Jusqu'à l'infini,
et après ?

Exposé du 5 avril 2023

Université Gustave Eiffel

Bernard Maurey

Collaborateur bénévole, IMJ–PRG, Sorbonne Université

https://webusers.imj-prg.fr/~bernard.maurey/

Cet exposé, destiné aux étudiants de Licence de mathématiques de l'Université Gustave Eiffel à Champs-sur-Marne, a pour thème central l'introduction par Georg Cantor de la *notion d'ordinal*, autour de 1880. Dans une première partie, on suivra trente ans de la vie scientifique de Cantor, de sa thèse en 1867 à son dernier article mathématique, paru en 1897. On évoquera ensuite les développements axiomatiques que les travaux de Cantor ont, au moins en partie, suscités. Au premier rang de ceux-ci, on trouvera les axiomes pour la théorie des ensembles, développés dans les trente premières années du 20e siècle.

L'exposé devant les étudiants a duré deux heures (il ne serait pas raisonnable de présenter la version allongée qui suit en deux heures !). Ce jour-là, il faisait très beau autour du bâtiment Copernic de l'Université, j'en remercie très chaleureusement Marco Cannone, organisateur de l'événement.

Georg Cantor, vers 1870

(1845–1918)

Das Wesen der Mathematik
liegt in ihrer Freiheit.
G. Cantor

(1883)

L'essence des mathématiques
réside dans leur liberté

page précédente **:** on peut voir ce portrait du jeune Cantor dans le livre de Purkert et Ilgauds dont la référence est donnée à la fin de ce document. Ces auteurs ont aussi choisi la belle phrase de Cantor sur la « *liberté en mathématiques* » comme titre d'une de leurs sections (p. 121 de leur livre).

---

page suivante **:** un nuage de mots qui peuvent s'appliquer à ce qu'on verra des travaux de Cantor dans la suite. L'expression « objets de la pensée » sera souvent applicable **;** au contraire, la présence dans ses textes de fonctions explicites des mathématiques, telles que sin(x) par exemple, deviendra plutôt rare après ses articles « trigonométriques » de 1870 et 1872.

Suites infinies arbitraires d'objets

Tiers exclu (logique classique)

succession infinie d'opérations

objets de la pensée

bijections, cardinalité

ensembles dénombrables

Karl Weierstrass

points limites, *ensemble dérivé*

ensemble parfait

Topologie de la droite (et de l'espace)

Berlin

ensemble bien ordonné, *ordinaux*

Leopold Kronecker

Autour de 1870

DE AEQUATIONIBUS
SECUNDI GRADUS INDETERMINATIS.

DISSERTATIO INAUGURALIS
QUAM
CONSENSU ET AUCTORITATE
AMPLISSIMI PHILOSOPHORUM ORDINIS
IN
ALMA LITTERARUM UNIVERSITATE FRIDERICA GUILELMA
BEROLINENSI
PRO
SUMMIS IN PHILOSOPHIA HONORIBUS
RITE CAPESSENDIS
DIE XIV. M. DECEMBRIS A. MDCCCLXVII

H. L. Q. S.
PUBLICE DEFENDET
AUCTOR
GEORGIUS CANTOR
PETROPOLITANUS.

(1867)

ADVERSARII ERUNT:
M. SIMON, DR. PHIL.
M. HENOCH, DR. PHIL.
E. LAMPE, DR. PHIL.

BEROLINI
TYPIS CAROLI SCHULTZII
KOMMANDANTEN-STRASSE.

Université de Berlin

Ernst Kummer (1810–1893)

Leopold Kronecker (1823–1891)

Karl Weierstrass (1815–1897)

Christoffel (56), Fuchs (58),
Schwarz (64), Cantor (67),
Frobenius (70), Killing (72),
Kovalevskaya (74), Schoenflies (77),
Runge (80), Kneser (84),
Hensel (84), Lerch (85),

page précédente, à gauche : la page de garde de la thèse de Cantor ; la thèse est en latin, une pratique alors sur le déclin. À droite en haut, les trois « grands professeurs » de Berlin à cette époque ; en bas, quelques noms d'étudiants de l'Université de Berlin qui sont devenus des mathématiciens notables.

---

deux pages suivantes : un CV de Cantor, et une déclaration des « thèses » ou « propositions » qu'il entend défendre ; la page de Cantor est découpée ici en deux morceaux. Né à Saint-Pétersbourg en 1845, Georg Cantor vient en Allemagne à l'âge de 11 ans, avec ses parents ; ils s'installent à Francfort, dans la province de Hesse. Il fait ses études secondaires à Darmstadt, aussi en Hesse, jusqu'en 1862.

Il commence un semestre universitaire à Zürich, mais la mort de son père en 1863 le fait revenir. La même année il s'inscrit à Berlin où il poursuit jusqu'à la thèse, avec une interruption d'un semestre passé à Göttingen.

La thèse de Cantor est dans le domaine de l'algèbre. On notera la proposition que Cantor entend soutenir : « *en mathématiques, l'art de poser un problème est plus important que de le résoudre* ».

Vita.

(1845)

Ego, Georgius Cantor, mense Martio anni MDCCCXLV patre Georgio, matre Maria, e gente Böhm, Petropoli natus sum, quo pater meus, negociator Copenhagcniensis commigraverat. — Fidei addictus sum evangelicae. Primis litterarum elementis in schola S. Petri imbutus, puer undecim annorum Germaniam cum parentibus petivi.

Darmstadiae scholam realem, deinde scholam polytechnicam, quae rectore beato Pr. Dr. Külp florebant, quatuor per annos frequentavi, ubi anno
(1862) MDCCCLXII testimonium maturitatis adeptus sum.

Sub hiemem ejusdem anni Turicum profectus sum, unde tamen morte patris mei, pia memoria per vitam colendi, jam ineunte vere anni MDCCCLXIII
(1863) revocatus sum. Auctumno ejusdem anni, inter cives universitatis litterariae Friderico Guilelmae a rectore magnifico Beseler receptus, a decano spect. Müllenhoff philos. ordini adscriptus sum.

Per octo semestria studiis mathematicis me dedi fere continuo Berolini, nisi quod per semestre aestivum anni MDCCCLXVI in civitate Georgia Augusta Gottingensi versatus sum. Legentes audivi Berolini viros ill. Arndt, Dove, Kronecker, Kummer, Magnus, Trendelenburg, Weierstrass; Gottingiae viros ill. Lotze, Minnigerode, Schering, Weber.

Excrcitationibus seminarii mathematici, quas moderantur viri ill. Kummer et ill. Wcierstrass per quattuor semestria interfui.

Quibus viris omnibus maxime de me meritis, imprimis ill. Kronecker, ill. Kummer, ill. Weierstrass, qui benevolentissime tironem adjuverunt, summas, quas possum, gratias ago.

Anno proximo disquisitionibus arithmeticis cel. Gauss et theoriae numerorum cel. Legendre operam navavi, unde materiam dissertationis depromsi.

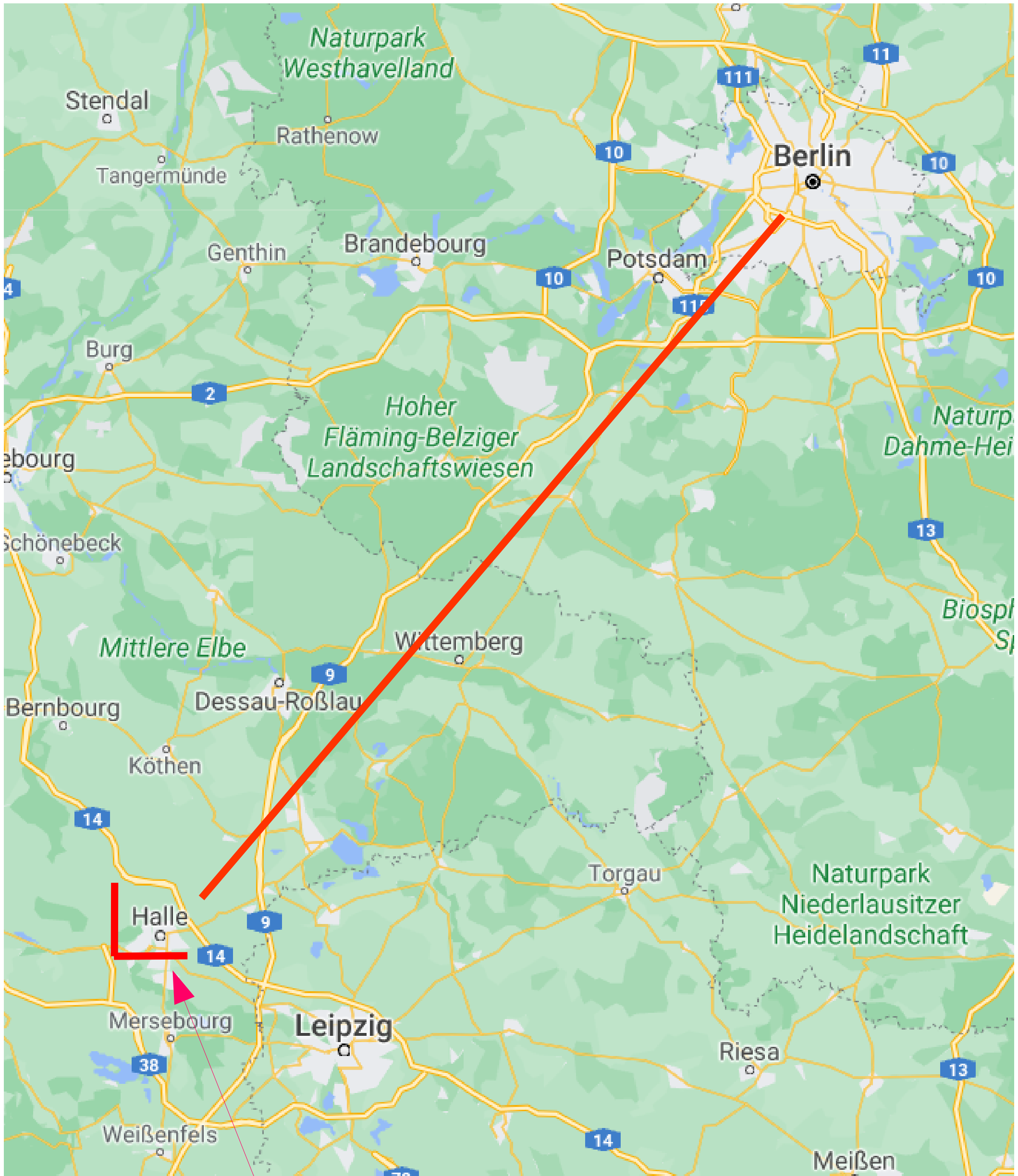
Theses.

I. In arithmetica methodi mere arithmeticae analyticis longe praestant.

II. Num spatii ac temporis realitas absoluta sit, propter ipsam controversiae naturam dijudicari non potest.

III. In re mathematica ars proponendi quaestionem pluris facienda est quam solvendi.

(document visible dans le volume des Œuvres de Cantor)



Halle (environ 150 km de Berlin
à vol d'oiseau)

À Halle : habilitation de Cantor au printemps 1869

Eduard Heine (1821–1881)

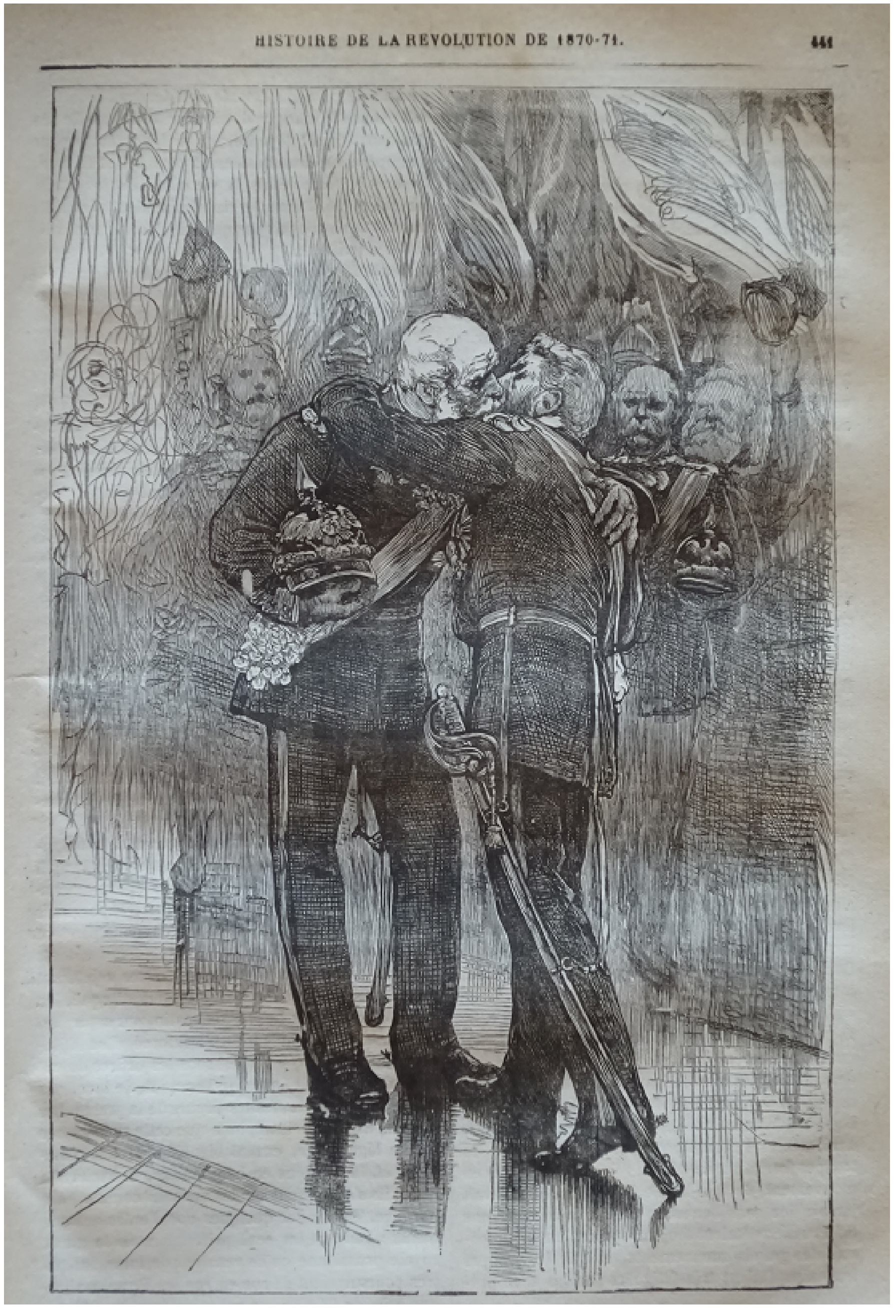
professeur titulaire à Halle

page précédente : pour devenir enseignant universitaire, une « Habilitation » est l'étape nécessaire après la thèse ; Cantor écrit un mémoire d'Habilitation, encore dans le domaine algébrique, et obtient un poste d'enseignant (non titulaire : « extraordinarius ») à l'Université de Halle, bien moins prestigieuse pour les sciences que celle de Berlin. Il espérera longtemps obtenir un poste à Berlin, mais cela ne se produira jamais ; Cantor suspectera que l'opposition à ses travaux professée par Kronecker sera en bonne partie responsable de son « échec berlinois ».

---

deux pages suivantes : en 1870, Napoléon III déclare la guerre à la Prusse. Il est battu, les Prussiens et leurs alliés envahissent la France et font le siège de Paris. Sur la gravure de la page suivante, provenant d'un livre de Jules Claretie, on les voit combattre à proximité de l'église de Bry-sur-Marne, bien reconnaissable sur l'image (suivre la flèche blanche). En 1871, le roi de Prusse fait le voyage de Versailles et y est proclamé Empereur d'Allemagne (deuxième page suivante, une image prise dans le même livre).

Le siège de Paris – Prise de Bry-sur-Marne, le 30 novembre 1870 (J. Claretie)



Le roi de Prusse Guillaume Ier
devient Empereur d'Allemagne
à Versailles dans la
Galerie des Glaces du château,

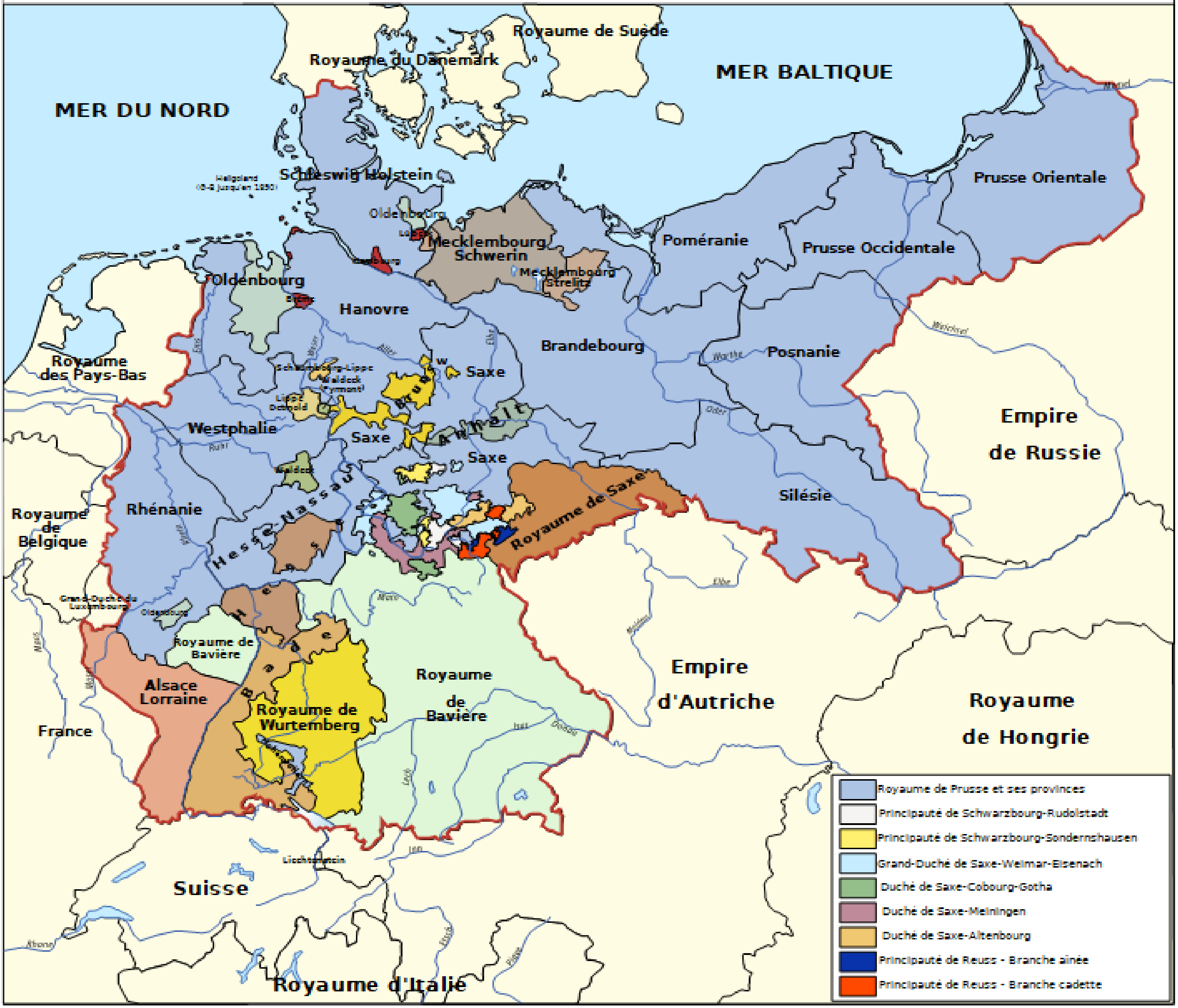
le 18 janvier 1871

Memel (de nos jours : Klaipėda en Lituanie)
Königsberg (Kaliningrad)
LE SECOND REICH
1871-1918
Royaume de Prusse
Danemark
Schleswig Holstein
Mecklembourg
Poméranie
Prusse Occidentale
Prusse Orientale
Hanovre
Brandebourg
Berlin
Posnanie
Pays Bas
Westphalie
Saxe
SAXE
Belgique
Rhénanie
Nassau
Hesse
Thuringe
Royaume de Saxe
Silésie
Empire Russe
FRANCE
Alsace
Lorraine
Bavière
Bade
Wurtemberg
Royaume de Bavière
Suisse
Empire Austro-Hongrois
1 Oldenbourg
2 Lauenbourg
3 Lubeck
4 Hambourg
5 Brème
6 Schaumbourg
7 Lippe Detmold
8 Waldeck
9 Braunschweig
10 Hesse Kassel
11 Anhalt
12 Strelitz
Darmstadt
Zürich
Göttingen
Berlin

page précédente : la carte (que j'ai volée sur internet) montre l'étendue de l'Empire, qui inclut alors l'Alsace et la Lorraine après la défaite française, et qui s'étend à l'est jusqu'à la Lituanie actuelle, en passant par Königsberg (aujourd'hui Kaliningrad, « exclave » russe située entre Pologne et Lituanie). On a indiqué les quatre villes où Cantor a étudié. Berlin occupe une place assez centrale dans cet Empire.

page suivante : une autre carte de l'Empire, peut-être plus lisible. On pourra y noter que la Pologne et la Tchéquie n'existent pas !

L'EMPIRE ALLEMAND
1871 - 1918



Séries trigonométriques

Série trigonométrique,

$$\text{(T)} \qquad \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)), \qquad x \text{ variable réelle,}$$

coefficients a_n, b_n réels.

ci-dessus **:** l'écriture classique des séries trigonométriques, avec des coefficients réels a_n et b_n **;** on préfère en général de nos jours l'écriture qui utilise les fonctions exponentielles complexes ainsi que des coefficients c_n complexes **:**

$$\sum_{n \in \mathbb{Z}} c_n \, \mathrm{e}^{\mathrm{i}\, n\, x}$$

page suivante, quelques noms qu'on peut attacher aux séries trigonométriques avant 1870 **:** au 18e siècle, Daniel Bernoulli a proposé d'utiliser des combinaisons linéaires de fonctions trigonométriques pour trouver des solutions générales du *problème des cordes vibrantes*. Vers 1810, dans ses études sur la propagation de la chaleur, Fourier a systématisé l'association d'une série trigonométrique à toute fonction raisonnable. En 1829 (après avoir été en contact avec Fourier à Paris), Dirichlet a démontré l'un des premiers « théorèmes au sens moderne » concernant la convergence de la série de Fourier d'une fonction. On reparlera des travaux trigonométriques de Riemann plus loin. Et on indique un bel exemple de série dont le calcul de la somme est dû à Euler.

Euler vers 1750 :

$$\sum_{n=1}^{\infty} \frac{\sin(nx)}{n} \longrightarrow$$

Daniel Bernoulli (1700–1782)

Leonhard Euler (1707–1783)

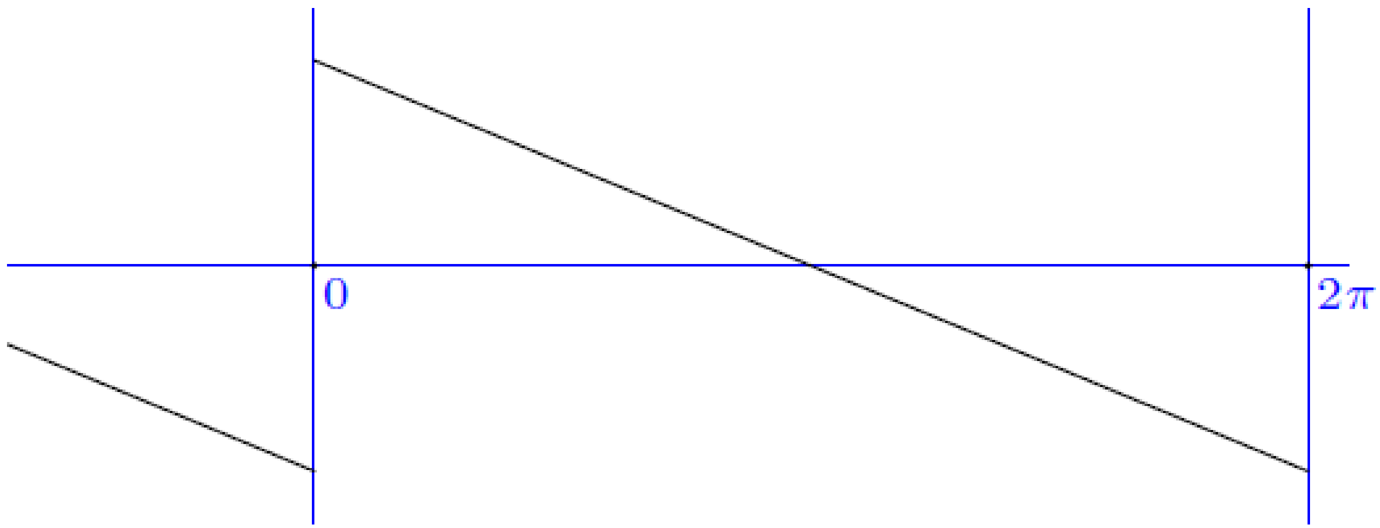
Joseph Fourier (1768–1830)

Gustav Lejeune-Dirichlet (1805–1859)

Bernhard Riemann (1826–1866)

$$\sin(x) + \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \cdots + \frac{\sin(nx)}{n} + \cdots = \frac{\pi - x}{2}$$

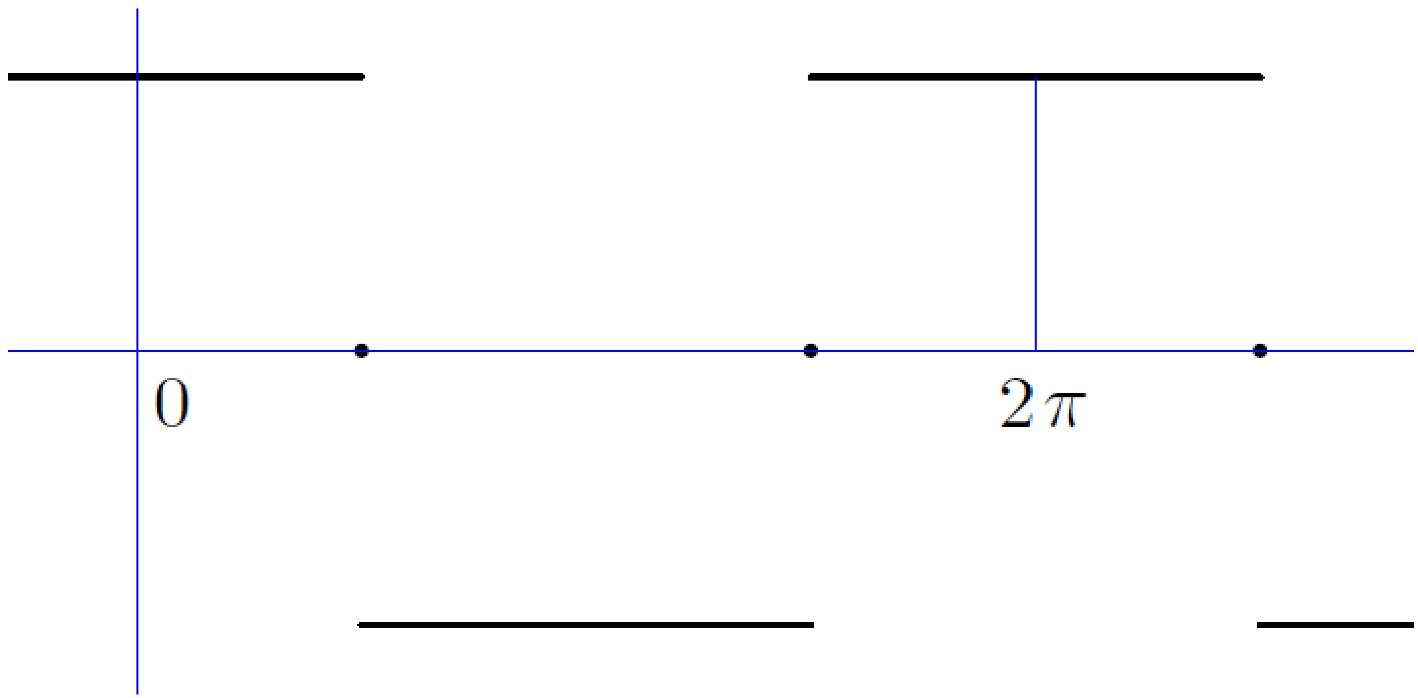
$$(0 < x < 2\pi)$$



Euler (exemple repris par Fourier)

Joseph Fourier (1805, 1812, 1822) :

$$\cos(x) - \frac{\cos(3x)}{3} + \frac{\cos(5x)}{5} - \cdots + (-1)^n \frac{\cos(2nx+x)}{2n+1} + \cdots$$



fonction créneau

page précédente : Fourier a donné le développement trigonométrique d'une « fonction créneau » périodique, qui lui sert dans le cadre de son étude de l'évolution de la température dans une plaque rectangulaire homogène dont les bords sont maintenus à des températures fixes, plaque qu'on peut imaginer représentée sous la forme

$$\{(x, y) \ : \ |x| \leq \pi / 2, \ y \geq 0 \}.$$

On remarquera que la série « créneau », tout comme la série qu'on a attribuée à Euler, *converge en tout point* : ce sera le cadre pour les premiers théorèmes de Cantor. On notera aussi que la somme de la série, aux points de discontinuité, est égale à la demi-somme des limites à droite et à gauche : cette propriété fait partie de l'énoncé du *théorème de Dirichlet* qu'on a mentionné plus haut.

Weierstrass utilise une série trigonométrique du type

$$\sum_{n=0}^{\infty} \frac{\cos(5^n x)}{2^n}$$

pour donner le premier exemple d'une fonction continue qui n'est dérivable en aucun point. Il n'utilise pas les valeurs 5 ni 1/2 mais des nombres réels a et b vérifiant une condition de la forme $a\,b >$ K, pour un certain K qu'on ne va pas donner ici ; ce qui importe est que 5 . (1/2) > 1 (largement).

(Weierstraß, 1872 ; dans *Werke II*, p. 71 – 74.)

Il est impossible de dessiner le graphe de cette fonction ! C'est l'un des objets mathématiques de cette période qu'on peut *imaginer* et traiter par les outils de l'Analyse, mais qui échappe à la représentation géométrique traditionnelle.

Cantor, G.: *Über einen die trigonometrischen Reihen betreffenden Lehrsatz*. Journal f. reine und angew. Math. *72* (1870), 130–138.

signé **:** *à Berlin, le 20 Mars 1870*

Cantor, G.: *Beweis, daß eine für jeden reellen Wert von x durch eine trigonometrische Reihe gegebene Funktion f (x) sich nur auf eine einzige Weise in dieser Form darstellen läßt*. Journal f. reine und angew. Math. *72* (1870), 139–142.

signé **:** *à Berlin, le 6 Avril 1870*

« Preuve du fait qu'une fonction $f(x)$ qui est donnée pour toute valeur de x par une série trigonométrique ne peut se représenter sous cette forme que d'une seule manière ».

page précédente **:** les titres des deux premiers articles importants de Cantor, qui sont parus l'un à la suite de l'autre dans une très bonne revue allemande, qu'on surnommait « Journal de Crelle », du nom du fondateur, ou « Journal de Borchardt », du nom de l'éditeur entre 1855 et 1880.

---

page suivante **:** Cantor mentionne d'entrée un texte dû à Riemann, consacré à la représentation des fonctions en série trigonométrique, qui jouera un rôle essentiel dans son travail. En note de bas de page, Cantor explique que ses deux articles ont été suscités par des discussions avec Heine, professeur à Halle, qui s'intéressait à ce moment-là aux séries trigonométriques.

«
J'ai été incité à écrire les articles qui suivent par [*des contacts avec*] M. Heine. Ce dernier a eu la bonté de me communiquer la primeur de ses recherches sur les séries trigonométriques. Mes deux articles sont issus de la tentative d'étendre ses résultats en évitant toute hypothèse sur le type de convergence de ces séries.
»

Ueber einen die trigonometrischen Reihen betreffenden Lehrsatz.

(Von Herrn *G. Cantor* in Halle *).)

Riemanns Forschungen im Gebiete der trigonometrischen Reihen sind in der Abhandlung „Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe, Göttingen 1867“ bekannt geworden.

note de bas de page :

---

*) Zu den folgenden Arbeiten bin ich durch Herrn *Heine* angeregt worden. Derselbe hat die Güte gehabt, mich mit seinen Untersuchungen über trigonometrische Reihen frühzeitig bekannt zu machen. Aus dem Versuche seine Resultate in der Richtung zu erweitern, dass jedwede Voraussetzung über die *Art* der Convergenz bei den auftretenden Reihen vermieden wird, sind beide hervorgegangen.

Théorème d'unicité de Cantor

Si la série (T) converge pour tout x réel et

$$\text{si} \quad \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big) = 0$$

pour tout x réel, alors tous les coefficients sont nuls,

$$a_0 = a_1 = \ldots = a_n = \ldots = 0, \qquad b_1 = \ldots = b_n = \ldots = 0.$$

Il s'agit bien d'un théorème d'unicité **:** si une fonction avait *deux* développements en série trigonométrique qui la représentent *en tout point*, la série obtenue en faisant la différence de ces deux séries aurait une somme nulle en tout point, et les différences des coefficients de ces deux séries seraient donc nulles d'après le théorème, autrement dit, *les deux séries seraient identiques*.

Lemme de Cantor :

si pour tout x d'un intervalle ouvert non vide (u, v) on a

$$a_n \cos(nx) + b_n \sin(nx) \underset{n}{\longrightarrow} 0,$$

alors les coefficients tendent vers 0,

$$a_n \underset{n}{\longrightarrow} 0, \qquad b_n \underset{n}{\longrightarrow} 0.$$

La preuve de ce « lemme de Cantor » occupe la totalité du premier article qu'on a mentionné précédemment. Cette première preuve du lemme n'avait pas encore la simplicité qu'on saura lui donner plus tard (nous choisirons la version simple !).

page suivante **:** on prépare la preuve du lemme, une preuve par l'absurde. Pour simplifier cet exposé, on suppose que la « suite trigonométrique » étudiée ne contient que des fonctions sinus. On fait l'hypothèse absurde que la suite $b_n \sin(n\,x)$ tend vers 0 pour tout x d'un intervalle non vide (u, v), mais que b_n ne tend pas vers 0. Il existe alors une sous-suite des coefficients b_n qui reste « loin » de 0, et on peut supposer que les indices n_k dans cette sous-suite sont très espacés...

Cas plus simple *à écrire* :

$$b_n \sin(nx) \longrightarrow 0$$

Si

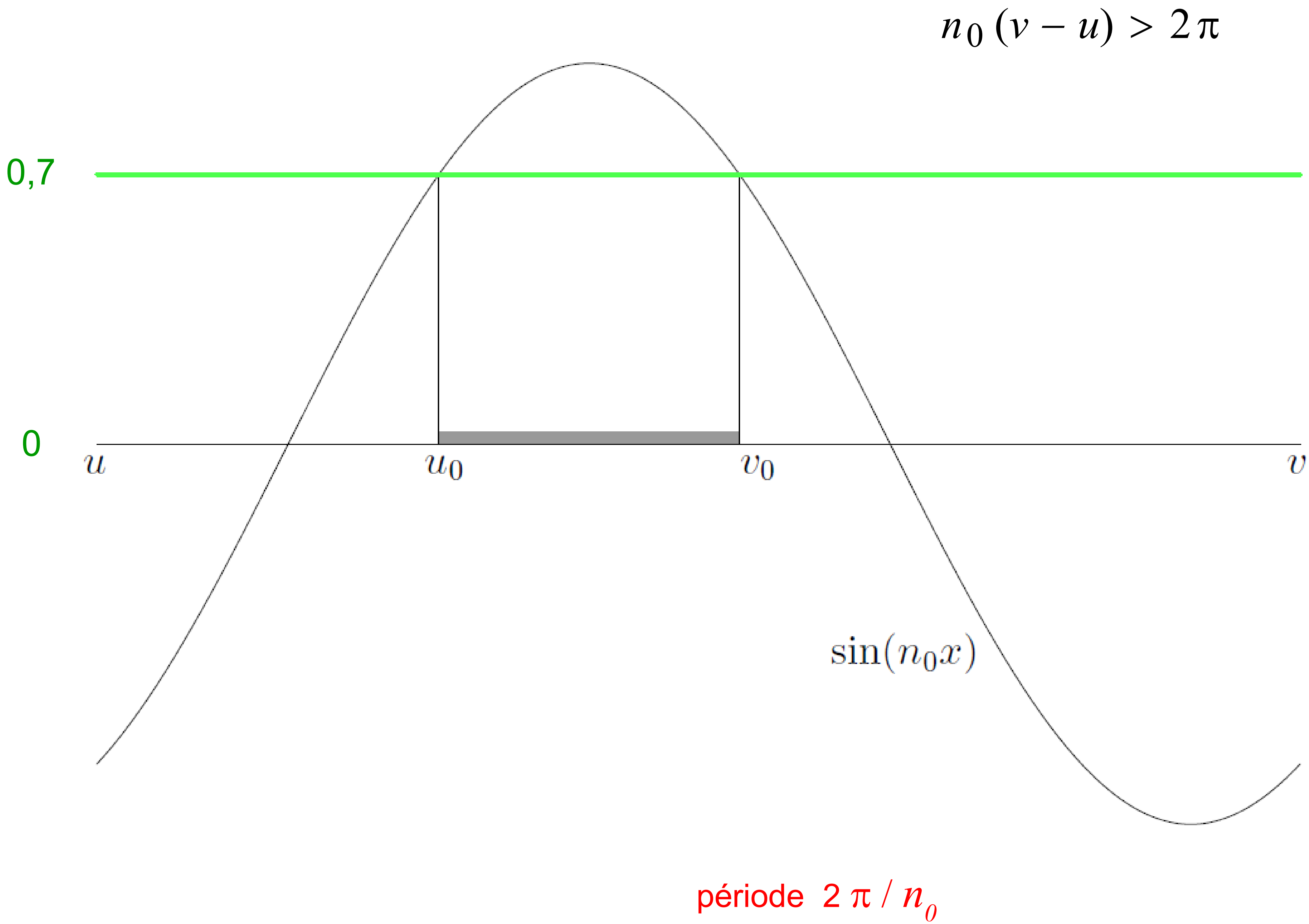
$$|b_{n_k}| \geqslant \delta > 0$$

alors

$$\sin(n_k x) \longrightarrow 0$$

mais c'est impossible :

$n_{k+1} \geqslant 100\, n_k$ et n_0 grand

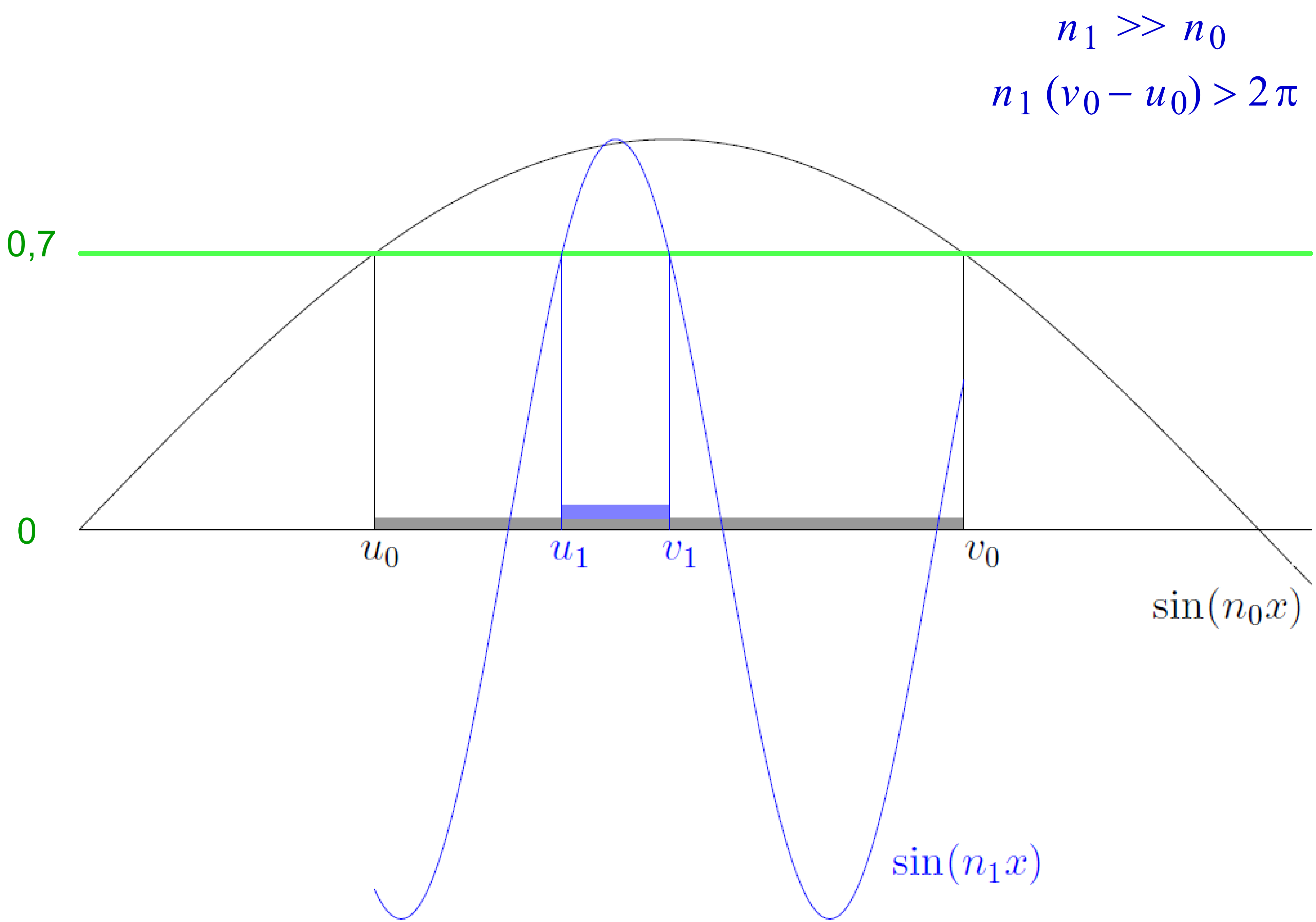


$n_0 (v - u) > 2\pi$
0,7
0
u
u_0
v_0
v
$\sin(n_0 x)$
période $2\pi / n_0$

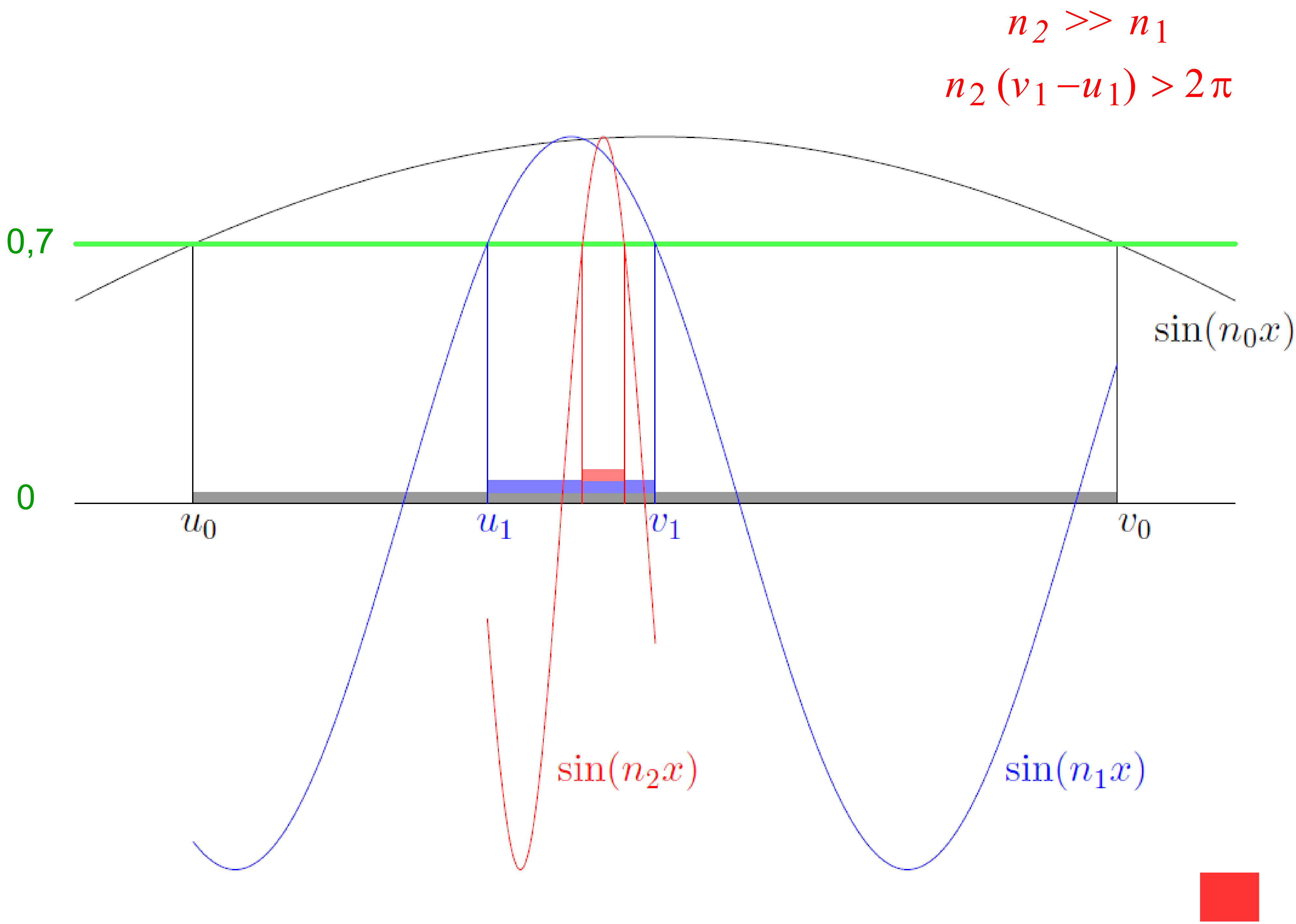
page précédente : on a pu choisir n_0 assez grand pour que la fonction $\sin(n_0 x)$ accomplisse une période complète à l'intérieur de l'intervalle donné (u, v). Il y a donc un segment $[u_0, v_0]$ intérieur à (u, v) (avec bien sûr $u_0 < v_0$) dans lequel cette fonction est « grande » : sur la figure, supérieure ou égale à 0.7.

---

page suivante : on a pu également choisir n_1 nettement plus grand que n_0, de façon que la nouvelle fonction $\sin(n_1 x)$ accomplisse une période complète à l'intérieur de l'intervalle (u_0, v_0). Il y a donc un segment $[u_1, v_1]$, $u_1 < v_1$, intérieur à (u_0, v_0) dans lequel cette nouvelle fonction est « grande », etc.



$n_1 >> n_0$
$n_1 (v_0 - u_0) > 2\pi$
0,7
0
u_0
u_1
v_1
v_0
$\sin(n_0 x)$
$\sin(n_1 x)$



$n_2 >> n_1$
$n_2 (v_1 - u_1) > 2\pi$
0,7
0
u_0
u_1
v_1
v_0
$\sin(n_0 x)$
$\sin(n_2 x)$
$\sin(n_1 x)$

page précédente : on a pu introduire de proche en proche une suite décroissante de segments, sur lesquels une des fonctions de la sous-suite est « grande ». En un point x de l'intersection de cette suite de segments (il y en a : on le sait bien en 1870), *toutes* les fonctions sinus de la sous-suite sont grandes, il n'y a donc pas convergence vers 0 en ce point x. Cette contradiction termine la démonstration du lemme de Cantor.

---

page suivante : on va passer à la preuve du théorème. Si la série de sinus dont la somme est $f(x)$ converge en tout point — ou bien seulement en tout point d'un intervalle (u, v) —, on saura en tout cas que son terme général tend vers 0. C'est précisément l'hypothèse du lemme de Cantor : on déduit donc que les coefficients b_n tendent vers 0, et en particulier qu'ils sont bornés, ce qui justifie l'existence de la fonction F dont on va beaucoup reparler plus loin. Le fait plus précis que les coefficients b_n tendent vers 0 interviendra aussi, mais pas tout de suite.

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$$

La somme $f(x)$ est définie pour certaines valeurs de x, la somme $F(x)$ est définie pour *toutes* les valeurs de x.

On va garder en mémoire ces deux séries.

$$F(x) = \sum_{n=1}^{\infty} \frac{b_n}{n^2} \sin(nx)$$

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) \qquad\qquad \mathrm{F}(x) = \sum_{n=1}^{\infty} \frac{b_n}{n^2} \sin(nx)$$

Bernhard Riemann (1826–1866) a montré en 1854 que : *si la série qui définit $f(x)$ converge au point x, alors*

$$(*) \qquad \lim_{h\to 0} \frac{\mathrm{F}(x+h)+\mathrm{F}(x-h)-2\mathrm{F}(x)}{h^2} = -f(x)$$

La limite (*) peut être considérée comme une *dérivée seconde généralisée* de F au point x. Ce résultat de Riemann est un substitut pour ce qu'on aimerait faire, sans en avoir le droit : si on dérivait brutalement *terme à terme* deux fois la série de F, « on voit bien » qu'on obtiendrait $-f(x)$.

page précédente : les résultats trigonométriques qui vont être cruciaux pour Cantor ont été prouvés par Riemann longtemps avant, en 1854 ; ils faisaient partie de son « Habilitationsschrift » et n'avaient jamais été publiés de son vivant ; ils ne sont parus qu'en 1867, un an après la mort de Riemann.

page suivante : si $f(x)$ vaut 0 en tout point d'un intervalle (u, v), on pense que la dérivée seconde de F est nulle sur cet intervalle, donc F est affine sur l'intervalle : ce n'est pas tout à fait aussi simple puisque ce n'est pas la *vraie* dérivée seconde de F qui est nulle, il faut un certain *travail*, pour lequel Cantor a reçu l'aide de Hermann Schwarz, autre ancien de l'Université de Berlin, ami de Cantor à ce moment (ils seront brouillés plus tard).

Hermann Schwarz (1843–1921)

si « $f(x)$ est défini » ET $f(x) = 0$ pour tout $x \in (u, v)$,
alors F est affine sur (u, v).

Et si la « nullité » de f a lieu pour tout x ?

Si $f(x)$ vaut 0 en tout point x de la droite réelle, on déduit que F est affine sur $\mathbf{R}$. Mais la fonction F est périodique, elle est donc constante, et en fait identiquement nulle puisqu'on voit que F(0) = 0.

(à suivre) →

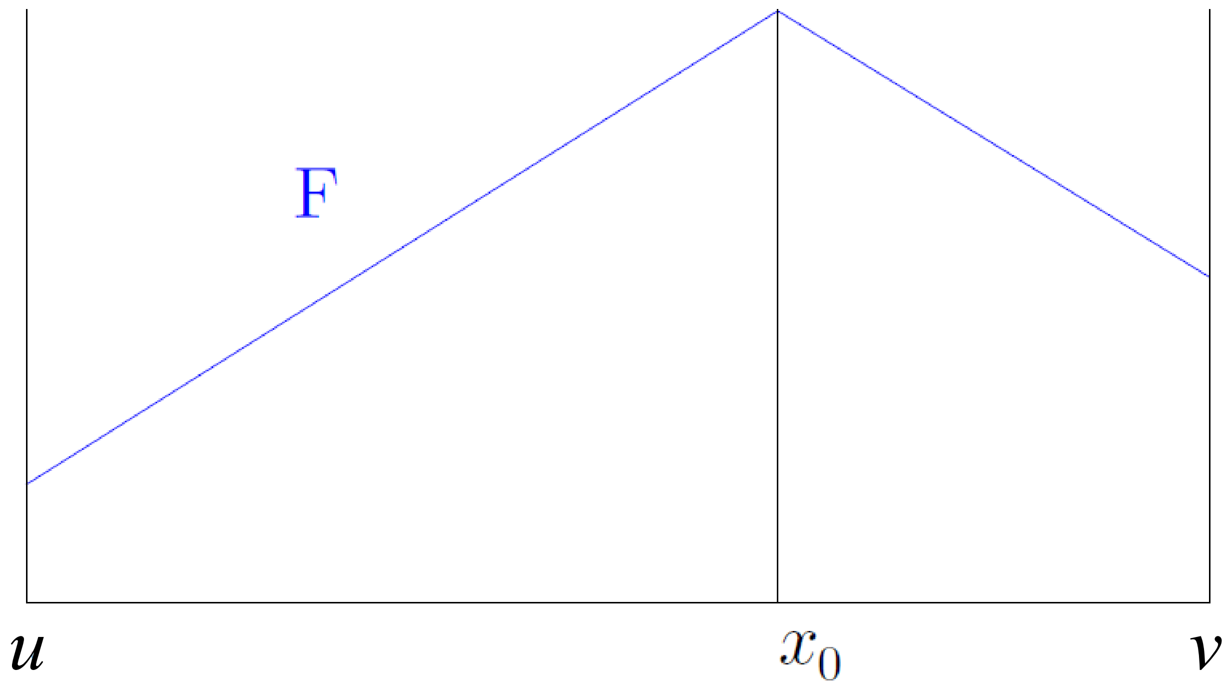
Comme F est la somme d'une série « *normalement convergente* », il est facile de justifier le calcul des coefficients de la série de F par interversion de la sommation de la série et du calcul de l'intégrale qui fournit les « coefficients de Fourier ». Puisque F est nulle, on déduit que les coefficients b_n / n^2 sont nuls, donc les b_n sont nuls aussi. Le théorème d'unicité est démontré.

page suivante : il y a beaucoup d'exemples de séries trigonométriques qui convergent en tout point *sauf* pour *un* point dans chaque période, ou *sauf* en *un nombre fini* de points dans chaque période. Le premier résultat d'unicité de Cantor ne s'applique pas à ces séries, on peut imaginer qu'il a rapidement voulu étendre son résultat à ce cas un peu plus général.

Et si l'hypothèse est vraie sauf pour un point exceptionnel x_0 ?

alors

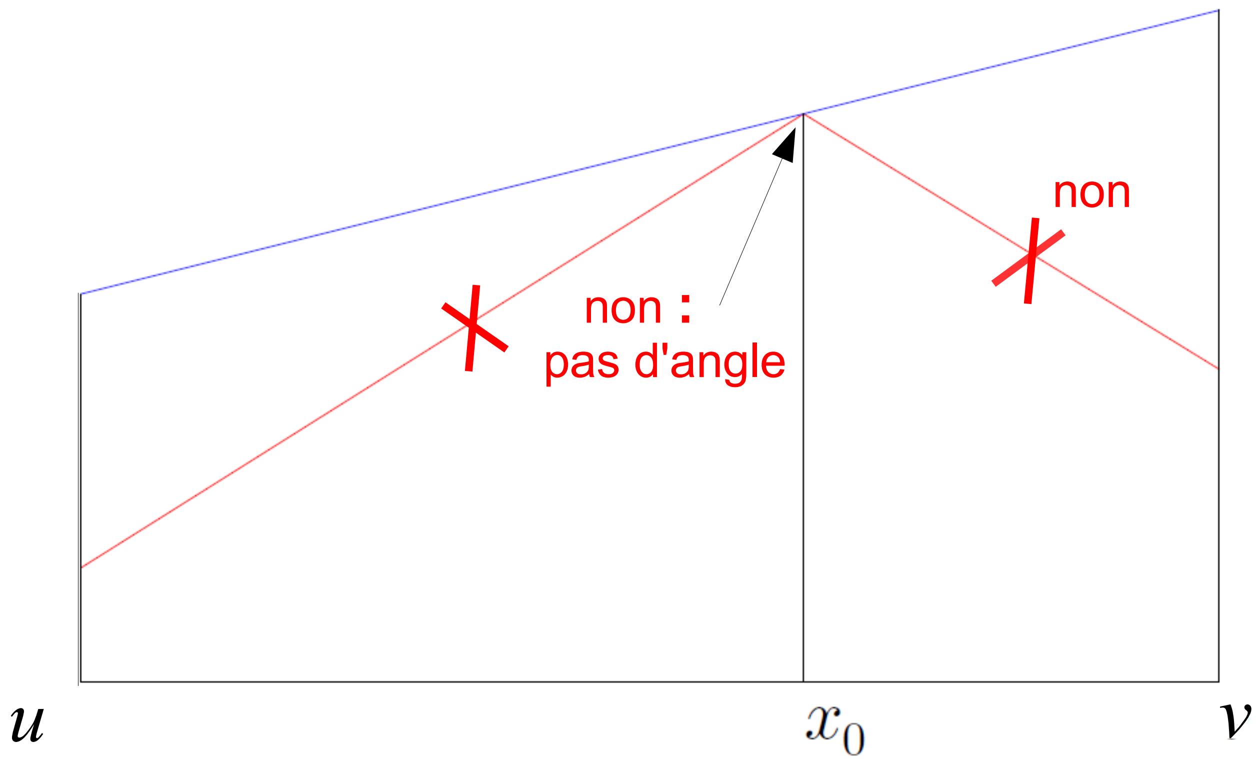
D'après ce qui précède, la fonction F de Riemann est affine sur (u, x_0) ET sur (x_0, v), et de plus elle est continue.



encore Riemann en 1854 (et ce n'est pas très difficile) **:**

si les coefficients b_n *tendent vers* 0, *alors*

$$(**) \qquad \lim_{h \to 0} \frac{\mathrm{F}(x+h) + \mathrm{F}(x-h) - 2\mathrm{F}(x)}{h} = 0$$



page précédente : Cantor est plutôt chanceux ! Dans le même article de Riemann, on trouve une autre limite (**), valable sans hypothèse de convergence en x pour la série de f ; la nullité de cette limite entraîne l'absence de « coin » dans le graphe de F, c'est-à-dire que F est affine sur la totalité de l'intervalle considéré. On se retrouve alors dans la même situation que quand $f(x)$ existe et est nul partout, et on conclut de la même manière. Il est facile d'étendre le résultat au cas où il y aurait un *ensemble fini* de points où $f(x)$ n'existe pas.

---

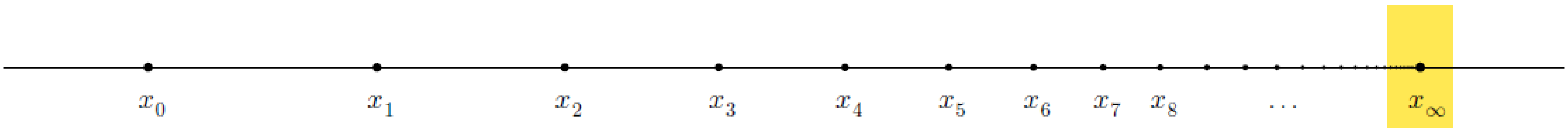
page suivante : en 1872 Cantor va faire un progrès décisif ; il va réussir à traiter des séries trigonométriques qui, dans chaque période, sont convergentes et ont une somme nulle, *sauf* aux points d'une suite convergente et en sa limite : en ces points *exceptionnels*, il est possible que la série ne converge pas, ou bien qu'elle converge mais que sa somme ne soit pas nulle.

Cantor en 1872

Supposons qu'il existe un « ensemble exceptionnel » infini :

(en un *point exceptionnel*, il est possible que la série de f ne converge pas, ou bien qu'elle converge mais que sa somme ne soit pas nulle)

On suppose que ces points sont situés dans une même période :



$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$$

$$\mathrm{F}(x) = \sum_{n=1}^{\infty} \frac{b_n \sin(nx)}{n^2}$$

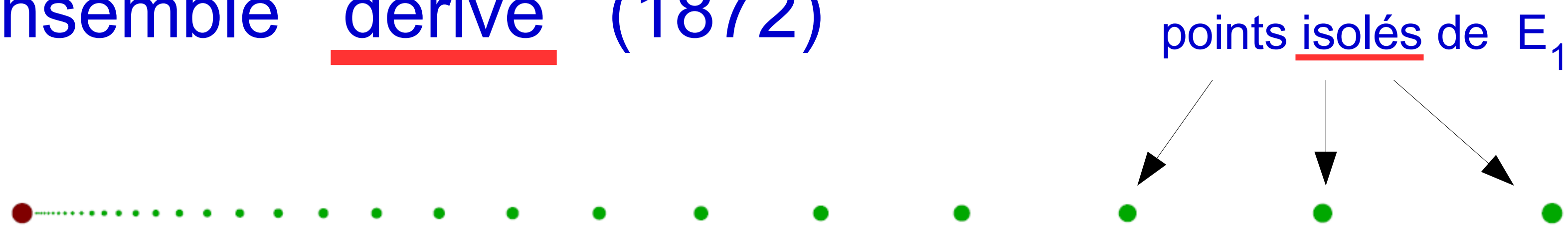
page précédente : si on efface un petit intervalle « jaune » autour de la limite de la suite de points exceptionnels, on obtient de part et d'autre de cette limite deux intervalles qui ne contiennent *qu'un nombre fini* de points exceptionnels : on sait alors que la fonction F est affine dans ces deux intervalles. Mais comme cet intervalle jaune est arbitrairement petit, on déduit que F est affine jusqu'au point limite, de chaque côté, et on retrouve le cas d'un point exceptionnel unique, déjà traité. On peut traiter aussi le cas où l'ensemble exceptionnel est formé *d'un nombre fini* de suites convergentes et de leurs limites.

---

page suivante : les points de l'ensemble E (les « points exceptionnels ») qu'on vient de traiter se divisent en deux classes : Cantor dit que le *point limite* fait partie de *l'ensemble dérivé* de E ; les points exceptionnels distincts de la limite sont des *points isolés* : on peut trouver un petit intervalle ouvert contenant un point isolé et ne contenant *aucun autre* point de l'ensemble E. L'ensemble dérivé de l'ensemble E_1 de la page suivante est réduit au point limite, c'est un *singleton*.

point limite (point d'accumulation), voisinage d'un point ;
point isolé

Ensemble dérivé (1872)



si E_1 = la suite ET le point limite, alors

le dérivé $(\mathrm{E}_1)'$ n'a qu'un seul point

un ensemble fini « n'a pas » d'ensemble dérivé (1872)

plus tard (années 1880) :

un ensemble est *fermé* s'il contient son dérivé

un ensemble est *parfait* s'il est non vide et égal à son dérivé page précédente : l'ensemble dérivé E' d'un ensemble de points E est constitué des *points d'accumulation* de E. Un ensemble fini n'a pas de point d'accumulation, on dit aujourd'hui que son dérivé est *l'ensemble vide*, Cantor préfère dire en 1872 qu'un ensemble fini *n'a pas de dérivé*.

L'ensemble dérivé E' n'est pas nécessairement contenu dans E : le dérivé de l'ensemble R formé des rationnels de [0, 1] est le segment [0, 1] tout entier. Mais le dérivé E' est toujours un ensemble fermé, il contient donc ses points d'accumulation et par conséquent E'' = (E')' est contenu dans E' ; plus généralement, les dérivés successifs E', E'', E''' = (E'')', . . . forment une suite décroissante de fermés.

---

page suivante : un exemple numérique d'ensemble E_1 qui est formé d'une suite convergente et de sa limite. L'ensemble dérivé $(E_1)'$ est égal à {0}.

L'ensemble E_1 contient tous les points

$$\frac{1}{n}, \quad n \text{ entier } \geqslant 1$$

et le point limite 0,

$$\frac{1}{n} \xrightarrow[n]{} 0$$

son ensemble dérivé est un *singleton* :

$$(\mathrm{E}_1)' = \{0\}$$

(voir le fichier d'illustrations gentillettes :

LimitesB.pdf ← (lien cliquable)

à propos de la notion d'ensemble dérivé et des dérivés successifs)

Nombres réels

Nombres réels (autour de 1870)

Karl Weierstraß (1815–1897) — nombres réels et propriétés des fonctions continues

Charles Méray (1835–1911)

Georg Cantor (1845–1918) — réels définis par des suites de Cauchy

Richard Dedekind (1831–1916) — réels par coupures

Eduard Heine (1821–1881) — fait le point, dans un article assez connu page précédente **:** en même temps qu'il traite la question des ensembles exceptionnels infinis, Cantor présente sa définition des nombres réels au moyen de suites de Cauchy. Il a été devancé dans cette voie par le français Méray (en 1869), mais la contribution de ce dernier est passée inaperçue pendant un long moment. Dedekind pour sa part a présenté la *méthode des coupures*. Dans ses cours à Berlin, Weierstrass avait aussi son système pour introduire les nombres réels et prouver les propriétés des fonctions continues sur un segment. Heine a écrit un article qui fait le point sur ces résultats de Weierstrass (essentiellement non publiés), et il utilise « les réels de Cantor » dans son exposé.

---

page suivante, bordure rouge **:** Heine reconnaît ce qu'il doit à Cantor :

« Je dois des remerciements particuliers à M. Cantor pour ses communications orales, qui ont eu une influence notable sur la conception de mon travail. Je lui dois l'idée d'introduire les nombres généralisés [*les nombres réels*] au moyen de certaines suites adaptées qui seront appelées ici « Zahlenreihen » [*suites de nombres*]. »

handlung auf die Fundamentalsätze der Functionenlehre zu beziehen, welche mich dennoch zur Veröffentlichung der gegenwärtigen veranlasste, in der ich schliesslich diese Sätze beweise.

Zu besonderem Danke bin ich dem Herrn *Cantor* in Halle für seine mündlichen Mittheilungen verpflichtet, welche einen bedeutenden Einfluss auf die Gestaltung meiner Arbeiten ausübten, indem ich von ihm den Gedanken entlehnte, die allgemeinen Zahlen vermittelst jener besonders geeigneten Reihen einzuführen, die hier (A, §. 1, Def. 1) Zahlenreihen genannt werden. Es scheint mir dies eine, besonders für die Anwendungen auf die Functionenlehre (B, §. 2, Lehrs. 1), glückliche Fortbildung der ursprünglichen Einführungsart, bei welcher die allgemeineren Zahlen durch die in ihnen enthaltenen Vielfachen gewisser Grössen in unendlicher Anzahl bestimmt werden. Die Berechtigung, das durch die Reihen Eingeführte als Zahlengrösse zu betrachten, findet Herr *Cantor* darin, dass es möglich sei, auch hier die Begriffe des Grösser-, Kleiner- und Gleichseins festzustellen.

Die Frage, was eine Zahl sei, beantworte ich, wenn ich nicht bei den rationalen positiven stehen bleiben will, nicht dadurch dass ich die Zahl be-

Eduard Heine, « Die Elemente der Functionenlehre » (1872)

(c'est l'article qui contient le « théorème de Heine » sur la continuité uniforme)

«

Les nombres rationnels constituent le fondement pour établir le concept plus large de « Zahlengrösse » [*grandeur numérique*] ; je désignerai par A le domaine des rationnels (en y incluant 0).

Je parle de *Zahlengrösse* au sens élargi quand est donnée par une certaine loi une suite (a_n) de rationnels qui a la propriété que la différence $a_{n+m} - a_n$ devient infiniment petite quand n croît, m étant un entier positif quelconque. Autrement dit, étant donné ε arbitraire (rationnel, strictement positif) il existe n_1 tel que $\underline{a_{n+m} - a_n < \varepsilon}$ quand $n \geq n_1$ et quand m est un entier positif quelconque.

»

Cantor (1872, en allemand ; traduction française dans *Acta Math*. 1883) page précédente **:** la définition des nombres réels selon Cantor **;** on peut noter que le nom de Cauchy n'est pas mentionné, et que Cantor a omis la valeur absolue $|a_{n+m} - a_n| < \varepsilon$ que nous mettrions dans la définition. Cantor poursuit en définissant l'égalité de deux « Zahlengrösse » = « nombres réels », l'ordre et les opérations élémentaires d'addition et de multiplication.

Il est devenu tellement habituel de parler de « la droite réelle », ce qui revient à identifier la notion de droite géométrique à l'ensemble des nombres réels, que nous avons peut-être du mal à saisir la portée de l'axiome dit de « Cantor–Dedekind », formulé séparément par Cantor et Dedekind, qui postule que la « droite » de la géométrie classique (d'Euclide, par exemple) peut être mise en correspondance bijective avec les « nouveaux nombres », les nombres réels.

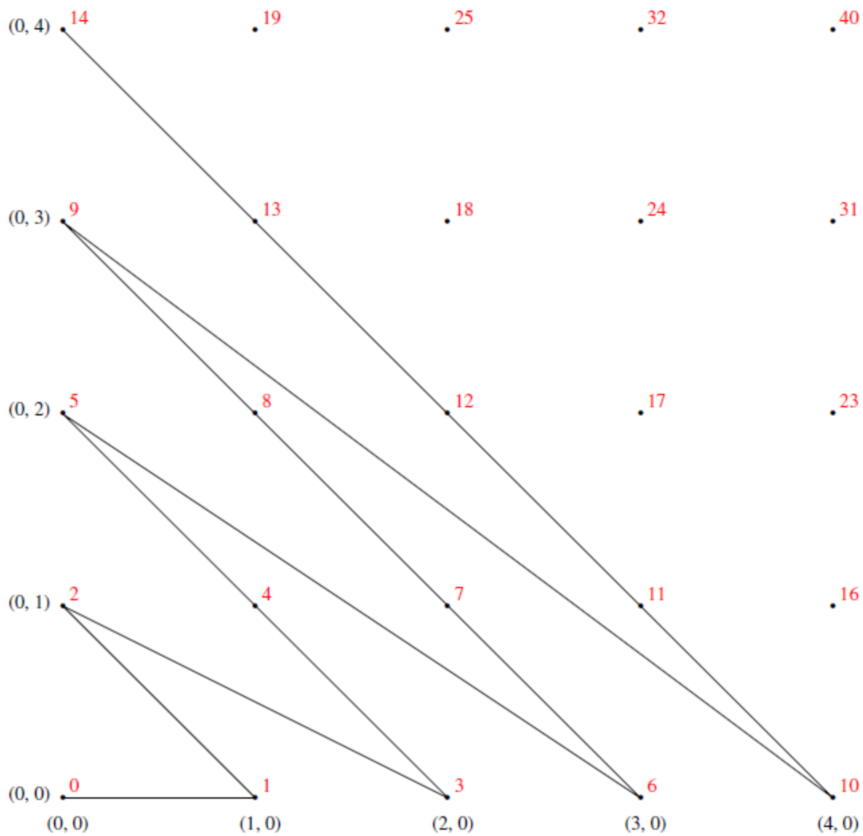
Axiome de Cantor–Dedekind

Version de Cantor **:** les points de la droite de la géométrie classique correspondent bijectivement aux « Zahlengrösse ».

Dénombrable,
ou pas...

Cantor vers 1874

Un parcours habituel de $\mathbf{N} \times \mathbf{N}$:



$\mathbf{Q}$ est dénombrable

L'ensemble de toutes les suites finies d'entiers (où la longueur k varie) est dénombrable :

$$(n_0, n_1, \ldots, n_k)$$

$$n_j \in \mathbb{N}, \quad k \in \mathbb{N}$$

dénombrable = en bijection avec l'ensemble $\mathbf{N}$ des entiers page précédente : en 1874, Cantor expose au monde mathématique la notion de *dénombrabilité* ; il connaît les propriétés des ensembles dénombrables, toutes celles qu'on enseigne aujourd'hui au début d'un cycle universitaire : réunion dénombrable de dénombrables, produit de deux dénombrables, l'ensemble des rationnels est dénombrable, etc.

page suivante : Cantor commence à établir sa théorie de la cardinalité. Il donne sa première preuve du fait qu'il ne peut pas exister de bijection entre l'ensemble des entiers naturels et l'ensemble des nombres réels : l'ensemble des réels n'est pas dénombrable, il a la *puissance du continu*. Cette première preuve n'est pas la preuve fameuse par *argument diagonal* qu'il donnera en 1891 ; on va voir que le ressort de cette première preuve est proche de celui qu'on a vu intervenir dans la preuve du « lemme de Cantor » pour les séries trigonométriques.

Non dénombrabilité de la droite réelle (1874)

on considère une suite de nombres réels :

$$(x_n)_{n \geqslant 0}$$

On va laisser venir les points successifs de la suite donnée, en commençant avec x_0, x_1 :

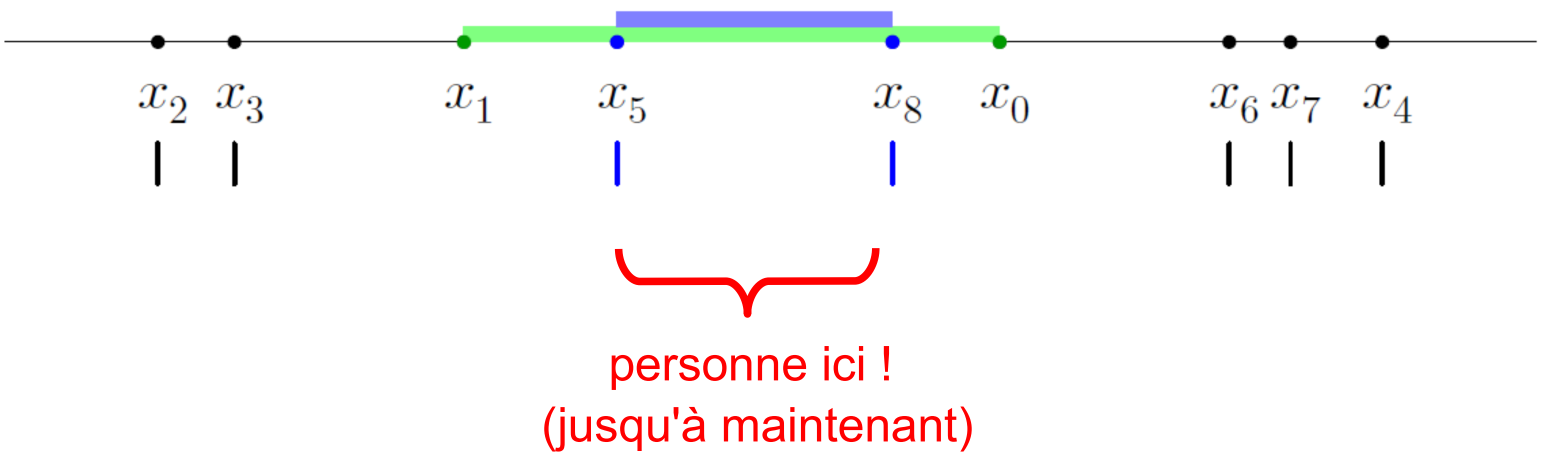
$x_1 \neq x_0$, etc ...

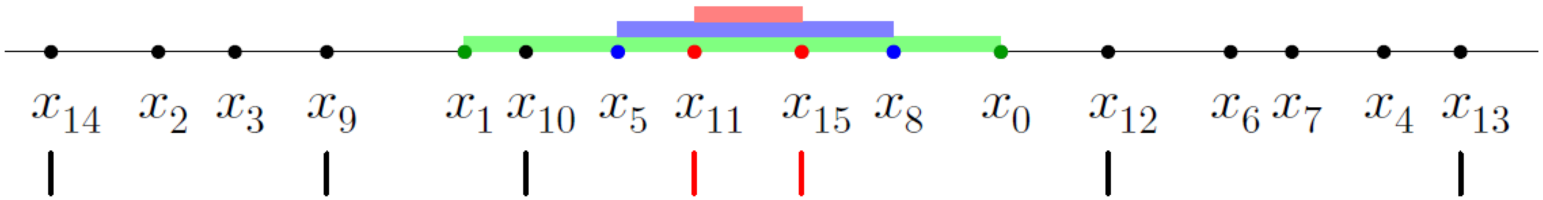
page précédente **:** rappelons qu'il s'agit de montrer que les points de la suite donnée « n'épuisent pas » l'ensemble des nombres réels **;** pour simplifier on suppose que les éléments de cette suite sont deux à deux distincts. Les deux premiers éléments x_0 et x_1 de la suite donnée définissent un segment, indiqué en vert. On va attendre que deux des éléments successifs de la suite « tombent » dans ce premier segment.

---

page suivante **:** dans la figure, ce sont les points x_5 et x_8 qui sont les deux premiers à tomber dans le segment vert, et ils déterminent un nouveau segment, indiqué en bleu. Si aucun point de la suite au delà de l'indice 1 ne tombait dans le segment vert (ou si seulement un seul point y tombait), la procédure s'arrêterait là, et dans ce cas il serait évident de trouver dans ce segment un nombre réel différent de tous les termes de la suite proposée. Sinon on continue **. . .**


x_2 x_3 x_1 x_5 x_8 x_0 x_6 x_7 x_4


x_2 x_3 x_1 x_5 x_8 x_0 x_6 x_7 x_4
personne ici !
(jusqu'à maintenant)



personne ici !
(jusqu'à maintenant)

suite de segments emboîtés

En poursuivant suivant le même principe on a identifié une suite décroissante de segments. Il existe au moins un réel y qui appartient à tous ces segments. On voit facilement que tous les points x_n de la suite donnée sont, soit strictement à gauche de y, soit strictement à droite : le nombre réel y ne fait pas partie de la suite proposée, cqfd.

$\blacksquare$

page suivante : Cantor a expliqué au début de l'article que l'ensemble des *nombres algébriques* est dénombrable. Il tient donc une preuve du fait qu'il existe une immense quantité de *nombres transcendants* (par définition : ceux qui ne sont pas algébriques).

Prouver qu'un nombre particulier, tel que e ou π, est transcendant est une toute autre affaire. Hermite a donné une preuve pour e en 1872 et Lindemann pour π en 1882 (dans les années 1770, Lambert avait déjà eu beaucoup de mal à seulement prouver que π est *irrationnel*). Liouville avait proposé en 1844 des nombres, non « explicites » en ce qu'ils n'avaient pas été considérés précédemment en mathématiques, mais définis comme somme de certaines séries numériques explicites, dont il pouvait prouver assez facilement la transcendance.

Les nombres algébriques

Le nombre θ est *algébrique* quand

$$a_n\theta^n + \cdots + a_1\theta + a_0 = 0, \quad a_n \neq 0, \qquad n \geq 1$$

avec $a_0, a_1, \ldots, a_n$ entiers relatifs ($a_i \in \mathbb{Z}$)

$\sqrt{2}$: $\theta^2 - 2 = 0$, les racines cubiques d'entiers, . . .

les « quantités » constructibles à la règle et au compas sont algébriques

sinon le nombre θ est *transcendant*

Johann Heinrich Lambert (1728 –1777)

Charles Hermite (1822–1901)

Joseph Liouville (1809 –1882)

Ferdinand von Lindemann (1852 –1939)

$$\{0,1\}^{\mathbb{N}} \underset{\text{bijec}}{\equiv} [0,1)$$

L'existence d'une bijection entre l'ensemble des réels de l'intervalle semi-ouvert [0, 1) et l'ensemble des suites de 0 et de 1 résulte assez simplement de la notion de développement en base 2 d'un nombre réel ; les *développements impropres* sont un petit obstacle qu'il faut franchir (en base 2, un développement impropre ne comporte que des 1 à partir d'un certain rang ; en base 10, il ne comporte que des 9 à partir d'un moment, comme dans 0,12349999999... qui représente le même nombre que 0,12350000000...)

un ensemble *parfait* dans $\mathbf{R}^n$ a la puissance du continu

$$\mathbb{N}^{\mathbb{N}} \underset{\text{bijec}}{\equiv} [0,1)$$

Julius König (1849 –1913)

La bijection qui est mentionnée en bas de la page précédente, entre l'ensemble des suites d'entiers positifs ou nuls et l'intervalle [0, 1), provient d'une observation qu'on peut attribuer à Julius König ; cette bijection associe par exemple la suite d'entiers :

$$(3, 1, 4, 0, 1, 0, 2, \ldots)$$

au nombre réel qui est écrit en base 2 sous la forme :

$$0,1\,1\,1\,0\,1\,0\,1\,1\,1\,1\,0\,0\,1\,0\,0\,1\,1\,0\ldots$$

Le lecteur est invité à deviner la règle (simple) qui définit cette bijection ; mais il ne gagnera aucune récompense.

Le développement en base 2 obtenu au moyen de cette correspondance n'est jamais *impropre* ; cette approche évite la complication causée par ces développements impropres.

deux pages suivantes : à partir d'une bijection entre $[0, 1)$ et l'ensemble des suites d'entiers (pages précédentes), on imagine facilement aujourd'hui qu'en découpant une suite infinie d'entiers en deux suites, la suite des termes d'indice pair et la suite des termes d'indice impair, on va déduire une bijection d'un segment sur un carré ; c'est la découverte que fait Cantor en 1877. À la deuxième page suivante, on sera surpris de voir, dans cette lettre de Cantor à Dedekind, à quel point cette propriété devenue banale avec le temps a semblé alors extraordinaire à Cantor.

$$(n_0, n_1, n_2, n_3, \ldots) \longleftrightarrow x \in [0,\ 1)$$

$$\updownarrow$$

$$\left\{ \begin{array}{l} (n_0, n_2, n_4, n_6, \ldots) \longleftrightarrow x_1 \in [0,\ 1) \\ (n_1, n_3, n_5, n_7, \ldots) \longleftrightarrow x_2 \in [0,\ 1) \end{array} \right\} \longleftrightarrow (x_1, x_2) \in [0,\ 1)^2$$

Le continu sous plusieurs formes

$$[0,1) \times [0,1) \underset{\text{bij}}{\equiv} [0,1)$$

$$\mathbb{R}^2 \underset{\text{bij}}{\equiv} \mathbb{R}$$

$$\mathbb{R}^d \underset{\text{bij}}{\equiv} \mathbb{R}^n$$

Mais il n'existe pas de bijection *continue* entre un ouvert non vide de $\mathbf{R}^d$ et un ouvert de $\mathbf{R}^n$ quand d est différent de n (Brouwer 1911 ; Cantor a tenté une preuve en 1879, qu'il a publiée, mais qui n'a pas été reconnue correcte).

Luitzen Egbertus Brouwer (1881 – 1966)

Cantor n'en revient pas :

«Entschuldigen Sie es gütigst meinem Eifer für die Sache, wenn ich Ihre Güte und Mühe so oft in Anspruch nehme; die Ihnen jüngst von mir zugegangenen Mittheilungen sind für mich selbst so unerwartet, so neu, dass ich gewissermassen nicht eher zu einer gewissen Gemüthsruhe kommen kann, als bis ich von Ihnen, sehr verehrter Freund, eine Entscheidung über die Richtigkeit derselben erhalten haben werde. Ich kann, so lange Sie mir nicht zugestimmt haben, nur sagen: je le vois, mais je ne le crois pas.»

Cantor à Dedekind, lettre du 29 juin 1877
(dans le livre de Walter Purkert et Hans Joachim Ilgauds, p. 50)

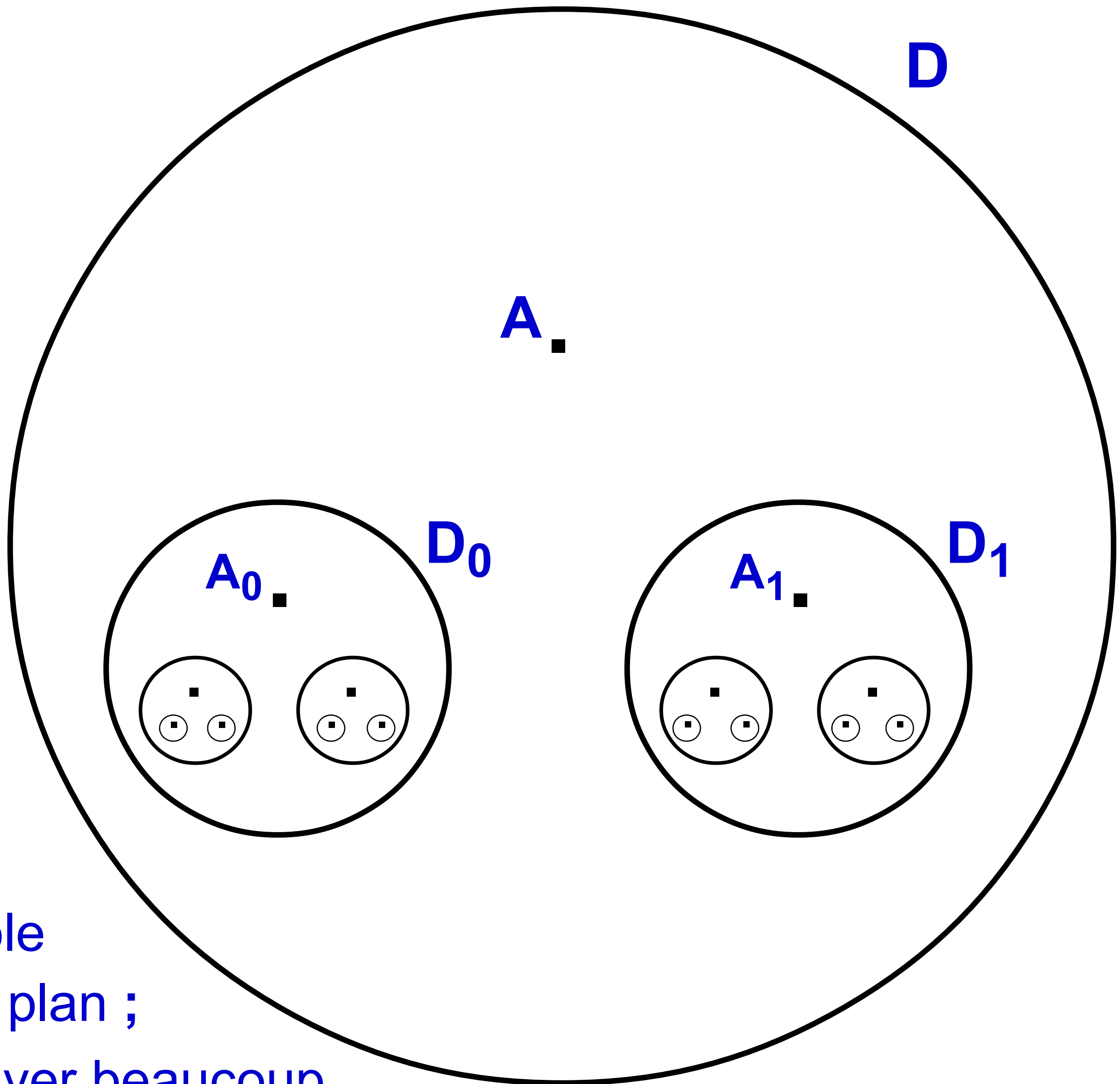
Ayez la bonté d'excuser mon ardeur impatiente à propos de cette question, alors que je fais déjà si souvent appel à votre gentillesse et à vos efforts ; les résultats que je vous ai communiqués récemment sont pour moi-même si inattendus, si nouveaux, que je ne pourrai retrouver une certaine sérénité que quand j'aurai obtenu de vous, très cher ami, une confirmation de leur exactitude. Tant que vous ne vous êtes pas déclaré d'accord avec moi, je ne peux que dire : « je le vois, mais je ne le crois pas » [*en français dans le texte !*].

Cantor peut prouver qu'un ensemble *fermé* de la droite ou de $\mathbf{R}^d$, *ou bien* est dénombrable, *ou bien* a la *puissance du continu*, c'est-à-dire la cardinalité du segment $[0, 1]$. En revanche, il se cassera les dents sur une conjecture qui lui semblera naturelle, qu'on peut formuler en disant que **:** *tout sous-ensemble de la droite réelle ou bien est dénombrable, ou bien possède la puissance du continu*.

C'est *l'hypothèse du continu*, qui ne sera définitivement élucidée que dans les années 1960.

Dans les pages suivantes, on va expliquer l'affirmation qu'on a faite un peu plus tôt **:** *un ensemble parfait de la droite ou de* $\mathbf{R}^d$ *a la puissance du continu* (rappelons qu'un ensemble est *parfait* quand il est non vide et qu'il n'a aucun point isolé **:** il est non vide et égal à son dérivé).

Un ensemble parfait de la droite ou de $\mathbf{R}^d$ *a la puissance du continu*



Le point A
est un point
d'un ensemble
parfait P du plan ;
on va lui trouver beaucoup
de voisins A_0, A_1, etc. qui seront aussi des points de P.

page précédente : on considère un disque ouvert D qui contient un point A d'un ensemble parfait P. Puisque P est parfait, le point A n'est pas isolé dans P : il existe donc une suite de points de P, distincts de A, et qui tendent vers A. Dans cette suite convergente de points on pourra trouver deux points A_0 et A_1 qui sont dans P, qui sont distincts et qui sont dans le disque ouvert D qui contient A.

Ensuite, on peut trouver autour de A_0 et A_1 deux disques ouverts D_0 et D_1 avec les propriétés suivantes :

— pour $j = 0, 1$, le disque D_j contient le point A_j, et son rayon est au plus égal à la moitié du rayon de D.

— les *disques fermés* adhérences de D_0 et D_1 sont *disjoints* et ils sont contenus dans D.

La situation de D_0 et A_0, ainsi que celle de D_1 et A_1, est identique à la situation initiale de D et A ;

(à suivre) →

on peut reprendre le raisonnement depuis le début avec chacun des deux couples (D_j, A_j) et construire un *arbre* de points,

$$A_0 \longrightarrow (A_{00}, A_{01}), \quad A_1 \longrightarrow (A_{10}, A_{11}), \quad A_{00} \longrightarrow (A_{000}, A_{001}), \text{ etc.}$$

Tous les points limites de cet « arbre » seront des points de P puisque $P = P'$ est *fermé* **;** les points limites des « branches » qui sont passées dans D_0 resteront dans le disque *fermé* adhérence de D_0 et seront donc distincts des limites des branches qui sont passées dans D_1. On pourra voir ainsi que l'ensemble des points limites est en bijection avec l'ensemble des suites de 0 et de 1, il a donc la *puissance du continu,* de même que l'ensemble P.

On a appelé « branche » une suite infinie de points qui commence avec A *et qui continue avec* A_0 ou A_1 **;** *ensuite, à partir d'une position déjà atteinte, telle que* A_1 *par exemple, on continue avec* l'un des deux *points qui lui « succèdent » dans la construction, par exemple*

$$(A, \quad A_1, \quad A_{10}, \quad A_{100}, \quad A_{1001}, \quad A_{10010}, \ldots)$$

Ordinaux

réunion et intersection d'une suite d'ensembles (Cantor 1880)

~~Q, $\bar{Q}$ ein Multiplum von P.~~ Sind $P_1, P_2, P_3, \cdots$ irgend welche Punktmengen in endlicher oder unendlicher Anzahl, so gehört zu ihnen sowohl ein kleinstes gemeinsames Multiplum, welches wir mit:

$$\mathfrak{M}(P_1, P_2, P_3, \cdots) \qquad \text{(réunion)}$$

bezeichnen und welches die Menge ist, die aus allen verschiedenen Punkten von $P_1, P_2, P_2, \cdots$ besteht und sonst keine anderen Punkte als Elemente besitzt, — wie auch ein grösster gemeinsamer Divisor, den wir $\equiv$

$$\mathfrak{D}(P_1, P_2, P_3, \cdots) \qquad \text{(intersection)}$$

setzen und welcher die Menge der Punkte ist, die allen $P_1, P_2, P_3, \cdots$ gemeinsam sind. Beispielsweise können wir, wenn $P', P'', P''', \cdots$

Cantor introduit la *réunion* et *l'intersection* d'une suite infinie d'ensembles, qu'il nomme « commun multiple » pour la réunion ou « diviseur commun » pour l'intersection. On voit qu'il *introduit* une notation pour chacune de ces deux notions, on doit supposer qu'il n'en existait pas à l'époque.

deux pages suivantes : Cantor a déjà envisagé en 1872 d'appliquer un nombre fini de fois l'opération de dérivation à un ensemble P. Il considère maintenant *l'intersection* de la suite des dérivés d'ordre fini $P^{(n)}$ et l'appelle le *dérivé d'ordre infini*.

On notera que lorsque P est borné, la famille des $P^{(n)}$ est une suite décroissante de fermés bornés ; si tous les $P^{(n)}$ sont non vides, on sait alors que leur intersection (le dérivé d'ordre infini) n'est pas vide (compacité).

Cantor envisage un pas de plus : en dérivant une fois le dérivé d'ordre infini, il arrive à la notation « infini plus un » et aux suivantes,

$$\infty + 1, \quad \infty + 2, \quad \infty + 3, \ \ldots$$

Ce ne sont pour l'instant que des « Unendlichkeitssymbole », des « symboles d'infinité », mais on peut considérer qu'on a ici le début de la *théorie des ordinaux*. À la deuxième page suivante, on a récrit ces symboles de Cantor de manière plus lisible que dans les photos du vieil article.

die andere R umfasst diejenigen Punkte, welche in *allen* Gliedern der Folge P', P'', P''', $\cdots$ erhalten bleiben, es ist also R definirt durch die Formel:

$$R = \mathfrak{D}(P', P'', P''', \cdots).$$

Diese aus der Menge P hervorgehende Punktmenge R werde nun durch das Zeichen:

$$P^{(\infty)}$$

ausgedrückt und Ableitung von P der Ordnung ∞ genannt.

Die erste Ableitung von $P^{(\infty)}$ werde mit $P^{(\infty+1)}$, die n^{te} Ableitung von $P^{(\infty)}$ mit $P^{(\infty+n)}$ bezeichnet; $P^{(\infty)}$ wird aber auch eine, im Allgemeinen von O verschiedene Ableitung von der Ordnung ∞ haben, wir nennen sie $P^{(2\infty)}$. Durch Fortsetzung dieser Begriffsconstructionen kommt man zu Ableitungen, die consequenterweise durch:

$$P^{(n_0\infty+n_1)}$$

zu bezeichnen sind, wo n_0, n_1 positive ganze Zahlen sind. Wir kommen aber auch darüber hinaus, indem wir:

$$\mathfrak{D}(P^{(\infty)}, P^{(2\infty)}, P^{(3\infty)}, \cdots)$$

bilden und dafür das Zeichen $P^{(\infty^2)}$ festsetzen.

$$\mathrm{P}^{(\infty)} = \bigcap_{n \geqslant 0} \mathrm{P}^{(n)}, \qquad \mathrm{P}^{(\infty+1)} = \left(\mathrm{P}^{(\infty)}\right)'$$

$$\infty, \quad \infty + 1, \quad \infty + 2, \quad \ldots, \quad \infty + n, \quad \ldots, \quad 2.\infty$$

$$\infty, \quad 2.\infty, \quad 3.\infty, \quad \ldots, \quad n.\infty, \quad \ldots, \quad \infty^2$$

$$\infty, \quad \infty^2, \quad \infty^3, \quad \ldots, \quad \infty^n, \quad \ldots, \quad \infty^\infty$$

$$\infty, \quad \infty^\infty, \quad \infty^{\infty^\infty}, \quad \ldots$$

Es ist ferner zweckmässig ein Zeichen zu haben, welches die Abwesenheit von Punkten ausdrückt, wir wählen dazu den Buchstaben O; $P \equiv O$ bedeutet also, dass die Menge P *keinen einzigen* Punkt enthält, also streng genommen als solche gar nicht vorhanden ist. Um auch

(vide)

und durch Fortsetzung dieses Verfahrens kommt man zu:

$$P^{(n_0 \infty^\nu + n_1 \infty^{\nu-1} + \cdots + n_\nu)},$$

wo $n_0, n_1, \cdots, n_\nu$ positive ganze Zahlen sind. Zu weiteren Begriffen gelangt man, indem man ν variabel werden lässt; man setze:

$$P^{(\infty^\infty)} \equiv \mathfrak{D}(P^{(\infty)}, P^{(\infty^2)}, P^{(\infty^3)}, \cdots).$$

Durch consequentes Fortschreiten gewinnt man successive die weiteren Begriffe:

$P^{(n\infty^\infty)}$, $P^{(\infty^{\infty+1})}$, $P^{(\infty^{\infty+n})}$, $P^{(\infty^{n\infty})}$, $P^{(\infty^{\infty^n})}$, $P^{(\infty^{\infty^\infty})}$ u. s. w.;

wir sehen hier eine dialektische Begriffserzeugung*), welche immer weiter führt und dabei frei von jeglicher Willkür in sich nothwendig und consequent bleibt.

page précédente : Cantor indique qu'il serait utile d'avoir un signe pour représenter le cas où un « ensemble » P ne contiendrait aucun élément. Il choisit pour cela la lettre O mais il insiste pour dire que si on écrit P = O, « au sens strict l'ensemble n'existe pas ». Il explique aussi comment poursuivre les opérations de dérivation au delà de l'infini et laisse un peu à son lecteur le soin de deviner ce que signifieront les symboles tels que

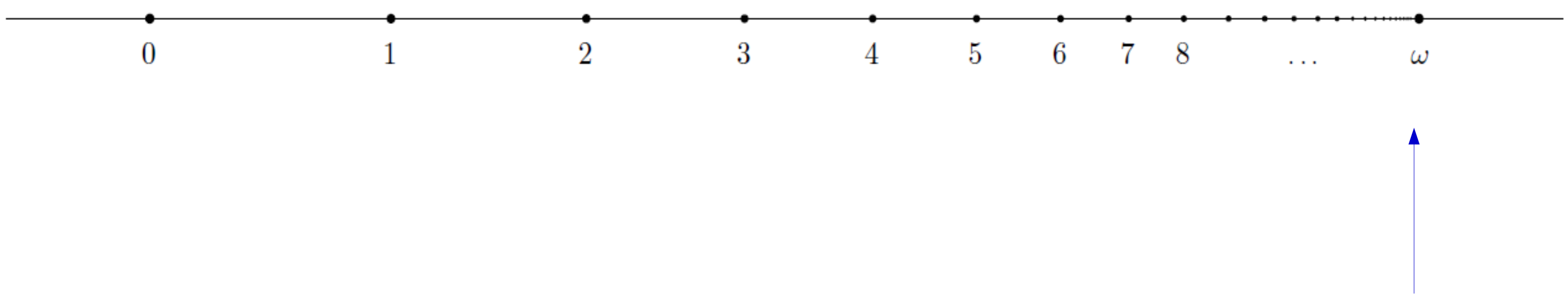
$$\infty^{\infty}, \quad \infty^{\infty^{\infty}}, \quad \text{etc.}$$

---

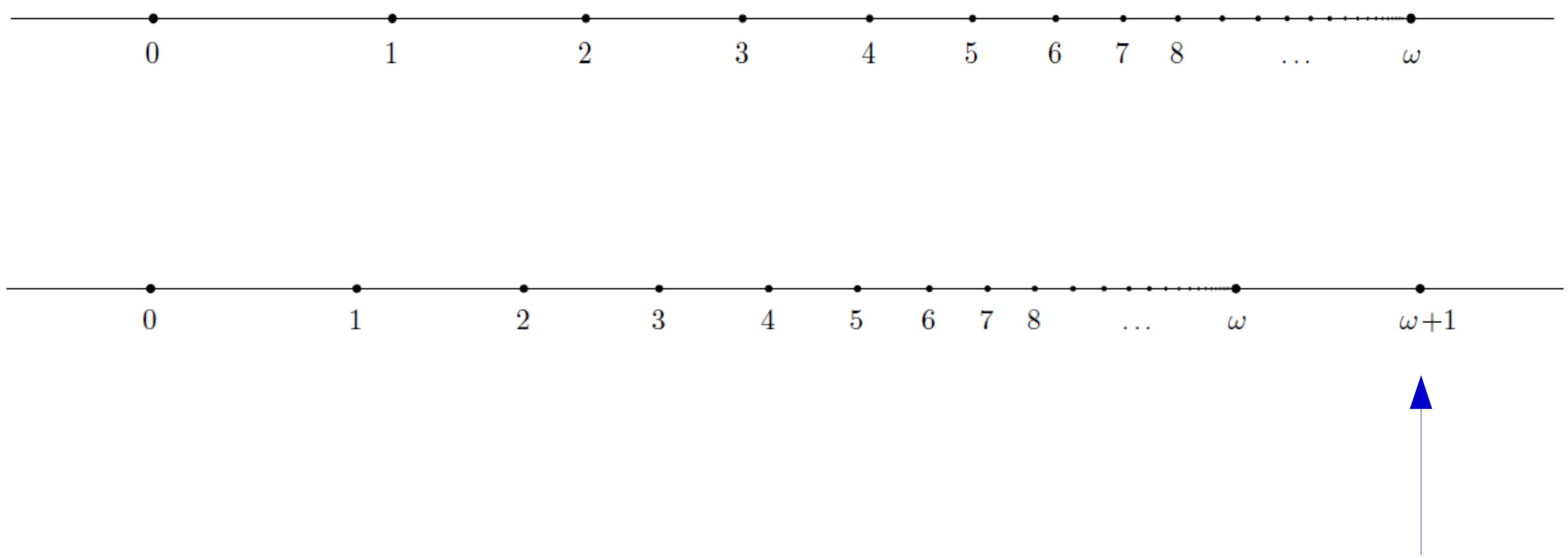
page suivante : anticipant un peu sur Cantor, on va commencer à considérer les simples « symboles d'infinité » comme de nouveaux nombres qu'on pourra présenter dans une liste ordonnée. En 1883, Cantor décidera d'abandonner le symbole ∞, vraiment trop passe-partout, au profit de la lettre grecque ω (omega) pour désigner le « premier nombre infini », le premier qui est plus grand que tous les entiers naturels.

Pendant un petit moment, on va représenter les *symboles d'infinité* au moyen d'ensembles de points d'un segment ; ce seront d'ailleurs des exemples d'*ensembles exceptionnels* admissibles pour le théorème *d'unicité* trigonométrique : Cantor le prouve en 1872 pour les exemples de la forme ω^n qu'on va voir bientôt (plus précisément, il le prouve pour les ensembles P dont un dérivé d'ordre fini devient vide).

Ne regardons plus que la *structure d'ordre*
d'une famille de points tendant vers une limite :



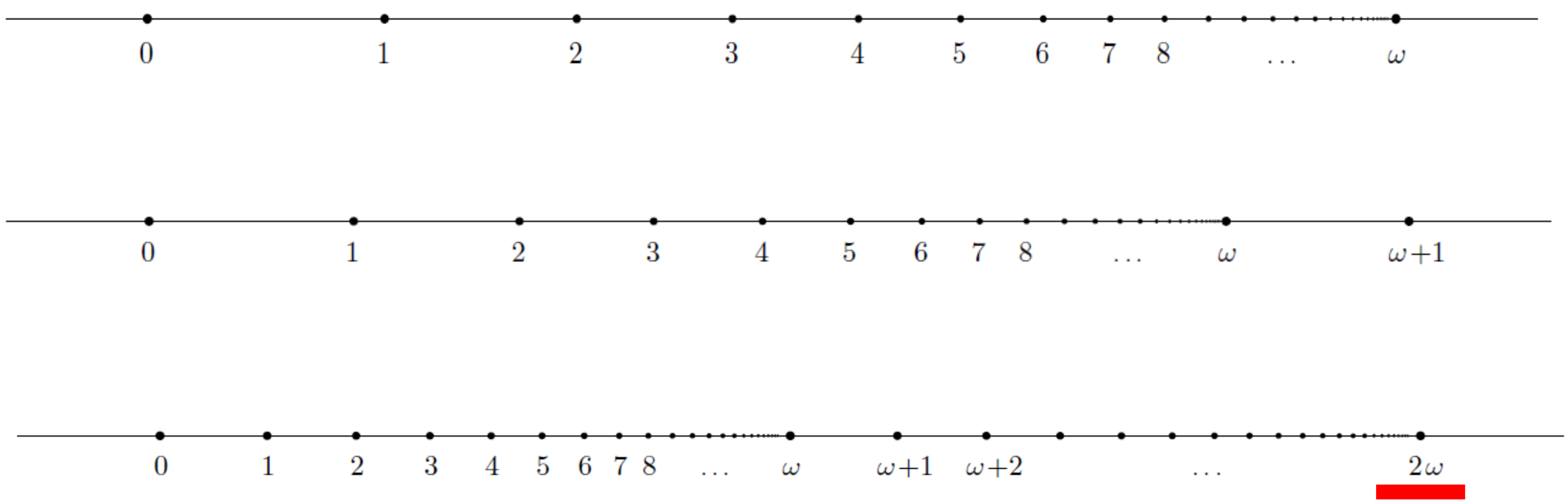
(1883)



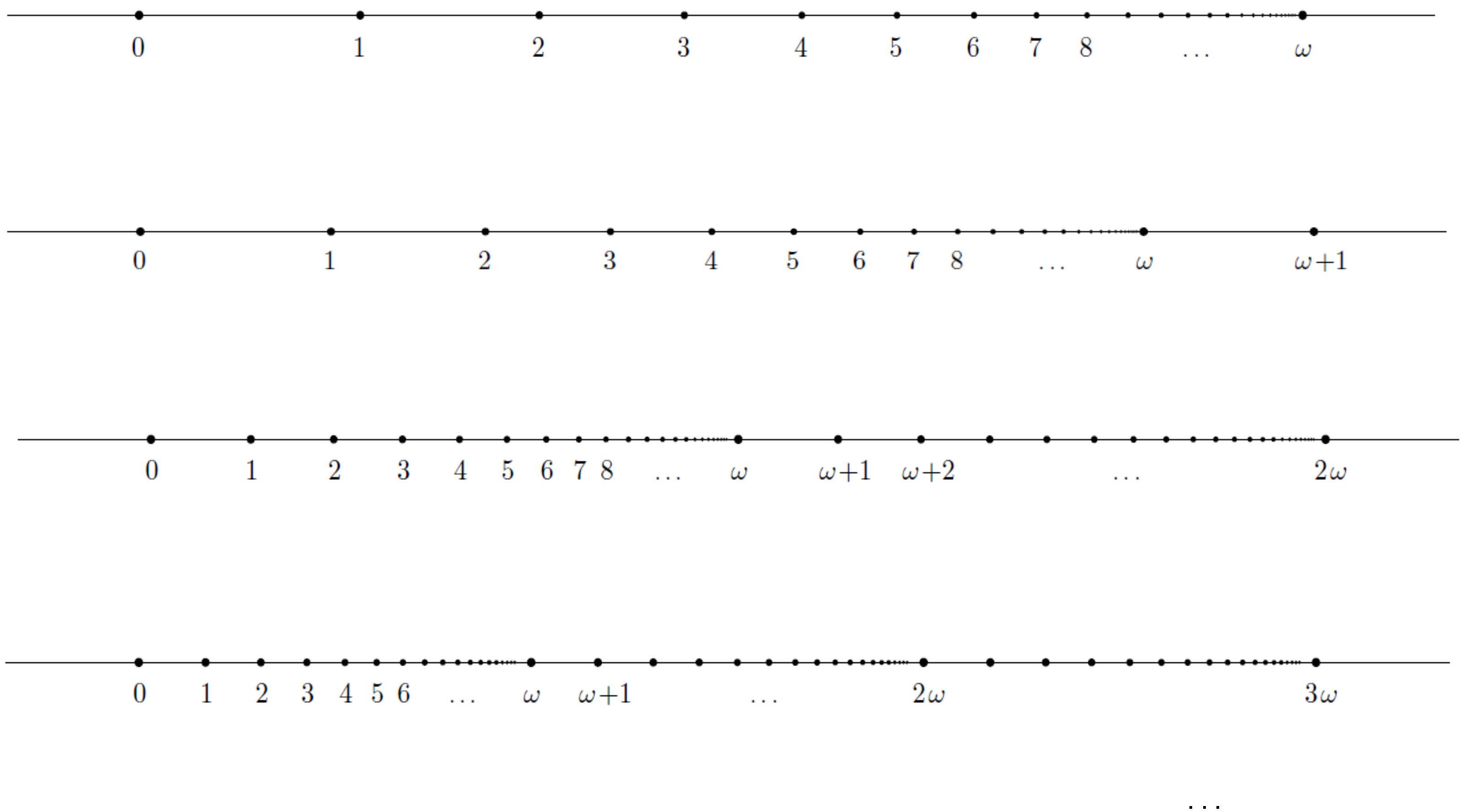
Le premier « principe de formation » de nouveaux symboles d'infinité (Cantor ne dit pas encore *nombre ordinal*) est aussi le plus simple **:** si un de ces symboles est déjà en place, on peut envisager son successeur immédiat. Ainsi, après ω vient $\omega + 1$ puis $\omega + 2$ et ainsi de suite **. . .**

page suivante : le « deuxième principe de formation » consiste à introduire un nouveau symbole qui sera le premier à être plus grand que tous les $\omega + n$; en 1882 et 1883 Cantor adopte la notation $2\,\omega$ pour ce nouveau symbole. Il aura de bonnes raisons plus tard pour changer d'avis, il écrira alors $\omega \cdot 2$ pour ce même symbole ; on y reviendra.

Le *premier principe* conduit à $2\,\omega+1$, $2\,\omega+2$, $2\,\omega+3$, etc. Naturellement, on aura un « nouveau symbole » $3\,\omega$ qui sera *le plus petit* qui soit au delà de tous les $2\,\omega+n$, où n décrit la suite des entiers naturels (deuxième page suivante).



2ω ou $\omega.2$ (en 1895)



On comprend bien comment continuer avec $4\,\omega$, $5\,\omega$, **. . .** , $n\,\omega$, pour tout entier naturel n.

page suivante : rapidement, on arrive à un moment où il y a trop de « points » pour qu'on puisse encore les dessiner. On peut alors essayer d'employer une notation symbolique pour représenter par une lettre unique des « exemplaires » ou « copies » d'un exemple déjà rencontré. Ainsi, on envisage un exemple E_2 qui est construit à partir d'une suite ordonnée de copies de l'exemple E_1 de la page d'avant, augmentée d'un point final qu'on a noté ω^2 et qui est considéré « plus grand » que tous les éléments $n\,\omega$ qui le précèdent.

On conçoit facilement comment on pourrait représenter *dans un segment* les exemples décrits par une suite de notations symboliques : supposant qu'on ait déjà expliqué comment représenter dans un segment $(a, b]$ chacun des exemples formant la suite de symboles, on pourra découper un intervalle, par exemple $(0, 1]$, en une suite (ordonnée) d'intervalles $(a_n, b_n]$, et représenter le $n$ème symbole par un ensemble qui sera placé dans le $n$ème segment.

$$\overbrace{\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \;\cdot\;\cdot\;\cdot\;\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot \quad \ldots \quad \cdot}^{E_1}$$

$$0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \qquad \ldots \qquad \omega$$

« E_1 » représente toute la suite (y compris la limite).

E_2 :

$$E_1 \quad \ldots \quad .$$

$$\omega \quad 2\omega \quad 3\omega \quad 4\omega \quad 5\omega \quad 6\omega \quad 7\omega \qquad \ldots \qquad \omega^2$$

$$\overbrace{\begin{array}{cccccccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots\cdots\cdots & \ldots & \cdot \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \ldots & & \omega \end{array}}^{E_1}$$

$$\begin{array}{cccccccccccccccccccccccc} E_1 & \ldots & \cdot \\ \omega & 2\omega & 3\omega & 4\omega & 5\omega & 6\omega & 7\omega & & & & & & \ldots & & & & & & & & & & \omega^2 \end{array}$$

$$\overbrace{\begin{array}{cccccccccccccccccccccccc} E_1 & \ldots & \cdot \\ \omega & 2\omega & 3\omega & 4\omega & 5\omega & 6\omega & 7\omega & & & & & & \ldots & & & & & & & & & & \omega^2 \end{array}}^{E_2}$$

$$(E_2)' \sim E_1$$

$$E_3 : \quad \begin{array}{cccccccccccccccccccccccc} E_2 & \ldots & \cdot \\ \omega^2 & 2\omega^2 & 3\omega^2 & 4\omega^2 & 5\omega^2 & 6\omega^2 & 7\omega^2 & & & & & & \ldots & & & & & & & & & & \omega^3 \end{array}$$

$$(E_3)' \sim E_2 \; ; \qquad E_4,\ E_5,\ E_6,\ \ldots$$

page précédente : on peut introduire une suite d'exemples E_n, définis pour tout entier naturel n.

---

page suivante : à partir d'une suite d'exemples de plus en plus compliqués, on peut toujours imaginer un nouvel exemple encore plus compliqué que tous les éléments de la suite donnée, un exemple qu'on complétera encore par un « point final ». Dans le cas de l'exemple donné page suivante, ce point final viendra après tous les ω^n et sera noté ω^ω.

$$(\mathrm{E}_{n+1})' \simeq \mathrm{E}_n,$$

$$(\mathrm{E}_3)' \simeq \mathrm{E}_2, \qquad (\mathrm{E}_2)' \simeq \mathrm{E}_1$$

$$(\mathrm{E}_1)' \simeq \mathrm{E}_0 := \{\omega\}, \qquad (\mathrm{E}_0)' = \emptyset$$

$$\mathrm{F}^{(0)} = \mathrm{F}$$

$$\mathrm{F}^{(n+1)} = \left(\mathrm{F}^{(n)}\right)'$$

$$\mathrm{F}^{(\omega)} = \bigcap_{n \geqslant 0} \mathrm{F}^{(n)}$$

$$\underbrace{\begin{array}{cccccccccccc} \mathrm{E}_1 & \mathrm{E}_2 & \mathrm{E}_3 & \mathrm{E}_4 & \mathrm{E}_5 & \mathrm{E}_6 & \mathrm{E}_7 & \ldots & \mathrm{E}_n & \ldots & \bullet \\ \omega & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 & \omega^7 & \ldots & \omega^n & \ldots & \omega^\omega \end{array}}_{\mathrm{E}}$$

L'ensemble dérivé d'ordre ω de l'ensemble E ci-dessus est réduit à un point unique et par conséquent

$$\mathrm{E}^{(\omega+1)} = \left(\mathrm{E}^{(\omega)}\right)' = \emptyset$$

page précédente : considérons l'ensemble E de la page précédente, qui est donné par la suite des E_n ; dans le calcul du dérivé de l'ensemble E, le premier ensemble E_1 de cette suite donne un point d'accumulation unique, et l'ensemble E_2 fournit pour E un ensemble de points d'accumulation qui est « isomorphe » à E_1. À la deuxième dérivation, la contribution à E'' du premier E_1 disparaît, celle de E_2 devient réduite à un point. À la troisième dérivation, la contribution de E_2 disparaît, etc. Le dérivé d'ordre infini de E ne contient que le « point final » qui a été placé après la suite (E_n).

---

page suivante : tout d'un coup, sans prévenir, Cantor passe des exemples de dérivés relativement explicites, tels que $P^{(\infty^\infty)}$ par exemple, au dérivé « d'ordre général » $P^{(\alpha)}$ où α est un *symbole d'infinité* quelconque. On verra plus loin des raisons de trouver la démarche de Cantor un peu précipitée.

*Ist P eine **nicht abzählbare** Punktmenge, so ist auch $P^{(\alpha)}$ **nicht abzählbar**, sowohl wenn α eine endliche ganze Zahl, wie auch wenn es eines der Unendlichkeitssymbole ist.* —

«

Si P est un ensemble de points non-dénombrable, alors l'ensemble dérivé $P^{(\alpha)}$ est lui aussi non-dénombrable, que α soit un entier fini ou bien l'un des symboles d'infinité.

»

Cantor, *Ueber unendliche, lineare Punktmannichfaltigkeiten* n° 4,

paru en 1883, signé **:** *Harzburg, le 1er septembre 1882*.

En 1882, le mathématicien suédois Mittag-Leffler crée le journal *Acta Mathematica*, qui deviendra rapidement et restera jusqu'à nos jours un des plus grands journaux mathématiques. Il y invite à publier des mathématiciens européens importants (Henri Poincaré, entre autres). Mittag-Leffler est passé par Berlin en 1875 pour étudier (avec Weierstrass), il parle allemand ; Cantor et lui nouent une relation amicale et échangent un courrier abondant, notamment dans les années 1882–1884.

Mittag-Leffler publie dans Acta des traductions en français de plusieurs articles de Cantor (dans le volume 2, six articles et un extrait d'une lettre, transformé en article par « l'éditeur ») : Cantor espère faire ainsi progresser ses idées auprès des mathématiciens français ; par ailleurs, Cantor a visité Paris en 1884, il y a rencontré (brièvement) Hermite, Picard, Poincaré, Appell.

Gösta Mittag-Leffler (1846–1927) Henri Poincaré (1854–1912)

Paul Appell (1855–1930) Émile Picard (1856–1941)

Dans l'exposition de mes recherches sur la théorie des ensembles, je suis maintenant arrivé à un point où il me faut développer une généralisation de la notion de nombre entier réel, et ce développement m'entraîne dans une direction où personne, à ma connaissance, ne s'est engagé jusqu'à présent.

Je me trouve contraint de développer cette notion de nombre au point que je pourrais à peine, sans cela, avancer dans la théorie des ensembles; que cette nécessité où je me trouve placé me serve de justification ou d'excuse, si cela était nécessaire, pour avoir introduit dans mon travail un ordre d'idées qui y paraît étranger. Car il s'agit de développer cette notion dans le but de continuer la série des nombres entiers réels au-delà de l'infini; si hardie que paraisse cette tentative, je puis exprimer non-seulement l'espoir, mais la ferme conviction qu'avec le temps on considérera ce développement comme très-simple, très-naturelle et parfaitement accessible. En même temps je ne me dissimule pas cependant que par cette entreprise, je me mets en contradiction, dans une certaine mesure, avec les idées généralement répondues sur l'infini mathématique et avec les opinions qu'on a souvent défendues sur l'essence de la grandeur numérique.

Cantor, Acta Math. 1883
(version française de Math. Ann. 21, 1883)

page précédente : début de l'article de Cantor paru dans Acta en 1883, la même année que sa version allemande. Cet article en français n'est pas la simple traduction de l'article en allemand : l'article allemand contient de très nombreuses considérations philosophiques, la plupart à propos de l'usage de *l'infini actuel*, qui ont été omises dans la version française. De plus, si les sections ont gardé leur numéro d'origine, leur ordre a été modifié.

Dans cet article, Cantor affiche clairement son intention de considérer que ce qui était « symbole d'infinité » doit être vu maintenant comme nombre à part entière, un *nombre entier infini*.

---

deux pages suivantes : Cantor indique les « principes de formation » des nouveaux nombres, principes qu'on a déjà évoqués plus haut. Il introduit aussi ses *classes de nombres,* la première classe étant formée des entiers naturels, les entiers finis ordinaires. La *deuxième classe de nombres,* la plus importante pour ce qui suivra, est formée des « nombres infinis » dont la famille des prédécesseurs est dénombrable, nous dirons : les *ordinaux dénombrables*.

Les *deux principes de formation*, à l'aide desquels on définit les nouveaux nombres infinis déterminés, comme on pourra s'en convaincre, sont tels qu'en les appliquant ensemble, on peut dépasser toutes les limites dans la formation abstraite des nombres entiers réels; mais heureusement on a d'autre part, comme nous le verrons, un *troisième* principe que j'appelle *principe d'arrêt* ou de *limitation*, et grâce auquel on peut donner certaines limites successives au procédé de formation qui est absolument sans fin; nous obtiendrons ainsi, dans la suite *absolument infinie* des nombres réels entiers, des *divisions naturelles*, que j'appellerai *classes de nombres*.

La *première classe de nombres* (I) est le système des nombres entiers finis $1, 2, 3, \ldots\ldots \nu, \ldots\ldots$; vient ensuite la seconde classe de nombres (II), composée de certains nombres entiers infinis α se suivant entre eux dans un ordre de succession déterminé:

$$\omega, \omega+1, \omega+2, \ldots, 2\omega, 2\omega+1, \ldots, \nu_0\omega^\mu+\nu_1\omega^{\mu-1}+\ldots+\nu_{\mu-1}\omega+\nu_\mu,$$

$$\ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots, \alpha, \ldots;$$

la seconde classe de nombres une fois définie, on arrive à la troisième, puis à la quatrième, et ainsi de suite.

(Cantor 1883)

que ω sera le *premier* nombre entier qui *suivra tous* les nombres ν, en sorte qu'il faut le déclarer supérieur à *tous* les nombres ν. En associant le nombre ω avec les unités primitives on obtient à l'aide du *premier principe* de formation les nombres plus étendus:

$$\omega + 1,\ \omega + 2,\ \ldots\ldots\ \omega + \nu,\ \ldots\ldots;$$

comme par là on n'arrive encore une fois à aucun nombre maximum, on imagine un nouveau, que l'on peut appeler 2ω et qui sera le *premier* après tous les nombres obtenus jusqu'à présent ν et $\omega + \nu$; si on applique

- -

Le deuxième principe de formation nous permet donc d'introduire un nombre qui suit immédiatement tous les autres $\mu\omega + \nu$, et que l'on peut appeler ω^2; à ce nombre se rattacheront dans un ordre de succession déterminé:

$$\lambda\omega^2 + \mu\omega + \nu$$

et l'on arrive alors évidemment, en suivant les deux principes de formation, à des nombres de la forme:

$$\nu_0\omega^\mu + \nu_1\omega^{\mu-1} + \ldots\ldots + \nu_{\mu-1}\omega + \nu_\mu;$$

mais alors le deuxième principe de formation nous amène à poser un nouveau nombre qui sera immédiatement supérieur à tous ces nombres et qu'on pourra désigner par ω^ω.

(Cantor 1883)

La « deuxième classe de nombres » est formée de tous les ordinaux infinis dénombrables.

La deuxième classe de nombres n'est pas dénombrable.

Un ensemble ordonné X est *bien ordonné* quand tout sous-ensemble non vide de X possède un plus petit élément, par exemple

$\mathbb{N}$, qui est l'ensemble des ordinaux < ω, ou encore chacun des ensembles d'ordinaux plus petits que :

$$\omega^2, \ \omega^3, \ \ldots, \ \omega^\omega, \ \ldots$$

À chaque ensemble bien ordonné correspond un ordinal page précédente : Cantor prouve que la deuxième classe de nombres n'est pas dénombrable, ce qui sera utilisé intensivement dans ce qui suivra. Il s'intéresse aux *ensembles bien ordonnés* et il explique que tout ensemble bien ordonné peut être « compté » au moyen de ses nombres entiers infinis : il associe ainsi un « nombre » unique fini ou infini (Anzahl) à chaque ensemble bien ordonné (voir une explication plus loin).

---

page suivante : après la famille de tous les ordinaux dénombrables vient *le premier ordinal non dénombrable* que Cantor désigne par Ω ; la deuxième classe de nombres est donc formée de tous les ordinaux *infinis* α tels que $\alpha < \Omega$.

Cantor définit l'addition $\alpha + \beta$ de deux ordinaux α et β en plaçant β *après* α ; ainsi, $1 + \omega = \omega$ est différent de $\omega + 1$. Pour la multiplication, il considère en 1883 que $\omega \cdot 2$ représente une suite ordonnée d'ensembles, dont chacun est ordonné et a deux éléments, ce qui du point de vue de l'ordre est identique à ω,

(en 1883) $\omega \cdot 2 = \omega.$

Noms donnés aux ordinaux par Cantor :

Unendlichkeitssymbol (1880-1882)
Unendliche (reale) ganze Zahl ; Anzahl (1883)
Ordinalzahl (1895)

Ω désigne le premier ordinal non dénombrable.

La « troisième classe » de nombres (qu'on ne verra pas intervenir dans la suite) est formée des ordinaux α tels que la famille des ordinaux qui sont inférieurs à α possède le même cardinal que la famille des ordinaux plus petits que Ω. Elle sera suivie d'une quatrième classe, etc.

Opérations sur les ordinaux :

$$\alpha + \beta \qquad\qquad \alpha\,\beta$$

En 1883 $\alpha\,\beta$ représente une « α-suite » de *copies* de β

« Exhaustion du dénombrable »

page précédente : *l'exhaustion du dénombrable* est le phénomène crucial qui sera beaucoup utilisé ; supposons qu'on considère avec Cantor la famille décroissante des dérivés $P^{(\alpha)}$ d'un ensemble fermé P de la droite réelle, pour tous les ordinaux $\alpha < \Omega$. On sait que les intervalles ouverts à extrémités rationnelles forment une famille dénombrable, on peut en faire une liste (I_n). Chaque fois qu'on a $P^{(\alpha+1)} \neq P^{(\alpha)}$, il existe dans $P^{(\alpha)}$ un point isolé x et on peut alors trouver un intervalle $I_{n(x)}$ de la liste qui contient x et qui ne rencontre pas $P^{(\alpha+1)}$. On voit qu'on ne reprendra jamais deux fois le même entier $n = n(x)$; comme l'ensemble des α est non-dénombrable, il y aura nécessairement un moment où on aura « épuisé les entiers naturels », et donc on ne pourra plus avoir que

$$P^{(\alpha+1)} = P^{(\alpha)}.$$

Cette égalité subsistera pour tous les ordinaux supérieurs. Ou bien l'ensemble « stabilisé » $P^{(\alpha)} = P^{(\Omega)}$ est vide, ou bien c'est un ensemble *parfait* — noter que $(P^{(\Omega)})' = P^{(\alpha+1)} = P^{(\Omega)}$ —.

page suivante **:** l'expression dans la première équation montre un possible *système de notation* pour la famille de *tous* les ordinaux qui sont $< \omega^\omega$. Étant donné que les notations qu'on peut envisager seront formées d'un nombre fini de signes pris dans un alphabet fini, on peut se convaincre qu'on ne saura produire qu'une *famille dénombrable de notations* et on arrive à une affirmation très troublante **:** puisque « *la deuxième classe de nombres de M. Cantor* » n'est pas dénombrable, il existe un premier « grand » ordinal pour lequel on ne peut plus trouver de système de notation pour la famille formée de tous les ordinaux plus petits que lui **!**

Les deux équations pour $\omega^{\omega+1}$ et $\omega^{1+\omega}$ nous montrent un effet désagréable de la convention que Cantor a adoptée pour la multiplication **:** l'ordre des termes est inversé entre *l'exponentiation* et l'addition,

$$\omega^{\omega+1} = \omega^1\,\omega^\omega, \qquad \omega^{1+\omega} = \omega^\omega\,\omega^1.$$

La convention finale (inversée) pour la multiplication, posée par Cantor dans ses derniers articles mathématiques, corrigera ce défaut.

Notation des ordinaux « modestes »

$$c_0\,\omega^m + c_1\,\omega^{m-1} + \cdots + c_{m-1}\,\omega + c_m$$

où m et les c_j sont des entiers finis, avec c_0 non nul.

En 1883 :

$$\omega^{\omega+1} = \lim_n n\,\omega^\omega = \omega\,\omega^\omega$$
$$\omega^{1+\omega} = \lim_n \omega^\omega\, n = \omega^\omega\,\omega = \omega^\omega$$

« Grands » ordinaux sans système de notation deux pages suivantes : tentative de traduction à la deuxième page, le texte est extrait de la *longue* version allemande de l'article de 1883. C'est ici qu'on retrouve la fière affirmation de la *liberté des mathématiques* qu'on a transcrite à côté du portrait du jeune Cantor, au début du présent document. Cantor sait que des mathématiciens importants et surtout influents vont s'opposer à ses travaux, ou même les dénigrer. On le voit sur la défensive, se sentant obligé d'assurer que ses conceptions ne vont pas « abîmer » les mathématiques.

La parution dans Acta de traductions en français de plusieurs de ses articles, dont certains assez anciens, est une tentative de Cantor pour intéresser à ses travaux la communauté des mathématiciens français, qui comptait de grands noms tels que Hermite et Poincaré, et au-delà, pour être lu par tous ceux, nombreux à l'époque, qui pouvaient lire des mathématiques en français ou même en écrire, sans être Français : on en verra un exemple plus loin.

Ueber unendliche, lineare Punktmannichfaltigkeiten. Math. Annalen 21, p. 563-564

En 1883, dans des considérations philosophiques d'une bonne dizaine de pages sur la nature des mathématiques et de l'infini mathématique, Cantor se réfère à Platon, Aristote, Démocrite, Thomas d'Aquin, Descartes, Locke, Spinoza, Leibniz, Kant, entre autres . . . avant de développer plusieurs points de sa théorie des ordinaux ; c'est à la fin de la « section philosophique » qu'on peut lire :

Es ist, wie ich glaube, nicht nöthig in diesen Grundsätzen irgendeine Gefahr für die Wissenschaft zu befürchten, wie dies von Vielen geschieht; einerseits sind die bezeichneten Bedingungen, unter welchen die Freiheit der Zahlenbildung allein geübt werden kann, derartige, dass sie der Willkür einen äusserst geringen Spielraum lassen; dann aber trägt auch jeder mathematische Begriff das nöthige Correctiv in sich selbst einher; ist er unfruchtbar oder unzweckmässig, so zeigt er es sehr bald durch seine Unbrauchbarkeit und er wird alsdann, wegen mangelnden Erfolgs, fallen gelassen. Dagegen scheint mir aber jede überflüssige Einengung des mathematischen Forschungstriebes eine viel grössere Gefahr mit sich zu bringen und eine um so grössere, als dafür aus dem Wesen der Wissenschaft wirklich keinerlei Rechtfertigung gezogen werden kann; denn das *Wesen* der *Mathematik* liegt gerade in ihrer *Freiheit*.

Es ist, wie ich glaube, nicht nöthig in diesen Grundsätzen irgendeine Gefahr für die Wissenschaft zu befürchten, wie dies von Vielen geschieht; ~~einerseits sind die bezeichneten Bedingungen, unter welchen~~

Je pense qu'il n'y a pas lieu de redouter – comme cela se produit pourtant chez beaucoup – que ces principes [*de formation de nouveaux nombres*] représentent un quelconque danger pour la science ;

~~wegen mangelnden Erfolgs, fallen gelassen.~~ Dagegen scheint mir aber jede überflüssige Einengung des mathematischen Forschungstriebes eine viel grössere Gefahr mit sich zu bringen und eine um so grössere, als dafür aus dem Wesen der Wissenschaft wirklich keinerlei Rechtfertigung gezogen werden kann; denn das *Wesen* der *Mathematik* liegt gerade in ihrer *Freiheit*.

En revanche, il me semble que toute limitation superflue imposée au mouvement de la recherche mathématique comporte un danger bien plus grand, d'autant que vraiment, aucune justification [*de telles limitations*] ne peut être tirée de l'essence de la science ; car *l'essence* des *mathématiques* réside précisément dans leur *liberté*.

page suivante : selon les historiens, Cantor souffre en 1884 d'une première crise dépressive. De retour au travail en 1885, il écrit un nouvel article qu'il soumet à Mittag-Leffler pour Acta ; alors que la publication était déjà « sur les rails », il a la fort désagréable surprise de recevoir une lettre de Mittag-Leffler qui lui conseille très vivement de ne pas encore publier ces résultats.

Cantor n'insiste pas et retire son article. Cet article de 1885 sur les *types d'ordre* ne sera jamais publié par Cantor ; en 1970, l'historien des mathématiques Grattan-Guinness va à la rencontre de descendants de Cantor et retrouve le texte de l'article dans la maison de celui-ci. Il peut ainsi publier dans Acta en 1970 l'article de Cantor « rejeté » (le mot est inexact) par Acta en 1885 !

Deux ensembles ordonnés X *et* Y *ont le même* type d'ordre *s'il existe entre eux une bijection qui préserve l'ordre. Par exemple, les ensembles* $\mathbf{N}$ *et* $\mathbf{Z}$ *munis de leurs ordres usuels n'ont pas le même type d'ordre.*

Été 1884

Grattan-Guinness (1970)

Lettre de Mittag-Leffler à Cantor au début 1885 :

~~von dieser Gesichtspunkt aus sehr viel erreichen können.~~ Aber ich will Ihnen nicht verhehlen dass es scheint mir es wäre Euer selbst wegen besser gewesen diese Untersuchungen nicht früher zu publicieren, als Sie neue sehr positive Resultate Ihrer neuen Betrachtungsweise darlegen können. Wäre es Ihnen z. B. gelungen durch die Typentheorie die Frage zu entscheiden ob das Linearcontinuum dieselbe Mächtigkeit hat oder nicht wie die zweite Zahlenclasse, dann würde gewiss

Mais je ne vais pas vous cacher qu'il me semble qu'il serait préférable pour vous de ne pas publier ces recherches avant que vous ne puissiez présenter de nouveaux résultats très positifs qui proviennent de votre nouvelle façon de voir les choses. Si vous étiez par exemple parvenu par votre théorie des types à décider si le continu linéaire possède ou non la même puissance que la deuxième classe de nombres, alors il serait certainement...

(*c'est l'hypothèse du continu*)

Ivor Grattan-Guinness (1941–2014)

Plus de dix ans après, à l'occasion de la traduction en italien de nouveaux travaux, Cantor évoque dans une lettre à Gerbaldi l'épisode de 1885 avec « M. L. », qu'il n'a visiblement toujours pas digéré : pour le caresser dans le sens du poil, Mittag-Leffler avait avancé dans sa lettre à Cantor que son travail était de cent ans en avance, et qu'il était donc urgent d'attendre...

Lettre de Cantor à Gerbaldi en 1896 :

Ich bekam plötzlich von Herrn M. L. einen Brief, worin er mir zu meinem grössten Erstaunen schreibt, er halte nach reiflicher Überlegung diese Publication für «*um hundert Jahre verfrüht*». Nach den Intentionen von Herrn M. L. hätte ich also noch bis zum Jahre 1984 damit warten sollen, was mir doch eine zu starke Zumuthung zu sein schien!

Je reçus tout à coup une lettre de Monsieur M. L. dans laquelle, à mon grand étonnement, il m'écrivait qu'après mûre réflexion il considérait que cette publication [*ce projet de publication de Cantor pour Acta*] était « *de cent ans trop en avance* ». Selon les intentions de Monsieur M. L., j'aurais dû attendre jusqu'en 1984, ce qui m'a semblé être une exigence bien trop forte !

Francesco Gerbaldi (1858–1934)

Ordinaux et topologie

On a déjà évoqué *l'exhaustion du dénombrable*, qui implique que le processus de dérivation d'un sous-ensemble de $\mathbf{R}$ (ou de $\mathbf{R}^d$, aussi bien) se « stabilise » à partir d'un certain ordinal dénombrable. On indique quelques noms de mathématiciens qui ont eu un rapport plus ou moins étroit avec les ordinaux de Cantor. Borel, par exemple, a mentionné les ordinaux dans sa thèse en 1895, mais ce fut plutôt anecdotique.

Si F *est un fermé de* $\mathbb{R}$ *(ou de* $\mathbb{R}^d$*), il existe un ordinal dénombrable* γ *tel que*

$$\mathrm{F}^{(\gamma)} = \mathrm{F}^{(\gamma+1)} = \cdots = \mathrm{F}^{(\Omega)}.$$

Si le dérivé $\mathrm{F}^{(\gamma)}$ *est vide, alors* F *est dénombrable, sinon l'ensemble* $\mathrm{F}^{(\gamma)}$ *est* parfait.

Ivar Bendixson (1861–1935)
René Baire (1874–1932)
Émile Borel (1871–1956)
Henri Lebesgue (1875–1941)

page suivante : une version de l'exhaustion permet de dire qu'une famille croissante de réels de $[0, 1)$, telle que $x_\alpha < x_{\alpha+1}$ pour tout ordinal α tel que $x_\alpha < 1$, finit par atteindre 1 : Lebesgue désigne ce procédé sous le nom de *chaîne d'intervalles*. Il l'utilise pour démontrer une version raffinée du *théorème des accroissements finis* : si f est continue, dérivable à droite avec $f'_\mathrm{d} > 0$ sauf peut-être en une famille dénombrable de points, alors on a $f(0) < f(1)$.

Il faut un argument spécifique, non ordinal, pour gérer l'ensemble dénombrable de points sans dérivée : supposons ici que la dérivée à droite existe en tout point de $[0, 1)$.

Posons $x_0 = 0$. Puisque $f'_\mathrm{d}(x_0) > 0$, on peut trouver $x_1 > x_0$ tel que $f(x_1) > f(x_0) = f(0)$. L'idée est que tant que $f'_\mathrm{d}(x_\alpha) > 0$, on pourra trouver $x_{\alpha+1} > x_\alpha$ tel que $f(x_{\alpha+1}) > f(x_\alpha)$; quand β est un ordinal limite, on définit x_β comme la limite des x_α précédents et la continuité de la fonction f implique $f(x_\beta) \geq f(x_1) > f(0)$. On continue « ordinalement » jusqu'à atteindre le point 1... Bourbaki a repris ce théorème, mais sans ordinaux !

plus généralement :

Si $(\mathrm{F}_\alpha)_{\alpha<\Omega}$ *est une famille décroissante de fermés de* $\mathbb{R}^d$*, il existe un ordinal dénombrable* γ *tel que*

$$\mathrm{F}_\gamma = \mathrm{F}_{\gamma+1} = \cdots = \bigcap_{\alpha<\Omega} \mathrm{F}_\alpha .$$

Borel, et surtout Lebesgue dans un livre de 1904 :

Si $(x_\alpha)_{\alpha<\Omega}$ *est une famille croissante de points de* $[0,1]$ *telle que*

$$x_\alpha < 1 \implies x_\alpha < x_{\alpha+1},$$

il existe un ordinal dénombrable γ *tel que* $x_\gamma = 1$.

application au théorème de recouvrement fini (dit de *Borel*–*Lebesgue*), versions du Théorème des Accroissements Finis

H. Lebesgue, Leçons sur l'intégration et la recherche des fonctions primitives, 1904 deux pages suivantes : si beaucoup de mathématiciens « installés » ne veulent pas entendre parler des idées de Cantor, le jeune suédois Bendixson (il a 22 ans en 1883) avale aisément le pavé allemand de Cantor paru en 1883 ; il y détecte une inexactitude qu'il corrige presque immédiatement, et son énoncé corrigé paraît dans Acta. Ce journal présente l'article comme étant une lettre de Bendixson à Cantor ; on remarquera bien sûr que cette « lettre-article » est écrite *en français,* et que si Bendixson cite un passage de Cantor en allemand, il n'éprouve pas le besoin de le traduire : on connaissait les langues, à l'époque !

Cantor sera « fair play » avec Bendixson : dans ses travaux ultérieurs, Cantor attribuera ce théorème, qu'il avait très largement contribué à mettre à jour, au seul Bendixson. On parle le plus souvent maintenant de *théorème de Cantor–Bendixson*.

QUELQUES THEORÈMES

DE LA THÉORIE DES ENSEMBLES

DE POINTS

Extrait d'une lettre adressée à M. Cantor à Halle

PAR

IVAR BENDIXSON

à STOCKHOLM.

.... Dans votre travail récemment publié: »Grundlagen einer allgemeinen Mannichfaltigkeitslehre», vous avez énoncé le théorème suivant, qui me paraît devoir être rectifié dans quelques parties. Vous y dites page 31:

»Hat $P^{(1)}$ die Mächtigkeit der zweiten Zahlenclasse (II), so lässt sich $P^{(1)}$ stets und nur auf einzige Weise in zwei Mengen R und S zerlegen, so dass

$P^{(1)} \equiv R + S$, wo R und S eine äusserst verschiedene Beschaffenheit haben:

R ist so beschaffen dass sie durch den wiederholten Ableitungsprocess einer fortwährenden Reduction bis zur Annihilation fähig ist, so dass es immer eine erste ganze Zahl γ der Zahlenclassen (I) oder (II) giebt, für welche $R^{(\gamma)} \equiv 0$; solche Punctmengen R nenne ich reductibel.

---

version « corrigée » par Bendixson :

»Si P est un ensemble de points situés dans un espace continu à n dimensions, et si P' a une puissance plus grande que la première, je puis toujours le diviser d'une seule manière

$$P' \equiv R + S$$

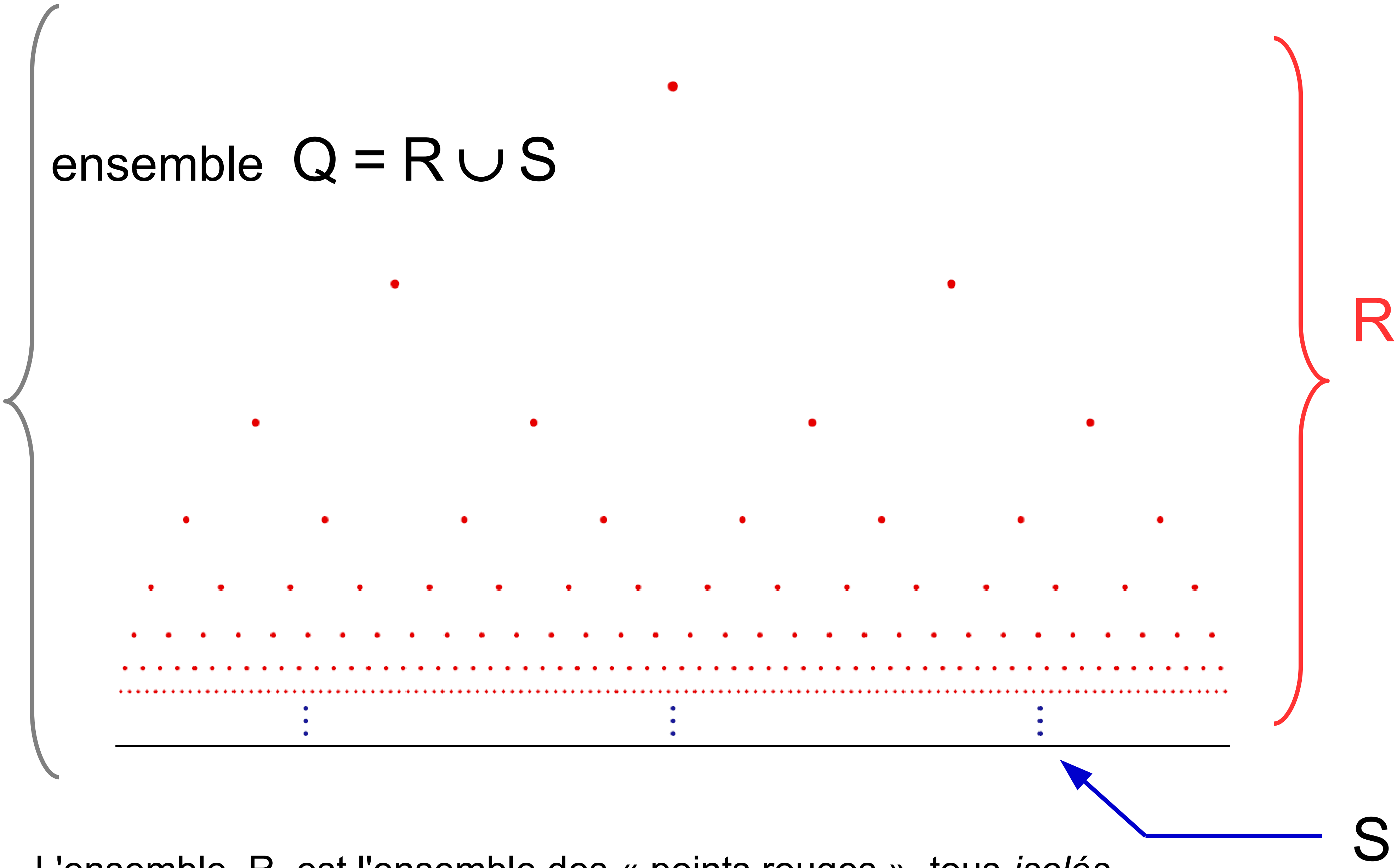
où S est un ensemble parfait et R est de la première puissance et tel qu'il existe toujours un γ, tel que

$$\mathfrak{D}(R,\ R^{(\gamma)}) \equiv 0, \qquad \mathrm{R} \cap \mathrm{R}^{(\gamma)} = \emptyset$$

γ étant un de vos symboles d'infini correspondant à ce que vous avez nommé nombre de la classe (I) ou (II).

(théorème de Cantor–Bendixson)

page précédente : il s'agit de décomposer un ensemble fermé non-dénombrable de $\mathbf{R}^d$, ayant la forme d'un *dérivé* P', en un ensemble dénombrable R et un ensemble parfait S. Cantor a affirmé un peu vite qu'il peut trouver un ensemble R tel que son dérivé $R^{(\gamma)}$ finisse par « s'annuler » pour un certain ordinal γ dénombrable. Bendixson présente un contre-exemple, situé dans un segment, ce qui rend les choses un peu difficiles ; un exemple dans le plan est plus facile à dessiner : on a vu un exemple d'ensemble Q du plan, formé d'un « arbre » dénombrable R de points isolés et de l'ensemble S = R' des points limites de R, où S est un segment, *parfait*. On a vu aussi que cet ensemble Q peut être considéré comme le dérivé P' d'un ensemble P (imaginer que P est formé en adjoignant à chaque « point rouge » de la figure page suivante une suite de points qui converge vers ce point rouge). L'ensemble dénombrable de l'énoncé de Cantor, appliqué à cet ensemble P' = Q, ne peut être que R, mais R' = S et il en résulte que $R^{(\gamma)}$ = S pour tout ordinal γ, donc $R^{(\gamma)}$ n'est jamais vide.



L'ensemble R est l'ensemble des « points rouges », tous *isolés.*

Les « points de suspension » bleus invitent à *imaginer* les points rouges absents, qu'on n'arriverait plus à discerner parce que trop proches les uns des autres.

Le segment noir S est l'ensemble des *points limites* des points rouges.

La correction de Bendixson revient à considérer R « en lui-même », comme un *espace topologique* en soi : faisant abstraction des points extérieurs à R, on ne considère alors que les *points limites* de points de R qui sont encore dans « l'espace R » lui-même (dans l'exemple de la page précédente, il n'y en a aucun). L'énoncé corrigé indique qu'il existe un ordinal dénombrable γ tel que le « dérivé en lui-même » d'ordre γ de R soit vide, c'est-à-dire que

$$\mathrm{R} \cap \mathrm{R}^{(\gamma)} = \emptyset .$$

Cantor raisonnera ainsi dans un article de 1885, son dernier article dans Acta, bien moins connu que le « fameux » article de 1883 :

Über verschiedene Theoreme aus der Theorie der Punctmengen in einem n-fach ausgedehnten stetigen Raume G_n. Zweite Mittheilung, Acta Math 7 (1885), p. 105–124.

Baire (1899)

Une fonction f définie sur une partie F de [0, 1] (ou de $\mathbf{R}^m$) est dite de *première classe* si elle est limite en tout point de F d'une suite de fonctions f_n *continues* sur F,

$$\forall x \in \mathrm{F}, \quad f(x) = \lim_{n \to \infty} f_n(x)$$

(on dit : f est *limite simple* des f_n). Si F est non vide *et fermé*, la fonction f a (au moins) un point de continuité dans F.

(définition et résultat de Baire)

Il s'agit de points de continuité $x \in F$, et *relatifs à* F **:** si y est proche de x **et si** $y \in$ F, alors $f(y)$ est proche de $f(x)$.

Le résultat s'étend aux fonctions sur un espace métrique F, à condition que F soit *complet* **:** si F était l'ensemble des rationnels et (x_n) une énumération de F, la fonction f qui serait définie par $f(x_n) = n$ pour tout n n'aurait aucun de point de continuité.

Baire va utiliser de façon cruciale les ordinaux de Cantor dans la preuve de son « grand théorème », qu'on va énoncer trois pages plus loin **;** dans son article (sa thèse), il rappelle les bases **:** →

Sur les fonctions de variables réelles.

(*Par* R. Baire *à Bar-le-Duc.*)

(publié en 1899)

[...]

Pour voir exactement jusqu'où nous pouvons aller dans cette voie, rappelons brièvement les définitions de M. Cantor relatives aux ensembles dérivés d'ordre supérieur (*). Si un ensemble P est de telle nature que le dérivé d'ordre n, P^n, existe, quel que soit n, il y a des points qui appartiennent à P^n, quel que soit n; on désigne l'ensemble de ces points par P^ω, et l'on convient de dire que P^ω est l'ensemble dérivé de P d'ordre ω. On désigne les ensembles dérivés successifs de P^ω par la notation $P^{\omega+1}, P^{\omega+2}, \ldots P^{\omega+n}, \ldots$; s'il en existe, quel que soit n, les points qui appartiennent à tous ces ensembles forment un nouvel ensemble qu'on appelle $P^{2\omega}$, etc.; on est ainsi conduit à la notion d'ensemble dérivé d'ordre α, α étant un quelconque des symboles définis par M. Cantor et qu'il appelle nombres de la deuxième classe.

Baire rappelle les notions introduites par Cantor, cependant . . . →

Je ferai remarquer à ce sujet, une fois pour toutes, que nous n'aurons jamais à nous préoccuper des difficultés que peut comporter en soi la notion abstraite de *nombre transfini*, bien que cette expression puisse être employée par nous dans la suite de ce travail. Dans le cas actuel, par exemple, l'ensemble P^α, α étant un nombre déterminé de la deuxième classe de nombres, représente quelque chose de parfaitement déterminé, indépendamment de toute considération abstraite relative aux symboles de M. Cantor; il n'y a donc, dans l'usage que nous pourrons faire de la locution *nombre transfini*, rien de plus que l'emploi d'un langage commode. Il en sera toujours de même dans la suite. (Baire 1899)

. . . on peut voir ici le souci de se démarquer un peu de Cantor et de ses théories, jugées fumeuses par bien des mathématiciens français dont l'avenir professionnel du jeune Baire peut dépendre. Mais si α est un ordinal si grand qu'il n'existe pour lui *aucun système de notation*, au sens qu'on a expliqué, Baire peut-il sérieusement prétendre que le dérivé d'ordre α de P est « quelque chose de parfaitement déterminé, indépendamment [. . . *des* . . .] symboles de M. Cantor » ?

Théorème (Baire, 1899).

> *Pour que f soit de première classe sur $[0,1]$, il faut et il suffit que pour tout ensemble parfait P contenu dans $[0,1]$, la restriction f_{P} de f à P possède un point de continuité.*

Une des directions de cette équivalence a déjà été mentionnée (l'existence de points de continuité) ; il n'y a pas d'ordinaux dans cette partie, non immédiate, qui est devenue « un classique ».

Pour l'autre direction, Baire explique d'abord l'idée *ordinale* de la preuve dans le cas très particulier où la fonction f à étudier ne prend que les valeurs 0 et 1 ; dans ce cas, si F est un fermé de [0, 1], si f_{F} désigne la restriction de f à F, un point $x \in \mathrm{F}$ est point de continuité de f_{F} s'il existe un intervalle ouvert I contenant x tel que

$$(*) \qquad y \in \mathrm{I} \cap \mathrm{F} \quad \text{implique} \quad f(y) = f(x),$$

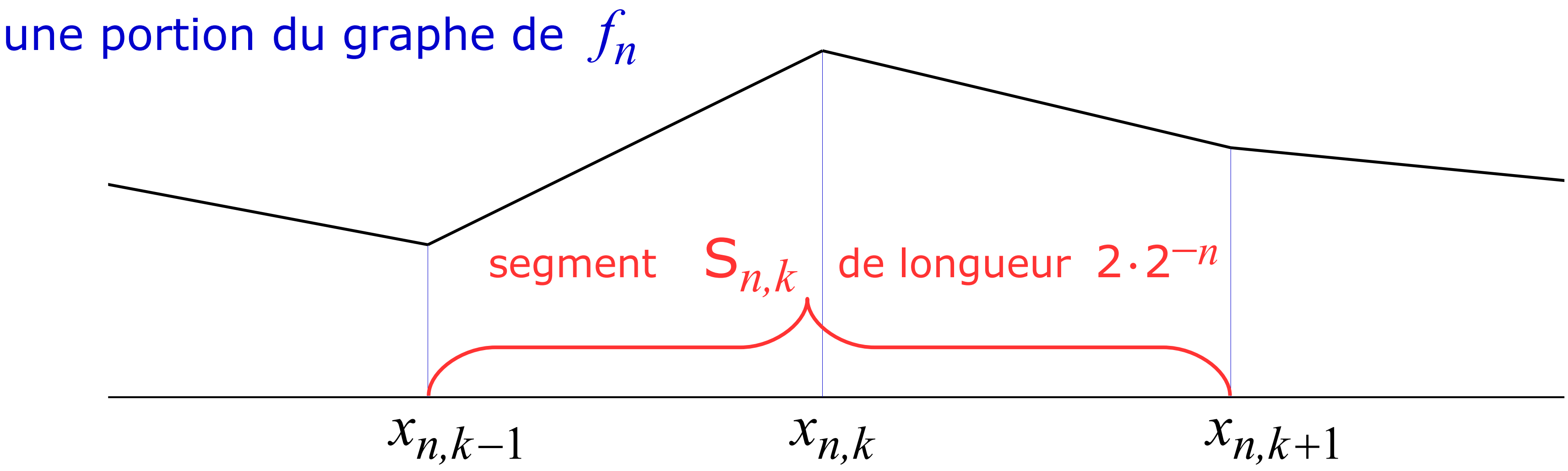
puisque les seules valeurs possibles sont 0 et 1, et dans ce cas les points de discontinuité de f_{F} forment un sous-ensemble *fermé* de F (c'est faux en général). Notons que si x est un point isolé de F, c'est un point de continuité de f_{F} : en effet, il existe un intervalle ouvert I tel que $\mathrm{F} \cap \mathrm{I} = \{x\}$, donc $(*)$ est vrai (on y a alors nécessairement $y = x$).

L'hypothèse de Baire peut être reformulée ainsi : *pour tout fermé* F non vide *contenu dans* [0, 1], *il existe un point de continuité* $x \in$ F *pour la restriction* f_{F}. En effet, ou bien F est parfait et on applique directement l'hypothèse, ou bien il existe un point x de F qui est isolé dans F, et ce point est alors point de continuité de f_{F}.

On introduit une famille ordinale décroissante de fermés F_α, *strictement* décroissante tant que F_α n'est pas vide, de la façon suivante : l'ensemble des points de discontinuité de la restriction de f à F_α est fermé ; on prend pour $F_{\alpha+1}$ cet ensemble fermé, qui est différent de F_α par hypothèse, si F_α n'est pas vide. Quand β est un ordinal limite, F_β est l'intersection des F_α précédents, intersection non vide si ces F_α sont non vides (compacité).

Le processus va se stabiliser (comme toujours) et on aura alors $F_\alpha = \varnothing$; à partir de cette information, Baire peut construire une suite de fonctions continues qui tend simplement vers f. On va voir cette déduction (mais pas le cas d'une fonction f générale).

On ne peut pas borner le premier ordinal α pour lequel F_α devient vide autrement qu'en disant que $\alpha < \Omega$.



La suite (f_n) de fonctions continues qui tend vers f est formée de fonctions affines par morceaux ; si on fixe l'entier n, on considère les points $x_{n,k} = k\, 2^{-n}$ pour $k = 0, \ldots, 2^n$ et on définit f_n sur $[0, 1]$ en fixant ses valeurs en ces points, puis en prolongeant « affinement ».

On désigne par $S_{n,k}$ le segment $[x_{n,k-1}, x_{n,k+1}]$. Puisque $F_\Omega = \varnothing$, il existe un plus petit ordinal β tel que $S_{n,k} \cap F_\beta$ soit vide, et par compacité on sait que β ne peut pas être un ordinal limite. On a donc $\beta = \alpha + 1$ pour un certain ordinal α. Alors $S_{n,k} \cap F_\alpha$ n'est pas vide et on pose (avec Baire)

$f_n(x_{n,k})$ = le minimum de f sur $S_{n,k} \cap F_\alpha$ (ici, la valeur est 0 ou 1).

On va montrer que f_n tend simplement vers f.

Fixons un point $x \in [0, 1]$; il existe un plus petit ordinal β tel que $x \notin F_\beta$, et cet ordinal est de la forme $\beta = \alpha + 1$. Alors x est dans F_α mais pas dans $F_{\alpha+1}$, ce qui signifie que x est un point de continuité de la restriction de f à F_α. Il existe donc un intervalle ouvert I contenant x tel que

$$(*) \qquad y \in I \cap F_\alpha \text{ implique } f(y) = f(x)$$

(parce que f ne prend que les valeurs 0 et 1), et on peut « rétrécir » l'intervalle I de façon qu'il ne rencontre pas le fermé $F_{\alpha+1}$ tout en contenant encore x.

Pour n assez grand, les segments $S_{n,k}$ (définis page précédente) qui contiennent x sont contenus dans l'intervalle I, donc ils rencontrent F_α (au moins en x) mais ils ne rencontrent pas $F_{\alpha+1}$. Si on revient à la définition de f_n, on constate que cet ordinal α est précisément celui qui a servi à poser

$$f_n(x_{n,k}) = \text{le minimum de } f \text{ sur } S_{n,k} \cap F_\alpha.$$

Ici f est constante égale à $f(x)$ dans $S_{n,k} \cap F_\alpha$ d'après $(*)$, ce qui permet de voir que $f_n(x) = f(x)$ pour n assez grand. □

Théorie des ensembles

page suivante : il s'agit d'une version de *l'argument diagonal* de Cantor. Dans l'article original, le résultat est formulé en termes de fonctions de l'ensemble X à valeurs dans la paire {0, 1}, ce qui revient au même que de parler de l'ensemble des parties de X : il suffit d'associer à chaque sous-ensemble Y de X la fonction χ_Y qui vaut 1 sur Y et 0 en dehors de Y.

Appliqué lorsque X = $\mathbf{N}$, l'argument diagonal fournit une autre preuve de la non-dénombrabilité de la droite réelle. Mais il est beaucoup plus général : si on utilise l'axiome du choix, l'argument diagonal montre que le cardinal de l'*ensemble des parties* $\mathscr{P}$(X) d'un ensemble X quelconque est (strictement) supérieur au cardinal de X (l'axiome du choix intervient quand on veut manipuler aisément un ordre des cardinaux). Sans faire appel à l'axiome du choix, on a en tout cas que :

il n'existe pas de surjection de X *sur* $\mathscr{P}$(X).

(si on a une surjection p de X sur Y, l'axiome du choix permet de trouver une injection j de Y dans X, *en choisissant* un point $x = \mathrm{j}(y)$ dans l'image inverse $\mathrm{p}^{-1}(y)$, *non vide*, de chaque point y de Y).

(Cantor 1891) :

Si X est un ensemble quelconque et si ϕ est une application de X dans $\mathcal{P}(X)$, le sous-ensemble

$$Y = \{x \in X : x \notin \phi(x)\} \subset X$$

$Y \in \mathscr{P}(X)$

n'est pas dans l'image de ϕ (ϕ n'est jamais surjective)

Si on avait $Y = \phi(y)$, alors : $y \in Y$? $y \notin Y$?

implique $y \in \phi(y) = Y$, contradiction

implique $y \notin \phi(y) = Y$, contradiction

$\mathcal{P}(X)$ . . . est toujours plus « gros » que . . . X

Derniers articles mathématiques de Cantor

« Mannigfaltigkeit » (encore en 1891)

devient définitivement « Menge » (1895)

Beiträge zur Begründung der transfiniten Mengenlehre[1].

[Math. Annalen Bd. 46 S. 481—512 (1895); Bd. 49, S. 207—246 (1897).]

„Hypotheses non fingo.“ [Newton.]

„Neque enim leges intellectui aut rebus damus ad arbitrium nostrum, sed tanquam scribae fideles ab ipsius naturae voce latas et prolatas excipimus et describimus.“

„Veniet tempus, quo ista quae nunc latent, in lucem dies extrahat et longioris aevi diligentia.“

« Unter einer “Menge” verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unsrer Anschauung oder unseres Denkens (welche die “Elemente” von M genannt werden) zu einem Ganzem. »

Tentative de traduction :
Par le mot « ensemble », nous entendons le rassemblement en un tout, que nous désignons par M, *de certains objets bien différenciés* m *(qu'on appellera les « éléments » de* M*), objets provenant de notre intuition ou de notre pensée.*

Cette définition des ensembles est bien vague, il ne peut pas en être autrement. Le contenu des articles de 1895 et 1897 reprend en partie celui de l'article de 1885 qui avait été retiré d'Acta par Cantor. Une notion principale est celle de *type d'ordre* ; Cantor caractérise, entre autres types, le type d'ordre des rationnels :

un ensemble ordonné dénombrable, sans plus grand ni plus petit élément, et tel *qu'entre* x et y distincts il existe toujours un autre élément z distinct de x et de y, est isomorphe pour l'ordre à l'ensemble des rationnels.

Cantor, en 1895 et 1897 **:** *types d'ordre* **;**

un *ordinal* représente une famille d'ensembles bien ordonnés, formée de tous les ensembles bien ordonnés qui sont *isomorphes pour l'ordre* à un membre (quelconque) de cette famille.

(un *cardinal* représente une famille d'ensembles formée de tous les ensembles qui peuvent être mis *en bijection* avec l'un des membres de cette famille)

Cantor montre que **:**

– étant donnés deux ensembles bien ordonnés, ou bien ils sont isomorphes pour l'ordre, ou bien l'un des deux est isomorphe pour l'ordre à une *section commençante* de l'autre,

$$\mathrm{S}_x = \{y \in \mathrm{X} : y < x\}, \qquad x \in \mathrm{X}\ ;$$

– si deux ensembles bien ordonnés sont isomorphes pour l'ordre, il existe un *unique* isomorphisme d'ordre entre eux.

page précédente : on peut noter un certain retournement entre les articles de Cantor des années 1880 et les derniers articles.

Précédemment, la *formation des ordinaux* était à la base des choses, et on pouvait ensuite assigner un ordinal à tout ensemble bien ordonné X, selon une procédure assez simple : par définition, l'ensemble bien ordonné X possède un plus petit élément, on l'associe au premier ordinal (pour Cantor, c'est 1). Tant qu'on n'a pas épuisé l'ensemble X, le sous-ensemble Y de X formé des éléments « non traités » est non vide, il a donc un plus petit élément auquel on associe le plus petit ordinal plus grand que ceux qu'on a déjà utilisés, etc.

Maintenant, Cantor part de la notion d'ensemble bien ordonné, et un ordinal représente la collection de tous les ensembles bien ordonnés qui sont deux à deux isomorphes pour l'ordre (autrement dit, qui ont le même *type d'ordre*) : on n'a plus besoin *d'engendrer* les ordinaux, ce sont les *ensembles existants* qui les amènent.

Comme on l'a déjà indiqué, Cantor « renverse » dans ses derniers articles la convention de notation pour le produit de deux ordinaux. Désormais, on a

$$2\,\omega = \omega, \qquad \omega\cdot 2 \neq \omega, \qquad \omega\cdot 2 > \omega+n,$$

où n est un entier fini quelconque. Cantor définit l'exponentiation des ordinaux, en particulier

$$\omega^{\alpha}$$

pour tout ordinal dénombrable α. Il définit cette exponentielle comme limite des ω^{α_n} pour une suite (α_n) « tendant » en croissant vers α. Avec la nouvelle convention pour la multiplication, on a bien que

$$\omega^{\alpha+\beta} = \omega^{\alpha}\,\omega^{\beta}.$$

Revenons à la lettre de Cantor à Gerbaldi en 1896, avec trois lignes qu'on n'avait pas encore vues **:**

«... Die Theorie der Ordnungstypen war bereits vor *elf* Jahren ... fertig, ... Über den *eigentlichen Grund*, warum der Druck damals sistirt wurde, bin ich heute nicht unterrichtet, er ist mir ein *Räthsel* !

Ich bekam plötzlich von Herrn M. L. einen Brief, worin er mir zu meinem grössten Erstaunen schreibt, er halte nach reiflicher Überlegung diese Publication für «*um hundert Jahre verfrüht*». Nach den Intentionen von Herrn M. L. hätte ich

[...] La théorie des types d'ordre était déjà prête depuis *onze* ans. **[…]** Je ne connais toujours pas aujourd'hui la *raison réelle* pour laquelle l'impression [*de l'article*] a été suspendue à ce moment-là, c'est pour moi un *mystère* !

Je reçus tout à coup une lettre de Monsieur M. L. dans laquelle, à mon grand étonnement, il m'écrivait qu'après mûre réflexion **[…]**

Axiomatique

page suivante : bien que ce paradoxe fameux porte le nom du logicien célèbre Russell, il est plus que vraisemblable qu'il n'a pas été le premier à le formuler. On pourra remarquer que le mécanisme du paradoxe fait penser fortement au mécanisme qu'on a vu dans le « procédé diagonal » de Cantor.

L'apparition de ces paradoxes a certainement renforcé la position des adversaires de la « théorie des ensembles ». On peut penser que les paradoxes n'ont pas bouleversé Cantor, qui, ayant montré que la famille des ordinaux dénombrables n'est pas dénombrable, pouvait donc bien imaginer que la collection de tous les ordinaux ne peut pas être envisagée pour former un nouvel ordinal ; aussi, il avait pris des précautions dans la formation de nouveaux ordinaux (le *troisième principe*). Mais Cantor n'avait pas de *système* pour clarifier les choses : les théories axiomatiques seront les bienvenues.

Le temps des paradoxes (1890–1900, environ)

le paradoxe dit « de Russell » :

Soit E l'ensemble des F tels que $F \notin F$.

Est-ce que $E \in E$?

Si F = E est un élément de E, alors $E = F \notin F = E$! impossible ;

Si F = E n'est pas élément de E, alors $E = F \in F = E$! impossible.

Bertrand Russell (1872–1970)

page suivante : on indique quelques noms de mathématiciens et de logiciens qui ont contribué aux débuts des préoccupations axiomatiques avant 1900. Dedekind et Peano ont travaillé sur la notion de nombre entier : les axiomes de Peano pour les entiers sont suffisamment célèbres, on les verra tout de suite ; on attribue à Frege l'introduction des quantificateurs en mathématiques, on les verra aussi, mais un peu plus loin dans ce qui suit. On ne donnera pas les premières notations de Frege, disons simplement qu'elles étaient assez compliquées et encore loin des actuelles (voir le livre de van Heijenoort, pages 1–82, en particulier p. 24 et suivantes ; la référence de ce livre est donnée à la fin de ce document).

Richard Dedekind (1831–1916),
« Was sind und was sollen die Zahlen ? » (1888)

Giuseppe Peano (1858–1932),
« Arithmetices principia, nova methodo exposita », en latin (1889)

Gottlob Frege (1848–1925),
« Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens » (1879) ;

Frege fonde le calcul des prédicats :

dans le calcul propositionnel on avait :

NON, ET, OU, =

dans le calcul des prédicats on ajoute : $\forall x\, \mathrm{P}(x)$ $\exists x\, \mathrm{P}(x)$

ARITHMETICES PRINCIPIA.

Peano (1889, en latin)

page 1, après la page XVI des préliminaires

§ 1. De numeris et de additione.

Explicationes.

Signo N significatur *numerus (integer positivus).*

» 1 » *unitas.*

» $a+1$ » *sequens a*, sive *a plus* 1.

» $=$ » *est aequalis*. Hoc ut novum signum considerandum est, etsi logicae signi figuram habeat.

Axiomata.

1. $1 \,\epsilon\, \mathrm{N}$.
2. $a \,\epsilon\, \mathrm{N} \,.\!\supset\!.\, a = a$.

Axiomata.

1. $1 \in N.$
2. $a \in N \,.\, \supset .\, a = a.$
3. $a, b, c \in N \,.\, \supset :\, a = b \,.\, = .\, b = a.$
4. $a, b \in N \,.\, \supset \therefore a = b \,.\, b = c : \supset .\, a = c.$
5. $a = b \,.\, b \in N : \supset .\, a \in N.$
6. $a \in N \,.\, \supset .\, a + 1 \in N.$
7. $a, b \in N . \supset : a = b \,.\, = .\, a + 1 = b + 1.$
8. $a \in N \,.\, \supset .\, a + 1 -= 1.$
9. $k \in K \therefore 1 \in k \therefore x \in N \,.\, x \in k : \supset_x .\, x + 1 \in k :: \supset .\, N \supset k.$

les *axiomes*

de Peano

Definitiones.

10. $2 = 1 + 1;\ 3 = 2 + 1;\ 4 = 3 + 1;$ etc.

PEANO, *Arithmetices principia.*

page précédente **:** l'axiomatique dans cet ouvrage de Peano *en latin* s'accompagne d'une quasi-disparition des mots. Les *points*, *deux-points* ou le symbole à *trois points* servent de séparateurs et de parenthèses. Un « signe moins » est utilisé pour la négation. Le symbole en forme de **C** renversé sert de signe d'implication, mais aussi, c'est plus troublant pour nous, il sert de symbole *d'inclusion* des « ensembles » que Peano appelle des *classes*. La lettre K majuscule désigne la famille des classes.

---

deux pages suivantes **:** on a traduit en notation moderne quelques uns des axiomes. À la deuxième page qui suit, Peano donne un exemple de preuve (sans parole !) pour le « théorème » suivant **:**

2, *défini comme* $1+1$, *est un entier*.

Axiomes de Peano (1889)

1. $1 \,\varepsilon\, \mathrm{N}.$ (pour Peano **1** est le plus petit entier)
2. $a \,\varepsilon\, \mathrm{N} \,.\!\supset\!.\, a = a.$
3. $a, b \,\varepsilon\, \mathrm{N} \,.\!\supset\!:\, a = b \,.\!=\!.\, b = a.$
4. $a, b, c \,\varepsilon\, \mathrm{N} \,.\!\supset\!\therefore\, a = b \,.\, b = c \,:\!\supset\!.\, a = c.$
5. $a = b \,.\, b \,\varepsilon\, \mathrm{N} \,:\!\supset\!.\, a \,\varepsilon\, \mathrm{N}.$
6. $a \,\varepsilon\, \mathrm{N} \,.\!\supset\!.\, a + 1 \,\varepsilon\, \mathrm{N}.$
7. $a, b \,\varepsilon\, \mathrm{N} \,.\!\supset\!:\, a = b \,.\!=\!.\, a + 1 = b + 1.$
8. $a \,\varepsilon\, \mathrm{N} \,.\!\supset\!.\, a + 1 \mathrel{-\!=} 1.$
9. $k \,\varepsilon\, \mathrm{K} \,.\!\therefore\!.\, 1 \,\varepsilon\, k \,.\!\therefore\!.\, x \,\varepsilon\, \mathrm{N} \,.\, x \,\varepsilon\, k \,:\!\supset_x\!.\, x + 1 \,\varepsilon\, k \,::\!\supset\!.\, \mathrm{N} \supset k.$

1. 6. 8.

$1 \in \mathbf{N}\,; \quad \forall a \in \mathbf{N} \,:\, a+1 \in \mathbf{N},\ a+1 \neq 1$

7. $\forall a, b \in \mathbf{N},\ a = b \Leftrightarrow a+1 = b+1$

9. c'est le principe de récurrence : k est une « classe » contenant 1 et qui contient $x+1$ chaque fois qu'elle contient un entier x. Alors cette classe k contient la classe N des entiers.

(l'extrait provient d'une traduction en anglais d'un passage de Peano, dans le livre de van Heijenoort cité à la fin)

Theorems

11. $2 \,\varepsilon\, \mathrm{N}.$

Proof:

P 1 .Ɔ:	$1 \,\varepsilon\, \mathrm{N}$	(1)
1 [a] (P 6) .Ɔ:	$1 \,\varepsilon\, \mathrm{N}$.Ɔ. $1 + 1 \,\varepsilon\, \mathrm{N}$	(2)
(1) (2) .Ɔ:	$1 + 1 \,\varepsilon\, \mathrm{N}$	(3)
P 10 .Ɔ:	$2 = 1 + 1$	(4)
(4) . (3) . (2, 1 + 1) [a, b] (P 5) :Ɔ:	$2 \,\varepsilon\, \mathrm{N}$	(Theorem).

Note. We have written explicitly all the steps of this very easy proof. For the sake of brevity, we now write it as follows:

P 1 . 1 [a] (P 6) :Ɔ: $1 + 1 \,\varepsilon\, \mathrm{N}$. P 10 . (2, 1+1) [a, b] (P 5) :Ɔ: Th.

or

P 1 . P 6 :Ɔ: $1 + 1 \,\varepsilon\, \mathrm{N}$. P 10 . P 5 :Ɔ: Th.

une preuve selon Peano

(histoire sans paroles)

David Hilbert (1862–1943)

(photo prise au début
des années 1900)

installé à Göttingen
depuis 1895.

Il publie
Grundlagen der Geometrie en 1899

page suivante : l'axiomatique de Hilbert pour la géométrie paraît en allemand en 1899, elle est très vite traduite en français. On notera un aspect crucial de cette axiomatique : les objets que Hilbert étudie *ne sont pas nécessairement* des points, des droites, des plans ; ce sont des « choses » faisant partie d'une certaine collection d'objets, qu'on peut décider d'appeler « points, droites, plans » s'il existe entre elles les relations prescrites par les axiomes : par exemple, *étant donnés deux « points », il doit exister une « droite » qui « passe par » ces deux points*, où toutes les expressions entre guillemets sont à prendre *au sens axiomatique*.

Si on veut présenter un exemple de géométrie plane, il faut d'abord expliquer qui sont les points, qui sont les droites et ce que signifie qu'un point est « sur » une droite ; ensuite, il faut voir si les axiomes sont satisfaits par ce candidat exemple.

Traduction française, publiée dès 1900

en allemand : « Dinge »
(choses)

Les éléments de la Géométrie et les cinq groupes d'axiomes.

Convention. — Concevons trois différents systèmes d'êtres : les êtres du PREMIER système, nous les nommerons *points* et nous les désignerons par A, B, C, ...; les êtres du DEUXIÈME système, nous le nommerons *droites* et nous les désignerons par a, b, c, ...; les êtres du TROISIÈME système, nous les nommerons *plans* et nous les désignerons par α, β, γ, ...; les points seront aussi nommés *éléments de la Géométrie linéaire* ; les points et les droites, *éléments de la Géométrie plane*; et les points, les droites et les plans, *éléments de la Géométrie de l'espace* ou *éléments de l'espace*.

Concevons que les points, droites et plans aient entre eux certaines relations mutuelles et désignons ces relations par des mots tels que : « SONT SITUÉS », « ENTRE », « PARALLÈLE », « CONGRUENT », « CONTINU » ; la description exacte et complète de ces relations a lieu au moyen des *axiomes de la Géométrie*

Le groupe d'axiomes I : Axiomes d'association.

Les axiomes de ce groupe établissent une *association* entre les notions précédemment indiquées, points, droites et plans. Ces axiomes sont les suivants :

I, 1. *Deux points distincts*, A, B, *déterminent toujours une droite* a ; *nous poserons* $AB = a$ *ou* $BA = a$.

Au lieu de « DÉTERMINENT », nous emploierons aussi d'autres tournures de phrase ; par exemple : A « EST SITUÉ SUR » a ; A « EST UN POINT DE » a ; a « PASSE PAR » A « ET PAR B » ; « a JOINT A ET B » ou « JOINT A A B ». Lorsque A est situé sur a et, en outre, sur une autre droite b, nous emploierons aussi le mode d'expression : « LES DROITES a ET b ONT LE POINT A EN COMMUN », et ainsi de suite.

I, 2. *Deux points distincts quelconques d'une droite déterminent cette droite, et sur toute droite il y a au moins deux points ; c'est-à-dire que, si l'on a* $AB = a$ *et* $AC = a$ *et* $B \neq C$, *on a aussi* $BC = a$.

I, 3. *Trois points* A, B, C *non situés sur une même droite déterminent toujours un plan* α ; *nous poserons* $ABC = \alpha$.

deux pages suivantes **:** la non-contradiction d'un système d'axiomes est évidemment une propriété éminemment souhaitable ! Hilbert montre que si son axiomatique de la géométrie était contradictoire, l'arithmétique serait déjà contradictoire. Il espère à cette époque qu'on parviendra à obtenir une preuve de la non-contradiction de l'arithmétique *avec les outils de l'arithmétique*. Cet espoir sera ruiné par les résultats de Gödel des années 1930.

Hilbert a en fait obtenu un résultat de *consistance relative* **:** si l'arithmétique n'est pas contradictoire, on peut lui ajouter les axiomes de Hilbert pour la géométrie sans introduire de contradiction. La logique mathématique du 20e siècle aura plusieurs résultats de ce type. Par exemple, on prouvera que si la théorie axiomatique des ensembles *sans axiome du choix* est non-contradictoire, on peut lui ajouter l'axiome du choix sans produire de contradiction.

ou « consistance » du système d'axiomes

La non-contradiction des axiomes.

Les axiomes des cinq groupes d'axiomes dont nous avons parlé dans le Chapitre I ne sont pas en contradiction, c'est-à-dire qu'il n'est pas possible d'en déduire par un raisonnement logique une proposition qui soit en contradiction avec un de ces axiomes. Pour le prouver il suffit d'assigner une géométrie où l'ensemble des cinq groupes soit vérifié.

A cet effet, considérons le domaine Ω de tous les nombres algébriques qui prennent naissance, lorsque, partant du nombre 1, l'on effectue un nombre fini de fois les quatre opérations, addition, soustraction, multiplication, division et une cinquième opération : $\sqrt{1+\omega^2}$, où ω désigne chaque fois un nombre ayant déjà pris naissance par le moyen de ces cinq opérations.

Nous regarderons un couple de nombres (x, y) du domaine Ω comme un point et le rapport $(u:v:w)$ de trois nombres quelconques de Ω, pourvu que u, v ne soient pas tous deux nuls, comme une droite ; enfin l'équation

$$ux + vy + w = 0$$

l'axiome d'Archimède V.

De tout cela on conclut que toute contradiction dans les conséquences tirées de nos axiomes devrait aussi apparaitre dans l'arithmétique du domaine Ω.

Les considérations analogues relatives à la Géométrie de l'espace ne présenteraient aucune difficulté.

Dans les développements qui précèdent, si l'on choisissait, au lieu du domaine Ω, le domaine de tous les nombres réels nous obtiendrions également une géométrie où l'ensemble des axiomes I-V serait aussi vérifié; mais pour notre démonstration il suffisait d'employer le domaine Ω qui renferme seulement un ENSEMBLE DÉNOMBRABLE d'éléments.

page suivante : préoccupation peut-être moins essentielle que celle de la non-contradiction du système d'axiomes, on souhaite ne pas avoir introduit d'axiomes *inutiles.* Inutiles, en ce qu'il résulteraient des autres axiomes et pourraient donc être omis : c'est la notion *d'indépendance* (ou pas) des axiomes d'une théorie.

On notera que pour prouver l'indépendance de « l'axiome d'Archimède », Hilbert envisage une « droite » où les points ne sont vraiment pas des points : ce sont des fonctions, qu'il prend aussi simples que possible pour réaliser son objectif, et qui sont telles que la différence de deux fonctions *distinctes* dans cette « géométrie » tend vers une limite *non nulle* à l'infini positif, limite finie ou infinie ; les « points » sont ordonnés sur la « droite » selon le signe de la limite en $+\infty$ de la différence. Ainsi, les multiples entiers de la fonction constante *positive* $\mathbf{1}$ ne dépassent jamais le « point » de la droite donné par la fonction $t \to t$; cela contrarie Archimède, mais confirme son indépendance.

Indépendance de l'axiome de la continuité V.
(Géométrie non archimédienne.)

Pour démontrer l'indépendance de l'axiome V dit *d'Archimède*, il nous faut construire une Géométrie où seront vérifiés tous les axiomes à l'exception de cet axiome en question (¹).

A cet effet, construisons le domaine $\Omega(t)$ de toutes les fonctions algébriques de t, qui proviennent de t au moyen des quatre opérations : addition, soustraction, multiplication, division, et de la cinquième opération $\sqrt{1+\omega^2}$, où ω désigne une fonction quelconque, déjà obtenue au moyen de ces cinq opérations. L'ensemble des éléments de $\Omega(t)$ — de même qu'il en était précédemment de Ω — est un ensemble dénombrable. Les cinq opérations peuvent être toutes effectuées d'une manière univoque et réelle. Le domaine $\Omega(t)$ ne renferme donc que des fonctions de t univoques et réelles.

Ernst Zermelo, en 1907

(1871–1953)

Zermelo fréquente la physique au début de sa carrière scientifique, notamment le domaine de Boltzmann. Il ne publie ses premiers résultats mathématiques notables qu'après l'âge de trente ans, en particulier **:** *tout ensemble peut être muni d'un bon ordre* (en 1904).

Dans sa preuve, Zermelo met en avant une forme particulière de l'axiome du choix, qui est l'existence d'une *fonction de choix* pour un ensemble X **:** une telle fonction ϕ est définie sur l'ensemble des sous-ensembles non vides de X et elle « choisit » un élément dans chacun de ces sous-ensembles Y non vides,

$$\phi(Y) \in Y.$$

On peut sentir à plusieurs endroits chez Cantor que ses conceptions très idéalistes sur les *principes de formation* des ordinaux lui faisaient considérer le « théorème de Zermelo » comme allant de soi. Avant même que des axiomes de la théorie des ensembles n'aient été mis en place clairement, Zermelo indique un chemin qui conduit de *l'axiome du choix* à l'existence d'un bon ordre sur un ensemble quelconque, et on ne trouve sur ce chemin que des constructions qui seront (presque) universellement admises plus tard.

Ludwig Boltzmann (1844–1906)

Théorème de Zermelo (1904)

Tout ensemble X peut être muni d'un bon ordre.

conséquence de *l'axiome du choix* :

$$\exists\, \phi : \mathcal{P}(\mathrm{X}) \setminus \{\emptyset\} \to \mathrm{X} \quad \text{t.q.} \quad \forall\, \mathrm{Y} \quad \phi(\mathrm{Y}) \in \mathrm{Y}$$

non vide contenu dans X,

(fonction de choix)

La « réciproque » du théorème de Zermelo est claire : si un ensemble X est bien ordonné, on obtient une fonction de choix en associant à tout sous-ensemble non vide Y de X son *plus petit élément*.

avec AC (Axiome du Choix) : tout ensemble peut être mis en bijection avec un ordinal pages suivantes : la première axiomatique pour la théorie des ensembles, due à Zermelo, est publiée en 1908. Comme Hilbert, Zermelo se propose de travailler sur une collection $\mathscr{B}$ de « choses », qu'on pourra décider de nommer « ensembles », au risque de confondre la notion vague et naïve *d'ensemble d'objets* avec les éléments de la collection $\mathscr{B}$, qui peut-être, ne sont pas plus de « vrais » ensembles au sens commun que les choses étudiées par Hilbert n'étaient de « vrais » points ou droites.

Pour essayer d'être aussi clair que possible dans tout ce qui suit, on écrira *ensemble* en *italique* et *en bleu* pour désigner un élément x de la collection $\mathscr{B}$ de « choses » envisagée par la théorie axiomatique des ensembles.

En parallèle, le « mathématicien normal » pourra aussi mentionner, s'appuyer sur, et manipuler les axiomes de la théorie axiomatique (en général sans le dire) quand il travaillera avec de « vrais » ensembles (le logicien dira plutôt *avec des ensembles au sens naïf*). Il y a des points de vue que le mathématicien normal n'adoptera pas je pense : le plus souvent, en théorie axiomatique des ensembles, tous les objets sont des *ensembles*, je ne crois pas que les entiers 9 ou 10 soient vus comme des ensembles par le mathématicien !

Zermelo 1908 : axiomes pour la théorie des ensembles

Grundlegende Definitionen und Axiome.

« choses »

1. Die Mengenlehre hat zu tun mit einem „*Bereich*" $\mathfrak{B}$ von Objekten, die wir einfach als „*Dinge*" bezeichnen wollen, unter denen die „Mengen" einen Teil bilden. Sollen zwei Symbole a und b dasselbe Ding bezeichnen, so schreiben wir $a = b$, im entgegengesetzten Falle $a \neq b$. Von einem Dinge a sagen wir, es „existiere", wenn es dem Bereiche $\mathfrak{B}$ angehört; ebenso sagen wir von einer Klasse $\mathfrak{K}$ von Dingen, „es gebe Dinge der Klasse $\mathfrak{K}$", wenn $\mathfrak{B}$ mindestens ein Individuum dieser Klasse enthält.

2. Zwischen den Dingen des Bereiches $\mathfrak{B}$ bestehen gewisse „*Grundbeziehungen*" der Form $a \,\varepsilon\, b$. Gilt für zwei Dinge a, b die Beziehung $a \,\varepsilon\, b$, so sagen wir, „a sei *Element* der Menge b" oder „b enthalte a als Element" oder „besitze das Element a". Ein Ding b, welches ein anderes a als Element enthält, kann immer als eine *Menge* bezeichnet werden, aber auch nur dann — mit einer einzigen Ausnahme (Axiom II).

3. Ist jedes Element x einer Menge M gleichzeitig auch Element der Menge N, so daß aus $x \,\varepsilon\, M$ stets $x \,\varepsilon\, N$ gefolgert werden kann, so sagen wir, „M sei *Untermenge* von N", und schreiben $M \mathrel{\underline{\in}} N$*). Es ist stets $M \mathrel{\underline{\in}} M$, und aus $M \mathrel{\underline{\in}} N$ und $N \mathrel{\underline{\in}} R$ folgt immer $M \mathrel{\underline{\in}} R$. „*Elementenfremd*"

page suivante : comme il existait dans la *géométrie de Hilbert* des *relations* entre points et droites, sur lesquelles portaient les axiomes, il existe ici une unique relation, la *relation d'appartenance* notée ε, sur laquelle porte les axiomes de la théorie des ensembles. Ainsi, Zermelo indique que sa collection $\mathscr{B}$ doit contenir une « chose » $\mathbf{0}$ telle qu'on n'ait jamais $x \,\varepsilon\, \mathbf{0}$, cette chose est « l'ensemble vide axiomatique » ; Zermelo le qualifie d'ensemble « impropre » : normalement, un « ensemble » décent se devrait de posséder des « éléments »...

Si x, y sont deux *ensembles* (noter *l'italique bleu* qu'on a expliqué deux pages plus haut) on écrira que x est *élément* de y, en utilisant encore « l'italique bleu », lorsqu'on aura $x \,\varepsilon\, y$ *au sens de* la théorie axiomatique des ensembles. On procédera de même pour les autres notions ensemblistes, telles que *sous-ensemble*, ou *ensemble des parties*, *ensemble vide*,...

die Frage, ob $M \subset N$ oder nicht.

Über die Grundbeziehungen unseres Bereiches $\mathfrak{B}$ gelten nun die folgenden „*Axiome*“ oder „*Postulate*“.

axiome d'extensionnalité :

Axiom I. Ist jedes Element einer Menge M gleichzeitig Element von N und umgekehrt, ist also gleichzeitig $M \mathrel{\underline{\in}} N$ und $N \mathrel{\underline{\in}} M$, so ist immer $M = N$. Oder kürzer: jede Menge ist durch ihre Elemente bestimmt.

(Axiom der Bestimmtheit.)

Die Menge, welche nur die Elemente $a, b, c, \cdots, r$ enthält, wird zur Abkürzung vielfach mit $\{a, b, c, \cdots, r\}$ bezeichnet werden.

(impropre) **Axiom II.** Es gibt eine (uneigentliche) Menge, die „*Nullmenge*“ 0, welche gar keine Elemente enthält. Ist a irgend ein Ding des Bereiches, so existiert eine Menge $\{a\}$, welche a und nur a als Element enthält; sind a, b irgend zwei Dinge des Bereiches, so existiert immer eine Menge $\{a, b\}$, welche sowohl a als b, aber kein von beiden verschiedenes Ding x als Element enthält.

(Axiom der Elementarmengen.)

5. Nach I sind die „Elementarmengen“ $\{a\}$, $\{a, b\}$ immer eindeutig bestimmt, und es gibt nur eine einzige „Nullmenge“. Die Frage

Axiome I. Si chaque élément d'un ensemble M est en même temps élément de N et inversement (on a donc en même temps $M \subset N$ et $N \subset M$), alors on a toujours $M = N$. En bref : chaque ensemble est déterminé par ses éléments.

Zermelo pour

Axiomes de la théorie des ensembles

Extensionnalité — X est déterminé par ses éléments

Existence des ensembles élémentaires — ensemble vide, singletons, paires

Axiome de la somme (concerne les réunions)

Ensemble des parties — $X \longrightarrow \mathscr{P}(X)$

de X on passe à $\mathscr{P}(X)$ (l'ensemble des parties)

Axiome de l'infini

Axiome du choix

Compréhension

$$Y = \{x \in X : P(x) \text{ est vraie}\}$$

où P(x) est une « propriété » de x

page précédente : *l'axiome de compréhension* est certainement l'un de ceux que le « mathématicien normal » utilise souvent consciemment : pour ne donner qu'un exemple, on a eu l'occasion de mentionner le sous-ensemble de $\mathbf{R}$ formé des réels qui ont *la propriété* d'être un nombre algébrique. La notion de propriété reste encore imprécise chez Zermelo, on y reviendra.

Pour les « techniciens de l'axiomatique », l'axiome de compréhension n'est pas UN axiome unique, puisqu'il y a autant de formules à satisfaire qu'il y a de *propriétés* P possibles : ils parlent de *schéma d'axiome*. Il en sera de même pour le *schéma d'axiome de remplacement* qu'on verra plus loin. Le « remplacement » implique la « compréhension », et il est en fait beaucoup plus puissant.

L'*axiome de l'infini* est indépendant des autres axiomes : sans lui, on pourrait se trouver dans un monde mathématique qui ne serait constitué que d'ensembles finis. La formulation de Zermelo est la suivante : il existe un *ensemble* Z qui admet *l'ensemble vide* $\mathbf{0}$ en tant qu'*élément* et qui, chaque fois qu'il admet un *élément* x, admet aussi le *singleton* $\{x\}$ comme *élément*. Ainsi, Z *contient*

$$\mathbf{0},\ \{\mathbf{0}\},\ \{\{\mathbf{0}\}\},\ \{\{\{\mathbf{0}\}\}\},\ \{\{\{\{\mathbf{0}\}\}\}\},\ \ldots$$

Skolem et Fraenkel montrent que **:**

les axiomes de Zermelo -1908 ne suffisent pas pour formaliser toutes les notions introduites par Cantor !

Thoralf Skolem (1887–1963) (en 1923)

Abraham Adolf Fraenkel (1891–1965) (en 1922)

John von Neumann (1903–1957) (en 1923, 1925)

Skolem et Fraenkel introduisent « l'axiome de remplacement »

Le bizarre nuage de lettres et de signes page suivante contient tous les éléments pour écrire des « énoncés mathématiques corrects ». On peut en particulier donner une notion précise (la plus couramment acceptée, mais pas la seule) de ce que doit être une « propriété », telle qu'elle apparaît dans *l'axiome de compréhension* : une propriété P(x) est un *énoncé correct* contenant une *variable libre* x. Le langage des ensembles contient aussi des *constantes*, la première à mentionner étant l'ensemble vide $\varnothing$.

Les énoncés corrects de théorie des ensembles commencent avec les *énoncés atomiques* de la forme u = v, u ∈ v où u, v sont des variables ou des constantes. On peut grouper deux énoncés au moyen du connecteur logique OU, et appliquer NON à un énoncé. Pour décrire le langage on a besoin de se permettre des « macro-variables » notées E, F page suivante, qui *représentent* des énoncés écrits strictement avec les éléments du langage. Ainsi, si E(x, y) est un énoncé correct contenant deux variables libres, on peut former un énoncé F(x) à une seule variable libre x qui serait : $\exists\, y$ E(x, y).

On a aussi besoin de parenthèses, on laisse le lecteur y penser.

OU

NON

$\exists$

x

y

$=$

$\emptyset$

z , . . .

$\in$

$x = y$

$x = \emptyset$

(. . .)

$\underline{\mathrm{F}}(x, y) \equiv x \in y$

E F

$\underline{\mathrm{E}}(x, y, z)$ OU $\underline{\mathrm{F}}(x, y)$

$\underline{\mathrm{E}}(x, y, z)$

$\underline{\mathrm{E}}(x, y, y)$

NON $\underline{\mathrm{F}}(x, y)$

$\underline{\mathrm{R}}(x, y) \equiv \exists z\, \underline{\mathrm{E}}(x, y, z)$

$\exists x \text{ NON}\big(\exists y\, (y \in x)\big)$

il existe un ensemble vide

page précédente, on n'a pas écrit le connecteur logique ET ni le quantificateur universel. On pourra les définir en jouant avec une (double) négation :

$$\text{E ET F} \equiv \text{NON}((\text{NON E})\,\text{OU}\,(\text{NON F})),$$

$$\forall x\ \text{E}(x) \equiv \text{NON}(\exists x\ \text{NON E}(x)).$$

page suivante : une relation fonctionnelle $\text{R}(x, y)$ est un « énoncé correct » qui définit une sorte de « fonction » $\Phi_{\text{R}} : x \to y$, qui peut être définie sur la collection entière des « choses » de la théorie, comme la relation fonctionnelle $y = x$ par exemple (très peu utile pour l'axiome de remplacement !), ou sur une sous-collection. Elle ne peut pas en général s'exprimer à partir d'un *ensemble*, c'est-à-dire à partir d'une des « choses » de la théorie. Par exemple, la collection formée de *tous* les *couples* (x, x), où x est un *ensemble* quelconque, n'est pas la collection des *éléments* ***d'un*** *ensemble* ; mais cette collection représente la relation fonctionnelle d'égalité mentionnée plus haut —on définit le *couple* (paire *ordonnée*) par

$$(x, y) \equiv \{x, \{x, y\}\} \text{ —.}$$

Axiomes de la théorie ZF (pour Zermelo–Fraenkel) :

Extensionnalité

Axiome de la somme

Ensembles des parties

Axiome de l'infini

Remplacement

$$a \longrightarrow \Phi_{\mathrm{R}}(a)$$

R est un énoncé correct qui est tel que :

$$\big(\mathrm{R}(x,y) \;\; \mathrm{ET} \;\; \mathrm{R}(x,y')\big) \;\Rightarrow\; y = y'$$

il est plus agréable de noter :

$$\big(y = \Phi_{\mathrm{R}}(x)\big) \quad \equiv \quad \mathrm{R}(x,y)$$

on « a le droit » de poser :

$$b \;=\; \{y \,:\, \exists x \in a,\; y = \Phi_{\mathrm{R}}(x)\}$$

On ajoute en général à ZF l'axiome du choix AC pour obtenir la théorie ZFC

Contrairement au caractère *correct* d'un énoncé, qui est une pure question de syntaxe, la vérification du caractère *fonctionnel* d'une relation peut demander l'application des axiomes de la théorie des ensembles. Une relation fonctionnelle peut aussi dépendre de *paramètres* qui sont des *ensembles* de la théorie, par exemple dans un énoncé de la forme $\mathrm{R}(x, y, a, b)$ comme ci-dessous.

Le (schéma d')axiome de remplacement permet de « projeter » un *ensemble* a de la théorie sur un autre *ensemble* $b = \Phi_{\mathrm{R}}(a)$. Pour donner un exemple élémentaire, ce qui est l'*axiome de la paire* $\{a, b\}$ chez Zermelo résulte du remplacement dans ZF **:** on constate d'abord que l'*ensemble des parties* $\mathscr{P}(\varnothing)$ est égal au *singleton* $\{\varnothing\}$, puis on voit que $\mathscr{P}(\{\varnothing\})$ a exactement deux *éléments*, distincts, qui sont $\varnothing$ et $\{\varnothing\}$ **;** la relation qui va « envoyer » $\varnothing$ sur a et $\{\varnothing\}$ sur b est une relation fonctionnelle, et la *paire* $\{a, b\}$ est l'image de $\mathscr{P}(\{\varnothing\})$ par cette relation R,

$$\mathrm{R}(x, y, a, b) \equiv (x = \varnothing \text{ ET } y = a) \text{ OU } (x = \{\varnothing\} \text{ ET } y = b).$$

page suivante : von Neumann amène une vision nouvelle des ordinaux en théorie axiomatique. Le (dernier) point de vue de Cantor, qui envisage un ordinal comme une sorte de « classe d'équivalence » formée de tous les ensembles bien ordonnés de même type d'ordre, n'est pas très satisfaisant en théorie axiomatique. Von Neumann propose de définir chaque ordinal en choisissant un *ensemble* particulier possédant des propriétés spéciales, et d'utiliser la *relation d'appartenance* pour définir l'ordre entre les *ordinaux*.

Un *ordinal* α plus petit qu'un *ordinal* β doit donc être *élément* de β, et l'*appartenance* doit définir un ordre (strict) sur la famille des *éléments* de β. De plus, cet ordre doit être un bon ordre : tout *sous-ensemble non vide* de β doit avoir un plus petit *élément* pour cet ordre.

L'ensemble vide est un *ordinal* : n'ayant aucun *sous-ensemble non vide*, il n'a aucune condition à satisfaire pour être admis dans la famille ! Et clairement, c'est *le plus petit ordinal.*

Ordinaux selon von Neumann

Un *ordinal* est un ensemble α qui a les trois propriétés suivantes :

- la relation $x \in y$ est une relation d'ordre strict sur l'ensemble α, c-à-d

  si x, y, z sont des éléments de α, alors :
  $x \notin x$; $(x \in y \text{ et } y \in z) \Rightarrow x \in z$

- tout sous-ensemble non vide de α a un plus petit élément pour la relation $\in$,

- si $\beta \in \alpha$, alors $\beta \subset \alpha$.

α est *bien ordonné* pour la relation d'appartenance (en particulier, il est *totalement ordonné*)

$$\underbrace{\emptyset, \quad \{\emptyset\}, \quad \{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ \ldots}$$

les premiers (les plus petits) ordinaux
pour l'inclusion

page suivante **:** on peut décider que *l'entier* 0 est *le plus petit ordinal* $\varnothing$, et définir les *entiers naturels* comme étant les *ordinaux finis* successifs (le « mathématicien normal » aura quand même du mal à considérer que 0 est élément de 1 qui est élément de 2 qui est élément de **. . .**) **;** un *ordinal* est dit *fini* si tous ses prédécesseurs non nuls sont des *successeurs*, $i.e.$, des *ordinaux* de la forme $\alpha + 1$.

Dans le point de vue de von Neumann, on a une caractérisation très agréable du successeur $\alpha + 1$ d'un *ordinal* α **:** l'*ordinal* $\alpha + 1$ doit avoir α comme *élément* et doit aussi avoir pour *éléments* les prédécesseurs de α, qui sont tous les *éléments* de α, donc

$$\alpha \in \alpha + 1, \quad \alpha \subset \alpha + 1 \; ; \quad \alpha + 1 = \alpha \cup \{\alpha\}.$$

Un *ordinal limite* est un *ordinal non nul* β qui n'est pas de la forme $\alpha + 1$ **:** l'axiome de l'infini revient à dire qu'il existe un *ordinal limite*. À la suite de Cantor, on a noté ω le plus petit *ordinal limite,* qui est aussi le plus petit *ordinal infini* (infini **:=** non fini).

L'ordinal représente l'entier : $\emptyset, \quad \{\emptyset\}, \quad \{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ \ldots$

0 1 = {0} 2 = {0, 1} 3 = {0, 1, 2} . . .

$$0 \in 1 \in 2 \in 3 \in \ldots \in n \in n+1 \in \ldots$$

L'ordre des ordinaux :

$$\alpha \leqslant \beta \Leftrightarrow \alpha \subset \beta\,; \qquad \alpha < \beta \Leftrightarrow \alpha \in \beta$$

$$\alpha + 1 = \alpha \cup \{\alpha\}$$

$$\emptyset \in \omega\,;$$

$$\alpha \in \omega \Rightarrow \alpha + 1 \in \omega$$

L'ensemble ω contient tous les ordinaux finis ci-dessus ; on pourra comparer à l'ensemble infini Z de Zermelo

une forme de l'axiome de l'infini : il existe un objet ω tel que . . .

Dans ce point de vue, les *cardinaux* sont des ordinaux β particuliers, ceux qui ne peuvent être mis en bijection avec *aucun* ordinal $\alpha < \beta$.

Et après…

page suivante : les ordinaux de von Neumann et l'axiome de remplacement valident l'utilisation du *principe de récurrence ordinale*, qui en quelque sorte complète le principe de récurrence classique. Le principe s'applique aussi à la *construction* d'objets mathématiques par récurrence ordinale : c'est précisément ce qui est apparu dans la construction des dérivés successifs d'un ensemble P. Quand certains « mathématiciens normaux » s'emparent de cette méthode, ils ne pensent pas je crois aux ordinaux de von Neumann : ils raisonnent plutôt avec les *ordinaux naïfs* de Cantor, mais ils peuvent se sentir « à l'abri », l'abri que procure l'autorité qu'on veut bien reconnaître aux axiomes de la théorie des ensembles.

Le plus souvent, on commence par effectuer une *construction ordinale*, avant d'appliquer une récurrence ordinale pour prouver des propriétés de tous les éléments construits.

On va donner deux exemples de récurrence ordinale.

Récurrence ordinale

Pour que la propriété $\mathrm{P}(\alpha)$ dépendant d'un ordinal arbitraire α soit vraie pour tout ordinal α, il suffit que (faut et il)

- la propriété $\mathrm{P}(0)$ soit vraie
- et que pour tout ordinal α, on puisse prouver que :

  si $\mathrm{P}(\beta)$ *est vraie pour tout* $\beta < \alpha$, *alors* $\mathrm{P}(\alpha)$ *est vraie.*

Par exemple : par récurrence ordinale, on peut « dériver » indéfiniment :

$$\mathrm{F}^{(\alpha)} = \bigcap_{\beta<\alpha} (\mathrm{F}^{(\beta)})' \qquad \mathrm{F}^{(\alpha+1)} = (\mathrm{F}^{(\alpha)})'$$

cas particulier d'un « *ordinal successeur* »

D'abord, un exemple de récurrence ordinale que Cantor aurait pu donner. C'est une généralisation de son théorème d'unicité, qu'il connaissait sans doute mais qu'il n'a pas écrite (je crois) **:** si P est un ensemble fermé dénombrable (non vide) contenu dans un intervalle ouvert (u, v) et si la fonction $f(x)$, somme d'une série trigonométrique, est définie et nulle en tout point x de (u, v) qui n'est pas dans P, alors la fonction F associée à f par Riemann est affine dans le segment $[u, v]$.

On va d'abord justifier l'existence de la fonction F de Riemann. Puisque P est dénombrable, il n'est pas parfait, il a donc au moins un point isolé x situé dans (u, v) **:** il existe un intervalle ouvert I, contenu dans (u, v), qui contient x et ne contient *aucun autre point* de P. On peut alors trouver un intervalle ouvert non vide J, contenu dans I, qui ne contient aucun point de P. D'après l'hypothèse, la série trigonométrique $f(x)$ converge en tout point de J, donc ses coefficients tendent vers 0 d'après le lemme de Cantor.

(à suivre) →

On sait que le dérivé itéré de P finit par devenir vide ; il existe donc un premier ordinal $\beta = \gamma(\mathrm{P})$, dénombrable, tel que le dérivé $\mathrm{P}^{(\beta)}$ soit un ensemble fini et non vide : on va raisonner par récurrence ordinale sur $\gamma(\mathrm{P})$. Supposons l'énoncé vrai pour tout ordinal $\alpha < \beta$, supposons que $\gamma(\mathrm{P}) = \beta$ et supposons pour simplifier que $\mathrm{P}^{(\beta)} = \{x_0\}$ soit un singleton (plutôt qu'un ensemble fini quelconque) ; donnons $\varepsilon > 0$. La partie Q de P qui est située dans le segment $[u, x_0 - \varepsilon]$ est telle que $\mathrm{Q}^{(\beta)} = \varnothing$, par conséquent $\gamma(\mathrm{Q}) < \beta$; l'hypothèse de récurrence implique que F est affine sur $(u, x_0 - \varepsilon)$. Comme $\varepsilon > 0$ est arbitraire, on en déduit que F est affine sur $[u, x_0]$, et de même sur $[x_0, v]$; on conclut comme Cantor avait fait (avec l'aide des résultats de Riemann).

Il reste à finir la preuve du théorème d'unicité correspondant : si la fonction f est définie et nulle en tout point de $[0, 2\pi]$, sauf aux points du fermé P, on sait alors que F est affine sur $[0, 2\pi]$, donc constante puisqu'elle est 2π-périodique.

On déduit (comme Cantor avait fait) en calculant les coefficients de Fourier de F que tous les coefficients de f sont nuls, ce qui finit la preuve.

William Young a pu montrer en 1909, par une preuve totalement différente, que tout ensemble dénombrable*, fermé ou non, a la* propriété d'unicité*.*

Le deuxième exemple de récurrence ordinale va concerner les tribus (on dit aussi σ-*algèbres*). Une *tribu* de parties d'un ensemble X est une famille de sous-ensembles de X qui est « stable » par réunion dénombrable, passage au complémentaire et qui contient X.

(on dit qu'une famille F de parties de X est *stable par réunion dénombrable* lorsque pour toute suite (A_n) d'éléments de F, la réunion $\cup A_n$ de cette suite est aussi un élément de F ; la famille F est *stable par passage au complémentaire* si pour tout élément A de F, le complémentaire de A dans X est aussi élément de F)

William H. Young (1863–1942)

La *tribu borélienne* de $[0, 1]$ (par exemple ; on pourrait traiter des espaces métriques plus généraux) est la *plus petite* tribu qui contienne les *ouverts* de $[0, 1]$ (ou aussi bien, la plus petite qui contient les *fermés* puisqu'on peut passer au complémentaire dans une tribu). La justification « par l'extérieur » de son existence invoque « l'inter-section de toutes les tribus qui contiennent la famille des ouverts » — et il en existe au moins une, la tribu formée de tous les sous-ensembles de $[0, 1]$ —.

On va maintenant approcher cette tribu borélienne « de l'intérieur », par une méthode ordinale, en commençant par la famille des ouverts. Pour avoir une tribu contenant les ouverts, il faudra au moins considérer les intersections dénombrables d'ouverts : ces ensembles ne sont plus nécessairement ouverts. Mais ce n'est pas assez : il faut continuer, et longtemps . . .

On introduit des familles successives G_α, où $G_{\alpha+1}$ est formée des ensembles qui sont des *intersections* dénombrables d'éléments de G_α, ainsi que des ensembles qui sont des *réunions* dénombrables ; quand β est un ordinal limite, on prend pour G_β la réunion de toutes les familles G_α précédentes. Toutes ces familles sont nécessairement contenues dans la tribu borélienne (preuve par récurrence ordinale !), et on montrera que G_Ω est une tribu : c'est donc la tribu borélienne.

La première famille G_0 est formée des ensembles ouverts, elle est stable par réunion dénombrable (et même par réunion quelconque) ; on peut donc simplifier la description de G_1 en disant seulement que cette famille est formée des *intersections* dénombrables d'éléments de G_0 ; maintenant G_1 est stable par intersection dénombrable (une intersection dénombrable d'intersections dénombrables est encore une intersection dénombrable), on peut donc décrire G_2 comme la famille des *réunions* dénombrables d'éléments de G_1. Ainsi de suite, en alternant . . .

page suivante : la notation classique des premières familles déduites de la famille des ouverts n'utilise pas un indice ordinal fini, mais des indices successifs cumulés, δ pour les intersections dénombrables et σ pour les réunions dénombrables ; on commence donc par la famille G_δ des intersections dénombrables d'ouverts, puis on a

$$G_{\delta\sigma} = (G_\delta)_\sigma, \quad G_{\delta\sigma\delta} = (G_{\delta\sigma})_\delta, \ldots$$

Dans le cas de $[0, 1]$ on a une propriété particulière (vraie pour beaucoup d'espaces métriques, mais pas tous) : un fermé y est une intersection *dénombrable* d'ouverts. La famille F des fermés est donc contenue dans la première famille G_δ construite à partir des ouverts. Si on construit une *échelle* de familles d'ensembles à partir de la famille des fermés, comme on a fait pour les ouverts, elle sera comparable à l'échelle construite avec les ouverts, à un rang près. Il y aura donc égalité des deux constructions au rang ω.

$$\mathcal{G} \subset \mathcal{G}_\delta \subset \mathcal{G}_{\delta\sigma} \subset \mathcal{G}_{\delta\sigma\delta} \subset \ \dots \ \subset \mathcal{G}_\omega$$

$$\mathcal{F} \subset \mathcal{F}_\sigma \subset \mathcal{F}_{\sigma\delta} \subset \mathcal{F}_{\sigma\delta\sigma} \subset \ \dots \ \subset \mathcal{F}_\omega$$

$$\mathcal{G}_\omega = \mathcal{F}_\omega \subset \mathcal{G}_{\omega+1} \subset \ \dots \ \subset \mathcal{G}_\Omega = \mathcal{B}$$

On montre par récurrence que le *complémentaire* d'un ensemble qui est élément d'une des familles construites à partir des ouverts se trouve aussi, au cran suivant, dans une famille issue des fermés (et inversement). On en déduit que $G_\omega = F_\omega$ est stable par passage au complémentaire.

page précédente : à propos de la construction par récurrence ordinale de la tribu borélienne, disons de $[0, 1]$, il faut encore citer Baire : aucune des familles G_α construites à une étape ordinale *dénombrable* n'est une tribu, *il faut* aller jusqu'à Ω. Le fait qu'on puisse arrêter la construction à Ω provient de l'argument suivant, qu'on n'a pas pu utiliser avant d'arriver à Ω : si des ensembles boréliens A_n sont dans G_Ω, ils sont apparus à des étapes respectives α_n de la construction ordinale, où chaque α_n est un ordinal $< \Omega$, donc un ordinal dénombrable ; on sait qu'il existe un ordinal dénombrable $\beta < \Omega$ qui est plus grand que tous les α_n ; les ensembles A_n étaient donc déjà tous présents à l'étape β de la construction, et on a pu leur appliquer une opération de réunion ou d'intersection dénombrable à l'étape $\beta+1$.

(à suivre) →

La réunion et l'intersection de la suite d'ensembles (A_n) de G_Ω sont donc présentes dans la classe $G_{\beta+1}$, et donc aussi dans G_Ω.

Par ailleurs, on pourra prouver par récurrence ordinale que :

— pour tout ordinal $\beta \geq \omega$, la classe G_β est stable par passage au complémentaire ;

en effet, on a dit que G_ω est stable par passage au complémentaire ; si A est un élément de G_β avec $\omega < \beta$, cet ensemble est, par exemple, *réunion* dénombrable d'ensembles A_n provenant d'une famille G_α avec $\omega \leq \alpha < \beta$; alors, d'après l'hypothèse de récurrence (ordinale), les complémentaires B_n des A_n sont aussi dans G_α, donc l'intersection dénombrable des B_n, qui est le complémentaire de A, est dans $G_{\alpha+1} \subset G_\beta$.

Il en résulte que G_Ω, stable par réunion dénombrable et passage au complémentaire, est une tribu. cqfd.

Certains lemmes de tribu engendrée ont une preuve plus intuitive par récurrence transfinie. Par exemple :

Si $\mathcal{A}$ est une algèbre de parties de X, qui est contenue dans une classe $\mathcal{M}$ de parties stable par limite des suites croissantes et des suites décroissantes, alors $\mathcal{M}$ contient la tribu engendrée par $\mathcal{A}$.

L'énoncé précédent constitue la « partie infinie » du classique *lemme des classes monotones*, la « partie finie » consisterait à passer d'un π-système à une algèbre ; la preuve habituelle saute par dessus cette étape (une *algèbre de parties* contient l'ensemble vide, et elle est stable par réunion finie et passage au complémentaire).

AC et lemme de Zorn (1935) : plutôt que d'appliquer directement l'axiome du choix, on utilise dans plusieurs branches des mathématiques le *lemme de Zorn* (qui est équivalent à l'axiome du choix), par exemple pour justifier l'existence d'*idéaux maximaux* dans certaines algèbres, ou pour montrer le *théorème de Hahn–Banach* en Analyse Fonctionnelle. On pourrait tout aussi bien prouver ces résultats par récurrence ordinale, mais la classe des mathématiciens qui acceptent d'utiliser des ordinaux est *strictement* (et nettement) incluse dans la classe des mathématiciens.

Max Zorn (1906–1993)

Modèles de la théorie des ensembles

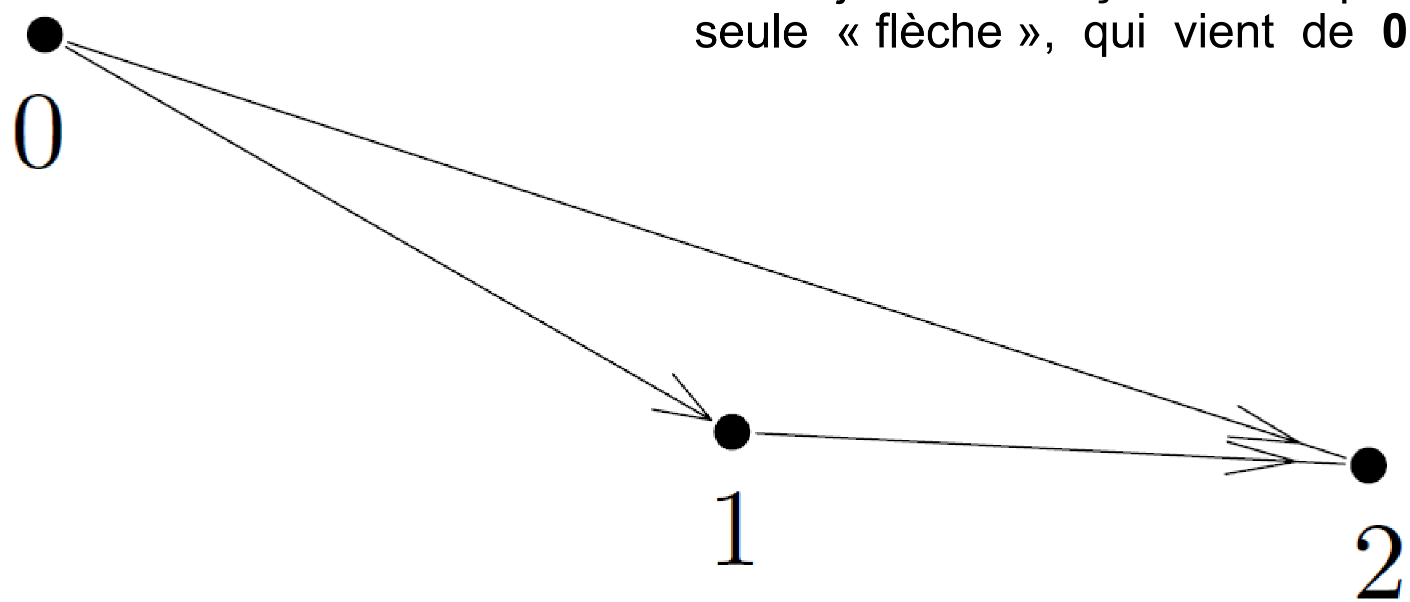
Un *modèle de la théorie des ensembles* suit les principes posés par Hilbert puis par Zermelo **:** on y travaille sur une collection d'objets, qu'on pourra décider d'appeler (avec précaution) « ensembles », comme Hilbert décidait d'appeler points, droites certaines des « choses » qu'il considérait.

On peut penser qu'un modèle est un *ensemble au sens naïf*, dont les éléments, toujours au sens naïf, seront les objets, les « choses » qui seront à l'étude dans la théorie axiomatique.

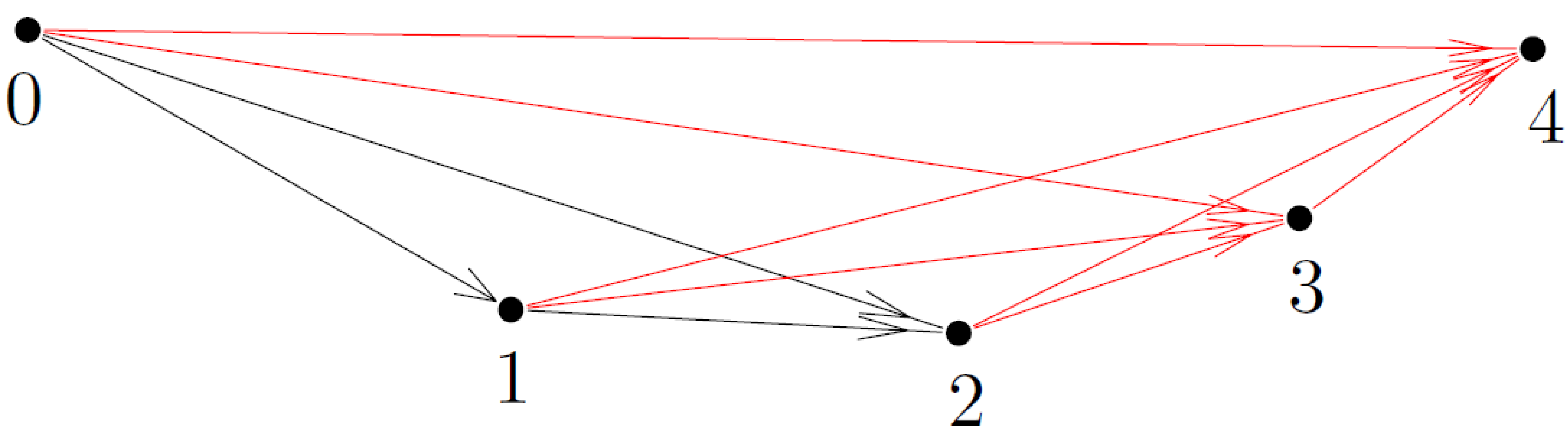
On peut aussi penser qu'un modèle de la théorie des ensembles est un *graphe orienté,* les sommets étant les objets de la collection, c'est-à-dire les « ensembles axiomatiques », et les arêtes (orientées) du graphe traduisant la relation d'appartenance.

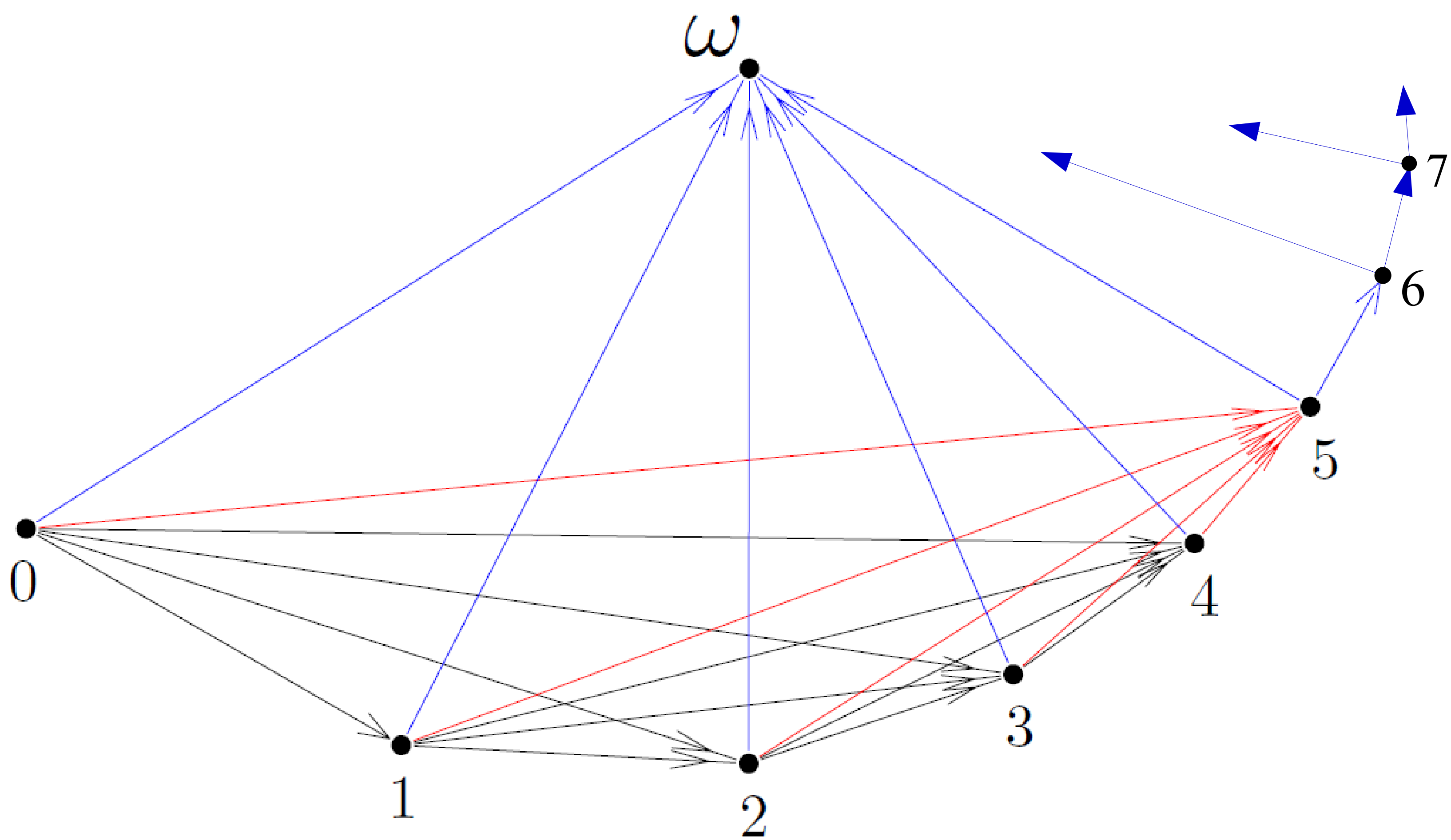
d'après les axiomes de ZF, il **doit** exister un objet b du modèle tel qu'on n'ait *jamais* a ∈ b **;** cet objet b noté **0** *représente* l'ensemble vide **:** aucune flèche ne pointe vers cet objet.

Il doit aussi exister un objet **1** qui représente l'ensemble des parties de l'ensemble vide **:** le seul élément de cet ensemble de parties est l'ensemble vide **;** cet objet **1** ne reçoit donc qu'une seule « flèche », qui vient de **0**.



avec deux entiers de plus, 3 et 4 :





C'est peut-être joli (?), mais ça ne va pas nous mener bien loin . . .

parce que :

on ne peut pas « fabriquer » un modèle de la théorie des ensembles ZF à partir des propres axiomes de ZF (cela résulte du *deuxième théorème d'incomplétude* de Gödel).

Dans un modèle de ZF, on définit des ordinaux (les *ordinaux* du modèle) en suivant la définition de von Neumann appliquée avec la notion d'appartenance propre au modèle.

L'axiome de l'infini implique l'existence d'un objet ω qui est le plus petit ordinal infini du modèle.

(fini au sens du modèle **!**)

Un ordinal α du modèle est dit ***fini*** si tout ordinal non nul β inférieur ou égal à α est de la forme $\beta = \gamma + 1$ (β est un *successeur*)

Pour ne pas mélanger les différents niveaux de langage, on va continuer d'écrire *ensembles* (en italique bleu) pour les objets x du modèle M, et écrire que l'*ensemble* y est un *élément* de x lorsqu'on a $y \varepsilon x$ au sens du modèle. Un *ensemble* α est un *ordinal* s'il vérifie la définition de von Neumann, au sens du modèle. L'*ensemble vide* du modèle est le « plus petit » *ordinal*.

Les *ordinaux* jouent un rôle très important en théorie axiomatique des ensembles. La *collection des ordinaux* du modèle M, qui est le sous-ensemble (au sens intuitif) de M formé de tous les *ordinaux*, « n'est pas » un *ensemble* **:** il n'existe pas d'*ensemble* x tel que tous les *ordinaux* soient *éléments* de x.

L'axiome de remplacement valide l'utilisation dans le modèle de constructions *par récurrence sur la collection des ordinaux*. On va donner un exemple dans les deux pages suivantes.

Une démarche qui a permis des avancées importantes est la construction de nouveaux modèles *à partir d'un* modèle dont on a supposé l'existence **;** page suivante, on mentionne un exemple fondé sur la récurrence ordinale **:** partant de l'*ensemble vide* du modèle, on va définir des *ensembles* successifs V_α, indexés par les *ordinaux*, en itérant l'opération qui consiste à prendre l'*ensemble des parties* d'un *ensemble* quand on passe de l'*ordinal* α à $\alpha+1$, et en prenant la *réunion* des *ensembles* précédents quand α est un *ordinal* limite.

En envisageant cette opération *pour tous les ordinaux*, on introduit une collection V, sous-ensemble (au sens naïf) du modèle initial M, et on peut montrer que V satisfait les axiomes de la théorie des ensembles **:** c'est un nouveau modèle. Ce modèle possède un très gros avantage **:** on a une « description » de ses objets, et on peut définir le « moment » où un *ensemble* est apparu dans la collection V. On peut montrer *par récurrence ordinale* que le modèle V possède certaines propriétés additionnelles, que l'on peut donc ajouter aux axiomes de la théorie des ensembles sans amener de contradiction (comme toujours **:** s'il n'y en avait pas déjà, ce qu'on a supposé en admettant l'existence d'un premier modèle M).

$$V_0 = \emptyset\,; \qquad V_{\alpha+1} = \mathcal{P}(V_\alpha)$$

$$V_\alpha = \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta)$$

La collection V formée de tous les *éléments* des V_α est encore un modèle de la théorie des ensembles. Ce modèle V satisfait *l'axiome de fondation*, qui implique en particulier que :

$$x \notin x$$

(dans ZF, on n'avait pas jugé indispensable d'interdire qu'un ensemble soit élément de lui-même).

page suivante : quelques résultats obtenus par la logique mathématique au 20e siècle. Le théorème de Löwenheim—Skolem peut paraître terriblement troublant de prime abord. Il se rattache un peu à la remarque qu'on a faite sur la dénombrabilité des systèmes possibles pour décrire les ordinaux ; ici, il s'agira de la dénombrabilité de l'ensemble des *énoncés* qu'on peut écrire.

On pourrait dire : *le modèle* (de Löwenheim-Skolem) *est dénombrable, mais le modèle ne le sait pas* ! « À l'intérieur » du modèle, il n'y a aucun *ensemble* qui serait le graphe d'une bijection entre *l'ensemble des entiers* et le modèle. Aussi, des objets que l'observateur extérieur voit dénombrables ne sont pas *dénombrables au sens du modèle*, parce que le modèle ne contient pas la *bijection* qu'il faudrait (c'est-à-dire, il ne contient aucun *ensemble* de *couples* qui serait le graphe d'une telle bijection).

Löwenheim–Skolem : si on dispose d'un modèle M de ZF, on peut construire un nouveau modèle M' qui est dénombrable.

Si on dispose d'un modèle M de ZF,
on peut construire un modèle M' qui contient des entiers non standard (des entiers infiniment grands).

Si on dispose d'un modèle M de ZF, on peut construire un modèle M' qui satisfait l'axiome du choix et l'hypothèse du continu (Gödel, 1940 ; c'est la preuve de la *consistance relative* de AC et HC) ;

on peut construire un modèle M' dans lequel l'hypothèse du continu est fausse (Cohen, 1963 ; indépendance de HC).

Leopold Löwenheim (1878–1957)

Kurt Gödel (1906–1978)

Paul Cohen (1934–2007)

page précédente : Cohen a montré que *l'hypothèse du continu* (celle que Mittag-Leffler voulait voir résolue par Cantor avant de publier son article de 1885 !) est *indépendante des axiomes de ZF* ; rappelons que Cantor avait formulé, et essayé sans succès de prouver l'hypothèse selon laquelle le segment $[0, 1]$ a la même cardinalité que l'ensemble des ordinaux $< \Omega$. On sait maintenant qu'on peut aussi bien supposer qu'elle est vraie, ou qu'elle est fausse ! Gödel avait prouvé qu'on pouvait supposer qu'elle est vraie, et Cohen a construit, par la délicate technique du *forcing* —qui est son invention—, un modèle où elle est fausse.

De nombreux résultats ont été obtenus par cette méthode de forcing, par exemple un théorème de Solovay qui pourra faire plaisir aux étudiants de théorie de la mesure : si on affaiblit (beaucoup) l'axiome du choix et qu'on ajoute un nouvel axiome (l'existence d'un *cardinal inaccessible*), on peut supposer que *tous les sous-ensembles de la droite réelle sont Lebesgue-mesurables* !

Robert Solovay, né en 1938

Page suivante, quelques références en rapport plus ou moins étroit avec le contenu du présent document. Il y a d'abord certains des articles de Cantor qui ont été mentionnés ici, ainsi que les textes qu'on a mentionnés de Riemann, Heine, Bendixson et Hilbert. De plus :

— le livre de Baire de 1905, qui explique une très grande partie de ce qui se trouve dans ces Notes ;

— le volume des Œuvres complètes de Cantor, avec des commentaires de Zermelo, le tout en allemand ;

— j'ai volé chez Krivine le peu que j'ai cru comprendre de la théorie axiomatique des ensembles ;

— le livre de Purkert et Ilgauds, malheureusement pour nous en allemand, est une mine d'informations sur Cantor ;

— le livre de van Heijenoort contient un grand nombre de textes des débuts de l'axiomatique : Frege, Peano, von Neumann entre autres ; les textes originaux y sont traduits en anglais.

René Baire, Leçons sur les fonctions discontinues, professées au Collège de France, Gauthier-Villars, Paris, 1905.

Ivar Bendixson, Quelques théorèmes de la théorie des ensembles de points, Acta Math. 2 (1883), 415-429.

Georg Cantor, en traduction française : Extension d'un théorème de la théorie des séries trigonométriques, Acta Math. 2 (1883), 336-348.

Georg Cantor, version française amendée : Fondements d'une théorie générale des ensembles, Acta Math. 2 (1883), 381-408.

Georg Cantor, en traduction française : Sur les fondements de la théorie des ensembles transfinis, Mémoires de la Société des Sciences physiques et natur. de Bordeaux, T. III (1899), 343-437.

Georg Cantor, Gesammelte Abhandlungen mathematischen und philosophischen Inhalts, éd. par Ernst Zermelo, 1932.

Eduard Heine, Die Elemente der Functionenlehre, J. reine angew. Math. 74 (1872), 172-188.

David Hilbert, en traduction française : Les principes fondamentaux de la géométrie, Gauthier-Villars: Paris, 1900

Jean-Louis Krivine, Théorie des ensembles, Cassini, 1998.

Walter Purkert et Hans Joachim Ilgauds, Georg Cantor: 1845-1918, Vita Mathematica 1, Birkhäuser Verlag, 1987.

Bernhard Riemann, en traduction française : Sur la possibilité de représenter une fonction par une série trigonométrique, Bulletin des sciences math. et astronomiques 5 (1873), 20-48, 79-96.

Jean van Heijenoort (ed.), From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931 (3rd ed.), Harvard University Press.

https://webusers.imj-prg.fr/~bernard.maurey/