Introduction aux ordinaux

1 Introduction

Les ordinaux sont apparus grâce à Cantor entre la fin du XIXè siècle et le début du XXè. Au début ils n'ont été pour lui qu'un simple outil mais en réalité ils sont extrêmement importants et font partie de la base de la recherche moderne en théorie des ensembles: ils sont essentiels pour comprendre la structure des modèles de ZF(C). Essentiellement, ce sont des "représentants canoniques du bon ordre" (on précisera cette pensée dans ce qui suit, mais c'est l'idée qu'on peut s'en faire), et sous certaines conditions, ils aident à décrire l'univers. Ils permettent par exemple des définitions auto-référentes (bien sûr, sous des conditions strictes) et ils rendent facile d'utilisation la notion de cardinal (qu'ils précisent aussi). Ils sont finalement le centre des problèmes modernes de théorie des ensembles (l'axiome du choix s'exprime très bien grâce à eux, l'hypothèse du continu aussi, etc.). C'est pourquoi je propose cette petite introduction à cet outil essentiel.

2 Les bons ordres

2.1 Notions, définitions

Ici on expose dans un premier temps les notions nécessaires à la compréhension des ordinaux. Je donne les définitions, même d'objets élémentaires pour être sûr qu'il n'y a pas de malentendu (il existe des divergences -négligeables- entre les auteurs, je m'assure donc qu'on part sur le même pied).
Relation: Soit E et F deux ensembles. Une relation de domaine E et de codomaine F est un sous-ensemble de $E \times F$. Si R est une relation, et si $(x; y) \in R$, on note aussi cela xRy.
Relation antiréflexive: Une relation est dite antiréflexive si et seulement si l'énoncé suivant est vérifié $\forall x \in E, \neg xRx$.
Relation transitive: Une relation définie sur $E \times E$ est dite transitive si et seulement si $\forall x, y, z \in E((xRy \wedge yRz) \implies xRz)$.
Relation d'ordre strict: Une relation définie sur $E \times E$ est une relation d'ordre strict si et seulement si c'est une relation antiréflexive et transitive.
Vous remarquerez qu'une relation antiréflexive et transitive est nécessairement antisymétrique, i.e. vérifie $\forall x, y \in E, ((xRy \wedge yRx) \implies x = y)$.
Bon ordre: Une relation $<$ de domaine et codomaine E est un bon ordre si et seulement si c'est un ordre strict qui vérifie $\forall X \in \mathcal{P}(E), (X \neq \emptyset \implies \exists y \in X, \forall x \in X, (x = y \vee y < x))$ (où $\mathcal{P}(E)$ désigne l'ensemble des parties de E). En français, c'est un ordre strict tel que toute partie non vide de E admette un

minimum (à ne pas confondre avec "élement minimal", ou "minorant": un minimum est un minorant et un élément minimal, mais on n'a pas la réciproque). On déduit immédiatement qu'un bon ordre est un ordre total, au sens suivant : si $<$ est un bon ordre sur E, alors $\forall x, y \in E, x = y \vee x < y \vee y < x$.
Ensemble ordonné: Si E est un ensemble et $<$ un ordre strict sur E, alors $(E, <)$ est un ensemble ordonné.
Morphisme, isomorphisme: Ce terme a énormément de sens différents, selon le contexte. Quand on le lira, il ne faudra donc pas oublier qu'ici on étudie des ordres, donc pas question de penser à des opérations algébriques. Si $(E, <_E)$ et $(F, <_F)$ sont deux ensembles ordonnés, un morphisme de $(E, <_E)$ dans $(F, <_F)$ est une application f de E dans F qui "conserve l'ordre", i.e. $\forall x, y \in E, (x <_E y \implies (f(x) <_F f(y) \vee f(x) = f(y)))$. Un isomorphisme de $(E, <_E)$ dans $(F, <_F)$ est une application bijective de E dans F qui vérifie $\forall x, y \in E, (x <_E y \iff (f(x) <_F f(y))$. Lorsque l'ordre est total sur E et F, une implication suffit (laissé en exercice, c'est très simple).
Cette remarque n'est pas à proprement parler mathématique mais il faut se rendre compte que s'il existe un isomorphisme entre deux ensembles ordonnés (ici on peut faire l'analogie avec les groupes, les corps, etc.), ces deux ensembles ordonnés sont "essentiellement le même". Cantor et d'autres personnes disaient d'ailleurs que si $(E, <_E)$ et $(F, <_F)$ étaient isomorphes, ils avaient le même "type d'ordre". Ils notaient par exemple η le type d'ordre de $\mathbb{Q}$, muni de l'ordre usuel (ordre important s'il en est, mais je digresse).
Segment initial: Soit $(E, <_E)$ un ensemble ordonné. Une partie X de E est un segment initial de $(E, <_E)$ si et seulement si il vérifie : $\forall y \in E, \forall x \in X, (y < x \implies y \in X)$. Un autre nom souvent utilisé est "section commençante": ce nom rend aussi bien l'idée du "début" d'un ensemble ordonné. On dit qu'un segment initial est propre s'il n'est pas égal à E.
Ceci conclut notre liste de définitions pour les ordres. Celles concernant les ordinaux viendront plus tard. On peut déjà commencer l'étude des bons ordres.

2.2 Propriétés générales des bons ordres:

Dans cette sous-section on se donne $(E, <_E)$ et $(F, <_F)$ deux ensemble bien ordonnés ("bien ordonnés" signifie que $<_E$ et $<_F$ sont des bons ordres).
Lemme 1: Pour tout $x \in E$, l'ensemble $S_x = \{y \in E \mid y <_E x\}$ est un segment initial propre de E.
Preuve: Le fait que si c'est un segment initial, il est propre, est évident (x n'y appartient pas). Le fait que ce soit un segment initial découle immédiatement de sa définition et de la transitivité de $<_E$.
Lemme 2: Tout segment initial propre de E est de la forme S_x pour un unique $x \in E$.
Preuve: On montre dans un premier temps l'unicité d'un tel x, sous réserve d'existence. Soit $x, y \in E$ tels que $S_x = S_y$. L'ordre étant total (remarque faite en dessous de la définition de bon ordre), on a $x = y$, $x <_E y$ ou $y <_E x$. La remarque faite dans la preuve du lemme 1 concernant la non appartenance de x à S_x suffit à montrer que seule la première option est possible: $x = y$, d'où l'unicité.
Montrons maintenant l'existence. Soit A un segment initial propre de $(E, <_E)$. $E - A$ est donc non vide, et par définition admet un minimum x. Montrons qu'alors $A = S_x$. Soit $y <_E x$. Par définition de x, $y \in A$. Donc $S_x \subset A$. Soit

maintenant $y \in A$. Nécessairement, $y \neq x$, par définition de x. Si $x <_E y$, par définition de segment initial, on a $x \in A$, ce qui est impossible par définition de x. Donc, par totalité de l'ordre sur E, on a $y <_E x$: $y \in S_x$. D'où $A \subset S_x$. Donc $A = S_x$, cela conclut la preuve.
Ce que montrent les lemmes 1 et 2, c'est qu'il y a une correspondance bijective entre les segments initiaux propres de E et ses éléments. Cela va nous servir à prouver un théorème extrêmement important pour la suite :
Théorème 1: Une et une seule des trois options suivantes est réalisée :
- il existe un unique isomorphisme g de $(E, <_E)$ sur $(F, <_F)$;
-il existe un unique segment initial propre Y de $(E, <_E)$ et un unique isomorphisme g de $(Y, <_Y)$ sur $(F, <_F)$ ($<_Y$ est la restriction à Y de $<_E$);
-il existe un unique segment initial propre Y de $(F, <_F)$ et un unique isomorphisme g de $(E, <_E)$ sur $(Y, <_Y)$.
Preuve: On montre d'abord que les trois cas s'excluent. On montre par exemple que les deux dernières options sont incompatibles, les autres preuves étant essentiellement les mêmes. Supposons qu'on ait un segment initial propre S_x de E et un isomorphisme g_1 de S_x sur F (j'omets les ordres, mais il ne faut pas les oublier) ainsi qu'un segment initial propre S'_y de F et un isomorphisme g_2 de E sur S'_y. Il est évident que g_1 et g_2 ne peuvent pas être égaux, ainsi l'ensemble $\{k \in S_x \mid g_1(k) \neq g_2(k)\}$ est non vide. On considère donc son minimum k_0. Donc $g_1(k_0) \neq g_2(k_0)$. Supposons par exemple que $g_1(k_0) <_F g_2(k_0)$. Alors $g_1(k_0) \in S'_y$. Donc il existe $k_1 \in E$ tel que $g_2(k_1) = g_1(k_0)$ (par surjectivité). On en déduit une contradiction, car si $k_1 > k_0$, alors $g_2(k_1) > g_2(k_0) > g_1(k_0)$, et si $k_1 < k_0$, par définition de k_0, $g_2(k_1) = g_1(k_1) < g_1(k_0)$. Dans tous les cas, c'est contradictoire. Donc les deux cas s'excluent bien. Il en est de même des autres possibilités, qui s'excluent aussi (la preuve est presque la même). D'ailleurs cette preuve montre aussi l'unicité de l'isomorphisme g dans les trois cas, sous réserve d'existence (je laisse le.a lecteurice y réfléchir, iel verra que c'est aussi presque la même preuve).
Il ne reste plus donc qu'à montrer qu'un des trois cas se produit, et on n'aura même pas à montrer l'unicité de l'isomorphisme.
Considérons l'ensemble $A = \{(x; y) \in E \times F \mid$ il existe un isomorphisme de S_x sur $S_y\}$. La démonstration précédente suffit aussi pour montrer que pour tout $x \in E$, s'il existe $y \in F$ tel que $(x, y) \in A$, alors ce y est unique (pour cela il suffit de remarquer que la relation d'ordre "être un segment initial de" est une relation d'ordre total sur l'ensemble des segments initiaux de F). A est donc une fonction définie sur un sous-ensemble de E. Montrons que son ensemble de définition $E_1 = \{x \in E \mid \exists y \in F, (x, y) \in A\}$ est un segment initial de E, et de même que son image $F_1 = \{y \in F \mid \exists x \in E, (x, y) \in A\}$ est un segment initial de F. Soit donc $x \in E$, et on suppose qu'il existe $a \in E$ et $b \in F$ tels que $(a, b) \in A$, et $x <_E a$. Soit g l'(unique) isomorphisme de S_a sur S'_b. On a $S_x \subset S_a$. Montrons que l'image directe de S_x par g est un segment initial de S_b: soit $t \in S_x$ et $h \in S'_b$ tel que $h <_F g(t)$. g étant un isomorphisme, il existe $r \in S_a$ tel que $g(r) = h$. Donc $g(r) <_F g(t)$. g étant un isomorphisme, on en déduit $r <_E t$, et donc $r \in S_x$. Donc h appartient à l'image directe de S_x par g. Donc celle-ci est un segment initial de S'_b. Elle est donc de la forme S'_z, et $(x, z) \in A$. Donc E_1 est un segment initial de E. En remarquant que la réciproque d'un isomorphisme est un isomorphisme, on en déduit qu'on peut mener le même raisonnement symétriquement et donc F_1 est un segment initial de F. De plus, en remarquant le même fait, on déduit immédiatement que A est injective (et

surjective sur F_1, par définition de ce dernier).D'après la démonstration qui a été faite pour montrer que E_1 est un segment initial on déduit aussi que A est strictement croissante (le fait que la croissance est stricte se montre en reprenant les détails et en reprenant les détails et en observant que l'image directe de S_x ne peut pas être égale à S'_b). Ainsi, A est un isomorphisme de E_1 sur F_1. Il ne reste plus qu'à montrer que $E_1 = E$ ou $F_1 = F$, et on aura conclu. Supposons donc que $E_1 \neq E$ et $F_1 \neq F$. On choisit alors x_0 le minimum de $E - E_1$ et y_0 celui de $F - F_1$. D'après la preuve du lemme 2, on a alors $E_1 = S_{x_0}$ et $F_1 = S'_{y_0}$. Or ils sont isomorphes (par A), donc $(x_0, y_0) \in A$, et donc $x_0 \in S_{x_0}$, ce qui est contradictoire. Donc $E = E_1$ ou $F = F_1$. On voit alors qu'une des trois options est réalisée (et d'après ce qui précède, une seule).
Ce théorème peut paraître anodin, mais en réalité il est très profond, il dit en gros "tous les bons ordres sont en gros les mêmes", ce qui est très important, et cela explique pourquoi on pourra (c'est la suite) trouver des représentants canoniques ! Ce théorème et ces lemmes sont suffisants pour commencer notre étude des ordinaux, mais j'introduis avant deux constructions qui seront importantes : la somme et le produit d'ensembles bien ordonnés.
Somme d'ensembles bien ordonnés: Soit $(E, <_E)$ et $(F, <_F)$ deux ensembles bien ordonnés. On suppose qu'ils sont disjoints (s'ils ne le sont pas on peut les rendre disjoints en utilisant $E \times \{0\}$ à la place de E et $F \times \{1\}$ à la place de F, et on se ramène donc à ce cas). On définit l'ensemble ordonné $(E, <_E) \oplus (F, <_F)$ ainsi: l'ensemble de base est $E \cup F$ et l'ordre dessus est le suivant $x < y$ si et seulement si ($x \in E$ et $y \in F$) ou ($x, y \in E$ et $x <_E y$) ou ($x, y \in F$ et $x <_F y$). Il est évident que cela fournit un bon ordre (vérification laissée aux lecteurices).
Produit d'ensembles bien ordonnés: Soit $(E, <_E)$ et $(F, <_F)$ deux ensembles bien ordonnés. On définit alors l'ensemble bien ordonné $(E, <_E) \otimes (F, <_F)$ ainsi: l'ensemble de base est $E \times F$, et l'ordre est l'ordre lexicographique inverse, i.e. $(x, y) < (x', y')$ si et seulement si ($y <_F y'$) ou ($y = y'$ et $x <_E x'$). C'est moins évident que pour la somme, mais il est aisé de vérifier qu'on obtient un bon ordre (vérification laissée aux lecteurices).
On vérifie sur ces deux constructions les propriétés suivantes : (on omet de noter les ordres, mais il ne faut pas oublier qu'on s'en est fixé un) si E est un ensemble bien ordonné à 1 élément, alors $E \otimes F$ est isomorphe à F qui est isomorphe à $F \otimes E$. De même, l'ensemble vide vu comme ensemble bien ordonné, est un neutre à droite et à gauche pour $\oplus$. Deux remarques: $\oplus$ et $\otimes$ ne sont pas commutatives, pas même à isomorphisme près. Cependant (laissé en exercice), on a la distributivité suivante : Si E, F et G sont trois ensembles bien ordonnés, alors $E \otimes (F \oplus G)$ est isomorphe à $(E \otimes F) \oplus (E \otimes G)$. Ce sont des petits exercices d'écriture, qu'il peut être bon de faire: laissés donc aux lecteurices. Dans le même genre, on laisse aux lecteurices en exercice la preuve du fait qu'à isomorphisme près, les opérations $\otimes$ et $\oplus$ sont associatives.
On achève ainsi l'étude préliminaire des bons ordres "simples".

3 Les ordinaux

3.1 Notions, définitions

On entre maintenant dans le vif du sujet, voici donc quelques notions à comprendre pour comprendre les ordinaux.

Ensemble transitif: Non sans rapport avec la transitivité d'une relation, un ensemble transitif est un ensemble X qui vérifie $\forall x \in X, x \subset X$.
On donne quelques exemples pour cette définition: l'ensemble vide $\emptyset$ est transitif, de même que $\{\emptyset\}$.
Voici maintenant la définition tant attendue:
Ordinal: α est un ordinal si et seulement si α est transitif, et si la relation d'appartenance $\in$ restreinte à α (i.e. l'ensemble $\{(x, y) \in \alpha \times \alpha \mid x \in y\}$) est une relation de bon ordre sur α.
Par exemple, $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}$ sont des ordinaux (vous remarquez peut-être quelque chose: ce n'est pas anodin).

3.2 Propriétés des ordinaux

Voici maintenant le plat de résistance. Avant de comprendre l'essence des ordinaux et leur manipulation, il faut passer par une série de petits résultats, que voici.
Lemme 3: Soit α un ordinal, et soit $\beta \in \alpha$. Alors β est un ordinal.
Preuve: Réglons dans un premier temps la transitivité: $\beta \subset \alpha$ par transitivité de α, donc si $\gamma \in \beta$, $\gamma \in \alpha$, donc $\gamma \subset \alpha$, donc pour tout $\delta \in \gamma$, $\delta \in \alpha$. Or $\in$ est une relation d'ordre sur α, donc pour tout $\delta \in \gamma$, on a $\delta \in \gamma$ et $\gamma \in \beta$. Ce sont trois éléments de α, donc par transitivité d'une relation d'ordre, $\delta \in \beta$, et donc finalement, $\gamma \subset \beta$: β est transitif. Réglons ensuite le bon ordre: $\beta \in \alpha$ donc $\beta \subset \alpha$, donc la restriction de $\in$ à β est égale à la restriction de la restriction de $\in$ à α à β. C'est donc une relation de bon ordre sur β (il est évident que si $(E, <_E)$ est un ensemble bien ordonné, et si $X \subset E$, alors la restriction de $<_E$ à X est une relation de bon ordre).
Lemme 4: Si α est un ordinal et X un segment initial de α, alors X est un ordinal.
Preuve: Si X n'est pas propre, c'est évident. Supposons maintenant que X est propre. Alors, d'après le lemme 2, $X = S_\beta$ pour un unique $\beta \in \alpha$. Revenons un instant à la définition: $S_\beta = \{x \in \alpha \mid x < \beta\}$. Mais ici, $<=\in_{|\alpha}$, donc en fait $S_\beta = \{x \in \alpha \mid x \in \beta\}$. Mais $\beta \subset \alpha$, puisque α est transitif, et donc finalement, $S_\beta = \{x \mid x \in \beta\} = \beta$. Donc $X = \beta$, et d'après le lemme 3, X est un ordinal.
Lemme 5: Si α, β sont deux ordinaux, et $\beta \subset \alpha$ alors $\beta = \alpha$ ou $\beta \in \alpha$.
Preuve: Supposons que $\beta \neq \alpha$. Alors β est un segment initial propre de α: en effet soit $x \in \alpha$ et $y \in \beta$ tel que $x \in y$. Alors par transitivité de β, $y \subset \beta$ et donc $x \in \beta$. Donc $\beta = S_\gamma$ pour un $\gamma \in \alpha$. Or $S_\gamma = \gamma$ comme remarqué précédemment, donc $\beta \in \alpha$.
Lemme 6: Si α, β sont deux ordinaux, alors une et une seule des trois options suivantes est vérifiée: $\alpha = \beta$, $\alpha \in \beta$ ou $\beta \in \alpha$.
Preuve: Les trois options sont incompatibles car l'ordre sur un ordinal est strict, i.e. antiréflexif. Il reste donc à montrer qu'au moins une des trois possibilités est vérifiée. On considère donc $\alpha \cap \beta = \xi$. On vérifie immédiatement que ξ est un ordinal. Supposons qu'aucune des trois possibilités n'est vérifiée, et donc $\alpha \cap \beta \neq \alpha, \beta$, d'après le lemme 5. On choisit donc le minimum de $\alpha - \xi$ λ et μ celui de $\beta - \xi$. D'après la preuve du lemme 2, on a donc, ξ étant un ordinal et donc un segment initial de α et β (lemme 5), $\xi = S_\lambda = S_\mu$. Or $S_\lambda = \lambda$ et $S_\mu = \mu$, donc $\xi = \lambda = \mu$. Donc $\lambda = \mu \in \alpha \cap \beta$, et donc $\xi \in \xi$. De même que précédemment, ceci est impossible, et donc on a une contradiction. Une des trois conditions est donc réalisée.

La proposition qui suit tient à montrer que les ordinaux sont des bons ordres non redondants: chacun est essentiellement différent.
Proposition 1: Deux ordinaux isomorphes sont égaux, et l'unique isomorphisme qui existe entre les deux est l'identité.
Preuve: Soit α, β deux ordinaux isomorphes, et g l'(unique) isomorphisme de α sur β. On considère alors $\{x \in \alpha \mid x \neq g(x)\}$. S'il est vide, alors la proposition est démontrée: g est l'identité sur α, et par bijectivité, $\beta = \alpha$. Sinon, on considère son minimum λ. Donc $\lambda \neq g(\lambda)$. D'après le lemme 6, on peut supposer par exemple $\lambda \in g(\lambda)$. De ce fait, $\lambda \in \beta$, et donc, par surjectivité, il existe $\delta \in \alpha$ tel que $g(\delta) = \lambda$. Si $\delta \in \lambda$, c'est contradictoire, car par définition de λ, $g(\delta) = \delta$. Si $\lambda \in \delta$, c'est contradictoire car alors $\lambda \in g(\lambda) \in g(\delta) = \lambda$. On obtient donc une contradiction. De même si on suppose $g(\lambda) \in \lambda$ (car alors $g(g(\lambda)) = g(\lambda)$ par définition de λ, ce qui contredit la stricte croissance de g). Donc, dans tous les cas on a une contradiction. Donc $g = id$ et $\alpha = \beta$.
Ainsi, les ordinaux ne sont pas "redondants": il y a au plus un ordinal par "type de bon ordre". Les quelques résultats qui suivent tendent à montrer qu'il y en a exactement un.
Lemme 7: Si X est un ensemble d'ordinaux, la relation $\in$ restreinte à X est une relation de bon ordre sur X.
Preuve: Soit X un ensemble d'ordinaux, et soit Γ un sous ensemble non vide de X. Γ est non vide donc il existe $\alpha \in \Gamma$. On considère $\alpha \cap \Gamma$. Si $\alpha \cap \Gamma$ est vide, alors α est le minimum de Γ: en effet si $\beta \in \Gamma$, d'après le lemme 6, $\beta = \alpha$, $\beta \in \alpha$ ou $\alpha \in \beta$. Si c'est le premier ou le troisième cas, c'est bon. Si c'est le deuxième cas, alors $\beta \in \alpha \cap \Gamma$, ce qui contredit le fait que c'est vide. Donc, α est bien le minimum de Γ. Sinon, $\alpha \cap \Gamma$ est non vide et est inclus dans α: elle admet donc un minimum (pour $\in$) dans α, qu'on note m. On va montrer qu'alors m est le minimum de Γ. Soit $\beta \in \Gamma$. Supposons $\beta \in m$. Alors, par transitivité, $\beta \in \alpha$ et donc $\beta \in \alpha \cap \Gamma$, ce qui contredit le fait que m est le minimum de cet ensemble. Donc $m \in \beta$ ou $m = \beta$ (lemme 6). Dans tous les cas, m est bien le minimum de Γ: X est bien ordonné.
D'après ce qui précède, on déduit immédiatement qu'un ensemble transitif d'ordinaux est un ordinal. La proposition qui suit est la clé du théorème annoncé:
Proposition 2: Il n'existe pas d'ensemble des ordinaux.
Preuve: Supposons qu'un tel ensemble X existe. Premièrement, X est bien ordonné par $\in$. De plus, soit $\alpha \in X$. Pour tout $\beta \in \alpha$, β est un ordinal (lemme 3). Donc $\beta \in X$, par définition de X. Donc $\alpha \subset X$: X est transitif. D'après la remarque qui suit le lemme 7, X est un ordinal, et donc $X \in X$. Or X est un ordinal, donc c'est impossible.
Cette proposition peut aussi s'exprimer ainsi: la classe des ordinaux est une classe propre (du point de vue de la théorie des modèles, ce théorème se comprend ainsi: si $\mathfrak{M}$ est un modèle de ZF, et si on note On le sous-ensemble de M qui contient les ordinaux de $\mathfrak{M}$, alors $On \subset M$, mais $On \notin M$). Elle nous permet de déduire le théorème suivant, l'un des plus importants de l'histoire des mathématiques (on comprendra mieux pourquoi plus tard):
Théorème 2: Tout ensemble bien ordonné $(E, <_E)$ est isomorphe à un unique ordinal.
Preuve: L'unicité (sous réserve d'existence) se déduit de la proposition 1. Montrons maintenant l'existence. Je ferai une remarque concernant cette preuve un peu après. D'aprèsle théorème 1, pour tout ordinal α, une et une seule des trois

conditions suivantes est vérifiée :
- α est isomorphe à $(E, <_E)$
-un segment initial de α est isomorphe à $(E, <_E)$
-α est isomorphe à un segment initial de $(E, <_E)$.
S'il existe un ordinal α tel que l'une des deux premières propriétés est vérifiée, alors c'est bon, car d'après le lemme 4 un segment initial d'un ordinal est un ordinal. On suppose donc que pour tout ordinal α, la troisième option est vérifiée, i.e. α est isomorphe à un segment initial propre de E, S_x, et l'on note alors $x = p(\alpha)$ (il est unique, d'après le théorème 1 et le lemme 2). De ce fait, on définit une relation fonctionnelle $\alpha = j(x)$ ainsi (pour celleux qui ne sont pas habitué.e.s, ça signifie: "à partir de maintenant, $\alpha = j(x)$ abrège ..."): "$x \in E$ et α est un ordinak isomorphe à S_x". Cette relation est bien fonctionnelle en x, en d'autres termes, si on a $\alpha = j(x)$ et $\beta = j(x)$, alors $\alpha = \beta$, en vertu de la proposition 1. D'après le schéma d'axiomes de remplacement, l'image directe de E par cette relation fonctionnelle (i.e. la collection des α tels qu'il existe $x \in E$ tel que $\alpha = j(x)$) est un ensemble. Or tout ordinal α est dans cette image directe d'après ce qui précède. De plus, par définition de cette relation fonctionnelle, tout élément de l'image directe est un ordinal. Donc les éléments de l'image directe sont exactement les ordinaux. Donc il existe un ensemble des ordinaux. Cela contredit la proposition 2. Donc la supposition est erronnée: $(E, <_E)$ est isomorphe à un (unique) ordinal.
Vous ne comprenez peut-être pas encore en quoi ce théorème est assez important pour que je dise que c'est "l'un des plus importants de l'histoire des mathématiques", mais vous le comprendrez sûrement dans les développements qui suivent : il est nécessaire pour obtenir une théorie des cardinaux satisfaisante.
Une petite remarque concernant cette preuve: vous remarquerez que vers la fin j'ai fait usage du schéma d'axiomes de remplacement, présent en théorie ZF, mais pas en théorie Z. On se dit alors qu'on peut peut-être s'en passer, mais il s'avère que non, ce théorème n'est pas démontrable en théorie Z (sa négation non plus, en tout cas si ZF est cohérente), et donc que ce schéma d'axiomes est nécessaire à cette preuve (pour celleux qui connaissent un peu déjà les ordinaux et la théorie des modèles, ou celleux qui relisent ce document une deuxième fois peuvent s'amuser à prouver que ce théorème est faux dans certains modèles de Z -et donc pas de ZF-, comme par exemple $V_{\omega+\omega}$. Je vous invite à revenir voir cette preuve et y réfléchir quand vous aurez fini de lire cette introduction).
On rentre maintenant dans les théorèmes qui justifient l'importance que j'accorde au théorème 2. Ils sont un peu "controversés", dans le sens où ils sont équivalents au fameux axiome du choix, mais n'importe quel mathématicien a de bonnes justifications pour utiliser cet axiome (notamment qu'il est irréfutable à propos des autres; et que finalement s'il choque l'intuition, ce n'est pas grave, c'est souvent le cas des mathématiques). Je l'énonce ici sous une forme parmi tant d'autres :
Axiome du choix: Quel que soit l'ensemble E, il existe une fonction h de $\mathcal{P}(E) - \{\emptyset\}$ dans E telle que $\forall X \in \mathcal{P}(E) - \{\emptyset\}, h(X) \in X$.
Le théorème qui suit (nommé "lemme", pour une raison historique; en plus il est mal attribué, mais je ne rentre pas dans les détails) est d'une extrême importance. C'est lui qui permet toutes les constructions "bizarres" dont on peut entendre parler (par exemple celle d'un ensemble non Lebesgue-mesurable), et il facilite grandement la vie dans de nombreux cas. Il est en fait équivalent à

l'axiome du choix (modulo les axiomes de ZF), mais on fera cette preuve plus tard. Il y a de nombreuses preuves de ce lemme, j'en propose ici une, mais rien ne vous empêche d'en chercher d'autres. On y utilise un procédé dont je reparlerai un peu après.

Lemme de Zorn (ou Théorème de Kuratowski-Zorn): Tout ensemble ordonné $(E, <)$ inductif admet un élément maximal.

Preuve: Avant d'en faire la preuve, il faut comprendre ce qu'il signifie. Un ensemble inductif est un ensemble dont toute chaîne (sous-ensemble totalement ordonné) admet un majorant. Un élément maximal est un élément qui n'est inférieur strictement à aucun élément.

On se donne donc $(E, <)$ un ensemble inductif, et on suppose qu'il n'a aucun élément maximal. En d'autres termes, pour tout $x \in E$, $M(x) = \{y \in E, y > x\}$ est non vide. On se donne h une fonction de choix telle que dans l'énoncé de l'axiome du choix. Pour tout $x \in E$, $h(M(x))$ est un majorant strict de x.Pour chaque chaîne T de E, on a alors $h(M(h(\{x \in E \mid x \text{ majore T }\}))) = m(T)$ un majorant strict de T. On utilise ici un procédé qui s'appelle "définition par récurrence transfinie": on définit l'image du premier ordinal $\emptyset$, puis on décrit comment obtenir l'image de α connaissant les images des $\beta \in \alpha$. C'est justifié, intuitivement (je n'entrerai oas trop dans les détails mais j'en parlerai après cette preuve) par le fait que si on prend le plus petit ordinal α tel que l'image de α n'est pas définie, alors son image est définie, et donc on a une contradiction. Ce sera précisé plus tard, et cette preuve vous permet d'avoir une idée de la méthode. E est non vide car inductif (preuve facile à faire), soit donc $x \in E$. On note $x = J(\emptyset)$. Soit α un ordinal quelconque, on suppose que $J(\beta)$ est définie pour tout $\beta \in \alpha$, de manière à ce que $\{J(\beta) \mid \beta \in \alpha\}$ soit une chaîne, et pour tout $\beta \in \alpha$, $J(\beta)$ soit un majorant strict de la chaîne $\{J(\gamma) \mid \gamma \in \beta\}$. On pose alors $J(\alpha) = m(\{J(\beta) \mid \beta \in \alpha\})$, de manière à ce que la propriété soit conservée. On définit ensuite la relation fonctionnelle $\alpha = r(x)$ ainsi: "α est un ordinal, $x \in E$, et $x = J(\alpha)$". Il s'agit bien d'une relation fonctionnelle car par construction les $J(\alpha)$ sont deux à deux distincts. De plus, elle a pour image directe la collection des ordinaux. De même que pour le théorème 2, on applique ensuite le schéma d'axiomes de remplacement, et on conclut qu'il existe un ensemble des ordinaux. C'est une contradiction. Donc E admet un élément maximal.

Revenons un peu sur la preuve. A première vue, tout semble correct sauf le passage où l'on définit J, qui semble farfelu. En fait, on peut le justifier de la manière suivante : soit $z = H(f)$ une relation fonctionnelle de domaine contenant la collection des fonctions f de domaine un ordinal α qui vérifient pour tout $\beta \in \alpha$ $f(\beta) = H(f_{|\beta})$. Alors il existe une unique relation fonctionnelle $z = F(\alpha)$ de domaine la collection des ordinaux telle que pour tout β, $F(\beta) = H(F_{|\beta})$. C'est un peu technique, mais la démonstration n'est pas si compliquée, et c'est plutôt intuitif comme je l'ai déjà expliqué. La relation $z = F(\alpha)$ s'écrit alors "il existe une fonction f de domaine α vérifiant pour tout $\beta \in \alpha$ $f(\beta) = H(f_{|\beta})$, et $z = H(f)$ ". On prouve assez facilement que cette relation est bien fonctionnelle, et on montre qu'elle a bien pour domaine la collection des ordinaux. Cela repose entièrement sur le bon ordre qui existe sur la collection des ordinaux. Je laisse la preuve aux lecteurices, et je les laisse adapter ce théorème pour vérifier que J est bien définie. Je ne l'ai pas précisé avant car il me semblait essentiel d'avoir vu un exemple avant d'avoir vu la théorie, ici un peu technique et absconse si on ne comprend pas ce qu'il

se passe. Vous verrez qu'on réutilisera cette technique.
Ce théorème est très important, et sera utilisé souvent. Lorsque son utilisation sera évidente, et lorsqu'assez d'exemples auront été vus, je n'écrirai plus que "on voit grâce à un Zornage que", ou "d'après le théorème de Zorn..." et je laisserai aux lecteurices le soin de vérifier les détails. Maintenant vient un théorème extrêmement important, qui est à la base de la théorie des cardinaux, et dont la démonstration repose essentiellement sur le lemme de Zorn (en tout cas celle qu'on propose ici):
Théorème de Zermelo: Tout ensemble peut être bien ordonné, i.e. pour tout ensemble X, il existe une relation $<$ telle que $(X, <)$ soit un ensemble bien ordonné.
Preuve: Soit X un ensemble quelconque. On s'apprête à "effectuer un Zornage", étudiez bien cette preuve et essayez de repérer les étapes essentielles. On note $\Omega = \{(F, <_F) \mid F \subset X, <_F \text{ est un bon ordre sur } F\}$. Il est évident que Ω est non vide (il suffit de prendre $(\emptyset, \emptyset)$ qui appartient bien à Ω). On ordonne Ω de la manière suivante : $(F, <_F) \preceq (G, <_G)$ si et seulement si ($F \subset G$, F est un segment initial de G pour $<_G$, et la restriction de $<_G$ à F est $<_F$). Il est alors évident que $\preceq$ est une relation d'ordre (large- la preuve est laissée aux lecteurices pour qui ça ne semble pas évident). On va alors montrer que $(\Omega, \preceq)$ est inductif, au sens donné dans la preuve du lemme de Zorn. Soit donc Γ une chaîne de Ω, c'est-à-dire un sous-ensemble de Ω totalement ordonné par $\preceq$. Montrons que Γ est majoré. Considérons le sous-ensemble de X défini ainsi: $\Delta = \bigcup_{(F,<_F)\in\Gamma} F$. On définit dessus la relation $<_\Delta$ ainsi: $x <_\Delta y$ si et seulement si il existe $(F, <_F) \in \Gamma$ tel que $x, y \in F$ et $x <_F y$. On va montrer que $<_\Delta$ est un bon ordre. Montrons premièrement que c'est un ordre strict (on a utilisé la notation $<_\Delta$ avant de prouver que c'est un ordre strict, ce n'est pas très grave, simplement il ne faut pas oublier de montrer que c'en est un). Montrons donc l'antiréflexivité. Soit $x \in \Delta$. Supposons $x <_\Delta x$. Alors il existe $(F, <_F) \in \Gamma$ tel que $x <_F x$. Ceci est contradictoire car $<_F$ est un ordre strict sur F. On passe alors à la transitivité: Soit $x, y, z \in \Delta$, on suppose $x <_\Delta y$ et $y <_\Delta z$. On se donne alors $(F, <_F), (G, <_G) \in \Gamma$ tels que $x, y \in F$, $y, z \in G$, $x <_F y$ et $y <_G z$. N'oublions pas que $\preceq$ est total sur Γ, et donc $(F, <_F) \preceq (G, <_G)$ ou $(G, <_G) \preceq (F, <_F)$. Supposons sans perte de généralité (la preuve est la même en supposant l'autre cas, vous verrez) que $(F, <_F) \preceq (G, <_G)$. Alors $F \subset G$ et $<_F$ est la restriction à F de $<_G$. En particulier, $x \in G$ et $x <_G y$. Par transitivité de $<_G$, on a donc $x <_G z$. Donc $x <_\Delta z$. DOnc $<_\Delta$ est transitive: c'est un ordre strict. Il ne reste plus qu'à montrer que cet ordre est bon. Soit donc $T \subset \Delta$ non vide. Soit $x \in T$. $x \in \Delta$ donc il existe $(F, <_F) \in \Gamma$ tel que $x \in F$. On considère $F \cap T$. Cet ensemble est non vide car il contient x. De plus c'est un sous-ensemble de F, et il admet donc un minimum pour $<_F$. Soit m ce minimum. On va montrer que m est un minimum de T pour $<_\Delta$: Soit $r \in T$ quelconque. On suppose que $r <_\Delta m$. Par définition, il existe donc $(G, <_G) \in \Gamma$ tel que $r, m \in G$ et $r <_G m$. Or $\preceq$ est total sur Γ donc $(F, <_F) \preceq (G, <_G)$ ou $(G, <_G) \preceq (F, <_F)$. Dans le premier cas, F est un segment initial de G donc $r \in F$ et $r <_F m$, et donc $r \in F \cap T$, ce qui contredit le fait que m est un minimum de cet ensemble. Dans le deuxième cas, $G \subset F$ et donc $r \in F$, et on a aussi $r \in F \cap T$, $r <_F m$. C'est donc aussi contradictoire. Donc m est bien le minimum de T: $<_\Delta$ est un bon ordre sur Δ. Donc $(\Delta, <_\Delta) \in \Omega$. Il

ne reste plus qu'à montrer que pour tout $(F, <_F) \in \Gamma$, $(F, <_F) \preceq (\Delta, <_\Delta)$. La première condition, à avoir $F \subset \Delta$ est évidente, et la troisième, par définition de $<_\Delta$, l'est aussi. Il ne reste en fait qu'à vérifier que F est un segment initial de $(\Delta, <_\Delta)$. Soit donc $x \in \Delta$, et $y \in F$ tels que $x <_\Delta y$. On va montrer que $x \in F$, et on aura conclu cette partie de la preuve. Soit $(G, <_G) \in \Gamma$ tel que $x <_G y$. Si $(G, <_G) \preceq (F, <_F)$, on a alors évidemment $x \in F$. Sinon, par totalité de l'ordre sur Γ, on a $(F, <_F) \preceq (G, <_G)$. De ce fait, F est un segment initial de $(G, <_G)$ et donc $x \in F$. Dans tous les cas, $x \in F$, donc cela permet de conclure que $(\Delta, <_\Delta)$ est un majorant de Γ: $(\Omega, \preceq)$ est inductif. D'après le lemme de Zorn, il admet donc un élément maximal. Soit $(F, <_F)$ cet élément maximal. Supposons que $F \neq X$ (si $F = X$, on a gagné car alors $<_F$ est un bon ordre sur X et le théorème est prouvé). Soit donc $x \in X - F$. On considère l'ensemble ordonné : $(F \cup \{x\}, <)$ ou $<$ est l'ordre tel que sa restriction à F soit $<_F$ et tel que $y < x$ pour tout $y \in F$. Il est alors aisé de vérifier que $<$ est un bon ordre sur $F \cup \{x\}$, et que F est un segment initial de $(F \cup \{x\}, <)$. Donc on a bien $(F, <_F) \prec (F \cup \{x\}, <)$, ce qui contredit la maximalité de $(F, <_F)$. Donc $F = X$, et il existe un bon ordre sur X.
Après avoir vu cette démonstration extrêmement peu constructive, on se rend compte que ledit bon ordre, on ne le connaît pas ! C'est ce que les gens veulent dire lorsqu'ils disent que l'axiome du choix permet d'obtenir des objets qu'on ne peut pas exhiber, car son utilisation repose essentiellement sur des raisonnements par l'absurde (lemme de Zorn). Ce théorème peut-il donc être démontré plus constructivement, sans axiome du choix ? Non: il lui est en fait équivalent, et c'est la petite parenthèse qu'on se permet ici:
Théorème 3: Modulo la théorie ZF, le théorème de Zermelo (et donc le lemme de Zorn) est équivalent à l'axiome du choix.
Preuve: La preuve du théorème de Zermelo que l'on vient de faire justifie (AC $\Longrightarrow$ Thm de Zermelo). On va désormais prouver l'axiome du choix dans la théorie ZF + théorème de Zermelo (l'axiome du choix est donc un théorème dans cette théorie, et le théorème de Zermelo un axiome). Soit donc E un ensemble quelconque. Soit $<$ un bon ordre sur E. L'application h que l'axiome du choix recquiert est en fait l'application qui à une partie non vide de E envoie son minimum. La propriété $h(X) \in X$ est alors évidente.

Revenons maintenant à nos moutons. Nous allons voir dans la prochaine sous-section en quoi le théorème de Zermelo, et finalement le théorème 2 constituent la fondation de la théorie des cardinaux.

3.3 Cardinaux

Avant de parler de cardinaux, on va introduire un peu la notion en expliquant en quoi elle est intéressante. Les quelques exemples d'ordinaux que j'ai donnés révèlent quelque chose qui n'est pas anodin. A la fin de cette explication, retournez les voir, et vous comprendrez. Soit α un ordinal.Existe-t-il des ordinaux plus grands? Evidemment, sinon la collection des ordinaux serait un ensemble, ce n'est pas possible. Donc il existe un plus petit ordinal plus grand que α. Un tel ordinal β vérifie : $\alpha \in \beta$ et $\alpha \subset \beta$. Le plus petit ensemble vérifiant cela est $\alpha \cup \{\alpha\}$. La vérification que $\alpha \cup \{\alpha\}$ est un ordinal est laissée aux lecteurices, elle est évidente. On note cet ordinal $\alpha + 1$ (ce n'est pas non plus anodin, vous comprendrez plus tard), on l'appelle le successeur. Voici donc les premiers

ordinaux : $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, ...$ On note généralement ceux-ci $0, 1, 2, 3, ...$, ce n'est pas non plus anodin. Maintenant, on peut se dire que tous les ordinaux sont le $\alpha + 1$ d'un α donné. Mais ce n'est pas la cas, sous-réserve de l'axiome de l'infini. Celui-ci dit "il existe un ordinal limite". Qu'est-ce qu'un ordinal limite? C'est un ordinal qui ne s'écrit $\alpha + 1$ pour aucun α. Etant donné qu'il en existe un d'après l'axiome de l'infini, on peut considérer le plus petit ordinal limite, généralement noté ω. En fait cet ordinal est l'ensemble $\{0, 1, 2, 3, ...\}$. Plus précisément: un ordinal α est dit fini s'il est successeur ou nul et si tous ses éléments sont des ordinaux successeurs. Sinon, il est dit infini. Un ordinal limite et donc nécessairement infini. De ce fait ω est l'ensemble de tous les ordinaux finis: la preuve est la suivante : si $n \in \omega$, n n'est pas limite et ne contient aucun ordinal limite, sinon cela contredirait la définition de ω. Donc n est fini. Réciproquement, si n est un ordinal fini, $\omega \notin n$ et $\omega \neq n$, donc $n \in \omega$. Pourquoi a-t-on besoin d'un axiome pour dire que ω existe ? Parce que sinon, l'expression "l'ensemble des ordinaux finis" n'a pas de sens: dans certains modèles de ZF sans l'axiome de l'infini, tous les ordinaux sont finis (pour celleux qui connaissent, ou celleux qui reviennent voir, V_ω par exemple ne contient que des ordinaux finis, et "à ses yeux" ω est la classe des ordinaux, ce n'est pas un ensemble). Donc l'axiome de l'infini est nécessaire. En présence de celui-ci, on peut donc parler de ω. Le successeur de ω, $\omega + 1$ est aussi infini, car il contient ω. Chose importante à remarquer : ω et $\omega + 1$ "ont le même nombre d'éléments": il existe une bijection de l'un sur l'autre. En reprenant la caractérisation de ω, on vous laisse la trouver. Ceci est plus général: pour tout ordinal infini successeur, il existe un ordinal strictement inférieur avec qui il est en bijection, cela est évident du fait que ω et $\omega + 1$ soit en bijection (une fois que vous aurez la preuve de ce fait, l'autre en découlera immédiatement). Ces remarques sur les ordinaux infinis, représentés notamment par le fameux paradoxe de l'hôtel de Hilbert, nous amène à poser la définition suivante :

Cardinal: Un cardinal est un ordinal qui n'est en bijection avec aucun ordinal strictement inférieur.

$\omega + 1$ n'est donc pas un cardinal, et un cardinal infini est nécessairement un ordinal limite. De plus deux cardinaux distincts ne sont pas en bijection, donc les cardinaux ne sont pas redondants. De plus, pour tout ordinal il existe un cardinal qui lui est équipotent: soit α un ordinal, on considère le sous-ensemble de $\alpha + 1$ défini ainsi: $\{x \in \alpha + 1 \mid x \text{ est en bijection avec } \alpha\}$. Cet ensemble est non vide car il contient α. Donc il admet un minimum: ce minimum est l'unique cardinal en bijection avec α. On peut donc noter pour tout ordinal α $card(\alpha)$ cet unique cardinal. Où interviennent le théorème de Zermelo et le théorème 2 là-dedans ? Quelle est cette théorie des cardinaux ? Tout cela est expliqué dans le théorème suivant :

Théorème 3: Pour tout ensemble X il existe un unique cardinal λ équipotent à X.

Preuve: L'unicité provient de la remarque faite précédemment, et de la transitivité de la relation d'équipotence. Montrons maintenant l'existence. D'après le théorème de Zermelo, il existe une relation $<$ telle que $(X, <)$ soit un ensemble bien ordonné. D'après le théorème 2, il existe un ordinal α isomorphe (en particulier en bijection avec) à $(X, <)$. Donc α est équipotent à X, et donc $card(\alpha)$, qui est un cardinal, est équipotent à X. Cela conclut la preuve du théorème.

Dans cette preuve on se rend compte en fait que l'ordinal α initialement associé à X est "arbitraire", car il dépend aussi de l'ordre $<$, qu'on ne sait pas décrire,

on en connaît juste l'existence. Le passage au cardinal assure que l'ordinal associé est "canonique". Ce théorème, extrêmement important, se déduit donc (modulo ZF) de l'axiome du choix. Il s'avère qu'il y est en fait équivalent. La preuve est laissée aux lecteurices (penser au théorème de Zermelo). L'énoncé suivant, qui semble moins fort, est aussi équivalent à l'axiome du choix : "pour tout ensemble X, il existe un ordinal α et une injection de X dans α", de même que le suivant "pour tout ensemble X il existe un ordinal α et une surjection de α sur X".
La remarque suivante est évidente: si λ et μ sont deux cardinaux, l'un s'injecte dans l'autre (penser au lemme 6 et à la transitivité des ordinaux). De ce fait évident découle le suivant, moins évident, équivalent à l'axiome du choix: si X et Y sont deux ensembles, l'un s'injecte dans l'autre. Nous allons montrer l'équivalence dans ce qui suit, et par là-même montrer qu'il existe une infinité de cardinaux :
Proposition 3: On suppose les axiomes de ZF. Supposons de plus que pour deux ensembles X et Y, l'un s'injecte dans l'autre. Alors l'axiome du choix est vérifié.
Preuve: Soit X un ensemble quelconque. On va montrer qu'il s'injecte dans un ordinal (en fait un cardinal), ce qui suffira d'après les remarques qui précèdent. On considère la collection $\Gamma(X)$ ainsi : un ordinal α est dans $\Gamma(X)$ si et seulement si il s'injecte dans X. On va montrer que $\Gamma(X)$ est un ensemble, et un ordinal (et un cardinal). Premièrement, c'est un ensemble: pour tout ordinal α et fonction f injective de α dans X, l'image de α par f est un sous-ensemble de X qui peut être bien ordonné. Réciproquement, à tout couple $(Y, <)$ où Y est un sous-ensemble de X et $<$ un bon ordre sur Y, on peut associer un unique ordinal d'après le théorème 2 et la proposition 1. Donc la relation $\alpha = J(Y, <)$ qui se définit par "α est un ordinal, Y un sous-ensemble de X, $<$ un bon ordre sur Y et α est isomorphe à $(Y, <)$" est une relation fonctionnelle, de domaine un sous-ensemble de $\mathcal{P}(X) \times \mathcal{P}(X \times X)$, et son image directe est donc un ensemble par le schéma s'axiomes de remplacement. Or son image directe est précisément la collection des ordinaux qui s'injectent dans X. Donc cette collection est un ensemble. Ses éléments sont des ordinaux, et il est évidemment transitif: c'est donc un ordinal. De plus, $\Gamma(X)$ ne s'injecte pas dans X, sinon on aurait $\Gamma(X) \in \Gamma(X)$, ce qui est impossible pour un ordinal. Donc X s'injecte dans $\Gamma(X)$. Cela conclut la preuve. Montrons simplement de plus que $\Gamma(X)$ est un cardinal: soit $\alpha \in \Gamma(X)$. Supposons que $\Gamma(X)$ sinjecte dans α. Alors $\Gamma(X)$ s'injecte dans X, ce qui est contradictoire. Donc $\Gamma(X)$ est un cardinal. Si X est un ordinal, $\Gamma(X)$ est le plus petit cardinal supérieur à $card(X)$ (vérification évidente). Pour un cardinal κ, le plus petit cardinal supérieur à κ est donc $\Gamma(\kappa)$, et il est généralement noté κ^+.
De cette preuve on déduit qu'il existe une infinité de cardinaux. On a à nouveau utilisé une instance du schéma d'axiomes de remplacement: ce théorème est faux dans la théorie Z (celui qui connait peut à nouveau considérer $V_{\omega+\omega}$). Une autre preuve, qui nécessite cependant l'axiome du choix, est le théorème de Cantor.
Théorème de Cantor: Soit X un ensemble. Il n'existe pas de surjection de X sur $\mathcal{P}(X)$.
Preuve: Soit f une fonction de X sur $\mathcal{P}(X)$. Montrons que f n'est pas surjective. Soit $M = \{x \in X \mid x \notin f(x)\}$. On a évidemment, par définition $M \subset X$ et donc $M \in \mathcal{P}(X)$. Montrons que M n'a pas d'antécédent par f: soit $y \in X$.

Si $y \in f(y)$ alors $y \notin M$, donc $M \neq f(y)$. Si $y \notin f(y)$, $y \in M$ donc $M \neq f(y)$. Donc M n'admet pas d'antécédent par f: f n'est pas surjective.
De ce fait, si on note pour tout ensemble X $card(X)$ l'unique cardinal dont l'existence est donnée par le théorème 3, on a $card(X) < card(\mathcal{P}(X))$. De ce fait vient la question naturelle suivante: si κ est un cardinal, a-t-on $\kappa^+ = card(\mathcal{P}(\kappa))$? Comme on le verra plus tard, c'est évidemment non pour les cardinaux finis, mais pas pour les cardinaux finis. L'hypothèse que la réponse est oui pour $\kappa = \omega$ est l'hypothèse du continu (on montrera plus tard que ω est un cardinal, quoique ça semble évident), et l'hypothèse que la réponse est oui pour tout cardinal κ est l'hypothèse généralisée du continu (respectivement HC ou CH en anglais, et HGC ou GCH en anglais). Cette hypothèse est indécidable (on ne le montrera pas ici: c'est hautement non trivial). Après avoir vu toute cette théorie, un peu de pratique: à quoi ressemble un cardinal ? La réponse se décompose en deux parties:
Proposition 4: Tout ordinal fini est un cardinal.
Preuve: La preuve repose sur une récurrence (usuelle), technique qu'on justifiera après la preuve. A nouveau, il vaut mieux la voir en action avant de voir la théorie. 0 est évidemment cardinal, car il n'existe aucun ordinal inférieur à 0. De même, il n'existe évidemment aucune bijection de 1 sur 0, donc 1 est un cardinal. Soit n un ordinal fini. On suppose que n est un cardinal. Montrons que $n+1$ en est aussi un. Supposons qu'il existe un ordinal $m < n+1$ et une bijection f de $n+1$ sur m. m est évidemment non vide, et comme il est fini, il s'écrit $p+1$. Supposons dans un premier cas que $f(n) = p$. Alors la fonction f restreinte à n est une bijection de n sur p, et $p < n$, ce qui contredit l'hypothèse de récurrence. Si $f(n) = j < p$, on compose f avec la transposition qui échange j et p, et on obtient une bijection f' de $n+1$ sur $p+1$ telle que $f'(n) = p$. On fait alors le même raisonnement. Ainsi, si n est un cardinal fini, alors $n+1$ aussi.
Après avoir vu cette preuve, le.a lecteurice n'a peut-être pas confiance en elle, car j'utilise une récurrence sans l'avoir justifiée. Je vais donc la justifier ici: soit P un propriété (un prédicat du premier ordre à une variable libre). On suppose que $P(0)$ est vraie et que pour tout ordinal fini n, $P(n) \implies P(n+1)$. Alors supposons qu'il existe un ordinal fini m tel que $\neg P(m)$. Soit alors m_0 le minimum de $\{n \in m+1 \mid \neg P(n)\}$. m_0 est nécessairement non nul, puisque $P(0)$ est vérifiée, et il est fini puisqu'il appartient à $m+1$, qui est fini. Donc par définition, m_0 est successeur, et $m_0 = k_0 + 1$. De plus, $k_0 < m_0$ donc par définition, $P(k_0)$ est vérifiée. De plus, $P(k_0) \implies P(k_0 + 1 = m_0)$ est aussi vérifiée par supposition, donc $P(m_0)$ l'est , ce qui est une contradiction. Donc pour tout ordinal fini n, $P(n)$ est vérifiée.
Il s'agit là d'une récurrence simple, et on verra plus tard qu'on peut faire des récurrences un peu plus générales, mais de ce fait plus compliquées, sur tous les ordinaux. Maintenant que cela est justifié, on sait que tous les ordinaux finis sont des cardinaux. Qu'en est-il des ordinaux infinis ? On a vu que $\omega + 1$ n'est pas un cardinal. De plus, on déduit immédiatement de la proposition 4 que ω est un cardinal. Donc certains ordinaux infinis sont des cardinaux, et pas d'autres. Pour les étudier, il faut d'abord comprendre les ordinaux infinis, et en particulier les ordinaux limites.
Lemme 8: Soit X un ensemble d'ordinaux. Alors X a une borne supérieure, c'est-à-dire un ordinal qui majore X, et tel que tout ordinal qui majore X soit plus grand que cet ordinal.

Preuve: Soit X un tel ensemble. On considère alors $\alpha = \bigcup_{\beta \in X} \beta$. Montrons premièrement que α est un ordinal. Ses éléments sont des éléments d'ordinaux, donc des ordinaux, et si $\gamma \in \alpha$, il existe $\beta \in X$ tel que $\gamma \in \beta$ et donc $\gamma \subset \beta$ et donc $\gamma \subset \alpha$: α est un ensemble transitif d'ordinaux, c'est donc un ordinal. Il est évident que α majore X, d'après le lemme 5. Il faut alors montrer que c'est la borne supérieure. Soit γ un autre majorant de X. Pour tout $\beta \in X$, $\beta \subset \gamma$. Donc $\alpha = \bigcup_{\beta \in X} \beta \subset \gamma$, donc d'après le lemme 5, $\alpha \leq \gamma$. α est donc bien la borne supérieure de X.
Si α est un ordinal limite, α majore α, et de plus si on se donne $\beta < \alpha$, $\beta+1 \in \alpha$ (en effet sinon $\alpha \leq \beta$, ce qui n'est pas possible), donc tout ordinal inférieur à α ne majore pas α. Donc α est la borne supérieure de α, et donc $\alpha = \bigcup_{\beta \in \alpha} \beta$. Cette égalité caractérise les ordinaux limites.
Ce lemme nous donne une manière de construire de nouveaux ordinaux: on prend un ensemble X d'ordinaux qui n'a pas de maximum, et on considère sa borne supérieure: c'est un ordinal qui n'appartient pas à X. On va maintenant voir d'autres manières de la faire pour notamment montrer un résultat qui peut ne pas apparaître comme évident: il existe des ordinaux limites qui ne sont pas des cardinaux. Repensons aux dernières constructions qu'on a vues, à la fin de la section sur les bons ordres, la somme et le produit de bons ordres. Soit donc deux ordinaux α et β vus comme ensembles bien ordonnés. On a vu que $\alpha \oplus \beta$ est un ensemble bien ordonné, de même que $\alpha \otimes \beta$. D'après le théorème 2, ils sont donc isomorphes à deux uniques ordinaux, qu'on note désormais $\alpha+\beta$ et $\alpha\beta$ (on omet le signe $\times$ qu'on pourrait choisir afin de ne pas confondre avec le produit cartésien). D'après ce qui avait été dit à la fin de la section sur les bons ordres, et d'après la proposition 1, on a : $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$, $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ et $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$. Attention à l'ordre pour la distributivité, car ni $+$, ni la multiplication ne sont commutatives en général. Cependant (exercice laissé aux lecteurices - utilisez une récurrence), les deux sont commutatives sur les ordinaux finis. De plus pour tout ordinal, $\alpha 1 = 1\alpha$ et $\alpha + 0 = 0 + \alpha$. Maintenant le.a lecteurice comprendra la notation $\alpha + 1$ pour l'ordinal successeur de α, je lui laisse démontrer que la notation est bien définie, i.e. que les deux coïncident. Petit exercice laissé aux lecteurices : montrer que $\omega + \omega$ est un ordinal limite, mais que ce n'est pas un cardinal, et donc que les deux notions sont réellement distinctes; puis montrer qu'un cardinal infini est nécessairement limite (on a donc une inclusion stricte).
Maintenant que l'on sait à peu près à quoi ressemblent les ordinaux limites, on peut s'attaquer à savoir à quoi ressemblent les cardinaux. La réponse est plutôt simple: essentiellement, la collection des cardinaux est "isomorphe" à celle des ordinaux. Les résultats suivants expliquent ce que je veux dire par là.
Lemme 9: Si X est un ensemble de cardinaux, sa borne supérieure est un cardinal.
Preuve: Soit X un ensemble de cardinaux. Si X a un maximum λ, sa borne supérieure est λ, et c'est un cardinal. Sinon, sa borne supérieure est $\bigcup_{\lambda \in X} \lambda$. D'après ce qu'on a vu avant, c'est un ordinal, que l'on note μ. μ ne s'injecte dans aucun des éléments de X : en effet soit $\kappa \in X$ tel que μ s'injecte dans κ. Soit $\delta \in X, \delta > \kappa$ (un tel δ existe car X n'a pas de maximum). Alors δ s'injecte

dans μ et donc dans κ, ce qui contredit le fait que c'est un cardinal. Soit maintenant γ un ordinal tel que μ s'injecte dans γ. Alors $card(\gamma)$ est un cardinal tel que μ s'injecte dedans. Donc chaque $\lambda \in X$ s'injecte dans $card(\gamma)$, et comme il s'agit de cardinaux, chaque $lambda \in X$ est inférieur à $card(\gamma)$ et donc $card(\gamma)$ majore X. Par définition de borne supérieure, on a donc $\mu \leq card(\gamma) \leq \gamma$. Donc μ est bien un cardinal.
De ce fait, on déduit la proposition suivante:
Proposition 5: Il n'existe pas d'ensemble des cardinaux.
Preuve: Soit X l'ensemble des cardinaux. Sa borne supérieure est un cardinal, et donc appartient à X. C'est donc un maximum. Soit κ ce maximum. κ^+ est un cardinal supérieur strictement à κ, il n'appartient donc pas à X, ce qui contredit la définition de X.
Cette propositon va nous permettre de déduire la "structure" de la classe des cardinaux infinis, une sorte de classification.
Théorème 4: Il existe une relation fonctionnelle $\lambda = \aleph_\alpha$ de domaine la collection de ordinaux et d'image la collection des cardinaux telle que : pour tout ordinal α, $\aleph_\alpha \geq \alpha$, si $\alpha < \beta$, $\aleph_\alpha < \aleph_\beta$; pour tout cardinal infini λ il existe α tel que $\lambda = \aleph_\alpha$.
Preuve: On applique la même technique que pour le lemme de Zorn, c'est-à-dire une définition par récurrence transfinie: on pose $\aleph_0 = \omega$, $\aleph_{\alpha+1} = \aleph_\alpha^+$ et si α est limite, donc $\alpha = \bigcup_{\beta\in\alpha} \beta$, alors $\aleph_\alpha = \bigcup_{\beta\in\alpha} \aleph_\beta$. Il est évident que cette relation fonctionnelle est croissante, sa stricte croissance provient du fait que si α est limite, alors $\{\aleph_\beta \mid \beta \in \alpha\}$ n'a pas de maximum, car pour tout $\beta \in \alpha$, $\beta + 1 \in \alpha$. On va appliquer un principe similaire à la définition par récurrence transfinie pour vérifier que $\aleph_\alpha \geq \alpha$ pour tout ordinal α: la preuve par récurrence transfinie. Je laisse aux lecteurices le soin de justifier que cette technique est valable, ça devrait être facile. $\aleph_0 \geq 0$, c'est évident. Soit α un ordinal quelconque. Si $\alpha = \beta + 1$, alors $\aleph_\alpha = \aleph_\beta^+$, or $\aleph_\beta \geq \beta$ par hypothèse de récurrence, donc $\aleph_\beta^+ \geq \beta + 1 = \alpha$. SI α est limite, alors pour tout $\beta \in \alpha$, $\aleph_\beta \geq \beta$, donc un majorant de $\{\aleph_\beta \mid \beta \in \alpha\}$ est aussi un majorant de $\{\beta \mid \beta \in \alpha\}$. Donc $\bigcup_{\beta\in\alpha} \aleph_\beta \geq \bigcup_{\beta\in\alpha} \beta$. Donc $\aleph_\alpha \geq \alpha$. Ce qu'on a prouvé est la chose suivante : notons $P(\alpha)$ la propriété $\aleph_\alpha \geq \alpha$; alors pour tout ordinal α, $(\forall \beta \in \alpha, P(\beta)) \implies P(\alpha)$. De ceci on déduit "pour tout ordinal α, $P(\alpha)$". La preuve est presque la même que pour la récurrence simple, je laisse aux lecteurices le soin de vérifier que cette déduction est valable.
Nous avons donc prouvé la stricte croissance de cette relation fonctionnelle, et sa supériorité à l'identité. Il ne reste plus qu'à montrer le dernier point. Soit λ un cardinal infini quelconque. On sait que $\aleph_\lambda \geq \lambda$. On peut donc considérer α le plus petit ordinal tel que $\aleph_\alpha > \lambda$. Montrons que α est un ordinal successeur: sinon on aurait $\bigcup_{\beta\in\alpha} \aleph_\beta > \lambda$, donc par définition de borne supérieure, il existe $\beta \in \alpha$ tel que $\aleph_\beta > \lambda$, ce qui contredit la définition de α. Donc $\alpha = \delta + 1$, et on a $\aleph_\alpha > \lambda \geq \aleph_\delta$. $\aleph_\alpha$ est le plus petit cardinal supérieur strictement à $\aleph_\delta$, donc $\lambda = \aleph_\delta$, cela conclut la preuve.
Ainsi, si on voit $\aleph$ comme une "méta-fonction", elle est injective (par stricte croissance), surjective (d'après le théorème 4) et strictement croissante de On sur Cn (on note maintenant On la collection des ordinaux et Cn celle des car-

dinaux infinis - lorsqu'on notera $x \in On$, où $On(x)$, il ne faut pas oublier que ce sont des abréviations pour "x est un ordinal", rien de plus). De ce fait, on peut qualifier $\aleph$ d'"isomorphisme de On sur Cn". On sait donc à quoi ressemblent les cardinaux infinis: ils ressemblent beaucoup aux ordinaux, et on sait à quoi ressemblent les cardinaux finis: ce sont exactement les ordinaux finis.
Souvent, quand on considère un cardinal en tant qu'ordinal, c'est-à-dire en oubliant pas que c'est un ensemble bien ordonné, on le notera plus λ, tandis que si on le considère en tant que "taille d'ensemble", on le notera plus $\aleph_\alpha$. Ce sont simplement des soucis d'écriture, mais c'est la convention. Par exemple, "ordinalement", ω et $\omega+1$ sont différents, mais "cardinalement", ils sont pareils (tous deux appartiennent à la "classe $\aleph_0$). Utilisons un abus de langage, et notons 2^X l'ensemble des parties de X, et si α est un cardinal, notons 2^α le cardinal de cet ensemble des parties. C'est évidemment un abus de langage car, sauf pour quelques cardinaux finis, 2^α n'est pas un ordinal. L'hypothèse du continu s'exprime alors comme suit:
Hypothèse du continu: $\aleph_1 = 2^{\aleph_0}$.
Hypothèse généralisée du continu: Pour tout ordinal α, $\aleph_{\alpha+1} = 2^{\aleph_\alpha}$.
Vous vous demandez sûrement pourquoi je parle ici de ω, alors qu'en général, quand HC est présentée, elle concerne $\mathbb{N}$. J'en parlerai dans la dernière section de cette introduction, dédiée aux rapprochements ordinaux/théorie des modèles.
Avant de conclure cette section sur les cardinaux, je présente maintenant un outil souvent utilisé: l'arithmétique des cardinaux.
Qu'entends-je par ce nom barbare ? Nous avons présenté précédemment des opérations (addition et multiplication) sur les ordinaux, mais ces opérations "gardaient en tête" le bon ordre. Comme je l'ai dit précédemment, quand on pense aux cardinaux, on oublie l'ordre, ce qui nous intéresse est la taille. Cependant il faut faire attention, car les opérations (addition et multiplication) que je vais présenter ici son différentes des précédentes, mais notées de la même manière. Donc à chaque fois qu'on écrit une telle opération, il faut préciser si on parle de la somme (resp. produit) ordinale, ou cardinale.
Etant donnés deux cardinaux λ et μ, on appelle somme cardinale et on note $\lambda + \mu$ le cardinal de la somme ordinale $\lambda + \mu$. On définit de même le produit cardinal.
L'arithmétique cardinale est l'étude des calculs d'opérations sur des cardinaux.
Proposition 6: (les opérations qui suivent sont au sens cardinal)Si λ et μ sont deux cardinaux, λ infini et $0 < \mu \leq \lambda$, alors $\mu + \lambda = \lambda + \mu = \lambda$ et $\mu\lambda = \lambda\mu = \lambda$.
On remarque que contrairement aux opérations ordinales, les opérations cardinales sont commutatives. C'est très facilement compréhensible.
Preuve: Remarquons dans un premier temps que si on prouve que $\lambda^2 = \lambda$, cela suffira. Pourquoi ? Simplement, λ s'injecte dans $\lambda + \mu$ qui s'injecte dans $\lambda + \lambda$, qui est égal à $\lambda 2$, qui s'injecte dans $\lambda\mu$, qui s'injecte dans λ^2, qui est donc égal à λ. Prouvons donc ceci : pour tout cardinal λ infini, $\lambda^2 = \lambda$. Si on réussit, cela conclura la preuve.
On procède par récurrence transfinie. Soit λ un cardinal infini. On suppose que pour tout cardinal infini $\mu < \lambda$, $\mu^2 = \mu$. On se permet ici l'utilisation d'une propriété évidente, qu'on laisse à démontrer en exercice aux lecteurices: si n est un ordinal fini, n^2 (ici c'est le produit ordinal, mais pour les ordinaux finis, les deux coïncident) en est aussi un. De ce fait, si μ est un cardinal quelconque inférieur strictement à λ, $\mu^2 < \lambda$. On considère alors $\lambda \times \lambda$ (produit cartésien),

qu'on ordonne ainsi : $(x, y) \prec (x', y')$ si et seulement si ($\sup(x, y) < \sup(x', y')$) ou ($\sup(x, y) = \sup(x', y')$ et $x < x'$) ou ($\sup(x, y) = \sup(x', y')$ et $x = x'$ et $y < y'$). Il s'agit d'un bon ordre: en effet si X est une partie non vide de $\lambda \times \lambda$, on prend d'abord le sous-ensemble X_1 de X constitué des couples dont le sup est minimal, puis le sous-ensemble X_2 de X_1 constitué des couples dont la première coordonnée est minimale, puis le sous-ensemble X_3 de X_2 constitué des couples dont la deuxième coordonnée est minimale. X_3 ne contient qu'un élément, qui est le minimum de X. Donc $(\lambda \times \lambda, \prec)$ est isomorphe à un unique ordinal α. On va montrer que cet ordinal est $\leq \lambda$, et cela permettra de conclure (on aura alors $\lambda \times \lambda$ qui s'injecte dans λ, la réciproque étant évidente).Supposons donc $\alpha > \lambda$. Soit f l'isomorphisme de α sur $\lambda \times \lambda$. Considérons $f(\lambda) = (\alpha_0, \beta_0)$. On pose alors $\delta_0 = \sup(\alpha_0, \beta_0) + 1$. Pour tout $x \in \lambda$, $f(x) \prec (\alpha_0, \beta_0)$, (car f est strictement croissante) donc l'image de $f_{|\lambda}$ (f restreinte à λ) est incluse dans $\delta_0 \times \delta_0$. Or $\delta_0 \in \lambda$, donc le cardinal de $\delta_0 \times \delta_0$ est un cardinal strictement inférieur à λ. Donc $f_{|\lambda}$ ne peut pas être injective, or, f l'étant, cela donne une contradiction. Donc $\alpha \leq \lambda$. Donc $\lambda \times \lambda$ s'injecte dans λ: λ est équipotent à λ^2. Cela conclut la preuve.
Pour respecter les conventions, ce résultat s'écrit généralement : pour tout ordinal α, $\aleph_\alpha^2 = \aleph_\alpha$ (pareil pour les autres propriétés, mais celle-ci est la plus forte). Il est évident que ce résultat se généralise à n'importe quels ensembles: on appelle ensemble fini un ensemble dont le cardinal est fini, sinon on dit que l'ensemble est infini. Alors, si E est un ensemble infini, et si F est un ensemble (non vide) qui s'injecte dans E, disjoint de E, alors $E \times F$ est équipotent à $E \cup F$, qui est équipotent à E. Cela découle immédiatement de la proposition qui précède. Il existe plusieurs autres résultats d'arithmétique cardinale, notamment concernant les sommes et les produits infinis (enfin, les sommes et produits en général, mais les résultats sont intéressants quand ils sont infinis), mais je ne les mentionne pas dans cette introduction, je laisse aux lecteurices le soin d'y réfléchir (ces résultats peuvent être établis sans utiliser les ordinaux, ils n'ont donc pas vraiment leur place ici).
Il y a encore beaucoup de choses à dire sur les cardinaux (qui sont un sujet de recherche active, encore aujourd'hui, en particulier les grands cardinaux), mais ceci n'est qu'une introduction, et ce qui a été développé suffit ici à se faire un aperçu correct de la question.
Cette étude des cardinaux étant faites, on peut passer à la dernière section, concernant les modèles. Dans cette partie sur les cardinaux, on voit que les ordinaux sont essentiels pour comprendre l'univers, car ils fournissent des références de taille et d'ordre canoniques et extrêmement pratiques. Dans la section qui suit, on verra qu'ils ont même des liens avec la structure de l'univers.

4 Modèles de Peano, de Z, de ZF-, et de ZF

Pour comprendre cette section, il faut un minimum de connaissances en théorie des modèles. , mais j'ai fait en sorte de ne pas trop rentrer dans les technicités de cette théorie: cette dernière partie est un peu informelle, mais elle apporte quand même des notions importantes, et des idées importantes à comprendre. Certaines preuves seront donc omises, d'autres un peu imprécises: c'est parce qu'il est possible que le formalisme de la théorie des modèles ne soit pas bien

connu des lecteurices. Celleux qui connaissent ce formalisme pourront essayer de compléter et préciser les preuves, c'est un bon exercice.

4.1 Les entiers et ω: un modèle de Peano ?

Vous l'avez remarqué, et ce n'est pas anodin, les notations qu'on utilise (+1, ou généralement +, etc.) rappellent celles qu'on utilise par exemple en arithmétique sur les entiers; de même la récurrence qu'on a utilisée ici rappelle fortement celle des entiers. Quel est donc le rapport entre les ordinaux et $\mathbb{N}$?
A l'origine, les ordinaux ont été conçus comme une généralisation (voire même initialement une définition) des entiers : tout ordinal est, en un sens, un entier. Ces "entiers" sont parfois très grands, et ne correspondent pas au sens qu'on donne généralement au mot. On va donc ici dire que "entier" abrège "ordinal fini". Comme je l'ai fait remarquer précédemment, pour deux entiers, l'addition et la multiplication on des comportements similaires avec les opérations qu'on a l'habitude d'utiliser (commutativité, associativité, 0 est un neutre pour +, 1 pour la multiplication,...). On laisse aux lecteurices le soin de prouver que si n et m sont des entiers alors $n+m$, nm et 2^n (cardinal de l'ensemble des parties de n) en sont aussi. Ceci étant fait, le.a lecteurice comprendra aisément que, lorsqu'il est muni de l'addition et de la multiplication ordinale, puis de $\emptyset$ comme 0 et $n \to n+1$ comme application successeur, ω est un modèle de la théorie du premier ordre Peano. En effet, il est évident de vérifier les premiers axiomes (concernant les successeurs, la neutralité de 0, etc.), et le schéma d'axiomes de récurrence est vérifié du fait de la propriété de récurrence qui est valable sur les entiers. Ainsi, dans ZF, on peut prouver la cohérence de Peano ! C'est pour ça qu'on dit que ZF est strictement plus forte que Peano (deuxième théorème de Gödel oblige: ZF ne peut pas prouver sa propre cohérence, donc la cohérence de Peano ne peut pas entraîner celle de ZF, si ZF est cohérente).
Du fait que ω (avec la bonne fonction d'interprétation) est un modèle de Peano, si on veut définir $\mathbb{N}$, on le définit souvent par ω: c'est d'une certaine manière un modèle canonique. Cela justifie en particulier les notations $0, 1, 2, ...$

4.2 Axiome de fondation, V_α et la structure de l'univers: modèles de Z, ZF- et ZF ?

L'axiome de fondation est l'axiome suivant : $\forall x \neq \emptyset, \exists y(y \in x \wedge y \cap x = \emptyset)$. Il n'est en général pas ajouté à ZF par les auteurs, et on note alors ZF+AF la théorie obtenue en l'ajoutant à ZF. Ici je ne m'attarderai pas à montrer que ZF +AF est cohérente (sous réserve de cohérence de ZF), premièrement car ce n'est pas le but de cette introduction, et deuxièmement parce qu'à partir de ce qui suit vous pourrez sûrement deviner comment cela se fait.
L'axiome de fondation a été introduit pour interdire des situations comme $x \in x$, $x \in y \in z... \in x$ ou plus généralement $...x_n \in ... \in x_1 \in x_0$ (je vous laisse montrer que ces situations sont impossibles sous l'axiome de fondation) qui semblent peu naturelles. Cet axiome est peu polémiqué, pourtant à mon avis, il est moins naturel que l'axiome du choix, du fait de ses conséquences, qu'on va décrire ici.
On va définir des ensembles V_α pour tout ordinal α par récurrence transfinie. Il y a une définition "moins au cas par cas" que ce que je vais présenter, mais elle est équivalente. On pose $V_0 = 0$, et pour tout ordinal α, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$, et si

α est limite, $V_\alpha = \bigcup_{\beta\in\alpha} V_\beta$. On note alors $V(x)$ la formule "il existe un ordinal α tel que $x \in V_\alpha$" (c'est, au sens intuitif, la "réunion des V_α"). On va montrer que (modulo ZF), l'axiome de fondation entraîne $\forall x, V(x)$.

Lemme 10: Pour tout ordinal α, V_α est un ensemble transitif, et si $\beta < \alpha$, $V_\beta \subset V_\alpha$.

Preuve: La preuve est laissée aux lecteurices: elle est évidente par récurrence transfinie.

Proposition 7: Pour tout ensemble x, il existe un plus petit ensemble transitif contenant x (au sens de l'inclusion), noté $Cl(x)$ (la clôture transitive de x).

Preuve: On pose $x_0 = x$, et pour tout entier n, on pose $x_{n+1} = x_n \cup \bigcup_{y\in x_n} y$. On considère ensuite $k = \bigcup_{n\in\omega} x_n$. Montrons que k est un ensemble transitif (il contient x, c'est évident). Soit $y \in k$. Il existe $n \in \omega$ tel que $y \in x_n$. On a alors $y \subset x_{n+1}$, et donc $y \subset k$. Donc k est transitif. Soit maintenant m un ensemble transitif contenant x. Il est aisé de montrer que m contient x_n pour tout $n \in \omega$, et donc $k \subset m$. On obtient bien ce qu'on veut: k est la clôture transitive de x, et on le note $Cl(x)$.

Lemme 11: Soit x un ensemble. Pour que $V(x)$, il faut et il suffit que $\forall y \in x, V(y)$.

Preuve: Soit x un ensemble quelconque, et supposons $\forall y \in x, V(y)$. Pour tout $y \in x$, on note $rg(y)$ (et on appelle "rang de y") le plus petit ordinal tel que $y \in V_{rg(y)}$. Par le schéma d'axiomes de remplacement, $\{rg(y) \mid y \in x\}$ est un ensemble, on peut donc considérer sa borne supérieure α. Donc $\forall y \in x, y \in V_\alpha$ (lemme 10). Donc $x \subset V_\alpha$, donc $x \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$. Donc $V(x)$.
Réciproquement, si $V(x)$: on se donne $y \in x$.Il existe un ordinal α tel que $x \in V_{\alpha+1}$, donc $x \subset V_\alpha$, donc $y \in V_\alpha$, et donc $V(y)$.

Théorème 5: Modulo les axiomes de ZF, l'axiome de fondation est équivalent à $\forall x, V(x)$.

Preuve: Premièrement, supposons $\forall x, V(x)$. Soit x quelconque, non vide. On choisit $y \in x$ de rang minimal (pour voir la définition de rang, voir la preuve du lemme 11). Il est évident que si $z \in y$, alors $rg(z) < rg(y)$. De ce fait, supposons $y \cap x$ non vide: il existe $z \in x$ tel que $rg(z) < rg(y)$, ce qui contredit la définition de y. Donc $y \cap x = \emptyset$: cela prouve l'axiome de fondation.
Réciproquement, supposons l'axiome de fondation. Soit x un ensemble quelconque. On va montrer $V(Cl(x))$, ce qui suffira pour obtenir $V(x)$ (pour la définition de $Cl(x)$, voir la proposition 7). Supposons $\neg V(Cl(x))$. D'après le lemme 11, l'ensemble $m = \{y \in Cl(x) \mid \neg V(y)\}$ est non vide. Il existe donc $t \in m$ tel que $t \cap m = \emptyset$ (axiome de fondation). Soit $d \in t$. $Cl(x)$ étant transitif, $d \in Cl(x)$. De plus on a $d \notin m$, donc $V(d)$. Donc $\forall d \in t, V(d)$. D'après le lemme 11, on a donc $V(t)$. Or $t \in m$: c'est contradictoire ! Donc $V(Cl(x))$, et donc $V(x)$.
Voilà pourquoi, à mon sens, l'axiome de fondation est plus contestable que l'axiome du choix (quoique les deux ne posent pas de problème de cohérence), parce que d'une certaine manière, il dit "tout ensemble est descriptible en fonction de $\emptyset$".
Cette description de l'univers (l'univers est égal à "l'union" des V_α) révèle bien que les ordinaux sont nécessaires à sa compréhension, et qu'ils sont importants

à sa "structure". On va justifier cela bien plus fortement dans les résultats qui suivent.
Pour ce qui suit, on note Inf l'axiome de l'infini, ZF- la théorie ZF à laquelle on a enlevé l'axiome de l'infini.
Théorème 6: ZF prouve la cohérence de ZF + ¬Inf.
Ceci dit premièrement que sans l'axiome de l'infini, on ne peut pas prouver l'existence d'un ensemble infini. Deuxièmemenent, ZF est strictement plus forte que ZF-+¬Inf, et que $ZF-$. Dans la preuve qui suit, certaines technicités ne sont pas abordées. On invite les lecteurices à vérifier les parties de la preuve qui sont négligées comme "évidentes".
Preuve: On va montrer que V_ω est un modèle de ZF-+¬Inf, ce qui suffira. Rappelons que $V_\omega = \bigcup_{n\in\omega} V_n$.
-Axiome d'extensionnalité: Soit $x, y \in V_\omega$ tels que leurs éléments dans V_ω soient les mêmes. Or V_ω est transitif, donc tous les éléments de x et de y sont dans V_ω: d'après l'axiome d'extensionnalité dans le "vrai" univers, $x = y$.
-Axiome de la paire: évident, car $\{x, y\}$ est une partie de V_n, avec $n = max(rg(x), rg(y))$.
-Axiome de la réunion: évident, par construction de V_ω
-Axiome de l'ensemble des parties: évident par construction
-Axiome de fondation : évident en utilisant le rang et la transitivité de V_ω
Il ne reste donc qu'à montrer deux choses: le schéma d'axiomes de remplacement, et ¬Inf. La preuve du schéma d'axiomes de remplacement dan V_ω est un peu technique, car elle nécessite l'utilisation d'énoncés restreints, que je n'ai pas introduits ici. Elle repose sur une combinaison des faits suivants: pour tout entier n, $\mathcal{P}(n)$ est équipotent à un unique entier; pour tout entier n, V_n est fini; un sous-ensemble d'un ensemble fini est fini; si X est fini et f est une fonction de domaine X à valeurs dans V_ω, alors l'ensemble $\{rg(f(x)) \mid x \in X\}$ a un maximum. Je laisse aux lecteurices le soin de les prouver (des récurrences suffisent).
Une fois que ces résultats sont acquis, intuitivement, le schéma d'axiomes de remplacement se prouve ainsi: si on a un énoncé $R(x; y)$ qui est fonctionnel en x lorsqu'interprété dans V_ω, alors en le restreignant à V_ω il est fonctionnel dans le "vrai" univers, mais comme son domaine et son codomaine sont inclus dans V_ω, c'est même une fonction f qu'on a. Si on se donne donc un ensemble $X \in V_\omega$ quelconque, que l'on restreint f à X, le rang des images des éléments de X par f est majoré, et donc l'image directe de X par f est incluse dans V_k, où k est un entier, donc l'image appartient à V_{k+1}, ce qui permet de montrer que l'image est dans V_ω. Les côtés techniques mis à part, c'est ce qu'il faut retenir de cette démonstration.
Maintenant, ¬Inf: Montrons qu'un ensemble est un ordinal de V_δ si et seulement si c'est un ordinal qui est dans V_δ: soit $X \in V_\delta$ un ensemble transitif, V_δ étant transitif, X est transitif dans V_δ ("le modèle V_δ le voit comme transitif"); de même si X est bien ordonné par $\in$, il l'est aussi dans V_δ. Donc si X est un ordinal qui appartient à V_δ, c'est un ordinal dans V_δ (plus formellement : l'univers satisfait "X est un ordinal et $X \in V_\delta$" implique que V_δ satisfait "X est un ordinal"). La réciproque est évidente aussi: si V_δ satisfait "X est un ordinal", alors l'univers satisfait "$X \in V_\delta$", et il satisfait aussi "X est un ordinal", par transitivité de V_δ. Il en découle assez immédiatement que V_δ satisfait "X est un ordinal fini" si et seulement si l'univers satisfait "X est un ordinal fini et $X \in V_\delta$".
De ce fait, V_ω satisfait Inf si et seulement si $\omega \in V_\omega$. Cela signifierait que

$\omega \in V_n$ pour un $n \in \omega$. C'est évidemment impossible d'après les remarques qui précèdent concernant le "comportement" des ensembles finis: V_n est fini, et un sous-ensemble d'un ensemble fini est fini, tandis que ω est infini. Plus généralement, on peut montrer que pour tout ordinal α, le rang de α est $\alpha + 1$, et donc $\omega \in V_{\omega+1} - V_\omega$. Je laisse aux lecteurices cette preuve, qui est intéressante à faire pour vérifier qu'on a compris le principe de récurrence transfinie.
Donc V_ω satisfait $\neg$Inf: d'après tout ce qui précède, c'est un modèle de ZF- + $\neg$Inf.
Maintenant, on va aller plus loin et montrer que ZF est strictement plus forte que Z (Z est la théorie de Zermelo, c'est ZF sans le schéma d'axiomes de remplacement, avec à la place le schéma d'axiomes de compréhension). Reprenons la preuve du théorème 6: si α est un ordinal quelconque, V_α satisfait l'axiome d'extensionnalité parce qu'il est transitif, et l'axiome de fondation grâce à l'astuce du rang. Que faut-il pour l'axiome des parties ? Il faut que si $\beta \in \alpha$, $\beta + 2 \in \alpha$, ce qui est le cas précisément quand α est limite. On suppose donc α limite. Cela suffit aussi pour que l'axiome de la réunion et celui de la paire soient satisfait dans V_α. On montrera plus tard qu'on ne peut pas démontrer le schéma d'axiomes de remplacement, mais seulement celui de compréhension. Soit donc $a \in V_\alpha$, où α est limite; et soit F un énoncé à une variable libre. Par compréhension dans le "vrai" univers, $z = \{x \in a \mid F'(x)\}$ existe (F' est l'énoncé F restreint à V_α, c'est un détail technique, je ne rentre pas dedans), et c'est une partie de a. Si $a \in V_{\beta+1}$, avec $\beta < \alpha$, alors $a \subset V_\beta$, et donc $z \subset V_\beta$, et donc $z \in V_{\beta+1}$, et donc $z \in V_\alpha$: le schéma d'axiomes de compréhension est vérifié. Supposons désormais que $\alpha > \omega$, en plus d'être limite. D'après la remarque sur le rang des ordinaux, on a donc $\omega \in V_\alpha$, et donc V_α satisfait Inf. Donc V_α satisfait la théorie Z. On a donc le théorème suivant :
Théorème 7: ZF prouve la consistance de Z.
Ainsi ZF est strictement plus forte que Z (à nouveau, second théorème d'incomplétude de Gödel oblige).
Le.a lecteurice qui se demande pourquoi on n'a pas prouvé le schéma d'axiomes de remplacement dans V_α pourra se satisfaire de la réponse suivante: dans ZF, on l'a prouvé, tout bon ordre est isomorphe à un ordinal. ω est en bijection avec $\omega + \omega$, donc il existe un sous-ensemble de $\mathcal{P}(\omega \times \omega)$ qui est un bon ordre sur ω, muni duquel ω est isomorphe à $\omega + \omega$ (on peut d'ailleurs expliciter ce bon ordre: tous les pairs sont inférieurs à tous les impairs, et les pairs sont rangés entre eux naturellement, les impairs aussi). Donc tout modèle de ZF contient un tel bon ordre. Or $V_{\omega+\omega}$ est un modèle de Z ($\omega + \omega$ est limite et supérieur à ω), auquel $\omega + \omega$ n'appartient pas: donc le schéma d'axiomes de remplacement n'y est pas vérifié : ce dernier n'est pas conséquence de Z.
Dans ce dernier exemple on voit le fameux "tel truc n'est pas conséquence de telle théorie", qui en dehors de la théorie des modèles peut sembler farfelu "comment ils savent ça, ils ont vérifié toutes les preuves ?", mais qui, à la lumière de cette théorie, devient clair comme de l'eau de roche.
Si vous êtes attentif.ve, vous aurez remarqué que le titre de cette sous-section (et de la section) parle de modèles de ZF, alors que jusqu'ici on n'a vu que des modèles de Peano, de ZF- + $\neg$ Inf, et de Z. Vous vous demandez aussi comment on peut avoir un modèle de ZF, alors que le théorème de Gödel interdit justement cela. Je ne détaillerai absolument pas l'exemple qui suit, mais il est intéressant, car il montre un peu ce en quoi la recherche moderne peut consister. On peut définir dans ZF ce qu'on appelle un cardinal inaccessible. Intuitive-

ment, c'est un cardinal si grand qu'on ne peut l'atteindre avec les opérations usuelles (ensemble des parties, limites): il faut le connaître pour en parler. Le problème avec ces cardinaux inaccessibles c'est qu'on ne sait pas s'ils existent. Pire encore, avec le théorème qui suit, que je ne démontre pas :
Théorème: Si ZF est consistante, alors ZF + "il n'existe pas de cardinaux inaccessibles" l'est aussi; ZF + "il existe un cardinal inaccessible" prouve "ZF est cohérente".
En d'autres termes, l'hypothèse de l'existence d'un cardinal inaccessible "ajoute une force stricte" à ZF. De ce fait, on ne peut pas montrer de résultats comme :"si ZF est cohérente, alors ZF + 'il existe un cardinal inaccessible' l'est aussi", car cela violerait le second théorème de Gödel. Donc quand on travaille avec les hypothèses de cardinal inaccessible (ou plus généralement, de grands cardinaux: ce sont des recherches menées aujourd'hui; on peut citer les cardinaux supercompacts, les Woodin-cardinals, etc.), contrairement avec l'axiome du choix ou l'hypothèse du continu par exemple, on ne peut pas être assuré de ne pas tomber sur une contradiction. Si on en trouve une, c'est simplement qu'on a supposé l'existence de trop grands cardinaux. Ce problème est le même qu'avec l'axiome Inf:on ne peut pas prouver "si ZF- est cohérente alors ZF l'est aussi", car "aux yeux des cardinaux finis", $\aleph_0$ est inaccessible.
Tout ceci est bien sûr informel, mais c'est un peu plus technique. La justification derrière tout ce que j'ai dit est que si κ est un cardinal inaccessible, alors V_κ est un modèle de ZF. Ceci éclairera peut-être tout ce qui vient d'être dit. Les recherches modernes s'orientent vers des grands cardinaux généralement, et c'est pour cela qu'il me semblait important de les mentioner. Le.a lecteurice qui en a le courage pourra vérifier que V_κ est un modèle de ZF si κ est inaccessible, la preuve est très similaire à celle du théorème 6.
Je finis donc sur cette ouverture cette introduction aux ordinaux.