THEOREME DE CANTOR-BERNSTEIN
(une preuve simple constructive)

Rémi SAINT-ROMAIN, Guy PHILIPPE

24 novembre 2004

Théorème : X et Y étant 2 ensembles non vides. S'il existe une injection de X dans Y et une injection de Y dans X alors il existe une bijection de X dans Y.

Cette propriété avait été conjecturée par Cantor (1845-1918) et fut prouvée par Félix Bernstein (1878-1956), à ne pas confondre avec Sergueï Bernstein (1880-1968) qui, lui, est célèbre pour ses polynômes. Elle fut aussi prouvée par Schroeder (1841-1902) d'où le nom de Théorème de Schroeder-Bernstein que l'on lui donne parfois.

Démonstration

La seule précaution à prendre, dans cette démonstration, quand on voudra restreindre une application $f : E \mapsto F$ à une partie H de E, sera de s'assurer que la partie H n'est pas vide !

Dans toute la suite la notation $\bigcup A_n$ désignera $\displaystyle\bigcup_{n\geq 0} A_n$, sinon on précisera en écrivant par exemple $\displaystyle\bigcup_{n\geq 1} A_n$ et la notation E^c, si E est une partie de X (respectivement de Y), désignera la partie complémentaire de E dans X (respectivement dans Y).

Soit f une injection de X dans Y et g une injection de Y dans X. On définit une suite $(A_n)_{n\geq 0}$ de parties de X et une suite $(B_n)_{n\geq 0}$ de parties de Y par les 2 conditions :

1) $A_0 = X - g(Y)$ et $\forall n \geq 1 \quad A_n = g(B_{n-1})$

2) $\forall n \geq 0 \quad B_n = f(A_n)$

Ces 2 conditions permettent de calculer successivement A_0, $B_0 = f(A_0)$, $A_1 = g(B_0)$, $B_1 = f(A_1)$, etc. On obtient ainsi tous les A_n et B_n pour $n \geq 0$.

Premier cas : si $A_0 = X - g(Y) = \emptyset$ alors g est surjective et g est donc la bijection cherchée. *Ce qu'il fallait démontrer.*

Second cas Si $A_0 \neq \emptyset$ alors il est clair que $B_0 = f(A_0) \neq \emptyset$, $A_1 = g(B_0) \neq \emptyset$, $B_1 = f(A_1) \neq \emptyset$ etc, les 2 suites (A_n) et (B_n) sont formées de parties toutes non vides. on en déduit que les ensembles $\bigcup A_n$ et $\bigcup B_n$ ne sont pas vides. Montrons que l'on peut restreindre f au départ à $\bigcup A_n$,*qui n'est pas vide*, et

à l'arrivée à $\bigcup B_n$. En effet $a \in \bigcup A_n \Longrightarrow \exists n \quad a \in A_n \Longrightarrow f(a) \in f(A_n) = B_n \Longrightarrow f(a) \in \bigcup B_n$. La restriction de f est ainsi légitimée. Notons f' cette restriction alors f' est une bijection car elle est injective en tant que restriction de l'injection f et surjective car $b \in \bigcup B_n \Longrightarrow \exists n \quad b \in B_n = f(A_n)$ donc $\exists a \in A_n \quad f(a) = b$ or $f(a) = f'(a)$ d'où $f'(a) = b$.
Ainsi f' établit bien une bijection de $\bigcup A_n$ sur $\bigcup B_n$.
Toujours dans le **second cas** on a l'alternative $(\bigcup B_n)^c = \emptyset$ ou $(\bigcup B_n)^c \neq \emptyset$.
• Si $(\bigcup B_n)^c = \emptyset$ i.e. $\bigcup B_n = Y$ alors $(\bigcup A_n)^c = \emptyset$. Sinon on aurait $(\bigcup A_n)^c \neq \emptyset$, on pourrait donc choisir $a \in X - \bigcup A_n$ et alors on aurait $f(a) \in Y = \bigcup B_n \Longrightarrow \exists n \quad f(a) \in B_n = f(A_n) \Longrightarrow \exists a' \in A_n \quad f(a) = f(a') \Longrightarrow a = a' \in \bigcup A_n$ d'où une contradiction. On a donc bien $(\bigcup A_n)^c = \emptyset$ i.e. $\bigcup A_n = X$ et la bijection précédente $f' : \bigcup A_n = X \mapsto \bigcup B_n = Y$ convient. *Ce qu'il fallait démontrer.*
• Si $(\bigcup B_n)^c \neq \emptyset$ alors montrons que l'on peut restreindre g au départ à $(\bigcup B_n)^c$, *qui n'est pas vide*, et à l'arrivée à $(\bigcup A_n)^c$. En effet $b \in (\bigcup B_n)^c \Longrightarrow g(b) \in (\bigcup A_n)^c$. Sinon on aurait $g(b) \in \bigcup A_n \Longrightarrow \exists n \quad g(b) \in A_n$. D'où 2 cas qui mèneraient tous les 2 à une contradiction :
ou $n = 0$ et alors on aurait $g(b) \in A_0 = X - g(Y)$ d'où une contradiction car $b \in (\bigcup B_n)^c \subset Y \Longrightarrow g(b) \in g(Y)$
ou $n \geq 1$ et alors on aurait $g(b) \in A_n = g(B_{n-1}) \Longrightarrow \exists b' \in B_{n-1} \quad g(b) = g(b')$ or g est injective d'où on aurait $b = b' \in B_{n-1}$ et par suite $b \in \bigcup B_n$ d'où une contradiction. Par conséquent on a bien $g(b) \in (\bigcup A_n)^c$ ce qui, par ailleurs, entraîne $(\bigcup A_n)^c \neq \emptyset$.
La restriction de g est ainsi légitimée. Notons-la g', donc $g' : (\bigcup B_n)^c \mapsto (\bigcup A_n)^c$. g' est injective en tant que restriction d'une injection et elle est surjective car $a \in (\bigcup A_n)^c = (A_0 \cup A_1 \cup A_2....)^c = A_0^c \cap A_1^c \cap A_2^c... = g(Y) \cap (A_1^c \cap A_2^c....) \Longrightarrow a \in g(Y)$ et $a \in \bigcap\limits_{n \geq 1} A_n^c$. Par suite $\exists b \in Y \quad g(b) = a$ et $g(b) = a \in \bigcap\limits_{n \geq 1} A_n^c = \bigcap\limits_{n \geq 1} (g(B_{n-1}))^c =$(décalage de l'indice n) $\bigcap (g(B_n))^c = (\bigcup g(B_n))^c = (g(\bigcup B_n))^c$. D'où $g(b) = a \in (g(\bigcup B_n))^c \Longrightarrow b \in (\bigcup B_n)^c$. Sinon on aurait $b \in \bigcup B_n$ et donc $g(b) = a \in g(\bigcup B_n)$ ce qui contredirait $g(b) = a \in (g(\bigcup B_n))^c$. Ainsi on a bien $b \in (\bigcup B_n)^c$ d'où $g'(b) = g(b) = a$ et g' établit bien une bijection de $(\bigcup B_n)^c$ sur $(\bigcup A_n)^c$.
Il ne reste plus qu'à "recoller les morceaux" pour expliciter une bijection F de X sur Y définie pour tout $a \in X$ par
$F(a) = f'(a)$ si $a \in \bigcup A_n$ et $F(a) = (g')^{-1}(a)$ si $a \in (\bigcup A_n)^c$.
En effet X est formé de 2 parties non vides disjointes $\bigcup A_n$ et $(\bigcup A_n)^c$ tout comme Y est formé des 2 parties non vides disjointes $\bigcup B_n$ et $(\bigcup B_n)^c$ et comme f' est une bijection de $\bigcup A_n$ sur $\bigcup B_n$ tout comme $(g')^{-1}$ est une bijection de $(\bigcup A_n)^c$ sur $(\bigcup B_n)^c$ il est clair que F est une bijection de X sur Y. *Ce qu'il fallait démontrer.*

Finalement on a bien montré dans tous les cas que X et Y sont

équipotents i.e. qu'il existe une bijection de X sur Y.

La relation binaire R définie dans la collection des ensembles NON VIDES par $XRY \iff$ "il existe une injection de X dans Y" est trivialement réflexive, transitive et elle permet de définir sans difficulté une relation r sur les classes d'équipotence des ensembles NON VIDES encore appelées cardinaux AUTRES QUE 0 en posant, si $x = card(X) \neq 0$ et $y = card(Y) \neq 0$, $xry \iff XRY$. Le théorème de Cantor-Bernstein montre que cette dernière relation r est ANTISYMETRIQUE et que finalement c'est une relation d'ordre dans la classe des cardinaux AUTRES QUE 0, que l'on peut prolonger par une relation ρ à la classe de tous les cardinaux (y compris 0) en posant $x\rho y \iff x = y = 0$ ou bien x et $y \neq 0$ et xry. Plus simplement, si on adopte la notation G_r(respectivement G_ρ) pour le graphe de la relation r(respectivement ρ) dans la classe des cardinaux non nuls(respectivement dans la classe des cardinaux) alors $G_\rho = G_r \bigcup \{(0,0)\}$. On vérifie facilement que la relation ρ, sur la classe des cardinaux, est bien une relation d'ordre. Grâce à la théorie des ordinaux on peut même montrer que cette relation ρ est un ordre TOTAL dans la classe des cardinaux, mais, cette fois, en recourant à l'axiome du choix.