si E et F sont au plus dénombrables alors $E \cup F$, $E \amalg F$ et $E \times F$ sont au plus dénombrables.
3. Si E et F sont indénombrables alors $E \cup F$, $E \amalg F$ et $E \times F$ sont indénombrables.

Démonstration

1. Supposons que E et F sont infinis dénombrables.

- On a alors $E \approx \mathbb{N}$ et $F \approx \mathbb{N}$.

On a alors

$$E \amalg F \approx \mathbb{N} \amalg \mathbb{N} \text{ d'après la prop. 94 p. 217}$$

$$\approx \text{card}(\mathbb{N} \amalg \mathbb{N}) \text{ par définition du cardinal}$$

$$= \mathbb{N} + \mathbb{N} \text{ par définition de l'addition cardinale}$$

$$= \max(\mathbb{N}, \mathbb{N}) = \mathbb{N} \text{ d'après la prop. 114 p. 272}$$

On a donc $E \amalg F \approx \mathbb{N}$, et donc $[E \amalg F \text{ est infini dénombrable}]$.

- On a $E \subseteq E \cup F$ donc $E \preccurlyeq E \cup F$ d'après la proposition 87 page 203.

De plus on a $E \cup F \preccurlyeq E \amalg F$ d'après la proposition 93 page 216.

Ainsi on a $\mathbb{N} \approx E \preccurlyeq E \cup F \preccurlyeq E \amalg F \approx \mathbb{N}$.

On a donc $\mathbb{N} \preccurlyeq E \cup F \preccurlyeq \mathbb{N}$ d'après la proposition 88 page 209.

On a donc $E \cup F \approx \mathbb{N}$ d'après le théorème de Cantor-Schröder-Bernstein.

Ainsi $[E \cup F \text{ est infini dénombrable}]$.

- On a dit que $E \approx \mathbb{N}$ et $F \approx \mathbb{N}$.

On a donc

$$E \times F \approx \mathbb{N} \times \mathbb{N} \text{ d'après la prop. 92 p. 215}$$

$$\approx \text{card}(\mathbb{N} \times \mathbb{N}) \text{ par définition du cardinal}$$

$$= \mathbb{N} \cdot \mathbb{N} \text{ par définition de la multiplication cardinale}$$

$$= \max(\mathbb{N}, \mathbb{N}) = \mathbb{N} \text{ d'après la prop. 114 p. 272}$$

On a donc $E \times F \approx \mathbb{N}$, et donc $[E \times F \text{ est infini dénombrable}]$.

2. Supposons que E et F sont au plus dénombrables.

- On a alors $E \preccurlyeq \mathbb{N}$ et $F \preccurlyeq \mathbb{N}$.

On a donc $E \amalg F \preccurlyeq \mathbb{N} \amalg \mathbb{N}$ d'après la proposition 94 page 217.

Or on a dit dans la démonstration de 1 que $\mathbb{N} \amalg \mathbb{N} \approx \mathbb{N}$.

On a donc $E \amalg F \preccurlyeq \mathbb{N}$ d'après la proposition 88 page 209.

Donc $[E \amalg F \text{ est au plus dénombrable}]$.

- On a $E \cup F \preccurlyeq E \amalg F$ d'après la proposition 93 page 216.

On a donc $E \cup F \preccurlyeq \mathbb{N}$ par transitivité de \preccurlyeq .

Donc $[E \cup F \text{ est au plus dénombrable}]$.

- On a dit que $E \preccurlyeq \mathbb{N}$ et $F \preccurlyeq \mathbb{N}$.

On a donc $E \times F \preccurlyeq \mathbb{N} \times \mathbb{N}$ d'après la proposition 92 page 215.

Or on a dit dans la démonstration de 1 que $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$.

On a donc $E \times F \preccurlyeq \mathbb{N}$ d'après la proposition 88 page 209.

Donc $E \times F$ est au plus dénombrable.

3. Supposons que E et F sont indénombrables.

- On a alors $\mathbb{N} \prec E$ et $\mathbb{N} \prec F$.

On a $E \subseteq E \cup F$ donc $E \preccurlyeq E \cup F$ d'après la proposition 87 page 203.

On a donc $\mathbb{N} \prec E \cup F$ d'après la proposition 88 page 209.

Donc $E \cup F$ est indénombrable.

- On a $E \cup F \preccurlyeq E \amalg F$ d'après la proposition 93 page 216.

On a donc $\mathbb{N} \prec E \amalg F$ d'après la proposition 88 page 209.

Donc $E \amalg F$ est indénombrable.

- Comme $\mathbb{N} \prec F$, F est en particulier non vide.

On a donc $E \preccurlyeq E \times F$ d'après la proposition 92 page 215.

On a donc $\mathbb{N} \prec E \times F$ d'après la proposition 88 page 209.

Donc $E \times F$ est indénombrable.

CQFD.

On retrouve comme annoncé plus tôt que toute union (au plus) dénombrable d'ensembles (au plus) dénombrables est (au plus) dénombrable.

Proposition 132 (Union d'ensembles dénombrables)

Soit I un ensemble **au plus dénombrable**.

Soit $(E_i)_{i \in I}$ une famille d'ensembles **au plus dénombrables**.

Alors $\bigcup_{i \in I} E_i$ est au plus dénombrable.



Démonstration

On applique la proposition 118 page 282 avec $\kappa = \mathbb{N}$.

CQFD.

On l'a dit juste avant d'énoncer la proposition 131 page 304, on n'y a pas fait mention des ensembles d'applications. La raison est simple : pour E et F deux ensembles au plus dénombrables, il n'y a pas de raison que $\mathcal{F}(E \rightarrow F)$ le soit aussi. Et oui, souvenons-nous du théorème de Cantor !

Proposition 133 (Indénombrabilité des suites à valeurs entières)

L'ensemble $\mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$ des suites à valeurs entières est indénombrable.

Démonstration

On a $\mathcal{F}(\mathbb{N} \rightarrow \{0, 1\}) \subseteq \mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$.

On a donc $\mathcal{F}(\mathbb{N} \rightarrow \{0, 1\}) \preccurlyeq \mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$ d'après la proposition 87 page 203.

Or on a $\mathcal{P}(\mathbb{N}) \approx \mathcal{F}(\mathbb{N} \rightarrow \{0, 1\})$ d'après la proposition 91 page 214.

De plus on a $\mathbb{N} \prec \mathcal{P}(\mathbb{N})$ d'après le théorème de Cantor.

Ainsi on a $\mathbb{N} \prec \mathcal{P}(\mathbb{N}) \approx \mathcal{F}(\mathbb{N} \rightarrow \{0, 1\}) \preccurlyeq \mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$.

On a donc $\mathbb{N} \prec \mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$ d'après la proposition 88 page 209.

Ainsi $\boxed{\mathcal{F}(\mathbb{N} \rightarrow \mathbb{N}) \text{ est indénombrable}}$.

CQFD.

Or, rappelons-nous que si pour tout $n \in \mathbb{N}$, on pose $E_n := \mathbb{N}$, alors chacun des termes de $(E_n)_{n \in \mathbb{N}}$ est dénombrable, mais pourtant $\prod_{n \in \mathbb{N}} E_n = \prod_{n \in \mathbb{N}} \mathbb{N} = \mathcal{F}(\mathbb{N} \rightarrow \mathbb{N})$, qui n'est donc pas dénombrable. Autrement dit, dans la proposition qui suit, on ne peut pas demander à l'ensemble des indices d'être infini dénombrable.

Proposition 134 (Produit cartésien fini de dénombrables)

Soit I un ensemble **fini**.

Soit $(E_i)_{i \in I}$ une famille d'ensembles **au plus dénombrables**.

Alors $\prod_{i \in I} E_i$ est au plus dénombrable.

Démonstration

On fait une démonstration par induction sur le cardinal de I .

Pour tout entier naturel n , on note $P(n)$ l'assertion suivante :

« Pour tout I tel que $\text{card}(I) = n$ et toute famille $(E_i)_{i \in I}$ d'ensembles au plus dénombrables, $\prod_{i \in I} E_i$ est au plus dénombrable ».

Initialisation

Soit I un ensemble tel que $\text{card}(I) = 0$.

Soit $(E_i)_{i \in I}$ une famille d'ensembles au plus dénombrables.

Par définition du cardinal on a $I \approx 0$, donc $I = \emptyset$ d'après la proposition 96 page 223.

On a donc $\bigcup_{i \in I} E_i = \bigcup_{i \in \emptyset} E_i = \emptyset$.

Donc $\mathcal{F}\left(I \rightarrow \bigcup_{i \in I} E_i\right) = \mathcal{F}(\emptyset \rightarrow \emptyset) = \{0\}$ qui est fini.

Or $\prod_{i \in I} E_i \subseteq \mathcal{F}\left(I \rightarrow \bigcup_{i \in I} E_i\right)$, donc $\prod_{i \in I} E_i$ est fini d'après la proposition 121 page 285.

En particulier $\prod_{i \in I} E_i$ est au plus dénombrable.

Ainsi on a $P(0)$.

Hérité

Soit n un entier naturel tel que $P(n)$.

Soit I un ensemble tel que $\text{card}(I) = n + 1$.

Soit $(E_i)_{i \in I}$ une famille d'ensembles au plus dénombrables.

Par définition du cardinal, on a $I \approx n + 1$.

Il existe donc $f : n + 1 \longrightarrow I$ une bijection.

Autrement dit on a $(E_i)_{i \in I} = (E_{f(k)})_{k < n+1}$.

Donc $\prod_{i \in I} E_i = \prod_{k < n+1} E_{f(k)} \approx \left(\prod_{k < n} E_{f(k)}\right) \times E_{f(n)}$ d'après la prop. 125 p. 290.

Or $\prod_{k < n} E_{f(k)}$ est au plus dénombrable d'après $P(n)$.

De plus $E_{f(n)}$ est au plus dénombrable par définition.

Donc $\left(\prod_{k < n} E_{f(k)}\right) \times E_{f(n)}$ est au plus dénombrable d'après la prop. 131 p. 304.

Donc $\prod_{i \in I} E_i$ est au plus dénombrable.

On a donc $P(n + 1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n + 1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$.

CQFD.

Concluons ce chapitre avec la promesse de fin du précédent chapitre. On avait alors défini

ε_0 comme étant $\omega^{\omega^{\dots}}$, un ordinal immensément grand dans l'infini. On avait surenchéri avec la définition de ζ_0 comme étant $\varepsilon_{\varepsilon\varepsilon\varepsilon\varepsilon\varepsilon\dots}$, ce qui semble dépasser l'entendement. Pourtant à ce moment-là, on a affirmé que le chapitre actuel allait nous fournir des ordinaux encore plus immense : il s'agit par exemple de $\omega_1 = \aleph_1$. Pourquoi ? Tout simplement parce que ε_0 et ζ_0 sont dénombrables ! La définition et la proposition qui suit ont pour objectif de le montrer.

Définition 43 (Assertion préservant les dénombrables)

Soit F une assertion fonctionnelle.

On dit que F **préserve les au plus dénombrable** si et seulement pour tout ordinal α au plus dénombrable, $F(\alpha)$ est aussi au plus dénombrable.

Proposition 135 (Fixation et dénombrabilité)

Soit $F : ON \rightarrow ON$ une assertion fonctionnelle **strictement croissante et continue**.

Supposons que F préserve les au plus dénombrables.

On a alors :

1. Pour tout entier naturel n , F^n préserve les au plus dénombrables.
2. F^ω préserve les au plus dénombrables.
3. F° préserve les au plus dénombrables.



Démonstration

1. Pour tout entier naturel n , posons $P(n)$ l'assertion

« F^n préserve les au plus dénombrables ».

Initialisation

Soit α un ordinal au plus dénombrable.

Alors $F^0(\alpha) = \alpha$ par définition de F^0 .

Donc $F^0(\alpha)$ est au plus dénombrable.

Ainsi F^0 préserve les au plus dénombrables donc on a $P(0)$.

Hérédité

Soit n un entier naturel tel que $P(n)$.

Soit α un ordinal au plus dénombrable.

Alors $F^n(\alpha)$ est au plus dénombrable d'après $P(n)$.

Donc $F(F^n(\alpha))$ est au plus dénombrable par hypothèse sur F .

Or $F^{n+1}(\alpha) = F(F^n(\alpha))$ par définition de F^{n+1} .

Donc $F^{n+1}(\alpha)$ est au plus dénombrable.

Ainsi F^{n+1} préserve les au plus dénombrables, donc on a $P(n + 1)$.

Ainsi pour tout entier naturel n , si $P(n)$ alors $P(n + 1)$.

Finalement P vérifie les deux conditions du principe d'induction chez les entiers naturels.

Donc pour tout entier naturel n , on a $P(n)$.

Ainsi pour tout entier naturel n , F^n préserve les au plus dénombrables.

2.

Soit α un ordinal au plus dénombrable.

Pour tout $n \in \mathbb{N}$, $F^n(\alpha)$ est au plus dénombrable d'après 1.

Autrement dit, $\forall n \in \mathbb{N}, F^n(\alpha) \preccurlyeq \mathbb{N}$ par définition.

De plus $\mathbb{N} \preccurlyeq \mathbb{N}$ par réflexivité de \preccurlyeq .

Donc $\bigcup_{n \in \mathbb{N}} F^n(\alpha) \preccurlyeq \mathbb{N}$ d'après la proposition 132 page 306.

Or on a $\sup_{n \in \mathbb{N}} F^n(\alpha) = \bigcup_{n \in \mathbb{N}} F^n(\alpha)$ d'après la proposition 11 page 29.

Donc $\sup_{n \in \mathbb{N}} F^n(\alpha)$ est au plus dénombrable.

Or $F^\omega(\alpha) = \sup_{n \in \mathbb{N}} F^n(\alpha)$ par définition de F^ω .

Donc $F^\omega(\alpha)$ est au plus dénombrable.

Ainsi F^ω préserve les au plus dénombrables.

3. Rappelons-nous qu'avec la proposition 83 page 195, on a montré que

$$\begin{cases} F^\circ(0) = F^\omega(0) \\ F^\circ(\alpha + 1) = F^\omega(F^\circ(\alpha) + 1) \text{ pour tout ordinal } \alpha \\ F^\circ(\gamma) = \sup_{\delta < \gamma} F^\circ(\delta) \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Pour tout ordinal α , posons $Q(\alpha)$ l'assertion

« Si α est au plus dénombrable alors $F^\circ(\alpha)$ est au plus dénombrable. »

Initialisation

0 est au plus dénombrable puisque $0 \leq \mathbb{N}$.

Donc $F^\omega(0)$ est au plus dénombrable d'après 2.

Donc $F^\circ(0)$ est au plus dénombrable d'après ce que l'on a dit plus haut.

On a donc $Q(0)$.

Héritéité

Soit α un ordinal tel que $Q(\alpha)$.

Supposons que $\alpha + 1$ est au plus dénombrable.

On a $\alpha \leq \alpha + 1$ donc $\alpha \preccurlyeq \alpha + 1$.

Donc α est au plus dénombrable d'après la proposition 129 page 302.

Donc $F^\circ(\alpha)$ est au plus dénombrable d'après $Q(\alpha)$.

Donc $F^\circ(\alpha) + 1$ est au plus dénombrable.

Donc $F^\omega(F^\circ(\alpha) + 1)$ est au plus dénombrable d'après 2.

Or on a dit que $F^\circ(\alpha + 1) = F^\omega(F^\circ(\alpha) + 1)$.

Donc $F^\circ(\alpha + 1)$ est au plus dénombrable.

On a donc $Q(\alpha + 1)$.

Donc pour tout ordinal α , si $Q(\alpha)$ alors $Q(\alpha + 1)$.

Hérité limite

Soit α un ordinal limite non nul tel que $\forall \beta < \alpha, Q(\beta)$.

Supposons que α est au plus dénombrable.

Soit β un ordinal tel que $\beta < \alpha$.

On a en particulier $\beta \leq \alpha$ donc $\beta \preccurlyeq \alpha$.

Donc β est au plus dénombrable d'après la proposition 129 page 302.

Donc $\forall \beta < \alpha, \beta \preccurlyeq \mathbb{N}$ par définition.

Donc $\forall \beta < \alpha, F^\circ(\beta) \preccurlyeq \mathbb{N}$ puisque $\forall \beta < \alpha, Q(\beta)$.

De plus $\alpha \preccurlyeq \mathbb{N}$ par hypothèse.

Donc $\bigcup_{\beta < \alpha} F^\circ(\beta) \preccurlyeq \mathbb{N}$ d'après la proposition 132 page 306.

Or $\sup F^\circ(\beta) = \bigcup_{\beta < \alpha} F^\circ(\beta)$ d'après la proposition 11 page 29.

Donc $\sup F^\circ(\beta) \preccurlyeq \mathbb{N}$ d'après la proposition 132 page 306.

Or on a dit que $F^\circ(\alpha) = \sup_{\beta < \alpha} F^\circ(\beta)$.

Donc $F^\circ(\alpha)$ est au plus dénombrable.

On a donc $Q(\alpha)$.

Donc pour tout ordinal limite non nul α , si $\forall \beta < \alpha, Q(\beta)$ alors $Q(\alpha)$.

Finalement Q vérifie les trois conditions du principe faible d'induction transfinie.

Donc pour tout ordinal α , on a $Q(\alpha)$.

Autrement dit, F° préserve les au plus dénombrables.

CQFD.

Nous pouvons conclure, en se rappelant la fin du précédent chapitre.

$$\text{Posons } F := \begin{pmatrix} ON & \longrightarrow & ON \\ \alpha & \longmapsto & \omega^{\mathcal{O}\alpha} \end{pmatrix}.$$

D'après la proposition 66 page 151, F est strictement croissante.

D'après la proposition 68 page 155, F est continue.

On peut donc considérer F° sa fixation d'après le théorème 10 page 192.

On peut même considérer $F^{\circ\circ}$ la fixation de sa fixation d'après la proposition 82 page 193.

On a alors posé $\varepsilon_0 := F^\circ(0)$ et $\zeta_0 := F^{\circ\circ}(0)$.

Il nous reste à montrer que F préserve les au plus dénombrables.

Proposition 136 (Les puissances d'omega)

$$\text{Soit } F := \begin{pmatrix} ON & \longrightarrow & ON \\ \alpha & \longmapsto & \omega^{\mathcal{O}\alpha} \end{pmatrix}.$$

Alors F préserve les au plus dénombrables.



Démonstration

Pour tout ordinal α , posons $R(\alpha)$ l'assertion

« Si α est au plus dénombrable, alors $\omega^{\mathcal{O}\alpha}$ est au plus dénombrable. »

Initialisation

On a $\omega^{\mathcal{O}0} = 1$ par définition, qui est bien au plus dénombrable, donc on a $R(0)$.

Hérédité

Soit α un ordinal tel que $R(\alpha)$.

Supposons que $\alpha + 1$ est au plus dénombrable.

On a $\alpha \leq \alpha + 1$ donc $\alpha \preccurlyeq \alpha + 1$.

Donc α est au plus dénombrable d'après la proposition 129 page 302.

Donc $\omega^{\mathcal{O}\alpha}$ est au plus dénombrable d'après $R(\alpha)$, c'est-à-dire $\omega^\alpha \preccurlyeq \omega$.

De même, $\omega \preccurlyeq \omega$ est au plus dénombrable par réflexivité de \preccurlyeq .

Donc $\omega \times \omega^{\mathcal{O}\alpha} \preccurlyeq \omega \times \omega$ d'après la proposition 92 page 215.

On a alors

$$\begin{aligned} \omega^{\mathcal{O}(\alpha+1)} &= (\omega^{\mathcal{O}\alpha})^\circledast \omega \text{ par définition de l'exponentiation ordinaire} \\ &= \text{type}(\omega \times \omega^{\mathcal{O}\alpha}) \text{ d'après le théorème 8 page 143} \\ &\approx \omega \times \omega^{\mathcal{O}\alpha} \text{ d'après la prop. 105 p. 239} \end{aligned}$$

$$\begin{aligned}
 &\preccurlyeq \omega \times \omega \text{ par ce qui précède} \\
 &\approx \text{card}(\omega \times \omega) \text{ par définition du cardinal} \\
 &= \omega \cdot \omega \text{ par définition de la multiplication cardinale} \\
 &= \omega \text{ d'après le théorème 19 page 268}
 \end{aligned}$$

et donc $\omega^{\mathcal{O}(\alpha+1)} \preccurlyeq \omega$.

Ainsi $\omega^{\mathcal{O}(\alpha+1)}$ est au plus dénombrable.

On a donc $R(\alpha + 1)$.

Ainsi pour tout ordinal α , si $R(\alpha)$ alors $R(\alpha + 1)$.

Hérité de limite

Soit α un ordinal tel que $\forall \beta < \alpha, R(\beta)$.

Supposons que α est au plus dénombrable.

Soit β un ordinal tel que $\beta < \alpha$.

On a en particulier $\beta \leq \alpha$ donc $\beta \preccurlyeq \alpha$.

Donc β est au plus dénombrable d'après la proposition 129 page 302.

Donc $\forall \beta < \alpha, \beta \preccurlyeq \mathbb{N}$ par définition.

Donc $\forall \beta < \alpha, \omega^{\mathcal{O}\beta} \preccurlyeq \mathbb{N}$ car $\forall \beta < \alpha, R(\beta)$.

Or $\alpha \preccurlyeq \mathbb{N}$ par hypothèse.

Donc $\bigcup_{\beta < \alpha} \omega^{\mathcal{O}\beta} \preccurlyeq \mathbb{N}$ d'après la proposition 132 page 306.

Or $\sup_{\beta < \alpha} \omega^{\mathcal{O}\beta} = \bigcup_{\beta < \alpha} \omega^{\mathcal{O}\beta}$ d'après la proposition 11 page 29.

Donc $\sup_{\beta < \alpha} \omega^{\mathcal{O}\beta} \preccurlyeq \mathbb{N}$.

Or α est un ordinal limite non nul.

Donc $\omega^{\mathcal{O}\alpha} = \sup_{\beta < \alpha} \omega^{\mathcal{O}\beta}$ par définition de l'exponentiation ordinaire.

Donc $\omega^{\mathcal{O}\alpha}$ est au plus dénombrable.

On a donc $R(\alpha)$.

Ainsi pour tout ordinal limite non nul, si $\forall \beta < \alpha, R(\beta)$ alors $R(\alpha)$.

Finalement R vérifie les trois conditions du principe faible d'induction transfinie.

Donc pour tout ordinal α , on a $R(\alpha)$.

Finalement F préserve les au plus dénombrables.

CQFD.

Ainsi F préserve les au plus dénombrables.

Donc F° et $F^{\circ\circ}$ aussi d'après la proposition d'avant.

Or 0 est au plus dénombrable.

Donc $\varepsilon_0 = F^\circ(0)$ et $\zeta_0 = F^{\circ\circ}(0)$ sont au plus dénombrables par ce qui précède.

En particulier $\varepsilon_0 < \omega_1$ et $\zeta_0 < \omega_1$. Fascinant !

Conclusion

L'aventure continue chez les ordinaux

Ainsi se termine notre aventure dans le pays des ordinaux. Elle se termine en tout cas à travers cet ouvrage, mais j'espère un jour la reprendre dans un tout nouvel ouvrage allant encore plus loin ! Et oui, dans le monde fabuleux des ordinaux, nous n'avons en réalité que gratté la surface.

Nous pourrions par exemple évoquer l'application \beth définie par récurrence par

$$\begin{cases} \beth_0 := \aleph_0 = \omega \\ \beth_{\alpha+1} := 2^{\beth_\alpha} \text{ pour tout ordinal } \alpha \\ \beth_\gamma := \sup_{\delta < \gamma} \beth_\delta \text{ pour tout ordinal limite non nul } \gamma \end{cases}$$

Nous pourrions aussi évoquer la topologie sur les ordinaux, issue de la topologie de l'ordre. Celle-ci donne alors pleinement son sens à la continuité et à la notion d'ordinaux limites, notions qui ont parsemé cet ouvrage.

Enfin, nous pourrions aussi évoquer la myriade d'applications des ordinaux en logique formelle ! Le lecteur est par exemple incité à se renseigner sur la célèbre suite de Goodstein, dont la stationnarité à 0 peut être montrée à l'aide des ordinaux.

L'aventure continue chez l'infini

La théorie des ordinaux nous a permis de parler de l'infini. On l'a fait en s'amusant à aller "toujours plus loin". Ce n'est cependant pas la seule façon de l'évoquer en mathématiques.

L'**analyse réelle**, que nous aborderons dans un prochain ouvrage, est un exemple de domaine mathématique où l'infini intervient fréquemment, sans que l'on ait besoin de parler d'ordinaux. On parle bien de suite tendant vers $+\infty$ par exemple. Il y a aussi le cas des infinitésimaux. Ces nombres infiniment petits ont perturbés les mathématiciens du fait de leur non rigueur : grâce aux travaux de Cauchy, Bolzano et Weierstrass, on a pu les balayer hors de l'analyse en les remplaçant par la fameuse méthode en $\varepsilon - \delta$. Pourtant en 1961, le mathématicien Robinson a enfin rendu rigoureux ces infinitésimaux à travers l'**analyse non standard** et les nombres **hyperréels**.

Enfin, évoquons l'incroyable théorie des nombres **surréels**, découverte par Conway en 1974. Cette théorie définit une classe de nombres dans lesquels se trouvent en particulier à la fois les nombres ordinaux et à la fois les infinitésimaux. Cette classe dispose même d'une structure de corps ordonné : il est possible d'ajouter, soustraire, multiplier et diviser les ordinaux sans soucis. C'est même la plus grande structure de corps ordonné qui soit. Absolument magnifique !

L'aventure continue chez les Barbuki

Enfin, même s'il est temps de dire au revoir aux ordinaux, notre aventure dans les Barbuki n'est pas terminée. Si les ordinaux étaient beaux à étudier en eux-même, ils nous ont offert au passage des outils puissants : les entiers naturels, la notion de cardinal et le lemme de Zorn.

Justement à propos des entiers naturels : leur étude plus en détails va concerner le prochain ouvrage : celui-ci portera sur l'**arithmétique dans \mathbb{Z}** . La suite de nos aventures va donc porter sur les ensembles de nombres.

Un peu d'histoire

Nous l'avons dit, c'est à Cantor que l'on doit la théorie des ordinaux. Dans quel cadre a-t-il été amené à les rencontrer ? Les explications qui suivent sont issues de l'article Wikipédia sur les nombres ordinaux.

Heine conseillera à Cantor de s'intéresser aux décompositions des fonctions périodiques en séries de Fourier. Cantor démontrera que si une série

$$\sum_{n \in \mathbb{Z}} a_n \cos(nx) + b_n \sin(nx)$$

est nulle sur \mathbb{R} , alors tous les coefficients a_n et b_n sont nuls. Cantor va chercher à affaiblir les hypothèses en réduisant le domaine sur lequel la série s'annule. Il montre que le résultat reste vrai si la série est nulle sauf en un nombre fini de points.

Il introduit alors la notion suivante : si P est une partie d'un segment $[a, b]$, il définit l'ensemble dérivé de P , noté P^1 , comme l'ensemble des points d'accumulation de P ou, de manière équivalente, comme l'ensemble P duquel ont été retirés tous les points isolés. Pour tout entier naturel n , il définit P^{n+1} comme étant l'ensemble dérivé de P^n . Il montre que, si la série trigonométrique est nulle sur $[0, 2\pi]$ en dehors d'un ensemble P pour lequel l'un des P^n est vide, alors les coefficients sont nuls.

Cherchant à prolonger ce résultat si les P^n sont tous non vides, il définit alors $P^\omega := \bigcap_{n \in \mathbb{N}} P^n$ puis $P^{\omega+1}$ comme étant le dérivé de P^ω . Les premiers ordinaux sont nés !

Plus généralement pour un ordinal α quelconque, il pose $P^{\alpha+1}$ l'ensemble dérivé de P^α , et pour tout ordinal limite non nul γ , il pose $P^\gamma := \bigcap_{\delta < \gamma} P^\delta$.

Baire reprendra cette démarche pour la convergence simple des suites de fonctions continues vers une fonction discontinue. Il définit une partie réductible P comme une partie pour laquelle il existe un ordinal α tel que P^α soit vide. Baire montre ensuite que si f est une fonction telle que l'ensemble des points où elle est discontinue est un ensemble réductible, alors f est limite simple d'une suite de fonctions continues.

Dans le cas contraire, la suite des P^α se stabilise avant l'ensemble P^{ω_1} . Il montre que P^{ω_1} est un ensemble parfait, c'est-à-dire sans point isolé.

Bibliographie

- ▶ Wikipédia
- ▶ Kenneth Kunen, *The Foundations of Mathematics*, 29 octobre 2007.
- ▶ Jean-Louis Krivine, *Théorie des ensembles*, 2007, éditions Cassini.
- ▶ Daniel Suratteau et Bertrand Hauchecorne, *Des mathématiciens de A à Z*, 2008, Ellipses.

Mathématiciens

Naissances au 12^{ème} siècle :

- (1170 – 1250) Leonardo Fibonacci page 76.

Naissances durant la première moitié du 19^{ème} siècle :

- (1831 – 1916) Richard Dedekind page 296.
- (1841 – 1902) Ernst Schröder page 208.
- (1845 – 1918) Georg Cantor page 179.

Naissances durant la deuxième moitié du 19^{ème} siècle :

- (1861 – 1931) Cesare Burali-Forti page 24.
- (1871 – 1953) Ernst Zermelo page 255.
- (1874 – 1943) Friedrich Moritz Hartogs page 258.
- (1878 – 1956) Felix Bernstein page 208.
- (1882 – 1935) Emmy Noether page 86.
- (1896 – 1980) Kazimierz Kuratowski page 246.

Naissances au 20^{ème} siècle :

- (1903 – 1957) John von Neumann page 16.
- (1906 – 1993) Max Zorn page 246.
- (1939 –) Jean-Louis Krivine page iv.
- (1943 – 2020) Kenneth Kunen page iv.