

Cantor-Bernstein-Schröder pour tous

Marc Roux^(*)

Le hasard de mes lectures en vue d'alimenter la rubrique « Matériaux » m'a fait rencontrer récemment plusieurs fois⁽¹⁾ le théorème appelé selon les sources « de Cantor- Bernstein » ou « de Schröder-Bernstein ». Il affirme :

S'il existe une injection d'un ensemble A dans un ensemble B et une injection de B dans A, alors il existe une bijection de A sur B.

Ce théorème est essentiel pour la théorie des ensembles, en particulier pour la théorie des cardinaux ; en effet, par définition⁽²⁾ :

$\text{Card}(A) = \text{Card}(B) \Leftrightarrow$ Il existe une bijection de A sur B ;

et :

$\text{Card}(A) \leq \text{Card}(B) \Leftrightarrow$ Il existe une injection de A dans B

Pour s'assurer que la relation \leq est une relation d'ordre, donc antisymétrique, il est nécessaire de démontrer le théorème cité.

Cette démonstration a une histoire compliquée :

- 1887. Cantor publie le théorème sans démonstration (il « n'aurait pas jugé nécessaire » de le démontrer).
- 1887. Dedekind le démontre, sans intervention de l'axiome du choix, mais ne publie pas sa preuve et n'informe pas Cantor. Cette preuve sera redécouverte par Zermelo en 1908.
- 1895. Cantor le démontre comme conséquence de l'existence d'un bon ordre sur les cardinaux ; mais cette existence se révélera être équivalente à l'axiome du choix (Hartogs 1915).
- 1896. Schröder publie une preuve qui se révèle erronée.
- 1897. Bernstein, jeune étudiant de Cantor (19 ans !), publie sa démonstration, qui est indépendante de l'axiome du choix. La même année, indépendamment, Schröder et Dedekind trouvent des démonstrations.
- 1898. La preuve de Bernstein est reprise par Borel ; le texte se trouve à l'adresse <http://archive.org/stream/leconstheoriefon00borerich#page/n115/mode/2up>

Parmi les diverses démonstrations trouvées dans [1], [2] et [3], j'ai une préférence pour la troisième de celles données par S. Nicolay, car elle est succincte et ne fait appel à aucun préalable ; elle est très proche de celle de Bernstein, quoique exprimée en termes différents :

Soient A et B deux ensembles, f une injection de A dans B, g une injection de B dans A. Si f ou g est surjective, alors on a tout de suite une bijection entre A et B. Dans la suite, nous supposons donc que $B \setminus f(A)$ et $A \setminus g(B)$ sont non vides⁽³⁾. On pose : $A_0 = A \setminus g(B)$, et pour tout naturel n , $B_n = f(A_n)$, $A_{n+1} = g(B_n)$.

(*) marc.roux15@wanadoo.fr

(1) Voir bibliographie

(2) Suivant en cela Samuel Nicolay, je ne donnerai pas de définition formelle du cardinal.

(3) J'utilise la notation $E \setminus F$ pour désigner l'ensemble des éléments de E qui ne sont pas éléments de F.

Une récurrence immédiate montre qu'aucun des A_n ni des B_n n'est vide ; on a donc

$\bigcup_{n \in \mathbb{N}} A_n \neq \emptyset$, $\bigcup_{n \in \mathbb{N}} B_n \neq \emptyset$ et $f\left(\bigcup_{n \in \mathbb{N}} A_n\right) \neq \emptyset$. De plus, $f\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \bigcup_{n \in \mathbb{N}} B_n \neq B$ car sinon

f serait surjective, donc bijective, et de même $g\left(\bigcup_{n \in \mathbb{N}} B_n\right) \neq \emptyset$ et $g\left(\bigcup_{n \in \mathbb{N}} B_n\right) \neq A$ puisque g est supposé non surjective.

Notons f' la restriction de f à $\bigcup_{n \in \mathbb{N}} A_n$, et g' la restriction de g à $B \setminus \bigcup_{n \in \mathbb{N}} B_n$.

f' est une bijection de $\bigcup_{n \in \mathbb{N}} A_n$ sur $\bigcup_{n \in \mathbb{N}} B_n$; g' est une injection de $B \setminus \bigcup_{n \in \mathbb{N}} B_n$ dans

$A \setminus \bigcup_{n \in \mathbb{N}} A_n$, mais elle est aussi surjective car si x n'appartient à aucun des A_n , il n'appartient en particulier pas à A_0 donc il appartient à $g(B)$, et il n'appartient à aucun des $g(B_n)$, donc son antécédent appartient à $B \setminus \bigcup_{n \in \mathbb{N}} B_n$.

L'application h de A dans B définie par :
$$\left\{ \begin{array}{l} \text{Si } x \in \bigcup_{n \in \mathbb{N}} A_n, h(x) = f(x) \\ \text{sinon } h(x) = g^{-1}(x) \end{array} \right\}$$
 est alors

une bijection.

Cette démonstration est parfaitement convaincante sur le plan de la logique, mais il lui manque à mon goût de « montrer » les différents objets qu'elle fait intervenir, de leur « donner du sens ». C'est pourquoi je vais essayer de l'illustrer par deux exemples.

Exemple 1 (arithmétique)

$A = 2\mathbb{N}^*$ = ensemble des naturels pairs non nuls, $B = 3\mathbb{N}^*$ = ensemble des naturels multiples de 3 non nuls.

Bien entendu, si le but du jeu était de montrer l'équipotence de ces deux ensembles,

la bijection $x \rightarrow \frac{3x}{2}$ ferait parfaitement l'affaire ; mais tel n'est pas notre propos.

Soit f l'injection, non surjective, de $2\mathbb{N}^*$ dans $3\mathbb{N}^*$ définie par $f(x) = 3x$, et g

l'injection, non surjective, de $3\mathbb{N}^*$ dans $2\mathbb{N}^*$ définie par $g(x) = 2x$.

L'illustration que je propose consiste à construire explicitement les ensembles A_n, B_n ,

$\bigcup_{n \in \mathbb{N}} A_n$, $A \setminus \bigcup_{n \in \mathbb{N}} A_n$ de la démonstration, et aller jusqu'à la construction effective de h

et de sa réciproque, pour lesquelles nous donnerons un algorithme simple.

Avec les notations de la démonstration précédente, A_0 est l'ensemble des nombres pairs non multiples de 3 :

$A_0 = \{2, 4, 8, 10, 14, 16, 20, 22, 26, 28, 32, 34, \dots\}$.

Alors $B_0 = 3A_0 = \{6, 12, 24, 30, 42, 48, 60, 66, 78, 84, 96, 102, \dots\}$

$$A_1 = 2B_0 = 6A_0 = \{12, 24, 48, 60, 84, 96, 120, 132, \dots\}$$

$$B_1 = 3A_1 = 6B_0 = \{36, 72, 144, 180, 252, 288, 360, 396, \dots\}$$

D'une façon générale on a $A_{n+1} = 6A_n$, $B_{n+1} = 6B_n$; donc $\min(A_{n+1}) = 6\min(A_n)$; si pour un certain k un entier pair n n'appartient à aucun des A_1, \dots, A_k et s'il est inférieur à $6\min(A_k)$, on peut affirmer qu'il n'est dans aucun A_n et qu'il appartient donc

$A \setminus \bigcup_{n \in \mathbb{N}} A_n$; ainsi par exemple 6 est dans $A \setminus \bigcup_{n \in \mathbb{N}} A_n$, d'où $h(6) = g^{-1}(6) = 3$; tandis que

2, 4, 8, 10, 12, 14, 16 sont dans $A \setminus \bigcup_{n \in \mathbb{N}} A_n$; d'où $h(2) = f(2) = 6$, $h(4) = f(4) = 12$, etc.

18 n'appartient ni à A_0 ni à A_1 et est inférieur à $\min(A_2) = 72$, donc $18 \in A \setminus \bigcup_{n \in \mathbb{N}} A_n$,

d'où $h(18) = g^{-1}(18) = 9$

On peut ainsi commencer à dresser une table de la bijection h :

$x \in A$	2	4	6	8	10	12	14	16	18
$h(x) \in B$	6	12	3	24	30	36	42	48	9

Essayons de déterminer de façon générale les éléments de $\bigcup_{n \in \mathbb{N}} A_n$, c'est-à-dire les éléments de A qui sont éléments d'au moins un des A_n .

Un entier p est élément de A_0 si et seulement s'il est de l'une des formes $6k + 2$,

$6k + 4$ ($k \in \mathbb{N}$); $A_n = 6^n A_0$, donc $p \in \bigcup_{n \in \mathbb{N}} A_n$ si et seulement si il existe deux naturels

k et n tels que $p = 6^n(6k + 2)$ ou $p = 6^n(6k + 4)$; ce qui revient à dire que le quotient de p par la plus grande puissance de 6 qui divise p est un nombre pair.

D'où on tire un algorithme très simple pour calculer l'image par h d'un naturel pair quelconque p :

y prend la valeur p

Tant que y est divisible par 6, remplacer y par y/6

Si y est pair, alors $p \in \bigcup_{n \in \mathbb{N}} A_n$, d'où $h(p) = f(p) = 3p$

Sinon $p \in A \setminus \bigcup_{n \in \mathbb{N}} A_n$, d'où $h(p) = g^{-1}(p) = p/2$

Exemples :

a) $p = 34\,992$: les divisions par 6 successives donnent 5 832, 972, 162, 27; d'où $h(34\,992) = 34\,992/2 = 17\,496$.

b) $p = 26\,652$: une division par 6 donne 4 442, qui n'est pas divisible par 6 et est pair : $h(26\,652) = 3 \times 26\,652 = 79\,956$.

Réciproquement, tentons d'expliciter la bijection h^{-1} . D'après la démonstration générale, h^{-1} est définie par :

$$\left\{ \begin{array}{l} \text{Si } x \in \bigcup_{n \in \mathbb{N}} B_n, h^{-1}(x) = f^{-1}(x) \\ \text{sinon } h^{-1}(x) = g(x) \end{array} \right\}$$

On a $B_n = 6^n B_0 = 6^n \cdot 3A_0$, autrement dit un naturel q multiple de 3 est élément de $\bigcup_{n \in \mathbb{N}} B_n$ si et seulement si il est de l'une des formes $q = 6^n \cdot 3(6k + 2) = 6^{n+1}(3k + 1)$ ou $q = 6^n \cdot 3(6k + 4) = 6^{n+1}(3k + 2)$; dans ce cas, $h^{-1}(q) = f^{-1}(q) = q/3$; sinon, $h^{-1}(q) = g(q) = 2q$. Comme plus haut, on peut donner un algorithme qui calcule $h^{-1}(q)$ pour tout naturel q multiple de 3 :

y prend la valeur p

Tant que y est divisible par 6, remplacer y par $y/6$

Si y n'est pas multiple de 3, alors $q \in \bigcup_{n \in \mathbb{N}} B_n$, d'où $h^{-1}(q) = f^{-1}(q) = p/3$

Si y est multiple de 3, alors $q \in B \setminus \bigcup_{n \in \mathbb{N}} B_n$, d'où $h^{-1}(q) = g(q) = 2q$

Exemples :

a) $q = 333$: pas de division par 6, $y = 333$ est multiple de 3 : $h^{-1}(333) = 2 \times 333 = 666$

Remarque : si q est impair, on est toujours dans le même cas : pas de division par 6, $y = q \in 3\mathbb{N}^*$ est multiple de 3, donc $h^{-1}(q) = 2q$

b) $q = 756$: après deux divisions par 6 on obtient 21, multiple de 3 : $h^{-1}(756) = 2 \times 756 = 1512$

c) $q = 180$: après deux divisions par 6 on obtient 5, non multiple de 3 : $h^{-1}(180) = 180/3 = 60$

d) $q = 48$: une division par 6 donne 8, non multiple de 3 : $h^{-1}(48) = 2 \times 48 = 96$

Pour chacun de ces exemples le lecteur vérifiera que le premier algorithme appliqué au résultat redonne bien le nombre de départ.

Exemple 2 (géométrique)

A est une demi-cercle de centre O , d'extrémités C et D (extrémités comprises).

B est une droite parallèle à (CD) , sans point commun avec A .

f est la projection orthogonale de A sur B : $M \rightarrow M'$.

g est l'application qui à chaque $P \in B$ fait correspondre P' intersection de la droite (OP) avec A . On note O' le projeté orthogonal de O sur B . Il est clair que f et g sont injectives et non surjectives. De la même manière que dans le premier exemple, il s'agit d'expliciter h et sa réciproque.

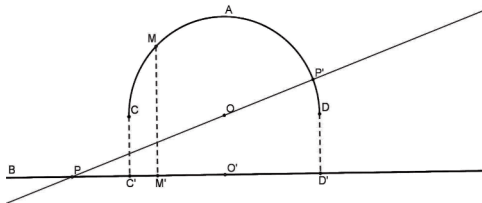


Fig 1

On a $g(B) = A \setminus \{C, D\}$ et donc $A_0 = \{C, D\}$; $B_0 = f(\{C, D\}) = \{C', D'\}$; $A_1 = \{g(C'), g(D')\}$

Une récurrence évidente montre que chacun des A_n est une paire de points de A ,

symétriques par rapport à (OO') , et chacun des B_n est une paire de points de B , symétriques par rapport à la même droite (OO') . Ces points se construisent de manière évidente :

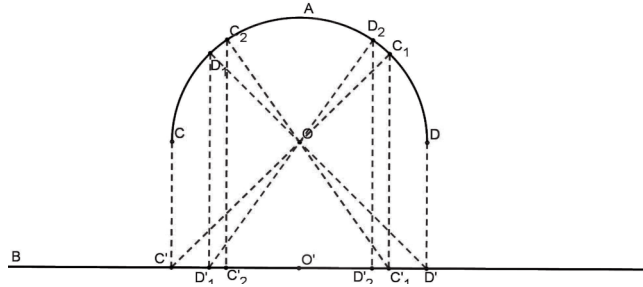


Fig 2

Les éléments de $\bigcup_{n \in \mathbb{N}} A_n$ sont donc les points des deux suites (C_n) et (D_n) définies par : $C_0 = C, C_{n+1} = g \circ f(C_n)$; $D_0 = D, D_{n+1} = g \circ f(D_n)$. La bijection h du demi-cercle A sur la droite B est alors définie par : Si M est un terme de l'une des suites (C_n) , (D_n) , alors $h(M) = f(M) =$ projeté orthogonal de M sur B ; sinon, $h(M) = g^{-1}(M) =$ intersection de la droite (MO) avec la droite B .

Remarque : a et b étant deux réels quelconques, $a < b$, on peut munir le plan d'un repère orthonormal dont le premier vecteur a la direction de B , et tel que les abscisses de C et D soient respectivement a et b ; en associant à chaque point de A ou de B son abscisse, la construction précédente fournit une bijection de l'intervalle fermé $[a, b]$ sur \mathbb{R} .

Ces exemples montrent que la bijection ainsi construite n'est en général ni simple ni naturelle. J'ai voulu faire partager le plaisir que j'ai eu, à travers le choix de ces cas particuliers, à faire vivre une démonstration abstraite et rébarbative à première vue, à montrer qu'on peut s'en forger une image mentale claire et précise, bien plus facile à mémoriser qu'une froide chaîne de pures déductions logiques.

Merci à Pierre Legrand qui m'a apporté les renseignements historiques.

Bibliographie – Sitographie :

- [1] Martin Aigner et Günter M. Ziegler : *Raisonnements divins – Quelques démonstrations mathématiques particulièrement élégantes*. Springer-Verlag France/Lavoisier. 2013
- [2] Samuel Nicolay : *Les nombres – Construction basée sur la théorie des ensembles en vue d'ériger les fondements de l'analyse*. Hermann, 2015
- [3] Wikipedia, article *Théorème de Cantor-Bernstein*.
- [4] Wikipedia en anglais : https://en.wikipedia.org/wiki/Schr%C3%B6der%E2%80%93Bernstein_theorem