Georg Cantor, de l'Analyse à l'hypothèse du continu

(d'après un exposé du 24 mars 2021)

Bernard Maurey

Collaborateur bénévole, IMJ–PRG, Sorbonne Université

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

Si par hasard vous êtes tombé sur cet objet,

vous aurez forcément repéré des erreurs !

Vos remarques seront les bienvenues

à l'adresse anti-robot (?) :

bernard . maurey

@ imj-prg . *fr*

Ce texte a été écrit à l'occasion d'une invitation à donner une conférence pour des étudiants de Licence de l'Université Gustave–Eiffel à Champs-sur-Marne, dans le cadre d'un séminaire *Histoire et Philosophie des Sciences,* dont le thème cette année 2020–2021 était **:** « l'infini en mathématiques ». À côté de thèmes que j'avais déjà fréquentés, j'ai découvert en préparant l'exposé de nombreux éléments que je ne connaissais pas. J'en remercie donc chaleureusement l'organisateur Marco Cannone, qui m'a permis de sortir d'une léthargie confinée et d'une inactivité tranquille.

Sommaire

(les liens vert sombre
sont « cliquables », ici
et dans la suite)

(suite)

Pour illustrer le thème de *l'infini en mathématiques*, j'ai choisi un sujet plutôt « facile », qui va un peu de soi pour un mathématicien, un sujet qui se prêterait bien aussi à l'écriture de romans pour le grand public : la vie et l'œuvre de Cantor.

Je dois tout de suite préciser que je ne suis pas historien : dans la suite, je mentionnerai des faits qui j'espère sont exacts, mais je crains de ne pas avoir les « outils méthodologiques » pour les mettre en perspective d'une manière historiquement satisfaisante...

1870,

et autour

Sans être historien, j'aime bien raconter des histoires ; celle que je vais essayer de traiter se déroule sur une petite centaine d'années, de la fin des années 1860 jusqu'aux années 1960, plus précisément, de la thèse de Cantor en 1867 jusqu'aux travaux de Paul Cohen vers 1963.

Page suivante, un beau portrait du jeune Cantor, et une magnifique « devise » extraite d'un de ses articles.

Georg Cantor, vers 1870

(1845–1918)

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

(1883)

La devise —on pourrait dire le slogan— de la page précédente signifie selon moi, compte tenu de mon expertise très relative en langue allemande **:**

« *L'essence des mathématiques réside dans leur liberté.* »

Cette phrase extraite d'un article de 1883 fait certainement écho aux oppositions à ses idées que Cantor a dû affronter dans les années 1880 **;** cette liberté en mathématiques, c'est celle qu'il entend exercer dans ses propres travaux.

En 1867, âgé de 22 ans, Cantor passe sa thèse de doctorat à l'Université de Berlin, en latin comme il se doit. Le thème de la thèse est de nature algébrique, le directeur de thèse principal était Ernst Kummer. Le titre de la thèse est imprimé en haut de la page de couverture de cette thèse.

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 (1867)
H. L. Q. S.
PUBLICE DEFENDET
AUCTOR
GEORGIUS CANTOR
PETROPOLITANUS.

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

BEROLINI
TYPIS CAROLI SCHULTZII
KOMMANDANTEN-STRASSE.

Université de Berlin

Les professeurs de maths :

Ernst Kummer (1810–1893)

Leopold Kronecker (1823–1891)

Karl Weierstrass (1815–1897)

quelques PhDs
sous leurs directions :

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

la page de couverture
de la thèse de Cantor

Vers 1870, Berlin est la plus importante université scientifique allemande. Les trois professeurs de mathématiques à Berlin ont formé, au long des années, beaucoup de mathématiciens dont quelques-uns (et une) sont mentionnés page précédente.

La page de couverture de la thèse de « Georgius Cantor » ne mentionne pas les professeurs, mais cite les « adversarii » Emil Lampe, Max Henoch, Max Simon ; ce sont de jeunes docteurs qui, je crois, étaient chargés de « cuisiner » le candidat pendant la soutenance de la thèse. Cantor est qualifié de « petropolitanus » : il est né à Saint-Pétersbourg en Russie.

On ne verra pas Saint-Pétersbourg sur la carte de la page suivante (volée sur internet), mais on verra que la ville de Berlin est bien située au centre du *Royaume de Prusse* (la zone en rose), on constatera la très grande étendue de la Prusse vers l'Est, en passant par Königsberg (Kaliningrad de nos jours) et jusqu'à Memel, aujourd'hui Klaipėda en Lituanie.

Königsberg (Kaliningrad)

Memel (de nos jours : Klaipėda en Lituanie)

Berlin

À l'Ouest, sur cette carte, l'Alsace et la Lorraine viennent d'être prises à la France. Que s'est-il passé ? À l'été 1870, après plusieurs différends avec les Prussiens, Napoléon III déclare la guerre à la Prusse. Mauvaise idée ; il est battu, début septembre 70 : c'est la défaite de Sedan, puis l'abdication de l'Empereur Napoléon III et la proclamation le 4 septembre d'une nouvelle république française, la *Troisième République* (après celles de 1792 et 1848).

Les Prussiens et leurs amis allemands envahissent rapidement la France et font le siège de Paris. Le roi de Prusse Guillaume et le chancelier Bismarck sont passés en septembre 70 par le château de Ferrières, à dix kilomètres (à vol d'oiseau) de Champs-sur-Marne.

Page suivante, une gravure de l'époque, qui rappelle la prise de Bry-sur-Marne : pour ceux qui connaissent Bry, remarquer l'église, vue à peu près du pont de Bry ; voir les casques à pointe des soldats Prussiens. Il y a eu des combats sévères à Villiers, à Champigny, pour ne mentionner que deux des très nombreux lieux de combats autour de Paris.

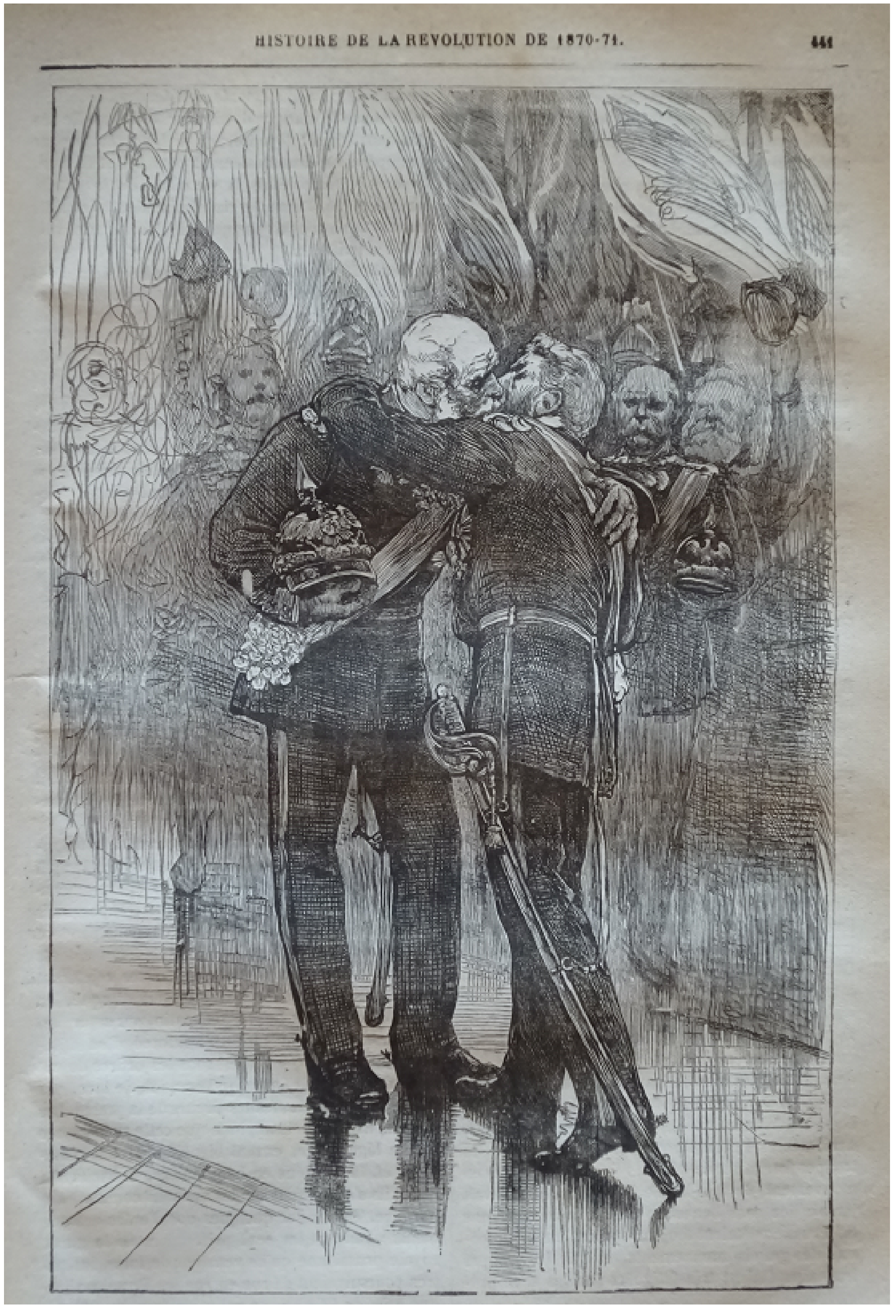
Le siège de Paris. — Prise de Brie-sur-Marne, le 30 novembre 1870.

Le siège de Paris – Prise de Bry-sur-Marne, le 30 novembre 1870

Pour bien humilier les Français, le roi de Prusse Guillaume se fait proclamer Empereur d'Allemagne dans la Galerie des Glaces à Versailles. Page suivante, le chancelier Bismarck et le nouvel Empereur à Versailles.

La Commune de Paris survient peu après, au printemps 71.

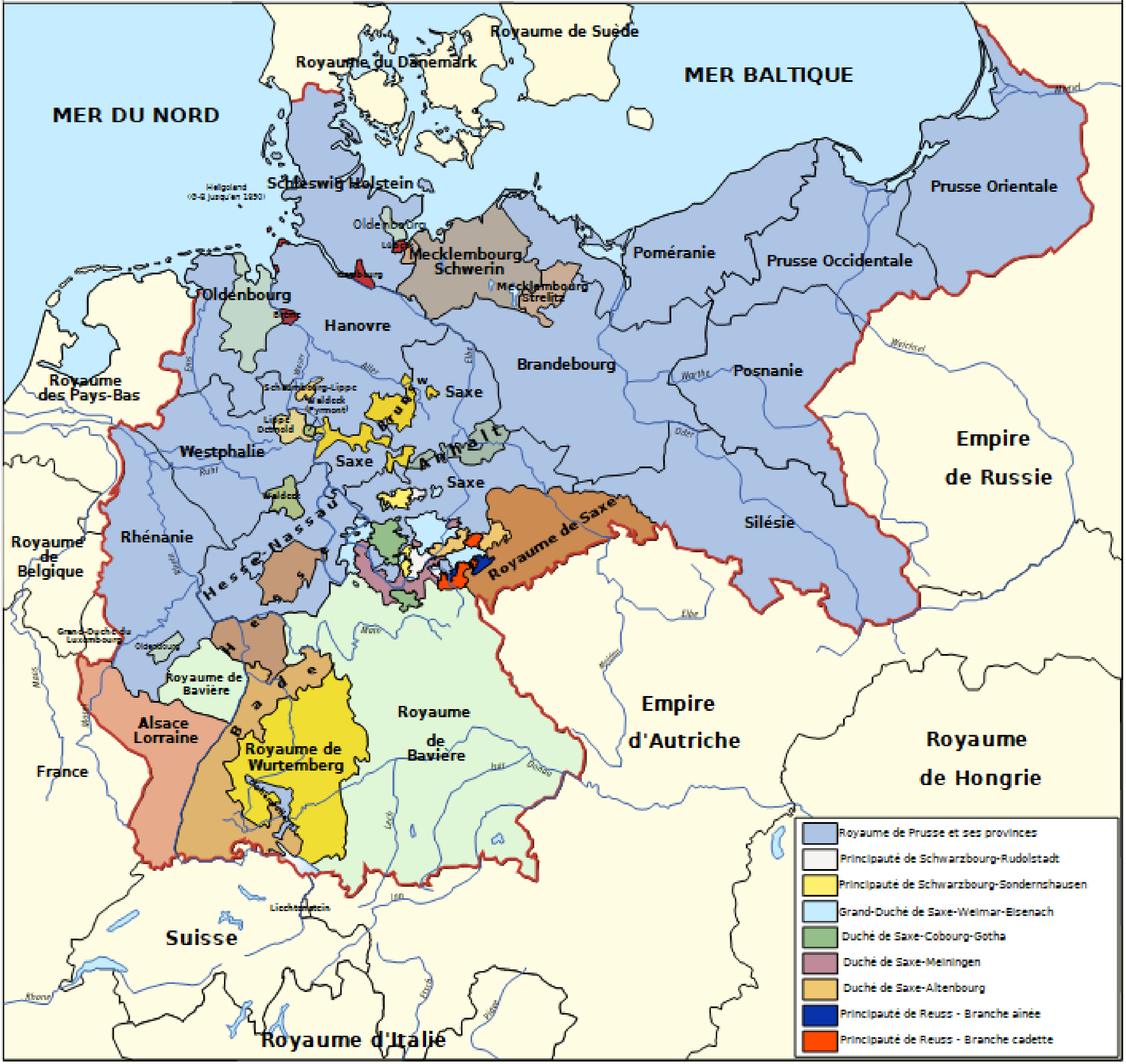
Deux pages plus loin, une autre carte de l'Empire allemand. On pourra remarquer l'absence de la Pologne : Varsovie se trouve alors dans l'Empire russe, et une très grande partie de la Pologne actuelle se trouve dans la Prusse de l'époque.



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

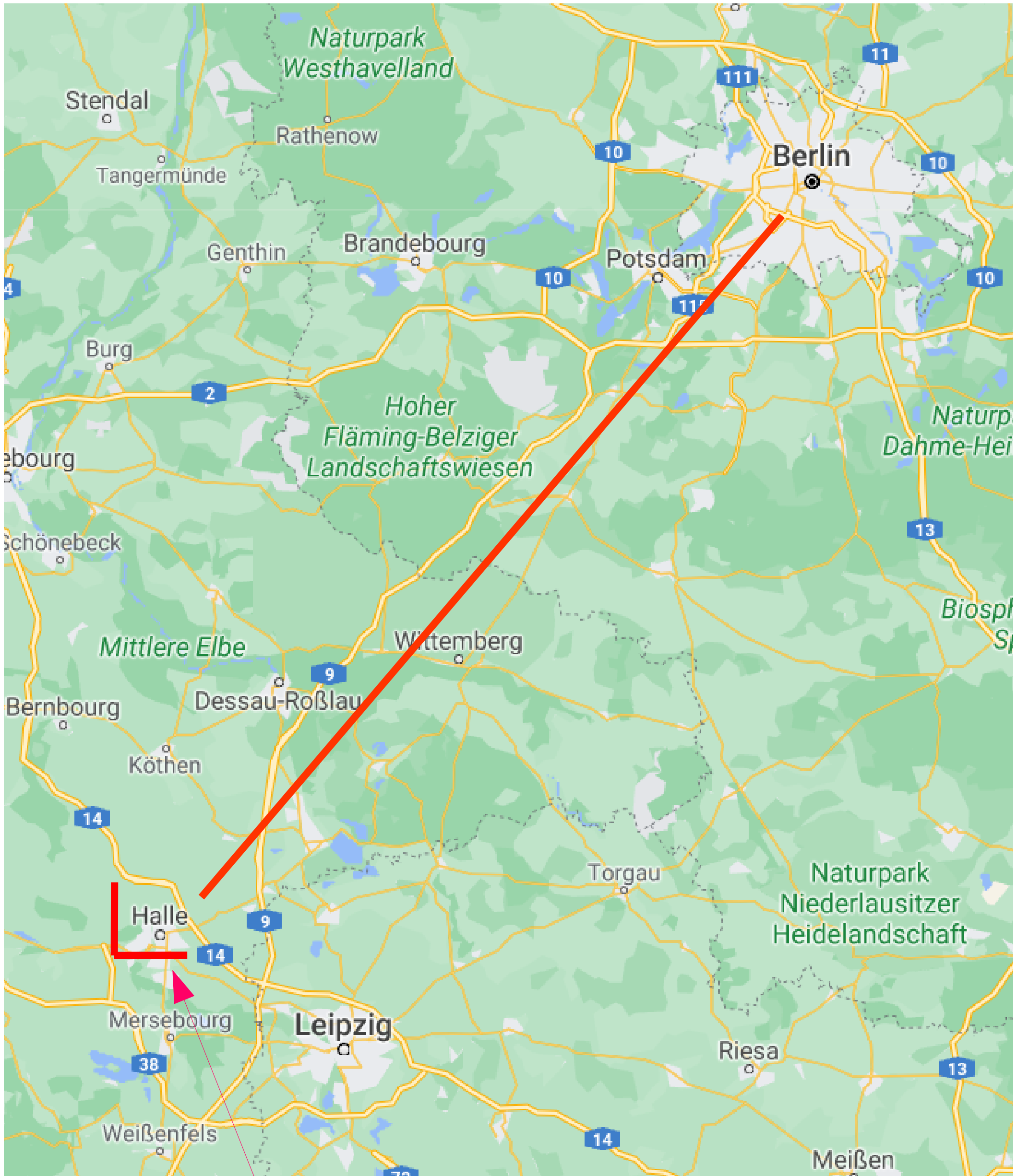
le 18 janvier 1871

L'EMPIRE ALLEMAND
1871 - 1918



Cantor présente son Habilitation à l'Université de Halle au printemps 1869, un travail algébrique dans le prolongement de sa thèse de doctorat. La ville de Halle est située à environ 150 km de Berlin à vol d'oiseau. Halle à cette époque n'est « pas top » en sciences, mais Eduard Heine, le professeur titulaire à Halle, est un interlocuteur valable (pour quelques années seulement **:** né en 1821, il n'a plus que douze ans à vivre). Cantor obtient un poste de « Privatdozent », enseignant non titulaire à Halle.

Plus tard, Cantor essaiera vainement d'obtenir un poste à l'Université bien plus prestigieuse de Berlin **:** il finira sa carrière à Halle (retraite en 1913, ses derniers cours ont eu lieu pendant l'année universitaire 1910/1911).



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,
il y accueille le jeune Cantor

Nombres réels

Un des sujets importants pour l'Analyse de l'époque est de mettre au point un traitement rigoureux de notions ou de résultats qui étaient traités jusque-là par l'intuition géométrique ou par l'appel à l'idée de mouvement, par exemple le théorème des valeurs intermédiaires ou le théorème des accroissements finis.

Un ingrédient nécessaire à cette mise au point est d'établir une notion solide de nombre réel : pour *démontrer* qu'il existe un nombre réel x tel que $f(x) = 0$ lorsque la fonction continue f satisfait $f(a) > 0$ et $f(b) < 0$, il faut bien pour commencer avoir une idée claire de ce qu'on entend par « un nombre réel x ».

Dans ses cours à Berlin, Weierstrass prouve les propriétés des fonctions continues sur un segment ; il a sa méthode, qui n'est peut-être pas la plus satisfaisante, pour introduire les nombres réels. Plusieurs autres méthodes apparaissent vers 1870 pour définir les réels : avec des suites de Cauchy (Méray en France, Cantor), ou par les *coupures* de Dedekind.

La naissance des nombres réels pose une question historique qui dépasse de beaucoup mes compétences. On savait depuis très longtemps que $\sqrt{2}$ n'est pas rationnel ; dans l'Antiquité, cette quantité était plutôt vue comme faisant partie du monde de la géométrie ; on pouvait cependant chercher à la *mesurer*.

D'ailleurs, on savait au 17^{e} siècle calculer les racines carrées avec autant de décimales qu'on voulait ; pour élaborer la théorie des logarithmes, Henry Briggs, dans les années 1620, calcule avec 30 décimales des racines carrées successives de 10 (partant de la valeur 10, il itère plus de cinquante fois l'opération de prise de racine carrée) : il me semble qu'il traite alors de pures quantités numériques sans support géométrique. N'avait-il pas déjà une certaine compréhension des nombres réels ?

Henry Briggs (1561–1630)

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

Eduard Heine (qui a 24 ans de plus que Cantor) a beaucoup travaillé sur les fonctions spéciales, les séries hyper-géométriques, les harmoniques sphériques, les fonctions elliptiques **;** son article sur la *théorie des fonctions* fait plutôt figure d'exception dans son œuvre mathématique.

Dans cet article, Heine présente une théorie des nombres réels et il démontre les propriétés des fonctions continues de variable réelle **;** l'article s'achève avec le résultat qu'on a souvent appelé « *théorème de Heine* », la propriété de continuité uniforme des fonctions réelles continues définies sur un segment.

Cependant, Heine insiste beaucoup dès la première page de son article sur le rôle de Weierstrass, auquel il attribue tout le mérite de la découverte des résultats sur les fonctions continues, et il mentionne Cantor pour la méthode d'introduction des nombres réels par les suites de Cauchy (c'est le nom que *nous* donnons à ces suites **;** Cauchy n'est pas mentionné par Heine, ni d'ailleurs par Cantor).

Page suivante, « côté-ligné » en rouge, le passage où Heine dit ce qu'il doit à Cantor.

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)

Traduction partielle :

« Je dois adresser des remerciements particuliers à M. Cantor, de Halle, pour ses communications verbales qui ont exercé une influence notable sur la conception de mon travail : je lui dois en effet l'idée d'introduire les nombres généralisés au moyen de certaines suites particulières, qui seront appelés ici *Zahlenreihen* [suites de nombres] ».

dans tout l'article, ce mot allemand garde le sens très particulier, très précis de : *suite de Cauchy définissant un nombre réel* ; le traduire par un vague « suite de nombres » ne me semble donc pas approprié, j'ai préféré garder le terme allemand tel quel.

En même temps, Cantor rédige sa propre version de la construction des nombres réels, qui constitue une partie de son article : « *Ueber die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen* ».

À cette époque, il était habituel que l'auteur d'un article indique à la fin de celui-ci le lieu où le travail avait été effectué et la date à laquelle il avait été achevé. Cantor indique « Halle, le 8 Nov. 1871 » et Heine pour sa part avait indiqué « Halle, en Octobre 1871 ».

Une version française de cet article de Cantor a paru plus de dix ans après dans Acta Mathematica **2**, p. 336–348 (1883) **:**
Extension d'un théorème de la théorie des séries trigonométriques. Je ne la suivrai pas tout à fait pour les passages suivants (où j'ai laissé quelques trous) **:**

«

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.

»

[*il faut comprendre ici que c'est la valeur absolue de la différence qui est plus petite que* ε]

«

Je traduis cette propriété par l'expression : « la suite (a_n) possède une limite déterminée b » ; cette expression n'a pas d'autre sens que celui de traduire la propriété précédente.

[...]

Si une autre suite (a'_n) possède une limite déterminée b', on sera dans l'un des trois cas suivants, qui s'excluent mutuellement : ou bien $a_n - a'_n$ devient infiniment petit quand n croît, ou bien $a_n - a'_n$ reste plus grand qu'une quantité positive (rationnelle) ε à partir d'un certain rang n, ou bien $a_n - a'_n$ reste plus petit qu'une quantité négative (rationnelle) $-\varepsilon$ à partir d'un certain rang n.

Dans le premier cas, je pose $b = b'$, je pose $b > b'$ dans le second cas et $b < b'$ dans le troisième.

[...]

On désigne par B la collection de toutes les *Zahlengrösse* b.

»

Pour Cantor, une « limite déterminée » b, élément du domaine B, est un *Zeichen*, un « signe », on pourrait dire aussi un « repère », une « marque » ou un « symbole ». Avoir *décidé de poser* $b = b'$ page précédente revient (pour nous, modernes) à considérer que B est l'ensemble des classes d'équivalence de suites de Cauchy de nombres rationnels. Il n'y a cependant pas ici de point de vue ensembliste, on semble envisager d'introduire autant de « symboles » b qu'il y aura de nombres réels.

Si a est un nombre rationnel, Cantor explique dans quel cas il *dira* que $b = a$, ou bien $b > a$, ou bien $b < a$. Puisqu'on peut comparer b à un petit rationnel ε donné, et en anticipant sur la possibilité de soustraire deux éléments de B (qui vient juste après dans l'article), on peut définir la notion de limite pour les suites d'éléments de B ; on voit ainsi que si b est *déterminé* par la suite de rationnels (a_n), cet élément b est bien la *limite,* en ce nouveau sens, de cette suite (a_n) qui a servi à le définir.

Cantor définit ensuite les quatre opérations élémentaires (addition, soustraction, multiplication, division) pour les éléments de B (et de A, assimilé à une partie de B), et considère les fonctions d'un nombre fini d'éléments de B obtenues par un nombre fini d'opérations élémentaires (polynômes par exemple).

Cantor a construit le « domaine » B à partir du domaine A des nombres rationnels. Il construit de façon analogue un domaine C à partir de B, et il indique **:**

« Alors que les domaines B et A se comportent entre eux de façon que chaque a puisse être déclaré égal à un b, mais qu'inversement un b ne puisse pas toujours être égal à un a, il apparaît que chaque c peut être déclaré égal à un b, et inversement chaque b à un c. »

C'est une façon de dire que le domaine B est « complet », que le passage de B à C n'ajoute rien. Ma vue baisse, mais je n'ai pas vu de preuve de cette affirmation dans l'article de Cantor.

Cependant, Cantor continue de construire de façon analogue un domaine D à partir de C, puis E à partir de D, etc. et il dit qu'il trouvera un intérêt pour sa théorie à garder la différence conceptuelle entre B et C, entre C et D, etc. J'avoue que j'ai du mal à comprendre cette affirmation.

Heine naturellement suit « en gros » le même chemin, il nous a prévenu. Mais il est peut-être plus clair que Cantor en explicitant la proposition suivante (p. 180) **:** « Les irrationalités d'ordre $m+2$ ne sont pas nouvelles, elles coïncident au contraire avec celles d'ordre 1 » **;** Heine en donne une preuve courte, claire et convaincante. Et il ne parle pas de l'intérêt qu'il y aurait à conserver les « irrationnalités d'ordre > 1 » **:** il embraye directement sur la théorie des fonctions, tout particulièrement la théorie des fonctions continues.

Chez Heine, les « irrationnalités d'ordre 1 » correspondent exactement au « domaine » B de Cantor, et les irrationalités d'ordre supérieur aux domaines successifs C, D, etc.

Cantor en vient ensuite à poser un « axiome » : les points de la droite (la bonne vieille droite de la géométrie antique ; Cantor mentionne les *Éléments* d'Euclide, livre 10) correspondent bijectivement aux *Zahlengrösse*.

Il passe ensuite à la notion de point limite, puis à celle d'*ensemble dérivé* P' d'un ensemble P de points de la droite : c'est l'ensemble des points p qui sont limite de suites de points p' de P, où les p' sont différents de p. Il envisage aussi d'itérer cette opération de dérivation, définissant ainsi les ensembles dérivés successifs $\mathrm{P}^{(k)}$. Et il en vient enfin à une question de séries trigonométriques que nous verrons plus loin, qui est une extension de ses premiers résultats trigonométriques.

Le dérivé d'un ensemble *fini* P ne contient aucun point.
Le dérivé de l'ensemble P_1 formé des points de la suite $(1/n)$ (où $n \geq 1$), suite qui converge vers 0, contient le seul point 0 ; le dérivé deuxième de P_1 est donc vide.

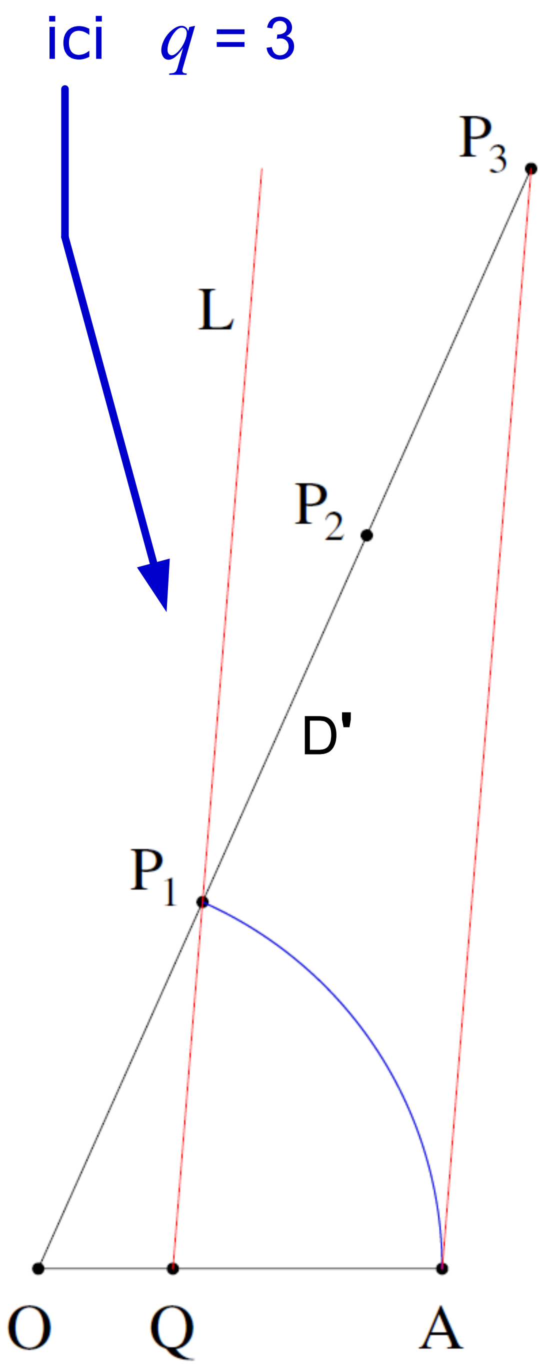
$$\mathrm{P}_2 = \left\{ \frac{1}{n+\frac{1}{k}} \; : \; n, k \geqslant 1 \right\} \longrightarrow$$

le dérivé de P_2 est P_1, le dérivé deuxième de P_2 est donc $\{0\}$, le troisième dérivé de P_2 est vide.

Revenons sur l'axiome dit *de Cantor–Dedekind* concernant la correspondance entre la droite et les nombres réels. Cet axiome peut nous paraître obscur, ou même sans intérêt ; c'est qu'il vient d'une époque où le monde de la *géométrie* et le monde du *nombre* étaient considérés comme distincts, ce qui n'est plus vraiment le cas aujourd'hui : nous parlons de la *droite réelle*, cette *droite* est formée de *nombres*. Faisons un peu de vieille géométrie.

Dans les constructions à la règle et au compas dans le plan, on suppose donné un segment unité sur une droite D, dont une extrémité sera notée O et l'autre A. À l'aide du compas, on peut reporter ce segment unité OA sur la droite D, un nombre fini de fois vers la gauche ou vers la droite. On peut donc produire sur la droite D des points dont les abscisses peuvent correspondre à un entier relatif quelconque.

Autres opérations : étant donnés deux points distincts, on peut avec la règle tracer une droite passant par ces deux points, et avec le compas et la règle, on sait tracer une parallèle à une droite donnée, passant par un point donné.



Si D' est une droite distincte de D, qui passe par l'origine O du segment unité OA, et si $q \geq 1$ est un entier donné, on peut reporter q fois le segment unité OA sur D', obtenant ainsi des points « successifs » $P_0 = O$, puis $P_1, \ldots, P_q$ sur la droite D'.

Si on trace une droite depuis le point P_q jusqu'à l'extrémité A du segment unité, puis sa parallèle L passant par P_1, le théorème de Thalès nous dira que l'intersection de L avec la droite D donne un point Q tel que le segment OQ représente la proportion $1/q$ du segment unité.

En reportant un nombre fini quelconque de fois avec le compas un tel segment OQ de longueur $1/q$, vers la droite ou la gauche, on pourra définir sur D des segments OR où l'abscisse de R sera n'importe quel nombre rationnel p/q.

Ainsi, la géométrie classique permet de définir sur une droite tous les « points rationnels ». L'axiome de Cantor–Dedekind consiste à compléter l'association entre points et nombres. Le point de vue des coupures de Dedekind est particulièrement approprié pour cela : une *coupure* consiste à classer *tous* les *nombres* rationnels dans deux familles A et B, de façon que tout élément a de A soit (strictement) inférieur à tout élément b de B ; une telle coupure définit le nombre réel x unique qui est plus grand que tous les éléments de A et plus petit que tous les éléments de B (si la coupure définit un nombre x qui est en fait *rationnel*, ce nombre x est, ou bien le plus grand élément de A, ou bien le plus petit élément de B).

Si on a classé tous les *points* rationnels de la droite D dans deux familles A et B, de façon que tout point a de A soit à gauche de tout point b de B, alors l'axiome consiste à dire qu'il existe un unique point X de la droite qui se trouve « entre » les familles de points A et B.

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.

de nos jours, on écrit plus souvent :

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

(avec même des coefficients c_n *complexes*)

L'étude des séries trigonométriques remonte au 18^e siècle (au 17^e, du temps de Pascal et Fermat, la courbe sinusoïde n'est encore que la « compagne de la cycloïde »). Dans les années 1750, Leonhard Euler découvre plusieurs développements en série trigonométrique, on en donnera un seul exemple plus loin ; Daniel Bernoulli propose d'utiliser les séries trigonométriques pour donner des solutions au fameux *problème des cordes vibrantes*, des solutions ainsi obtenues comme « superposition », peut-être infinie, de solutions élémentaires.

Leonhard Euler (1707–1783)

Daniel Bernoulli (1700–1782)

À partir de

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

$(= f(x))$

x variable réelle, coefficients a_n, b_n réels,

et en utilisant

$$\boxed{\sin(-u) = -\sin(u), \qquad \cos(-u) = \cos(u)}$$

on obtient

$$\text{(S)} \qquad \frac{f(x) - f(-x)}{2} = \sum_{n=1}^{\infty} b_n \sin(nx)\,;$$

pour simplifier l'exposé on pourra, à l'occasion,
se restreindre ici à une série de fonctions sinus.

Exemple (très joli je trouve) :

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

par le changement de variable

$$x \longrightarrow \pi - x$$

on obtient aussi

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

Je ne suis pas sûr de la date à laquelle Euler a découvert la somme de la série de sinus proposée page précédente (en tout cas, le résultat apparaît en 1755 dans l'un de ses livres ; la somme de cette série est rappelée page suivante). Cette série de fonctions sinus converge en tout point (de nos jours on justifierait la convergence « par la transformation d'Abel » ; la convergence en tout point sera l'hypothèse du *théorème de Cantor* qu'on verra plus loin). La fonction somme de cette série est périodique, elle est nulle quand $x = \pi$, la somme n'est pas continue au point π (puisque sa valeur est $x / 2$ quand $x < \pi$). C'est un excellent exemple pour étudier le *phénomène de Gibbs* (la somme n'étant pas continue au point π, la convergence de la série ne peut pas être uniforme : le graphe des sommes partielles reste éloigné du graphe limite au voisinage du point π).

La série numérique obtenue pour $x = \pi / 2$ était connue bien avant (Leibniz en 1674) : la représentation de $\pi / 4$ par cette *série de Leibniz* est assez célèbre.

Gottfried Leibniz (1646–1716)   Willard Gibbs (1839–1903) cette série converge pour tout x et la somme vaut

Euler :

(vers 1750 ?)

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

si $|x| < \pi$..

Une question pour Cantor : $\longrightarrow$ pourrait-on exprimer la même fonction avec une ***autre*** série trigo ?

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + \frac{(-1)^n}{2n+1} + \cdots$$

Leibniz vers 1675

ou bien :

$$\frac{\pi}{8} = \frac{1}{1.3} + \frac{1}{5.7} + \frac{1}{9.11} + \frac{1}{13.15} + \cdots$$

(pour obtenir la somme de cette seconde série, il suffit de regrouper les termes deux par deux dans la série alternée qui la précède)

Fourier utilise des séries trigonométriques pour sa *théorie de la chaleur*, dès 1805 ; il donne plusieurs développements en série trigonométrique de fonctions périodiques pas nécessairement continues, et il établit les formules intégrales qui permettent de calculer les coefficients de ces séries. Dirichlet a connu Fourier à Paris ; en 1829, il publie son *théorème de convergence*, première preuve suffisamment rigoureuse de convergence ponctuelle des séries de Fourier.

Riemann commence sa carrière avec Dirichlet ; dans son mémoire d'Habilitation de 1854, Riemann propose d'abord un historique des séries trigonométriques. On trouve aussi dans cette Habilitation la définition de *l'intégrale de Riemann*, en lien avec l'étude des séries de Fourier au moyen de formules intégrales ; noter que souvent, *l'intégrale de Cauchy* suffit : c'est l'intégrale des fonctions *continues*, qui est déjà définie dans les cours d'Analyse de Cauchy à l'École (royale) polytechnique vers 1820.

Joseph Fourier (1768–1830) (1807)
Louis Augustin Cauchy (1789–1857)
Peter Gustav Dirichlet (1805–1859) (1829)
Bernhard Riemann (1826–1866) (1854)

(pour les séries trigonométriques)

Un cas particulier

d' **Une fonction de Weierstrass (1872)**

$$x \mapsto \sum_{n \geqslant 0} 2^{-n} \cos(5^n x)$$

continue, nulle part dérivable,

impossible à dessiner :

ça vibre tout le temps.

La fonction de Weierstrass est impossible à dessiner ! Une intuition géométrique, ou mécanique, ou dynamique ne saurait l'expliquer : il se passe une infinité d'événements sur chaque intervalle ouvert non vide ; le point de vue analytique peut seul rendre compte des propriétés de cette fonction. Une somme partielle (*une somme finie*) de cette série trigonométrique n'est un exemple de rien du tout. Pas vraiment publiée par Weierstrass (une simple communication à l'Académie de Berlin, mais Weierstrass l'a aussi transmise à plusieurs de ses correspondants), on trouve cette communication qui contient l'étude précise de cette fonction dans les *Œuvres* de Weierstrass.

Dans plusieurs exemples apparus à cette époque, il faut embrasser par l'imagination, par la pensée, une infinité d'étapes : aux rangs finis, *rien* n'est encore fait ; on pourrait dire en revanche que si on approche un nombre réel, on a déjà une assez bonne idée quand on a calculé 20 décimales... (sauf s'il s'agit de savoir, par exemple, si le nombre réel en question est *rationnel ou pas*).

Quelque dix ans plus tard (en 1884), Cantor présente la *fonction de Cantor* ; elle croît de 0 à 1, mais on ne « voit pas » où ça se passe.

Premiers travaux trigonométriques de Cantor :

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

deux articles à la suite l'un de l'autre dans ce « Journal » :

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.

« 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 » et l'année suivante :

Notiz zu dem Aufsatze: 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.

[Crelles Journal f. Mathematik Bd. 73, S. 294—296 (1871).]

Cantor devait viser le record du titre le plus long. Mais la concurrence sur ces longs titres était très rude à cette époque, je ne crois pas qu'il ait gagné !

Théorème d'unicité de Cantor

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

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

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

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

Aucune autre hypothèse que la convergence pour tout x ;

par différence : si $f(x) = g(x)$ pour tout x et si f et g sont chacune représentable en tout point x par une série trigonométrique, alors les coefficients dans ces deux séries sont identiques ; c'est le titre de l'article : il n'y a qu'***une*** façon de représenter une fonction de cette manière.

Dans ce *théorème d'unicité de Cantor*, on ne suppose aucune propriété fonctionnelle de la série, on ne sait rien *a priori* sur les coefficients. Il faudra commencer par le *lemme de Cantor*, pour avoir un minimum d'information sur les coefficients de la série, avant de pouvoir utiliser des outils déjà existants.

Ce lemme va garantir que sous les hypothèses du théorème, les coefficients de la série sont *bornés*. Le lemme dit plus, mais c'est cette seule conséquence du résultat qui sera utile pour la suite de la preuve.

dans le premier des deux articles successifs mentionnés précédemment on trouve le

(notation pour un intervalle ouvert)

Lemme de Cantor :

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

$$a_n \cos(nx) + b_n \sin(nx) \xrightarrow[n]{} 0,$$

(*)

alors les coefficients tendent vers 0,

$$a_n \xrightarrow[n]{} 0, \qquad b_n \xrightarrow[n]{} 0.$$

(en particulier, ces coefficients sont donc bornés)

(noter que : si la série trigonométrique (T) converge au point x, son terme général tend vers 0)

(*)

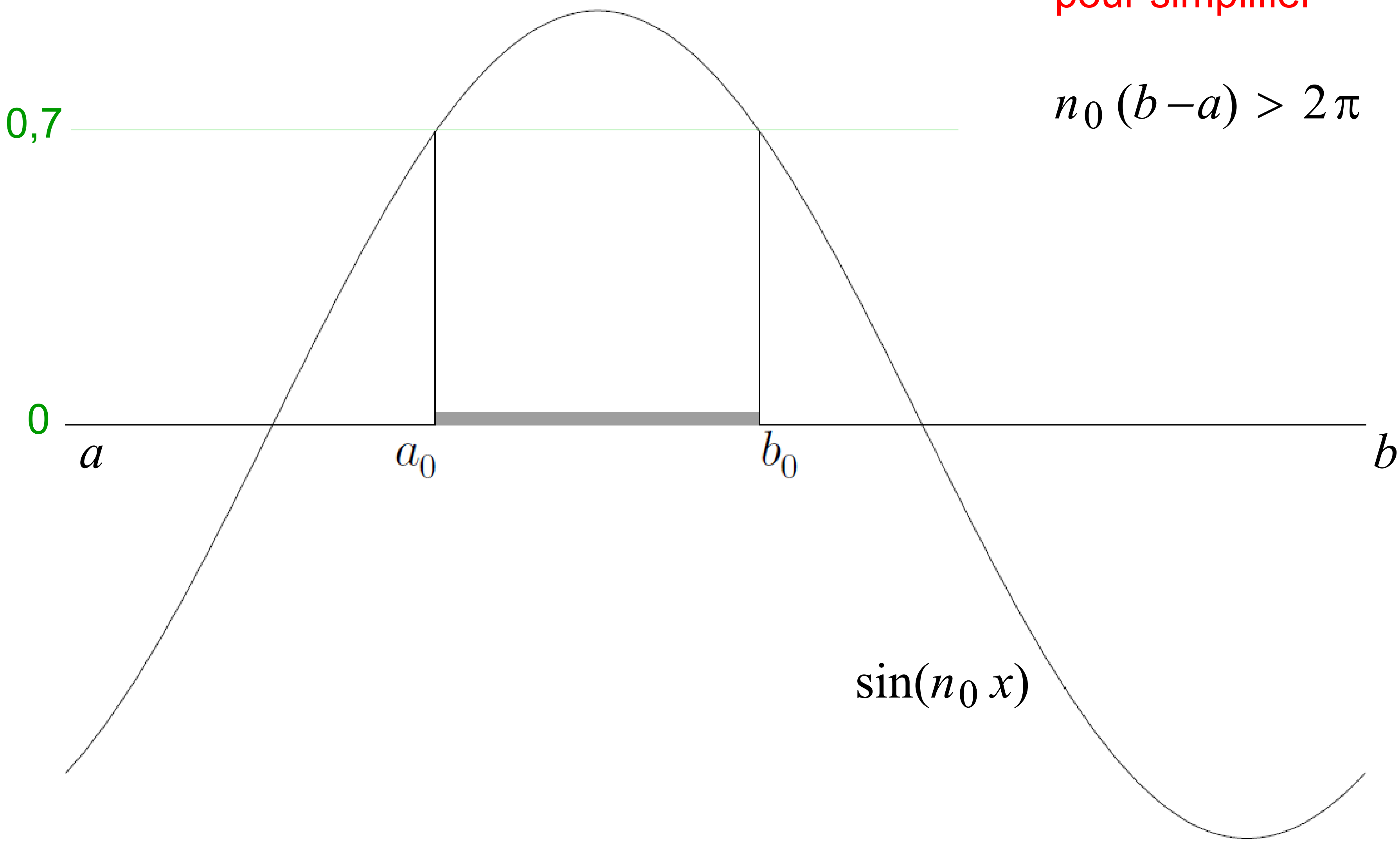
On illustrera la preuve du lemme avec des fonctions sinus « seules », pour simplifier ; quand n grandit, la période de la fonction $\sin(n\,x)$ diminue, le motif de base (l'arche sinusoïdale) se répète plus vite ; si on donne un petit segment quelconque et si on prend ensuite une fonction sinus $\sin(n\,x)$ qui « vibre suffisamment vite » (n assez grand), elle décrira toute une période dans le segment donné. En particulier, il existera un sous-segment sur lequel toutes les valeurs $\sin(n\,x)$ seront plus grandes que, disons, 0.7 (la valeur choisie page suivante).

On commence donc avec un segment $[a, b]$ et un entier n_0 assez grand.

Lemme de Cantor :

fonctions sinus
pour simplifier

$n_0\,(b-a) > 2\pi$



quand n est grand, la période de $\sin(n\,x)$ → période $2\pi/n$

devient petite, le motif de base se répète plus vite

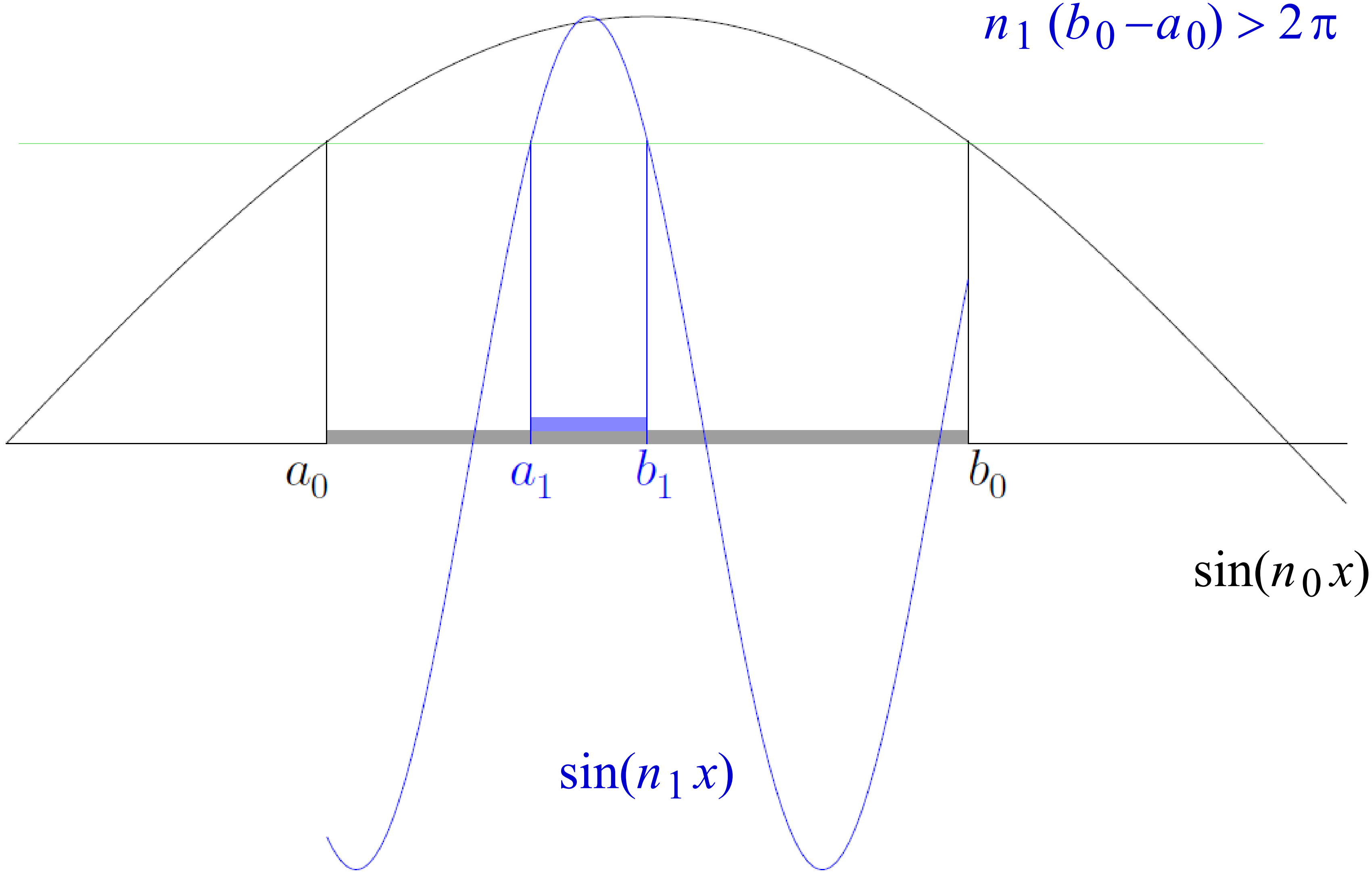
Ayant trouvé ce sous-segment $[a_0, b_0]$ de $[a, b]$ sur lequel la valeur de $\sin(n_0 x)$ reste plus grande que 0.7, on répète l'opération avec le nouveau segment $[a_0, b_0]$.

Pour n_1 assez grand, on trouvera un sous-segment $[a_1, b_1]$ de $[a_0, b_0]$ sur lequel ***les deux*** fonctions $\sin(n_0 x)$ et $\sin(n_1 x)$ seront grandes. Et on continuera...

Lemme de Cantor :

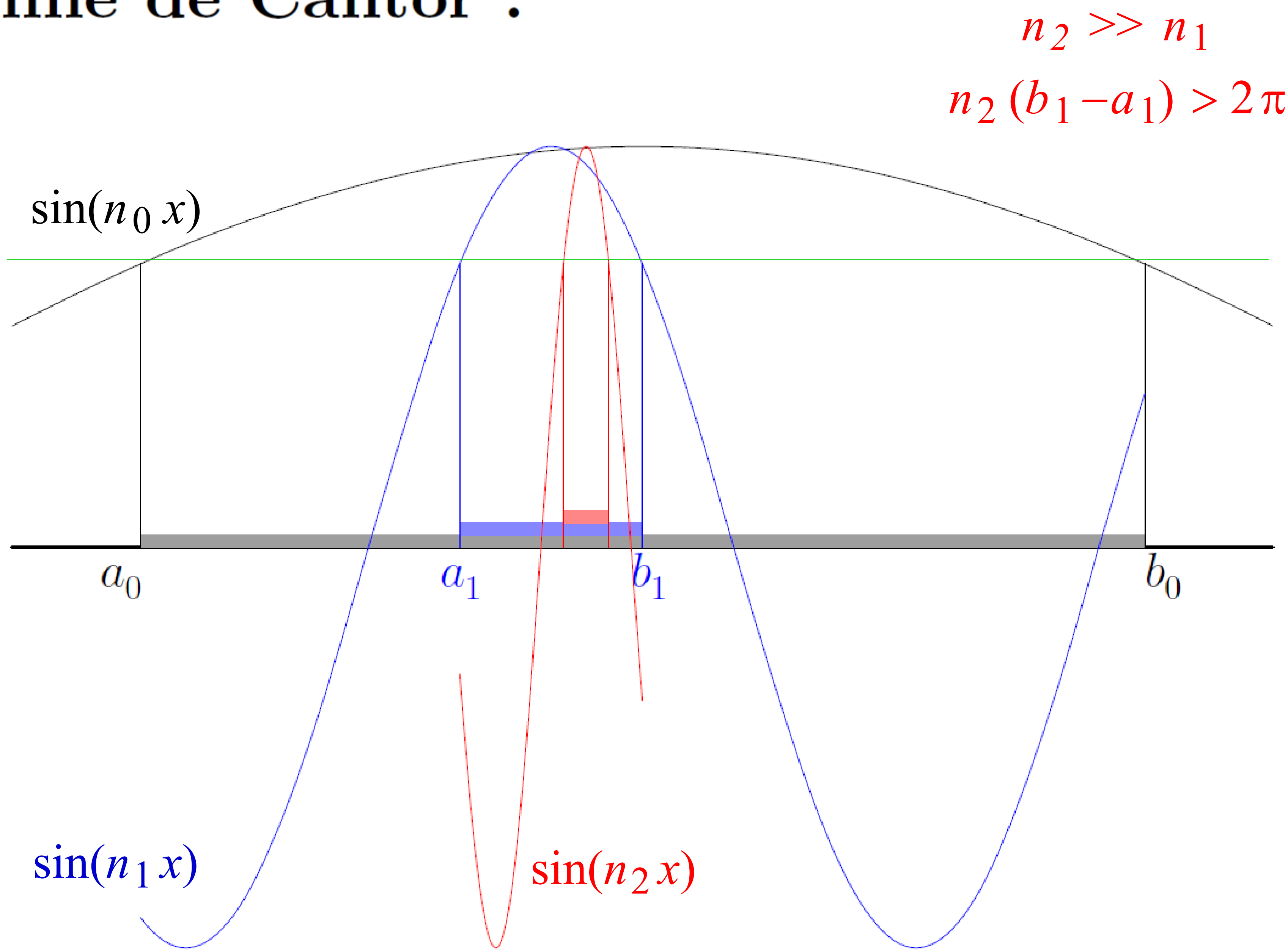
$$n_1 >> n_0$$

$$n_1\,(b_0 - a_0) > 2\pi$$



faisons encore un pas

Lemme de Cantor :



On obtient un troisième segment (rouge), emboîté dans les précédents.

La suite de segments emboîtés détermine un nombre réel Ω.

Preuve du

Lemme de Cantor : (limitée à des fonctions sinus)

il résulte de ce qui précède que pour tout intervalle (a, b) non vide et pour toute suite (n_j), il existe une sous-suite (n_{j_k}) et un nombre réel $\Omega \in (a, b)$ tels que

$$\sin(n_{j_k}\Omega) \geqslant 1/2 \quad \text{pour tout } k.$$

Si $b_n \sin(nx)$ tend vers 0 pour tout $x \in (a, b)$,

mais si b_n ne tendait pas vers zéro, on aurait

$$|b_{n_j}| \geqslant \delta > 0 \text{ pour une sous-suite,}$$

alors on déduirait pour une sous-sous-suite que

$$|b_{n_{j_k}} \sin(n_{j_k}\Omega)| \geqslant \delta/2$$ pour tout k, ce qui

serait une contradiction.

Fin de la preuve du lemme $\square$

Pour passer du *lemme de Cantor*
au *théorème de Cantor* **:** Cantor va utiliser des

Résultats obtenus par Riemann dans son Habilitationsschrift
de 1854, document publié peu après sa mort en 1866 :

(résultats qu'on restreindra ici à des fonctions sinus, pour simplifier)

Posons **:**

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

(avec des coefficients b_n bornés)

Pour chaque x fixé **:**

si la série $\sum_{n=1}^{\infty} b_n \sin(nx)$ converge, alors

$$« \mathrm{F}''(x) » = -\sum_{n=1}^{\infty} b_n \sin(nx).$$

On vient de dire que :

$$« \, \mathrm{F}''(x) \, » \;=\; -\sum_{n=1}^{\infty} b_n \sin(nx)$$

mais ce n'est pas la « vraie » dérivée seconde ! Ici,

$$« \, \mathrm{F}''(x) \, » \;=\; \lim_{h\to 0} \frac{\mathrm{F}(x-h) - 2\,\mathrm{F}(x) + \mathrm{F}(x+h)}{h^2} \,.$$

Remarques : si la « vraie » dérivée seconde existe, elle a la même valeur que la « pas vraie ».

Si F(x_0) est *maximal*, on voit que « F''(x_0) » est ≤ 0, puisque dans ce cas F($x_0 - h$) + F($x_0 + h$) $\leq$ 2 F(x_0) .

Preuve du théorème d'unicité de Cantor.

Supposons avec Cantor que

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) = 0$$

pour tout x d'un intervalle ouvert (a, b) non vide,

où $a < b$.

ce cas « général » sera utile plus loin ;
mais c'est [0, 2 π] qu'on veut vraiment ici
pour conclure la preuve du théorème,
au lieu de cet intervalle (a, b).

Supposons que

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) = 0, \qquad x \in (a, b).$$

(par le lemme de Cantor) Les (b_n) tendent vers 0, ils sont donc bornés. Considérons avec Riemann

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

La fonction F est continue et 2π-périodique.

On a supposé que

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) = 0, \quad x \in (a,b),$$

et posé

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

D'après Riemann, « $\mathrm{F}''(x)$ » $= -f(x) = 0$ pour tout $x \in (a,b)$.

Si $-f$ était la « vraie » dérivée seconde de F, et était de plus nulle sur l'intervalle (a, b), F serait évidemment affine sur (a, b), directement ; mais ici il faut une preuve, qui a été fournie à Cantor par Hermann Schwarz : pour obtenir le résultat, Schwarz utilise le fait qu'une certaine fonction continue atteint son maximum (un des théorèmes de l'enseignement de Weierstrass).

Voici l'idée de la preuve de Schwarz ; on montre que si la dérivée seconde généralisée de F est positive ou nulle sur (a, b), alors F est convexe sur (a, b) ; il suffit de voir, pour c et d quelconques, que sous cette hypothèse de dérivée seconde, si $F(c) = F(d) = 0$, alors F reste ≤ 0 sur l'intervalle (c, d). Il suffit encore de montrer que $G(x) = F(x) - \varepsilon(x-c)(d-x)$ reste ≤ 0 entre c et d pour tout $\varepsilon > 0$; sinon, $G(x)$ atteindrait un maximum > 0 en un certain point x_0 de (c, d) ; alors, la dérivée seconde généralisée de cette fonction G serait ≤ 0 en ce point, ce qui est impossible puisqu'elle vaut « $F''(x_0)$ » $+ 2\,\varepsilon > 0$.

Si f est nulle sur (a, b), F est à la fois convexe et concave, donc affine sur (a, b).

Récapitulons :

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) = 0, \quad x \in (a, b).$$

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

on a « $\mathrm{F}''(x)$ » $= 0$ pour tout $x \in (a, b)$;

(dans une lettre) Hermann Schwarz montre à Cantor que :

$$\text{F est affine sur } (a, b), \quad \underline{\mathrm{F}(x) = cx + d}.$$

Hermann Amandus Schwarz (1843–1921) ;

ami de Cantor à l'époque, ça ira moins bien entre eux plus tard . . .

Gardons en mémoire pour plus tard le cas a < b quelconque, mais revenons au véritable objectif, en supposant la convergence-nullité « partout » :

Si

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(nx) = 0$$

pour tout x, alors

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

est affine sur $\mathbb{R}$, $\mathrm{F}(x) = cx + d$.

Mais F est aussi périodique ; affine et périodique implique que F est constante sur **R**, donc F est nulle !

(prendre $x = 0$)

Pour tout x,

$$\mathrm{F}(x) = \sum_{n=1}^{\infty} \frac{b_n}{n^2} \sin(nx) = 0\,;$$

(Fourier, Dirichlet)

dans le cas d'une fonction continue F on sait calculer ses « coefficients de Fourier » par une intégrale :

$$\frac{b_n}{n^2} = \frac{1}{\pi} \int_0^{2\pi} \mathrm{F}(x) \sin(nx)\, \mathrm{d}x = 0, \qquad n \geqslant 1.$$

$\square$

Au vu de la preuve précédente, un anti-Cantor pourrait bien nous dire **:** Cantor ne s'est pas trop fatigué pour prouver son théorème, il s'est contenté de ramasser divers éléments connus, et pour ce qui n'était pas connu, il se l'est fait démontrer par son ami Schwarz ! Mais après tout, c'est comme ça qu'une bonne quantité de théorèmes ont été prouvés de par le monde, et en plus, le reproche précédent ne s'applique pas selon moi au *lemme de Cantor,* qui me semble incontestablement original.

On a dit qu'on voulait garder en mémoire ce résultat intermédiaire,

Retenons :

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

où on a posé :

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

Un nouveau départ.

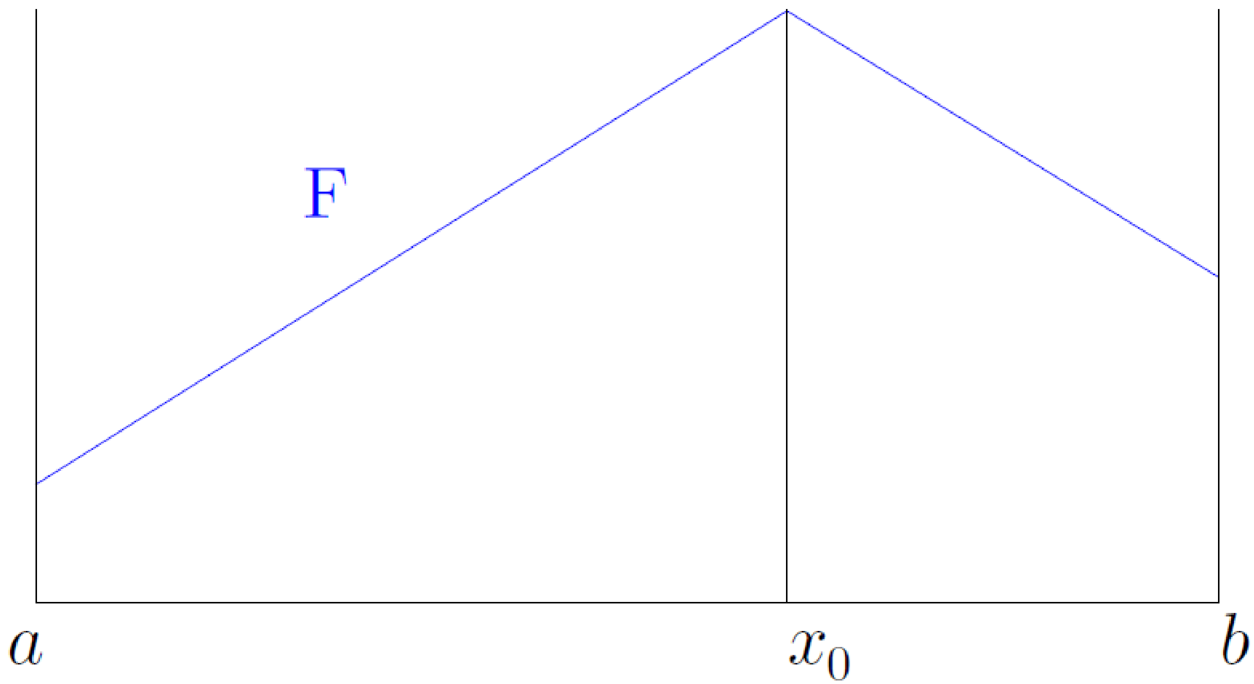
Si $f(x) = 0$ pour $x \in (a, b)$, F est affine sur (a, b).

Et si l'hypothèse $f(x) = 0$ est vraie SAUF pour un point exceptionnel $x_0 \in (a, b)$?

Cette question, après une généralisation convenable, va nous entraîner loin ailleurs...

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

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



En supposant seulement les b_n bornés, on a

$$\left(F(x-h)-2F(x)+F(x+h)\right)/h \longrightarrow 0$$

quand $h \to 0$

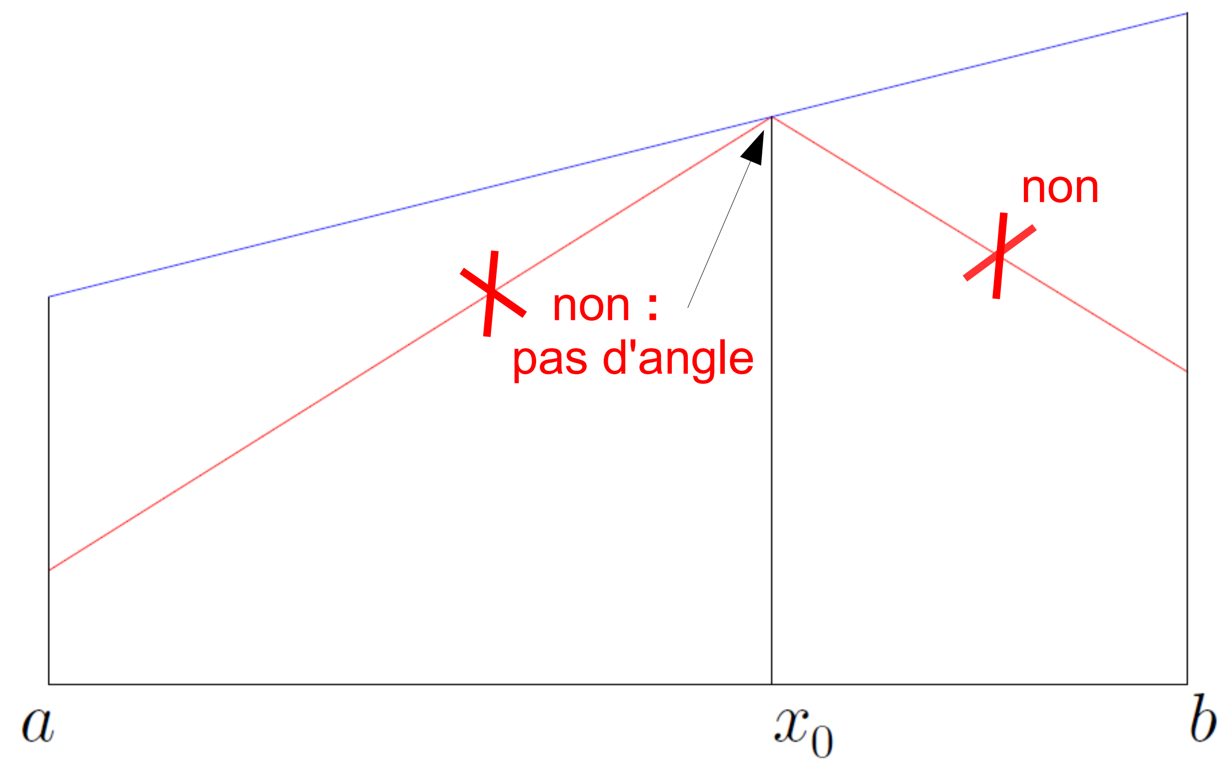
(un résultat de l'Habilitation de Riemann, encore)

il en résulte que le graphe de la fonction F ne peut pas présenter de « coins » :

la fonction F est donc affine partout sur (a, b),

et on est ramené au cas précédent.

(pour s'en convaincre, tester $F(x) = |x|$ en $x = 0$)



Par récurrence on obtient :

si l'hypothèse $f(x) = 0$ est vraie SAUF pour des points exceptionnels $x_0, x_1, \ldots, x_n \in (a, b)$ alors ...
F est affine sur (a, b).

Considérons (a, x_n) et (x_n, b) ;

on suppose le résultat établi quand il y a au plus n points exceptionnels dans un intervalle ;
sur chacun des deux intervalles ci-dessus on est dans ce cas déjà vu, donc F est affine sur chacun d'eux, et comme il n'y a pas de « coin » dans le graphe de la fonction F, elle est affine sur (a, b). $\square$

S'il y a un nombre fini de points exceptionnels dans $[0, 2\pi]$, la fonction F est affine sur $[0, 2\pi]$, $2\pi-$périodique continue et nulle en 0, donc F est nulle sur **R** et on conclut comme avant que les coefficients de f sont nuls.

cette histoire n'est pas finie . . .
mais on va avoir un entracte.

Dénombrable,
ou pas...

Non dénombrabilité de la droite réelle (1874)

on considère une suite de nombres réels :

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

il s'agit de voir que ces nombres
« n'épuisent pas » la droite réelle ;

pour simplifier on supposera
que ces nombres sont
deux à deux distincts.

Non dénombrabilité de la droite réelle

On va laisser venir les points successifs de la suite donnée, en commençant avec x_0, x_1 qui délimitent un premier segment qui sera marqué en vert page suivante.

$$x_1 \neq x_0, \quad \text{etc} \ldots$$

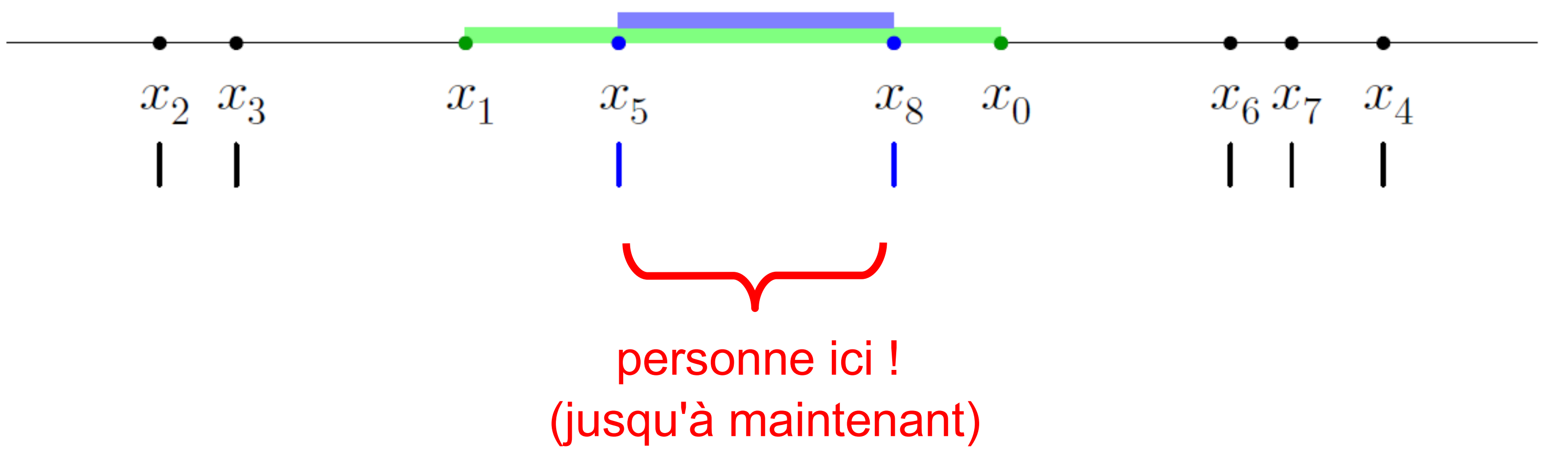
Non dénombrabilité de la droite réelle

On poursuit avec les points suivants, $x_2, x_3, \ldots$

Non dénombrabilité de la droite réelle

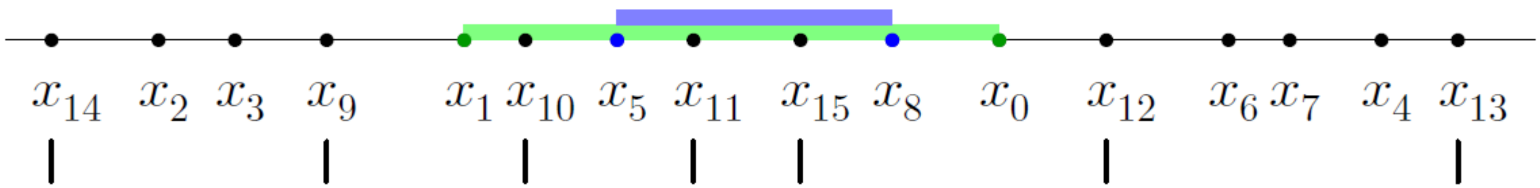
Deux nouveaux points sont arrivés
dans le segment vert ; ils délimitent
un nouveau segment, qui
sera marqué en bleu

Non dénombrabilité de la droite réelle



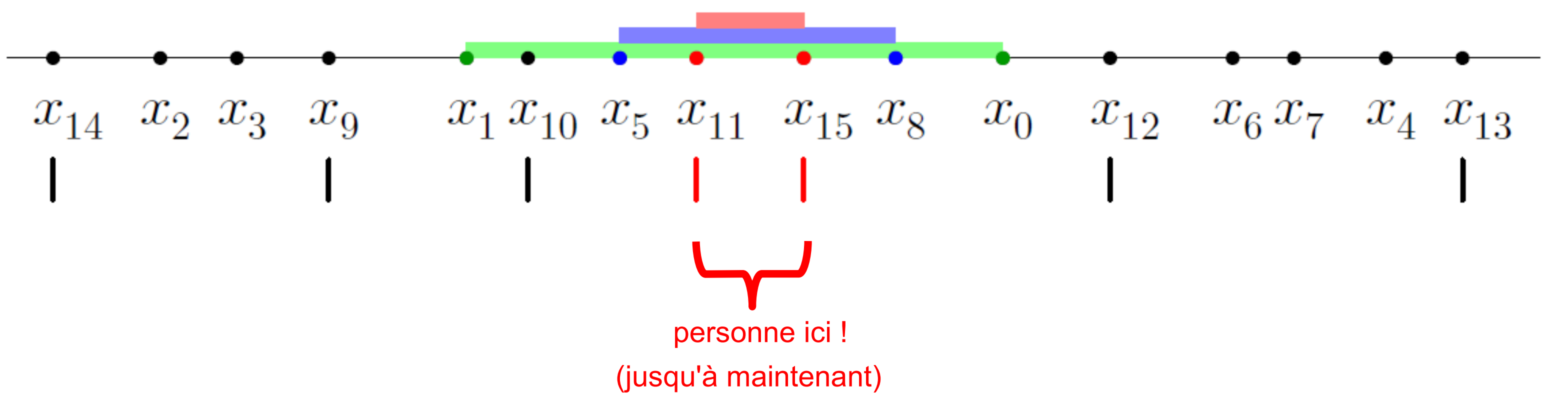
On continue

Non dénombrabilité de la droite réelle



Deux nouveaux points sont arrivés
dans le segment bleu ; ils délimitent
un nouveau segment, qui
sera marqué en rouge

Non dénombrabilité de la droite réelle



suite de segments emboîtés :

les points communs conviennent !
(et on sait qu'il y en a)

Ils ne peuvent pas faire partie de la suite donnée au départ

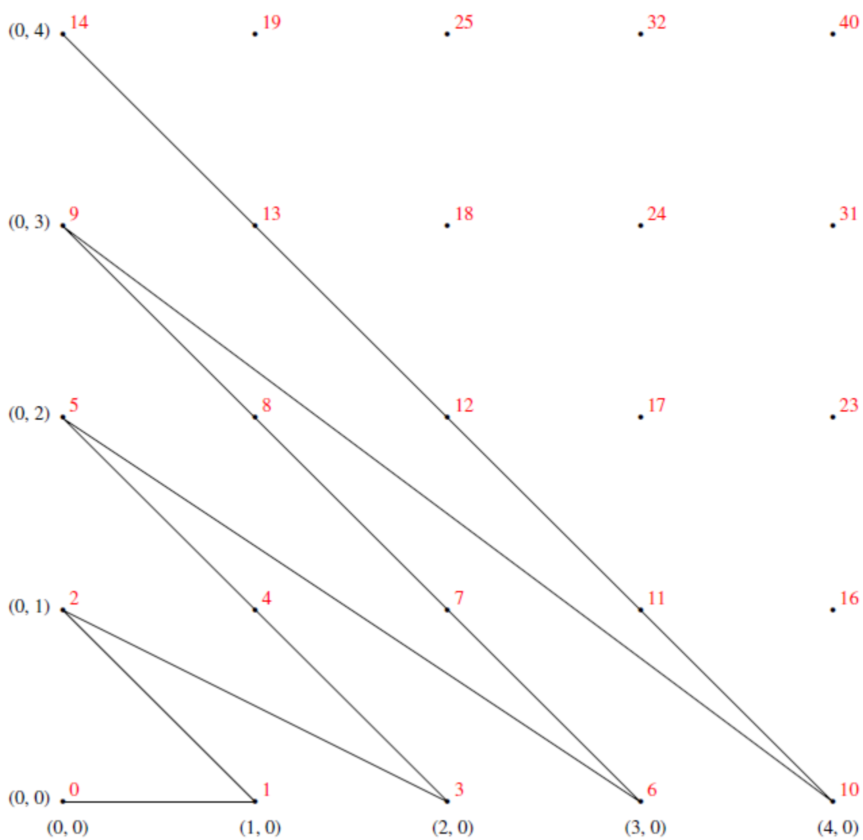
$\square$

La suite des segments successifs emboîtés obtenus par le processus précédent a une intersection non vide, et il n'est pas trop difficile de vérifier qu'un point de cette intersection *ne peut pas* faire partie de la suite donnée au début.

Il est possible que le processus se soit arrêté à un certain moment **:** dans ce cas, le dernier segment apparu dans le processus ne contient pas *deux* points de la suite, ce qui bloque le processus. Puisque ce dernier segment (qui n'est pas réduit à un point) ne contient qu'un seul point ou bien aucun point de la suite, ce segment contient aussi quantité de points qui n'appartiennent pas à la suite qui a été donnée au départ.

En 1874 on sait (en tout cas Cantor sait) jouer avec le dénombrable ;

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



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 **N** des entiers

Les nombres algébriques

Le nombre θ est *algébrique* quand

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

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

(à partir d'un segment de longueur 1)

Les suites finies d'entiers $(a_0, \ldots, a_n)$ forment un ensemble dénombrable, l'ensemble des polynômes à coefficients entiers est dénombrable, chacun d'eux n'a qu'un nombre fini de racines.

Il en résulte que :

Les nombres algébriques

forment un ensemble dénombrable.

Cantor a ainsi prouvé :

il y a (vraiment) <u>beaucoup</u> de réels non algébriques !

Nombres transcendants :

Ceux qui ne sont pas algébriques.

Il est en général très difficile de montrer qu'un nombre particulier donné est transcendant.

Le mathématicien suisse Johann Heinrich Lambert a eu déjà beaucoup de mal pour prouver (en 1767) que le nombre π n'est pas rationnel.

ce résultat est en rapport avec le problème de la ***quadrature du cercle*** (mais il ne le résout pas)

construire à la règle et au compas un carré ayant la même surface qu'un cercle de rayon 1

Nombres de Liouville : exemples relativement simples de nombres transcendants

Johann Heinrich Lambert (1728–1777)

Joseph Liouville (1809–1882)

Le caractère transcendant d'un nombre de Liouville x provient d'une idée assez simple, qu'on va évoquer de façon exagérément vague : les décimales non nulles de x qui ont un rang m assez élevé deviennent extrêmement « espacées », de telle sorte que la contribution à $\mathrm{P}(x)$ du 10^{-m} de la décimale m sera séparée des contributions des autres décimales (P étant un polynôme à coefficients entiers), et il n'y aura donc pas d'annulation possible.

Pour prendre un exemple très simple, si $\mathrm{P}(x) = x^3 - a\,x + 1$ avec a entier, la contribution à $x^3 - a\,x$ de la décimale m se trouve essentiellement entre les rangs m et $3\,m$, donc, dans un intervalle entier qui sera disjoint des autres intervalles analogues, pour m grand : cette contribution ne pourra pas être annulée. Ainsi, $x^3 - ax$ possédera une infinité de décimales non nulles et ne pourra pas être l'entier -1, donc $\mathrm{P}(x) \neq 0$.

On fera un (tout petit) peu mieux deux pages plus loin.

Nombres de Liouville (1844) : exemples relativement simples de nombres transcendants

en gros on a $n\,!$ zéros successifs après le n ème **1**

$$x = 0{,}\mathbf{11}000\mathbf{1}\underbrace{00000000000000000\mathbf{1}00000000000000000000000000000000000000}_{}$$
$$000000000000000000000000000000000000000000000000000000000\mathbf{1}\ldots$$

$$x = 10^{-1} + 10^{-2} + 10^{-6} + 10^{-24} + 10^{-120} + \cdots$$

$$\boxed{x = \sum_{n\geqslant 1} 10^{-n!}.}$$

Il n'est pas essentiel que les « places » des $\mathbf{1}$ dans le développement soient *exactement* les factorielles successives : il suffit que les « trous » successifs formés de $\mathbf{0}$ soient *au moins aussi grands* que ceux qu'on a ci-dessus.

Complétons l'exemple très élémentaire commencé il y a deux pages.

On considère un polynôme $P(X) = \overbrace{X^3 - aX + 1}$, avec a entier > 0 ; on veut montrer que si x est le nombre de Liouville de la page précédente, on a $P(x) \neq 0$. On choisit q_0 entier tel que $a + 7 \leqslant 2^{q_0}$. Écrivons pour le nombre x, et pour m assez grand, la décomposition

$$x = y + 10^{-m} + z = t + z, \qquad t = y + 10^{-m},$$

où y est une somme de termes 10^{-k} avec $k \leqslant k_0 < m$, et z une somme (infinie) de termes 10^{-p} distincts avec $p \geqslant p_0 > m$,

$$y = 10^{-1} + 10^{-2} + \cdots + 10^{-k_0}, \quad z = 10^{-p_0} + 10^{-p_1} + \cdots .$$

On peut voir facilement sur les développements décimaux que $z < 2.10^{-p_0}$ et qu'on a $z < t < x < 1$. Alors

$$\begin{aligned} x^3 - a\,x = \; & -a\,y - a\,10^{-m} + y^3 + 3y^2 10^{-m} + 3y10^{-2m} \\ & + 10^{-3m} \\ & - a\,z + 3t^2 z + 3\,t\,z^2 + z^3. \end{aligned}$$

La première ligne est une somme de termes $b\,10^{-k}$ avec $k \leqslant 3k_0 + 2m$ et b entier, sa somme S peut s'exprimer sous la forme $\mathrm{S} = c\,10^{-3k_0 - 2m}$ avec c entier. La troisième ligne est bornée par

$$a\,z + 3t^2 z + 3t^2 z + t^2 z < (a+7)z < 2.10^{q_0 - p_0}.$$

On voit que $\mathrm{P}(x) \neq 0$ si on peut garantir que

$$(*) \qquad 3\,k_0 < m \;\text{ et }\; 3\,m + q_0 < p_0.$$

En effet, $10^{3m}\mathrm{P}(x) = d\,10^{m-3k_0} + 1 + u10^{3m+q_0-p_0}$, avec d entier et $|u| \leqslant 2$, ne peut pas être nul dans ce cas $(*)$; de plus, la taille énorme des « trous » dans le développement du nombre de Liouville x garantit cette condition (*), dès que m est un peu grand.

Écrivons la suite des rangs m des décimales non nulles d'un nombre x sous la forme

$$m_0 < m_1 < \ldots < m_j < \ldots$$

La preuve précédente montre (presque) que x n'annule aucun polynôme de degré 3 à coefficients entiers dès que

$$m_{j+1} - 3m_j \to +\infty.$$

Le rôle des factorielles dans les nombres de Liouville est de permettre de traiter les polynômes de tous les degrés.

Nombres transcendants

Charles Hermite (1822–1901) : e transcendant, 1873

Ferdinand Lindemann (1852–1939) : π transcendant, 1882

Cantor est « referee » de l'article de Lindemann pour la revue *Mathematische Annalen* (à la demande du « directeur » Felix Klein)

Felix Klein (1849–1925)

La transcendance de e et de π, ça c'est du sérieux. Si plusieurs des résultats de Cantor sont devenus « faciles » de nos jours, ce n'est pas le cas de ces résultats de transcendance de Hermite ou Lindemann.

Chargé par Felix Klein d'examiner l'article de Lindemann avant publication dans *Mathematische Annalen*, Cantor a fait un travail consciencieux **:** on possède une lettre de Cantor à Klein où Cantor propose de suggérer à Lindemann des améliorations de tel ou tel énoncé de son article. Je ne suis pas assez savant pour pouvoir dire si ces suggestions de Cantor figurent effectivement dans la version finale de l'article de Lindemann.

Cette lettre de Cantor à Klein se trouve dans un volume intitulé « Georg Cantor Briefe », publié en 1991 (c'est la lettre 26 de ce recueil) **;** ce volume contient une bonne partie de la correspondance de Cantor (notamment celle avec Mittag-Leffler, mathématicien dont on parlera plus loin).

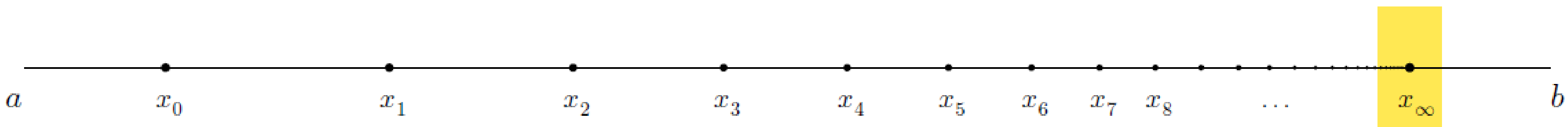
Ordinaux, cardinaux

Revenons à l'unicité de la représentation trigonométrique ;

on va voir apparaître
un « ensemble exceptionnel » infini :

(une suite convergente et sa limite x_∞)

Un ensemble exceptionnel infini



Supposons que $f(x) = 0$ sur (a, b) sauf aux points de la suite ;
alors : il n'y a qu'un nombre *fini* de points exceptionnels en dehors de l'intervalle ouvert « jaune » autour de la limite, on sait d'après ce qu'on a vu précédemment que la fonction F est affine hors de cet intervalle, et ce, quel que soit l'intervalle jaune autour de la limite . . .

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

Si une fonction g est définie et de plus, est nulle sur $(a,\, c-\varepsilon)$ pour tout $\varepsilon \in (0,\, c-a)$, il est clair que g est définie et est nulle sur $(a,\, c)$; en appliquant cette remarque au cas où g serait la dérivée seconde de F, on déduit que si F est affine sur $(a,\, c-\varepsilon)$ pour tout $\varepsilon \in (0,\, c-a)$, elle est affine sur $(a,\, c)$.

Ainsi, dans l'exemple qui nous intéresse, la fonction F étudiée par Riemann sera affine sur $(0,\, x_\infty)$ et sur $(x_\infty,\, 2\,\pi)$; comme les « coins » dans le graphe de F sont exclus, la fonction continue F sera « affine tout-court ».

Il en résultera, comme dans les cas déjà traités, que tous les coefficients de la série dont la somme est $f(x)$ sont nuls : une suite convergente et sa limite forment donc un ensemble exceptionnel « admissible » ou « acceptable » pour énoncer une généralisation du théorème d'unicité de Cantor.

Un ensemble exceptionnel infini

$$\mathrm{E} = \{x_\infty \,;\; x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, \ldots\}$$

Si la série (T) converge vers 0 en tout point de $[0, 2\pi]$, sauf peut-être aux points de E, alors tous les coefficients sont nuls.

On a procédé à une sorte de récurrence topologique **. . .**

si F est affine sur $(a, c-\varepsilon)$ pour tout $\varepsilon > 0$,

alors, elle est affine sur (a, c).

Dans les années plus proches de nous, on a appelé « *ensemble d'unicité* » U tout sous-ensemble de $[0, 2\pi]$ ayant la propriété des ensembles que nous avons appelés « exceptionnels » jusqu'ici **:**

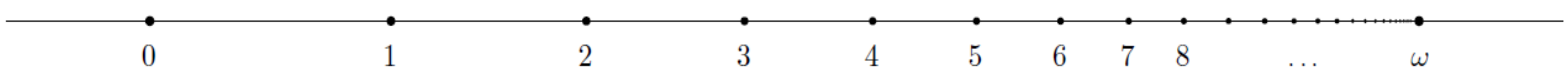
si une série trigonométrique converge en tout point x de $[0, 2\pi]$ avec une somme $f(x)$ ***nulle***, *sauf peut-être* aux points de U, alors tous ses coefficients sont nuls.

La terminologie est un peu bizarre **:** le tout premier résultat trigonométrique de Cantor s'exprime dans ce langage en disant que « *l'ensemble vide est un ensemble d'unicité* ».

Au début du 20^{e} siècle, on a découvert des ensembles d'unicité qui sont du type de *l'ensemble triadique de Cantor*, des exemples qui ont la *puissance du continu,* c'est-à-dire des ensembles qu'on peut mettre en bijection avec $[0, 1]$.

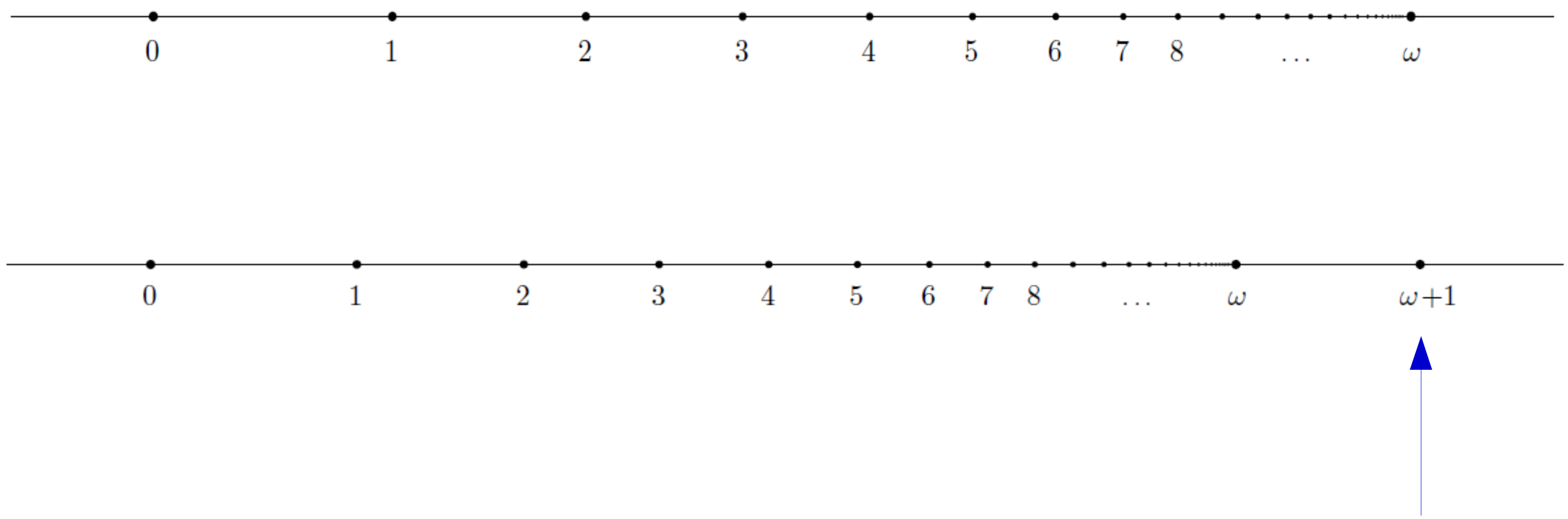
Ordinaux ;

on ne va plus regarder que la structure d'ordre
de la famille des points de l'exemple précédent :



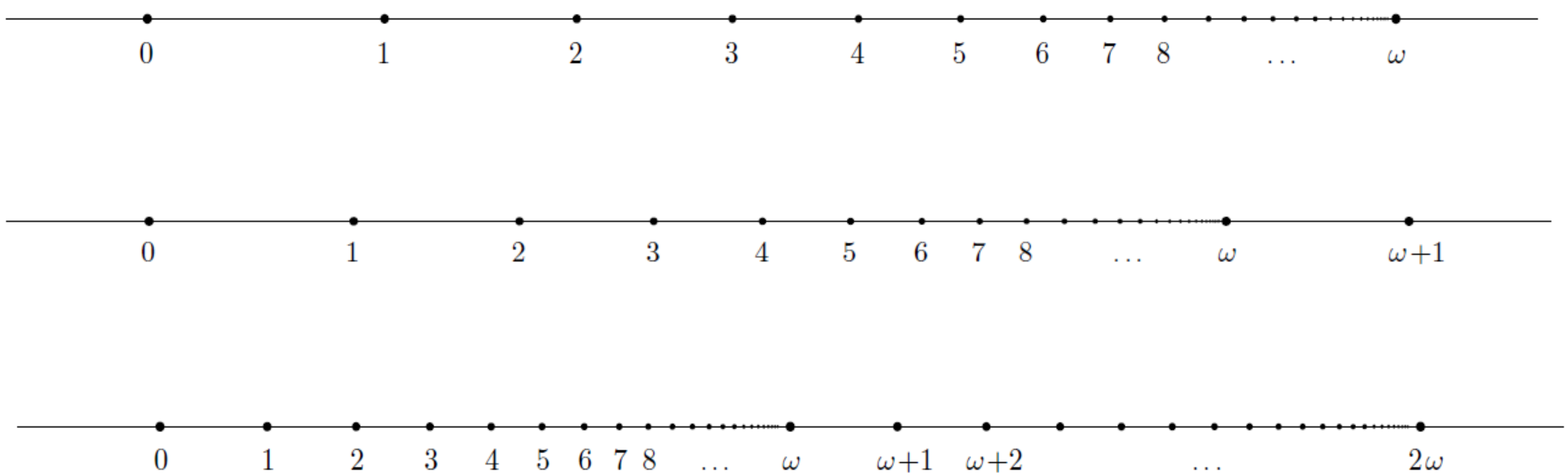
On pourra cependant observer que cet ensemble de points, ainsi que ceux qui apparaîtront dans les pages immédiatement suivantes, peuvent être « tracés » dans le segment [0, 2 π].

On ne regarde plus que la structure d'ordre :



on peut toujours ajouter un point « à droite »,
puis deux, trois, etc . . .

On ne regarde plus que la structure d'ordre :



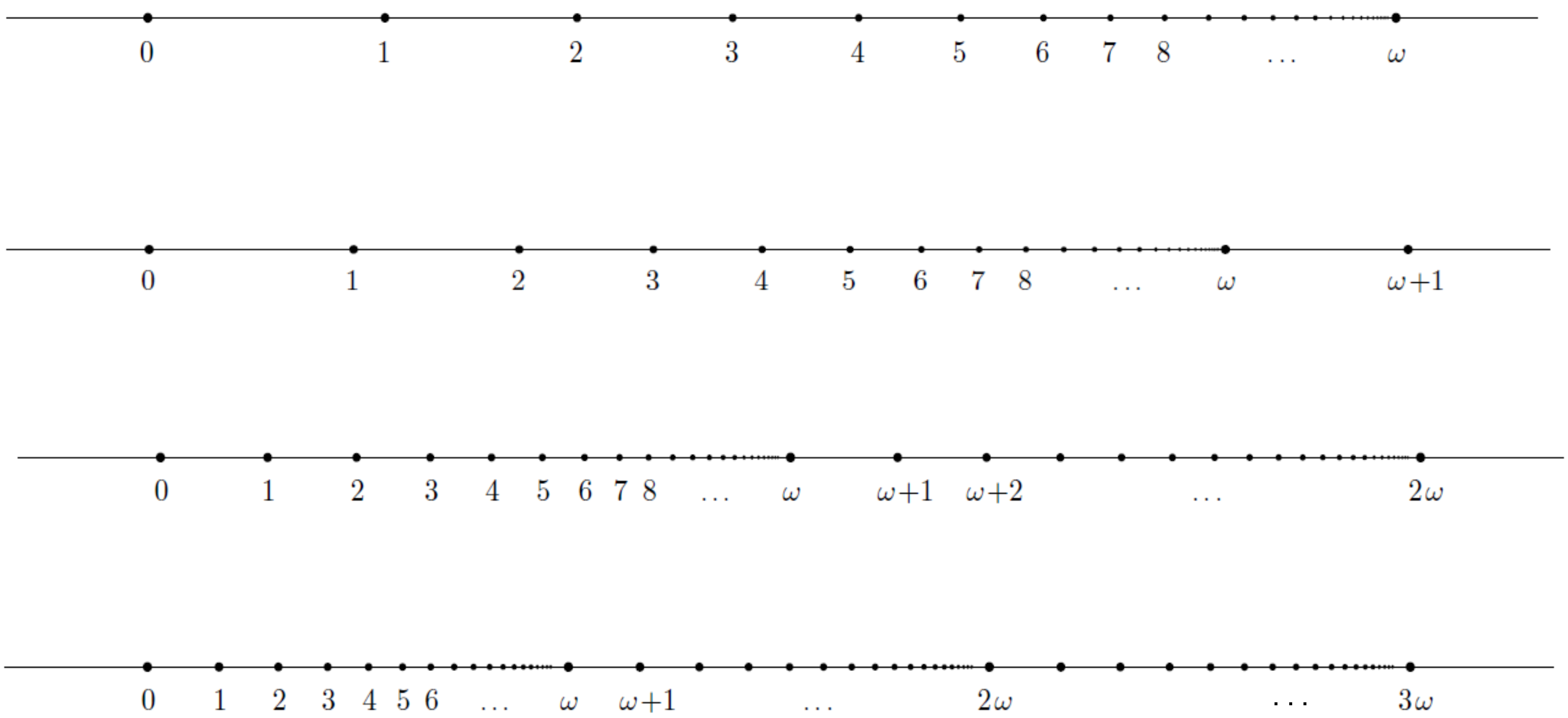
Il y a un problème de choix de notation avec 2ω : il faut pouvoir distinguer entre l'objet ci-dessus, dont l'ordre est visiblement d'un autre « type » que celui de ω, et l'objet qu'on obtiendrait en « doublant » chaque point entier, dans une suite $0, 0, 1, 1, 2, 2, \ldots$ qui elle, au contraire, aurait le même *type d'ordre* que ω.
Si j'ai dit 2ω pour le premier cas, je devrai dire $\omega.2$ pour le second, pour distinguer l'un de l'autre.

Cantor choisit la convention inverse dans ses articles de 1895 et 1897 :
pour lui alors (et aujourd'hui, j'en ai bien peur, pour la quasi-totalité des gens sérieux), notre exemple 2ω ci-dessus doit s'appeler $\omega.2$, cependant...

En 1880, Cantor n'a pas encore mis au point toutes ses notations, mais il a déjà les éléments qu'on a vus dans les pages précédentes et ceux qu'on verra dans la suite, qui apparaissent d'abord pour lui dans le cadre de la notion de *dérivation des ensembles*, et plus précisément, dans le cas d'une infinité de « dérivations » successives.

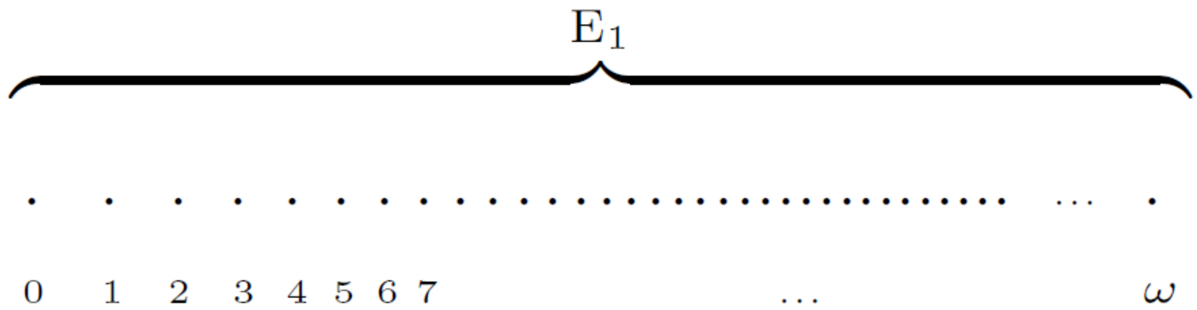
Au lieu de ω, qui n'apparaîtra qu'un peu plus tard en 1883, Cantor utilise en 1880 la notation ∞ ; après ∞, Cantor introduit $\infty + 1$, puis $\infty + n$, pour n entier, et après cette infinité des $\infty + n$, il note $2\,\infty$, comme nous avons fait avec $2\,\omega$ page précédente, puis il introduit $3\,\infty$, etc. En 1883, il note encore $2\,\omega$ (p. 560, p. 577) et explique que $\omega.2 = \omega$ (p. 559) ; il intervertira donc ces notations en 1895.

On ne regarde plus que la structure d'ordre :



La « récurrence topologique » s'applique encore ici pour avoir de nouveaux « ensembles exceptionnels » acceptables pour le théorème d'unicité du développement trigonométrique ;

ce sont de nouveaux « ensembles d'unicité ».



il y a maintenant trop de points,
on n'arrivera plus à dessiner,
écrivons de manière symbolique :
« E_1 » pour représenter d'un seul
coup toute la suite des points ci-dessus
(y compris la limite).

$$\overbrace{\bullet\quad \bullet\quad \bullet\quad \bullet\quad \bullet\ \bullet\ \bullet\ \bullet\,\bullet\,\bullet\cdots\cdots\cdots \quad \ldots\quad \bullet}^{E_1}$$

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

Considérons alors une nouvelle succession, notée symboliquement :

$$E_1\quad E_1\quad E_1\quad E_1\quad E_1\quad E_1\quad E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1\ E_1 \quad \ldots \quad \bullet$$

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

À propos de ces exemples, Cantor définit

l'ensemble dérivé F' d 'un ensemble F :

c'est l'ensemble des points limites de F

(points d'accumulation)

Le dérivé d'un ensemble fini est vide, mais :

$[0, 1]' = [0, 1]$,

ensemble *parfait*.

La « dérivation » remplace E_1 par un point unique : la limite.

Le deuxième ensemble dérivé de E_1 est vide.

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

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

$E_1 \quad \ldots \quad \cdot$

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

$$\overbrace{E_1 \quad \ldots \quad \cdot}^{E_2}$$

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

Le troisième ensemble dérivé $(E_2)'''$ est vide

$E_2 \quad \ldots \quad \cdot$

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

Cantor manipule déjà ces exemples en 1872.

si on a une *suite* d'exemples E_n de plus en plus compliqués,
on peut les rassembler en un exemple encore plus compliqué :

$$\underbrace{\begin{array}{ccccccccccc} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 & E_7 & \ldots & E_n & \ldots & \cdot \\ \omega & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 & \omega^7 & \ldots & \omega^n & \ldots & \omega^\omega \end{array}}_{E}$$

un ***ordinal*** est un « type d'ordre » (qui représente tous les ensembles ordonnés « du même type »).

La liste ci-dessus, de ω jusqu'à ω^ω, est une liste d'ordinaux.

Nous pouvons poursuivre pour tous les ordinaux dénombrables, représentés « concrètement » par un ensemble E de points de $[0, 2\pi]$, le raisonnement qui a fait de E un ensemble exceptionnel admissible pour le théorème d'unicité trigonométrique (un *ensemble d'unicité*). Autant que je me rappelle, Cantor ne l'a pas fait au delà des ω^n dans ses publications : il est passé à d'autres préoccupations, que nous allons voir.

En parallèle avec la conception des ordinaux, Cantor met en place plusieurs notions de la topologie de la droite, du plan ou plus généralement de l'espace. Il s'intéresse aux points d'accumulation d'un ensemble, qui forment son *ensemble dérivé*. Il introduit aussi la notion *d'intersection d'une suite d'ensembles*, qui permet −entre autres choses− de définir un ensemble dérivé d'ordre ω (pages suivantes).

Cantor possède les notions d'ensemble fermé et d'ensemble ouvert (le complémentaire d'un ensemble fermé). Il considère en particulier les ensembles fermés qu'il appelle *parfaits* **:** ce sont ceux dont tout point est un point d'accumulation, c'est-à-dire ceux qui sont égaux à leur dérivé (un exemple élémentaire est fourni par les segments $[a, b]$ où $a < b$ **;** un exemple moins élémentaire est *l'ensemble triadique* qu'on verra plus loin). Un ensemble parfait de la droite (ou de l'espace) possède toujours la puissance du continu.

On a dit qu'on pourrait poursuivre pour tous les ordinaux dénombrables le thème déjà vu des *ensembles d'unicité* trigonométrique. Mais ces ensembles ordonnés que nous avons commencé de décrire, Cantor va maintenant les utiliser d'une manière plus abstraite : ils vont servir à *énumérer* les itérations de l'opération de dérivation des ensembles ; après avoir considéré l'intersection $F^{(\omega)}$ des dérivés d'ordre fini d'un ensemble fermé F, il est possible que $F^{(\omega)}$ ait encore un dérivé non vide qui sera « compté » $\omega+1$, ainsi de suite, on pourra peut-être avoir à *compter* des ensembles dérivés jusqu'à ω^{ω}, et même au delà...

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

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

L'ensemble dérivé d'ordre ω d'un ensemble F est l'intersection de ses dérivés d'ordre fini

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

$$\underbrace{\begin{array}{ccccccccccc} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 & E_7 & \ldots & E_n & \ldots & \bullet \\ \omega & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 & \omega^7 & \ldots & \omega^n & \ldots & \omega^\omega \end{array}}_{E}$$

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

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

E_1	E_2	E_3	E_4	E_5	E_6	E_7	$\ldots$	E_n	$\ldots$	$\cdot$
ω	ω^2	ω^3	ω^4	ω^5	ω^6	ω^7	$\ldots$	ω^n	$\ldots$	ω^ω

un ordinal est un « type d'ordre » (qui représente tous les ensembles ordonnés « du même type »)

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

par exemple :

$$\mathbb{N}, \qquad \omega\,,$$

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

sont tous bien ordonnés.

Cantor envisage de continuer ***indéfiniment*** le processus de création ou de « production » de nouveaux ordinaux ;

« Erzeugungsprincip »

il envisage même d'épuiser la droite réelle de cette façon, en plaçant tous les réels dans une très lo...ongue liste bien ordonnée (mais pas avec l'ordre usuel évidemment).

Il faut alors imaginer une succession « continue » d'opérations, comportant autant d'étapes qu'il y a de nombres réels !

Beaucoup de ses contemporains ne sont pas d'accord.

Cantor énonce deux « principes de production » d'ordinaux : le premier principe permet de considérer le successeur $\alpha + 1$ d'un ordinal α existant ; le second principe introduit un *ordinal limite*, tel que ω, ω^n ou ω^ω, qui est le premier ordinal plus grand que ceux qu'on a déjà *produits*.

(c'est l'orthographe de l'article de 1883, « réformée » en *Erzeugungsprinzip* dans le volume des « Œuvres » qui a été publié en 1932)

Cantor invente la

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.*

Faut être «gonflé» pour prétendre inventer une nouvelle façon de démontrer et de raisonner, dans des mathématiques déjà bi-millénaires,

et de plus, une méthode fondée sur ces ordinaux qui sont de pures vues de son esprit.

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* »

Les preuves par récurrence ordinale, ainsi que la construction d'ensembles par récurrence ordinale, jouent encore aujourd'hui un rôle essentiel en logique mathématique, notamment dans les travaux de théorie axiomatique des ensembles (la définition des ensembles dérivés de Cantor est un premier exemple de construction d'ensembles par récurrence ordinale).

Les notions ordinales de Cantor rencontrent l'opposition des « finitistes », au rang desquels on peut compter Kronecker, très puissant à Berlin et également influent en Europe.

Page suivante, Cantor se sent obligé de se défendre de l'accusation qu'on pourrait peut-être formuler aujourd'hui comme « accusation de pollution de l'environnement mathématique » !

 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*.

Traduction partielle et très approximative :

« Je crois qu'il n'y a pas lieu de craindre, comme certains le font, que ces principes [*pour la création des nombres ordinaux*] représentent un quelconque danger pour la science. En effet, il n'y a pas vraiment de place pour l'arbitraire en ce domaine **:** si une notion est infructueuse ou inadéquate, elle le montrera très vite par son absence d'applicabilité, et elle sera abandonnée à cause de ce manque de succès.

Je pense en revanche que toute limitation superflue du champ des recherches mathématiques apporte avec elle un bien plus grand danger **;** car l'essence des mathématiques réside précisément dans leur liberté. »

Revenons à des questions de cardinalité →

Suites d'entiers et continu

ici 1 est exclus

Rappel : Développement en base 2 de $x \in [0, 1)$;

la notation

$$x = {}^{2}\overline{0, a_1\, a_2\, \ldots\, a_k\, \ldots}\,, \qquad a_k = 0 \text{ ou } 1$$

signifie que

$$x = \sum_{k \geqslant 1} \frac{a_k}{2^k} = \frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \cdots$$

développement propre : infinité de $a_k = 0$

$$ {}^{2}\overline{0, 0\,1\,1\,1\, \ldots\, 1\, \ldots} = {}^{2}\overline{0, 1\,0\,0\,0\, \ldots\, 0\, \ldots}$$

impropre :
que des 1 à partir
d'un certain rang

propre :
infinité de 0

Les problèmes liés aux *développements impropres* proviennent de l'égalité

$$\sum_{n=1}^{\infty} 2^{-n} = 1.$$

Si un nombre est représenté en base 2 par un développement (impropre) qui ne comporte que des **1** à partir d'un certain rang, il peut aussi être représenté par un autre développement, qui ne comporte que des **0** à partir d'un certain rang **:**

$$^{2}\overline{0,\,1\,1\,1\,\ldots\,1\,\ldots} \;=\; ^{2}\overline{1,\,0\,0\,0\,\ldots\,0\,\ldots}$$

Un développement impropre est caractérisé par la suite ***finie***, terminée par un **0**, qui précède l'infinité de **1 ;**
les développements impropres sont donc
en quantité dénombrable.

Suites d'entiers et continu

exemple :

$$x = \sum_{k \geqslant 1} \frac{a_k}{2^k} = \frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \cdots$$

$$\begin{aligned} {}^{2}\overline{0,\,0\,1\,0\,1\,\ldots\,0\,1\,\ldots} &= \frac{1}{4} + \frac{1}{16} + \cdots + \frac{1}{4^n} + \cdots \\ &= \frac{1}{4}\,\frac{1}{1 - 1/4} = \frac{1}{3}. \end{aligned}$$

Chaque réel x de $[0, 1)$ possède un unique développement propre

de la forme ci-dessus : $x = \overline{0, \ldots}$

Julius König (1849 – 1913)

Suites d'entiers (adapté d'une variante due à Julius König, apparue un
et nombres réels peu après la version donnée par Cantor en 1877)

Exemple : à la suite $3,\ 0,\ 1,\ 0,\ 2,\ 5,\ 0,\ 2,\ \ldots$

on associera

et inversement, si le développement est propre

$$x = {}^{2}\overline{0,\underbrace{1\ 1\ 1\ 0}_{3}\ \underbrace{0}_{0}\ \underbrace{1\ 0}_{1}\ \underbrace{0}_{0}\ \underbrace{1\ 1\ 0}_{2}\ \underbrace{1\ 1\ 1\ 1\ 1\ 0}_{5}\ \underbrace{0}_{0}\ \underbrace{1\ 1\ 0}_{2}\ \ldots}$$

le principe du codage :

entier k $\longrightarrow$ k chiffres 1 suivis de un 0

en particulier :

entier 0 $\longrightarrow$ 0 chiffre 1 « suivi » de un 0

$$x = {}^{2}\overline{0,\underbrace{1\;1\;1\;0}_{3}\;\underbrace{0}_{0}\;\underbrace{1\;0}_{1}\;\underbrace{0}_{0}\;\underbrace{1\;1\;0}_{2}\;\underbrace{1\;1\;1\;1\;1\;0}_{5}\;\underbrace{0}_{0}\;\underbrace{1\;1\;0}_{2}\quad\ldots}$$

ou encore, sous forme de série :

$$x = \underbrace{\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}}_{3} \underset{\underset{0}{\uparrow}}{+} \underset{\underset{1}{\uparrow}}{\frac{1}{2^6}} \underset{\underset{0}{\uparrow}}{+} \underbrace{\frac{1}{2^9}+\frac{1}{2^{10}}}_{2} + \underbrace{\frac{1}{2^{12}}+\frac{1}{2^{13}}+\frac{1}{2^{14}}+\frac{1}{2^{15}}+\frac{1}{2^{16}}}_{5} \underset{\underset{0}{\uparrow}}{+} \underbrace{\frac{1}{2^{19}}+\frac{1}{2^{20}}}_{2} + \cdots$$

Suites d'entiers

on vient de voir que :

ens. des suites d'entiers

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

(par ce signe on veut dire :
peuvent être mis en bijection)

Suites d'entiers

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

Les suites de couples d'entiers

ou aussi bien les couples de suites d'entiers,

$$(m_0, n_0),\ (m_1, n_1),\ (m_2, n_2),\ \ldots,\ (m_k, n_k),\ \ldots$$

tels que : $\big((m_0, m_1, m_2, \ldots),\ (n_0, n_1, n_2, \ldots) \big)$

sont en bijection avec les suites d'entiers

$$m_0,\ n_0,\ m_1,\ n_1,\ m_2,\ n_2,\ \ldots,\ m_k,\ n_k,\ \ldots$$

simplement en intercalant les éléments successifs
des ***deux*** suites pour former ***une*** nouvelle suite.

Le continu sous plusieurs formes

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

ens. des couples de réels

ens. des suites d'entiers

ens. des couples de suites d'entiers

Suites d'entiers et continu :

revenons sur la première idée : le mélange des « décimales » ;

à

$$x = {}^{2}\overline{0,\ a_1\, a_2\ \ldots\ a_k\ \ldots}\ ,$$

on associe le couple (y, z) :

$$y = {}^{2}\overline{0,\ a_1\, a_3\, a_5\ \ldots\ a_{2k+1}\ \ldots}\ ,$$

$$z = {}^{2}\overline{0,\ a_2\, a_4\, a_6\ \ldots\ a_{2k}\ \ldots}$$

Problème !

Le nombre $\longrightarrow$ $x_1 = {}^2\overline{0,\,0\,0\,0\,1\,0\,1\,0\,1\ldots 0\,1\ldots}$

est ainsi associé

au couple $\mapsto$ $y = 0, \quad z = {}^2\overline{0,\,0\,1\,1\,1\ldots 1\,1\ldots}$ (dév. impropre)

qui est identique au couple :

$$y = 0, \quad z = {}^2\overline{0,\,1\,0\,0\,0\ldots 0\,0\ldots}$$

ce second couple provenant de $\longrightarrow$ $x_2 = {}^2\overline{0,\,0\,1\,0\,0\ldots 0\,0\ldots}$ qui est différent de x_1

$x \longrightarrow (y, z)$ est surjective, mais pas injective (comme la courbe de Peano) ; il faut travailler plus, si on veut obtenir une *bijection* de cette façon.

Le statut du « continu » est clarifié :

Le continu sous plusieurs formes

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

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

du point de vue *cardinalité* :

c'est devenu très BANAL !

Pourtant . . .

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

(mais pas topologiquement : Brouwer le prouve en 1911)

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 !*].

Ein Beitrag zur Mannigfaltigkeitslehre.

[Crelles Journal f. Mathematik Bd. 84, S. 242—258 (1878)].

dénombrable

et entre les deux ?

continu

1878 : Cantor suppose qu'il n'y a pas de cardinalité entre ces deux-là,

c'est

l'hypothèse du continu.

Cantor a montré qu'il existe une bijection de $\mathbf{R}^m$ sur $\mathbf{R}^n$ **;** il pense cependant que dans le cas d'applications *continues*, la situation est différente, mais il n'arrivera pas à le démontrer (Cantor a publié une preuve en 1879, mais elle n'est pas correcte). Précisément, lorsque $m \neq n$, il n'existe pas d'application bijective bicontinue d'un ouvert non vide de $\mathbf{R}^m$ sur un ouvert de $\mathbf{R}^n$.

Ce résultat difficile ne sera prouvé qu'en 1911 par Brouwer. Le cas où $m = 1$ est facile **:** une preuve imprécise mais rapide consiste à remarquer que si on enlève un point à un intervalle ouvert I de la droite, on obtient deux « bouts » ouverts (connexes) disjoints, alors qu'après avoir enlevé un point à un ouvert de $\mathbf{R}^n$, où $n > 1$, il n'y a toujours qu'un seul « bout » (l'ensemble qui reste est encore « connexe »).

Donnons une preuve plus complète du cas $m = 1$ **:** si Q est une boule ouverte dans $\mathbf{R}^n$ et f une fonction réelle continue non constante sur Q, considérons deux points x et z tels que $f(x) \neq f(z)$, et supposons que $f(x) < 0 < f(z)$ pour fixer les idées.

Si on considère le « chemin » g_1 (une fonction continue de [0, 1] dans Q) qui parcourt le segment de x à z, chemin défini par $g_1(t) = (1-t)x + t\,z$, et si on pose $h_1(t) = f(g_1(t))$, on aura $h_1(0) < 0$ et $h_1(1) > 0$ **;** il existe donc par le théorème des valeurs intermédiaires une valeur t_1 telle que $h_1(t_1) = 0$, c'est-à-dire que $f(g_1(t_1)) = 0$. Mais dans $\mathbf{R}^n$, quand $n > 1$, on peut considérer un deuxième chemin g_2 de x à z tel que $g_2(t)$ ne soit *jamais* sur le segment $[x, z]$, sauf quand $t = 0$ et $t = 1$. Pour la même raison il existe t_2 tel que $f(g_2(t_2)) = 0$, t_2 n'est égal ni à 0 ni à 1, on a par conséquent $g_2(t_2) \neq g_1(t_1)$, la fonction f n'est donc pas injective.

Avant le théorème de Brouwer, un cas particulier a été obtenu par Jacob Lüroth **:** la bijection continue est impossible si $1 \leq m \leq 3$ et $m < n$ (résultat publié en 1907, mais obtenu bien avant, nous dit Lüroth dans son article).

Jacob Lüroth (1844–1910)

Cantor est aussi un des précurseurs de la théorie moderne de la mesure. Il sait qu'un ouvert U de [0, 1] peut se représenter comme réunion dénombrable d'intervalles ouverts deux à deux disjoints, et il considère bien naturellement que la longueur l(U) de cet ouvert U est la somme de la série des longueurs de ces intervalles disjoints qui composent U.

Pour un fermé F de [0, 1], Cantor définit *le contenu* de F comme étant $1 - l$(U), où U est le complémentaire de F dans [0, 1]. Ces deux « mesures », celle des ouverts et celle des fermés, coïncident avec les valeurs données par la théorie moderne (de Borel, de Lebesgue).

Cantor a considéré le fameux *ensemble triadique*, il sait qu'il est *parfait* et de contenu nul (de mesure nulle). Il invente ensuite la *fonction de Cantor,* fonction f qui est croissante au sens large mais qui est constante sur chaque composante ouverte du complémentaire de l'ensemble triadique. De plus $f(1) - f(0) > 0$, la fonction f varie, mais seulement sur un ensemble de mesure nulle : c'est encore un exemple qui a un graphe « indessinable ».

Émile Borel (1871–1956) Henri Lebesgue (1875–1941)

Si on ne peut pas tracer le graphe de la fonction de Cantor f (on peut tout de même *l'imaginer* à partir d'un dessin), la définition analytique de cette fonction f est simple et rapide, au moins pour les points x de l'ensemble triadique de Cantor : si

$$x = 2\sum_{n=1}^{\infty} u_n 3^{-n}, \qquad u_n = 0, 1,$$

$$x = \sum_{n=1}^{\infty} v_n 3^{-n}$$

$$v_n = 0, 2.$$

est un point quelconque de l'ensemble triadique, alors

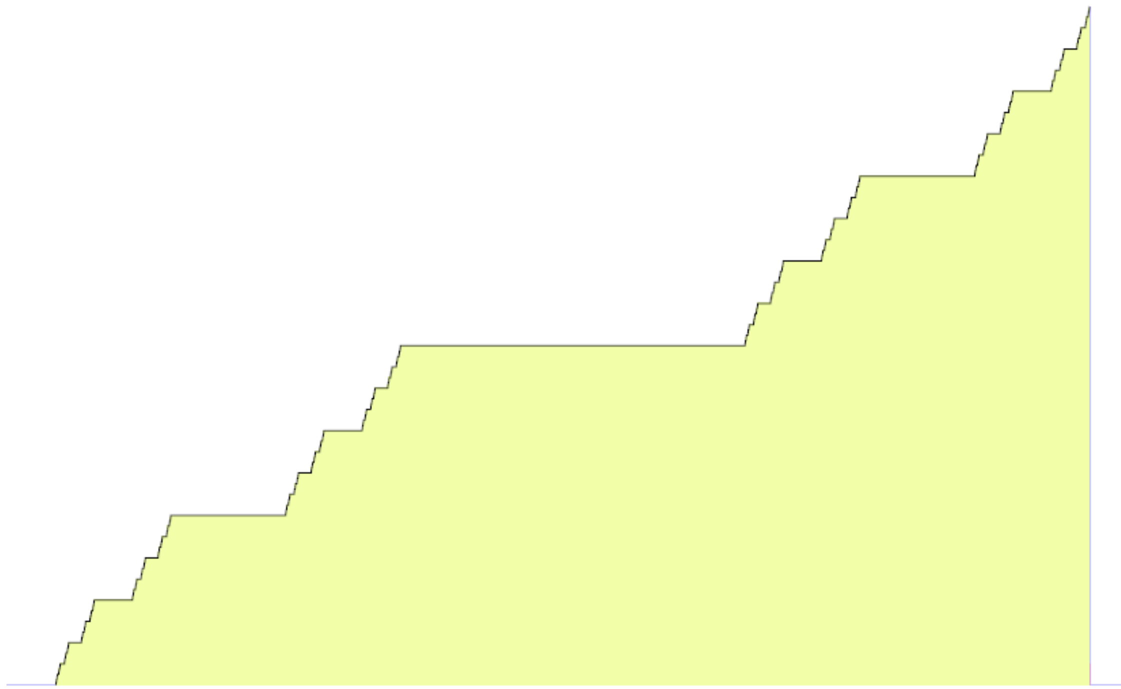
Si x est un point de l'ensemble triadique, son développement en base 3 ne comporte que des **0** et des **2**.

$$f(x) = \sum_{n=1}^{\infty} u_n 2^{-n}.$$

Si y n'est pas dans l'ensemble triadique, son développement comporte un premier $v_n = 1$ à un certain rang n_0. Si on annule toutes les « tri-males » v_n de y à partir de la place n_0 incluse, on obtient un point x de l'ensemble triadique, et la formule $f(y) = f(x)$ complète la définition de f.

Connaissez-vous la fonction de Cantor ?

découverte
fin 1883 ;
un complément
à « l'ensemble
(triadique)
de Cantor »



en tout cas, vous avez déjà vu « l'argument diagonal » :
on en reparle après une courte pause.

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. C'est dans l'une de ces lettres que Cantor explique à Mittag-Leffler « sa fonction », avec un croquis dans la marge de la lettre.

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)

(une observation avant de passer à l'argument diagonal **:** on lit souvent qu'Émile Borel a prouvé que l'ensemble des rationnels est de mesure nulle en remarquant que cet ensemble peut être contenu dans la réunion d'une suite d'intervalles dont la somme des longueurs a été choisie arbitrairement petite. Je ne trouve pas que cette présentation des choses soit vraiment correcte **;** en effet, on n'a pas attendu Borel pour faire cette remarque sur la somme des longueurs **:** Axel Harnack l'a faite dès 1885. Mais on a attendu Borel pour *décider* que cette information signifie que la mesure des rationnels *doit être* nulle)

Reprenons :

en tout cas, vous avez déjà vu « l'argument diagonal »

Axel Harnack (1851–1888)

Suite diagonale et cardinalités

Définir un sous-ensemble E de l'ensemble $\mathbb{N}$ des entiers revient à donner une suite $(x_n)_{n\geqslant 0}$ de 0 et de 1, par exemple en posant

$$x_n = 1 \iff n \in \mathrm{E}.$$

Exemple : si les entiers commencent avec l'entier 0, on associera

(pour certains c'est avec 1 que les entiers commencent)

E = l'ensemble des multiples de 3 : 0, 3, 6, 9, . . .

avec la suite :

(1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, . . .)

0 3 6 9 ← places numérotées occupées par des 1

La page suivante propose quelques petits jeux sur les cardinalités, en particulier le suivant **:** si X est un ensemble infini et D un ensemble dénombrable, la cardinalité de la réunion $X \cup D$ est égale à celle de X.

Il en résulte que l'ensemble $\mathscr{P}(\mathbf{N})$ des parties de **N** a la même cardinalité que [0, 1). On pourra donc par conséquent donner l'argument diagonal (qui doit montrer que [0, 1) n'est pas dénombrable) sous la forme suivante **:** l'ensemble $\mathscr{P}(\mathbf{N})$ n'est pas dénombrable, c'est-à-dire que l'ensemble formé des éléments d'une suite dans $\mathscr{P}(\mathbf{N})$ ne peut pas constituer la totalité de $\mathscr{P}(\mathbf{N})$.

Comme la donnée d'une suite d'éléments est la même chose que la considération d'une application définie sur **N**, il s'agira de voir qu'une application de **N** dans $\mathscr{P}(\mathbf{N})$ n'est *jamais surjective*.

Petits jeux de cardinalité

Ici, symbole pour : réunion disjointe

$$\mathbf{N} \equiv \mathbf{N} \oplus \mathbf{N}$$

(deux « copies » disjointes)

$$\mathbb{N} \equiv \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\right\} \subset [0, 1)$$

ce sont deux ensembles complémentaires dans [0, 1)

$$[0, 1) \equiv \mathrm{Y} \oplus \mathbb{N} \equiv \mathrm{Y} \oplus \mathbb{N} \oplus \mathbb{N} \equiv$$

$$\equiv (\mathrm{Y} \oplus \mathbb{N}) \oplus \mathbb{N} \equiv [0, 1) \oplus \mathbb{N}$$

tous les développements dyadiques : propres ou impropres

(ensemble des sous-ensembles de $\mathbf{N}$)

$$\mathcal{P}(\mathbb{N}) \equiv [0, 1) \oplus \mathbb{N} \equiv [0, 1)$$

$$[0, 1) \equiv (0, 1) \equiv \mathbb{R}$$

Suite diagonale et cardinalités

Définir un sous-ensemble E de l'ensemble $\mathbb{N}$ des entiers revient à donner une suite $(x_n)_{n\geqslant 0}$ de 0 et de 1, par exemple en posant

$$x_n = 1 \iff n \in \mathrm{E}.$$

Considérons $\quad \phi : \mathbb{N} \to \mathcal{P}(\mathbb{N}).$

Définir une application $\phi : k \mapsto \mathrm{E}_k$ de $\mathbb{N}$ dans l'ensemble $\mathcal{P}(\mathbb{N})$ des parties de $\mathbb{N}$ revient alors à donner une suite $(x^{(k)})_{k\geqslant 0}$ de suites formées de 0 et de 1 ; on a : $x^{(k)} = (x_n^{(k)})_{n\geqslant 0} \sim \mathrm{E}_k = \phi(k)$.
La suite diagonale (y_n), formée de 0 et de 1, est définie par

$$y_n \neq x_n^{(n)} \quad \text{c-à-d.} \quad y_n = 1 - x_n^{(n)}.$$

$$y_n \neq x_n^{(n)} \quad \text{c-à-d.} \quad y_n = 1 - x_n^{(n)}.$$

Si $\mathrm{F} \subset \mathbb{N}$ est le sous-ensemble associé à la suite (y_n),

$$n \in \mathrm{F} \;\Leftrightarrow\; y_n = 1 \;\Leftrightarrow\; x_n^{(n)} = 0 \;\Leftrightarrow\; n \notin \mathrm{E}_n \;\equiv\; n \notin \phi(n)$$

effet « contrariant »

$$n \in \mathrm{F} \;\Leftrightarrow\; y_n = 1 \;\Leftrightarrow\; x_n^{(n)} = 0 \;\Leftrightarrow\; n \notin \mathrm{E}_n \;\equiv\; n \notin \phi(n)$$

Cas général (1891) :

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

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

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

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

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

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

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

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

Si Y est un sous-ensemble non vide de Z, il existe une surjection s de Z sur Y (convenablement formulée, l'affirmation s'étend au cas $\mathrm{Y} = \emptyset$) : on choisit $y_0 \in \mathrm{Y}$, on pose $s(y) = y$ quand $y \in \mathrm{Y}$ et $s(z) = y_0$ quand $z \notin \mathrm{Y}$.

Si j est une injection de X dans Z et $\mathrm{Y} = j(\mathrm{X}) \subset \mathrm{Z}$, on peut considérer j comme une bijection f de X sur Y, et si s est une surjection de Z sur Y, l'application $f^{-1} \circ s$ est une surjection de Z sur X.

Cantor dit que le cardinal de X est *strictement inférieur* au cardinal de Z quand il existe une injection de X dans Z, mais pas d'injection de Z dans X.

On vient de voir que s'il n'existe pas de surjection de X sur Z, il n'y a pas non plus d'injection de Z dans X. Le cardinal de X est donc strictement inférieur à celui de Z quand il existe une injection de X dans Z mais pas de surjection de X sur Z.

L'application qui associe à chaque élément $x \in \mathrm{X}$ le singleton $\{x\} \in \mathcal{P}(\mathrm{X})$ est une injection de X dans $\mathcal{P}(\mathrm{X})$, et on a vu qu'il n'existe pas de surjection de X sur $\mathcal{P}(\mathrm{X})$: le cardinal de X est strictement inférieur à celui de $\mathcal{P}(\mathrm{X})$.

S'il existe une surjection s de Z sur X, on peut concevoir une injection j de X dans Z en associant à chaque x de X un élément $j(x)$ ***choisi*** dans l'image inverse $s^{-1}(x)$ (voir plus loin l'axiome du choix).

Cardinaux

Suite d'ensembles présentant une suite strictement croissante de cardinalités :

$$\mathbb{N}, \quad \mathcal{P}(\mathbb{N}), \quad \mathcal{P}(\mathcal{P}(\mathbb{N})), \quad \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))), \quad \ldots \quad ,$$

cardina-
lités de :

ens. des
réels

ens. des
fonctions sur $\mathbf{R}$,
quelconques

ça devient vite trop
gros pour des
mathématiciens !

et l'ensemble X réunion de cette suite d'ensembles est encore plus « gros » que chacun des termes de la suite ; et $\mathscr{P}$(X) encore plus gros, et ainsi de suite !

Les ensembles selon Cantor

Le terme **:**

employé pour désigner un ensemble

« Mannigfaltigkeit » (encore en 1891)

devient définitivement « Menge » (1895)

Ensembles selon Cantor ; la « définition » suivante :

« 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. »

apparaît comme première phrase de l'article :

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."

en 1895, il est temps pour Cantor de rassembler les résultats de sa théorie : opérations sur les ordinaux et sur les cardinaux, comparaison d'ensembles bien ordonnés, . . . Cette mise au point des résultats de Cantor sur la théorie des ensembles paraît en deux parties : un premier article en 1895 et un second en 1897.

Ensembles selon Cantor :

« 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.

C'est vague !

Traduction de deux des citations en tête d'article :

Je n'avance pas d'hypothèses (Newton)

The time will come when diligent research over long periods will bring to light things which now lie hidden (Sénèque, traduction Loeb)

La définition de la « puissance » ou « cardinalité »
n'est guère plus précise :

Jeder Menge M kommt eine bestimmte „Mächtigkeit" zu, welche wir auch ihre „Kardinalzahl" nennen.

„Mächtigkeit" oder „Kardinalzahl" von M nennen wir den Allgemeinbegriff, welcher mit Hilfe unseres aktiven Denkvermögens dadurch aus der Menge M hervorgeht, daß von der Beschaffenheit ihrer verschiedenen Elemente m und von der Ordnung ihres Gegebenseins abstrahiert wird.

Les deux longs articles de Cantor sont traduits ensemble en français en 1899 ; le passage précédent devient :

A tout ensemble M correspond une « puissance » bien déterminée que nous appelons aussi son « nombre cardinal ».

Nous appelons « puissance » ou « nombre cardinal » de M, la notion générale que nous déduisons de M à l'aide de notre faculté de penser, en faisant abstraction de la nature des différents éléments m et de leur ordre.

(G. Cantor, Sur les fondements de la théorie des ensembles transfinis, Mémoires de la Société des Sciences physiques et naturelles de Bordeaux, t. III, pp. 343–437, 1899)

Passée cette imprécision de départ, Cantor obtient de nombreux résultats qu'on apprend encore aujourd'hui **:** sur les types d'ordre des ensembles totalement ordonnés (ordre des entiers relatifs et ordre de l'ensemble des rationnels, en particulier), sur les ordinaux et les ensembles bien ordonnés.
Il énonce aussi, dès la quatrième page de l'article **:**

A. *Si* **a** *et* **b** *sont deux cardinaux, on a* **a** = **b** *ou bien* **a** < **b** *ou bien* **b** < **a**.

Ce n'est pas une évidence **:** il s'agit déjà (mais ce n'est pas tout) de voir qu'étant donnés X de cardinal **a** et Y de cardinal **b**, il existe une injection de X dans Y ou une injection de Y dans X.

On peut imaginer une preuve de cette affirmation par récurrence ordinale **:**

tant que ni X ni Y n'ont été épuisés, choisir deux *nouveaux* éléments

x_α dans X et y_α dans Y **;** quand l'un des deux ensembles sera épuisé,

la correspondance $x_\alpha \leftrightarrow y_\alpha$ fournira une injection dans l'un des deux sens.

Il semble que l'article de Cantor ne contienne pas la *preuve* de la propriété **A** de la page précédente. Les éditeurs des Œuvres complètes de Cantor (dont Zermelo) emploient à propos de **A** le participe passé « behauptet » (*affirmé, prétendu*).

Cantor déduit néanmoins de la propriété **A** que :

B. S'il existe une injection de X dans Y et une injection de Y dans X, il existe une bijection entre X et Y.

Ce résultat s'appelle de nos jours le *théorème de Cantor–Bernstein* **;** la preuve donnée par Bernstein en 1896 ne dépend pas du résultat **A** de Cantor **:** c'est une preuve « directe » qui n'utilise pas l'axiome du choix, mais une simple construction par récurrence sur les entiers naturels.

Il n'est pas clair qu'on puisse prouver **A** avant **B,** en revanche, à partir de **B** et de la remarque page précédente on obtient **A**.

Felix Bernstein (1878–1956) **;** on trouve sa preuve dans le livre de Borel « *Leçons sur la théorie des fonctions* », paru en 1898 (page 104 et suivantes).

Axiomatique

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 $E \in E$?

C'est le même mécanisme que celui de la non-surjectivité de l'application ϕ vue quelques pages auparavant ;
on ne peut avoir ni $E \in E$ ni $E \notin E$.

Il faut que les gens arrêtent de dire n'importe quoi . . .
Il est temps de mettre de l'ordre !

Cantor n'est pas bouleversé par l'apparition de ces paradoxes : il a déjà compris depuis un certain temps qu'il ne faut pas parler de « l'ensemble de tous les ensembles », ni de « l'ensemble de tous les ordinaux »... mais il n'a pas de « système » à proposer pour clarifier la situation.

Il est temps de travailler sur les « fondements » !

Quelques travaux sur les « fondements » :

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 : $\forall x \, P(x) \quad \exists \, x \, P\,(x)$

Page suivante, l'écriture assez peu agréable proposée par Frege en 1879 pour introduire le quantificateur universel. Cela dit, avec ces quantificateurs, les variables, les connecteurs logiques OU, ET, NON, la relation d'égalité, on a (presque) tous les ingrédients pour écrire des mathématiques formalisées, (presque) exactement comme on le fait encore aujourd'hui.

Frege (1879), traduit en anglais :

Generality

§ 11. In the expression of a judgment we can always regard the combination of signs to the right of $\vdash\!\!\!-\!\!\!-\!\!\!-$ as a function of one of the signs occurring in it. *If we replace this argument by a German letter and if in the content stroke we introduce a concavity with this German letter in it, as in*

$$\vdash\!\!-\!\overset{\mathfrak{a}}{\smile}\!\!-\!\!-\, \Phi(\mathfrak{a}),$$

aujourd'hui on écrit :

$$\boxed{\forall a \ \ \Phi(a)}$$

this stands for the judgment that, whatever we may take for its argument, the function is a fact. Since a letter used as a sign for a function, such as Φ in $\Phi(A)$, can itself be regarded as the argument of a function, its place can be taken, in the manner just specified, by a German letter. The meaning of a German letter is subject only to the . . .

Peano (1889,
traduit du latin
en anglais)

§ 1. Numbers and addition

Explanations

The sign N means *number* (*positive integer*).
The sign 1 means *unity*.
The sign $a + 1$ means *the successor of* a, or a *plus* 1.
The sign $=$ means *is equal to*. We consider this sign as new, although it has the form of a sign of logic.

Axioms

Ici,
les *axiomes*
de Peano

1. $1 \varepsilon \mathrm{N}$.
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. $a \varepsilon \mathrm{N} . \supset . a + 1 -= 1$.
9. $k \varepsilon \mathrm{K} \therefore 1 \varepsilon k \therefore x \varepsilon \mathrm{N} . x \varepsilon k : \supset_x . x + 1 \varepsilon k :: \supset . \mathrm{N} \supset k$.

Definitions

10. $2 = 1 + 1$; $3 = 2 + 1$; $4 = 3 + 1$; and so forth.

Theorems

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

Proof:

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

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

or

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

12. $3, 4, \ldots \varepsilon \mathrm{N}$.
13. $a, b, c, d \varepsilon \mathrm{N} . a = b . b = c . c = d : \supset : a = d$.

Proof: Hyp. P 4 $: \supset : a, c, d \varepsilon \mathrm{N} . a = c . c = d . \mathrm{P}\,4 : \supset :$ Thes.

14. $a, b, c \varepsilon \mathrm{N} . a = b . b = c . a -= c := \Lambda$.

Axiomes de Peano (1889)

Axioms

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. $a \;\varepsilon\; \mathrm{N} \;.\supset.\; a + 1 \;-\!\!= 1.$
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. $\downarrow$ 6. $\downarrow$ 8. $\downarrow$

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

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

9. c'est le principe de récurrence : k est un ensemble d'entiers contenant **1** et qui contient x + 1 chaque fois qu'il contient x. Alors . . .

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

Proof:

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

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

or

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

une preuve selon Peano

(histoire sans paroles)

Dans les années 1895, l'Université de Göttingen, sous l'impulsion du secrétaire d'État Friedrich Althoff (qui pousse l'université allemande en général) et de Felix Klein pour les mathématiques à Göttingen, est en train de dépasser l'Université de Berlin en importance et en prestige.

Un recrutement très important pour les mathématiques à Göttingen est celui de David Hilbert. Hilbert est né à Königsberg ; en 1893, il est devenu professeur titulaire à Königsberg, avant d'être recruté par Göttingen en 95. Un peu plus tard, il y fait venir son ami Hermann Minkowski.

Hermann Minkowski (1864–1909)

Friedrich Althoff (1839–1908)

David Hilbert (1862–1943)

(photo prise au début
des années 1900)

installé à Göttingen
depuis 1895.

Il publie :

Grundlagen der Geometrie, 1899

Le livre de Hilbert « Fondements de la géométrie », de 1899, est traduit en français dès 1900. Hilbert y propose un système d'axiomes pour la géométrie plane et la géométrie de l'espace. Il faut noter un aspect important de cette axiomatisation **:** Hilbert précise bien qu'elle s'applique à une collection d'objets qui sont simplement des « choses », que l'on *décide* d'appeler points, droites ou plans, mais qui ne sont peut-être pas les *vrais* objets de la géométrie habituelle. Il faut seulement que les axiomes ainsi énoncés pour la géométrie soient satisfaits par ces « choses ».

Dans les pages qui suivent, on verra quelques passages caractéristiques de ce livre.

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* $ABC = \alpha$.

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.

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.

Soit c une fonction quelconque du domaine $\Omega(t)$; la fonction c

Hilbert a posé dans les « Fondements de la géométrie » plusieurs des questions qui agiteront les théories axiomatiques dans les années qui suivront. La plus cruciale bien sûr à l'époque est celle de la non-contradiction ou *consistance du système d'axiomes*. Si le système n'est pas consistant, on peut en déduire n'importe quel énoncé **:** ça ne fait pas un système bien intéressant !

Une autre question, aux implications moins radicales peut-être, est celle de *l'indépendance* des axiomes **:** a-t-on formulé des axiomes inutiles, qui résulteraient des autres axiomes ?

Pour finir, on observera que Hilbert tient à faire remarquer, à tel ou tel endroit, qu'il vient de proposer un exemple *dénombrable* de géométrie. Ce thème de la dénombrabilité reviendra, beaucoup plus loin.

Page suivante, une référence à un énorme pavé logique, Russell–Whitehead, qui a été pour pas mal d'années le *Code Napoléon* de la logique mathématique. Mais comme les citoyens ordinaires, les citoyens mathématiques n'ont pas forcément pris la peine de lire le *Code* !

Une version préliminaire :

Bertrand Russell,
The Principles of Mathematics, Cambridge 1903.

Une référence incontournable pour la logique mathématique,
pour un bon nombre d'années :

Bertrand Russell et Alfred N. Whitehead,
Principia Mathematica, Cambridge 1910, 1912, 1913.

Ernst Zermelo (en 1907)

1871 – 1953

Zermelo passe sa thèse à Berlin en 1894, sous la direction de Hermann Schwarz ; la thèse porte sur le *calcul des variations*. Zermelo s'intéresse aussi à la physique, il devient assistant de Max Planck, toujours à Berlin.

Zermelo n'avait pas une bonne santé ; il a souffert de problèmes pulmonaires et a dû interrompre ses activités d'enseignement à plusieurs reprises, ce qui a certainement nui à sa carrière universitaire.

Venu à Göttingen en 1897 pour préparer son habilitation (qu'il a soutenue en 1899, encore dans le domaine des mathématiques appliquées), il embarque avec Hilbert dans les questions de fondements des mathématiques.

Zermelo donne un énoncé précis de *l'axiome du choix* dans un article important paru en 1904.

Max Planck (1858–1947)

Théorème de Zermelo (1904)

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

C'est une conséquence de *l'axiome du choix* :

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

pour tout sous-ensemble Y non vide de X

(fonction de choix)

(ensemble des sous-ensembles
de X qui sont non vides)

Un ensemble bien ordonné est isomorphe (pour l'ordre) à un certain ordinal (uniquement déterminé)

Une application de l'axiome du choix ; si s est une surjection de Y sur X, il existe une injection j de X dans Y : on définira j en *choisissant* pour tout x de X un élément $j(x)$ dans l'image inverse $s^{-1}(x)$, sous-ensemble non vide de Y.

$$\longrightarrow \quad j(x) = \phi(s^{-1}(x))$$

En 1908, Zermelo publie un article fondateur pour la *théorie axiomatique des ensembles*. On peut remarquer que comme Hilbert dans les « Fondements de la géométrie », Zermelo s'intéresse à une collection de « choses » **:**

La théorie des ensembles s'occupe d'un domaine $\mathcal{B}$ *formé d'objets que nous appellerons simplement des « choses »* [en allemand **:** Dinge]*, dont les « ensembles » formeront une partie*.

Entre deux objets quelconques a et b du domaine $\mathcal{B}$ peut exister une relation notée $a \, \varepsilon \, b$ qui symbolise la relation intuitive d'appartenance. Les *axiomes de Zermelo* fixent les propriétés de cette relation.

Zermelo 1908 : axiomes pour la théorie des ensembles

Grundlegende Definitionen und Axiome.

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.

« choses »

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 $M \Subset N$*). Es ist stets $M \Subset M$, und aus $M \Subset N$ und $N \Subset R$ folgt immer $M \Subset R$. „*Elementenfremd*“

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

axiome d'exten-sionnalité :

Axiom I. Ist jedes Element einer Menge M gleichzeitig Element von N und umgekehrt, ist also gleichzeitig $M ⋹ N$ und $N ⋹ M$, 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.

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 peut dire aussi qu'on a 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.

Axiomes de la théorie des ensembles

$$\mathfrak{B}$$

est le « domaine » (Bericht), collection de « choses » (Dinge)

$$a \,\varepsilon\, b$$

est la relation qui peut exister entre deux objets de la collection $\mathfrak{B}$

Axiomes de la théorie des ensembles
(Zermelo)

On dit qu'un objet b du domaine $\mathfrak{B}$ est un ENSEMBLE s'il existe a tel que

$$a \,\varepsilon\, b$$

et on dit que a est ÉLÉMENT de b.

Zermelo doit faire une exception pour l'*ensemble vide* noté **0**, qui a droit au titre « d'ensemble » mais ne satisfait pourtant aucune relation $a \,\varepsilon\, \mathbf{0}$

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

Une partie des axiomes de Zermelo de 1908 est restée pratiquement inchangée jusqu'à nos jours. Mais *l'axiome de compréhension* de la page suivante est encore plutôt imprécis en 1908, en ce que les « propriétés P » admissibles pour l'application de cet axiome ne sont pas parfaitement définies.

Elles le seront dans les années qui suivront, par d'autres acteurs, et dans un sens qui ne sera peut-être pas ce qu'aurait souhaité Zermelo à l'origine.

On doit remarquer que l'axiome de compréhension n'est pas un axiome unique **:** il prend autant de formes qu'il y a de propriétés P admissibles. On parle d'un *schéma d'axiome*.

Axiome de

Compréhension :

Si $\mathrm{P}(x)$ est une propriété bien définie que peut posséder ou non un objet x quelconque du domaine $\mathfrak{B}$,
et si un objet M du domaine est un ENSEMBLE,
il existe un ENSEMBLE N, objet de $\mathfrak{B}$ dont les ÉLÉMENTS sont précisément les ÉLÉMENTS de M qui ont la propriété P :

$$\mathrm{N} = \{x \in \mathrm{M} \ : \ \mathrm{P}(x) \text{ est vraie}\}.$$

$\varnothing = \{ x \in \mathrm{M} : x \neq x \}$ serait un exemple
(mais Zermelo ne
procède pas ainsi)

ci-dessus, la « propriété bien définie » $\mathrm{P}(x)$ n'est pas très bien définie !

Vers 1920

Le « programme de Hilbert » pour les nuls (pour moi)

On aimerait réduire les preuves mathématiques
d'une théorie donnée (par exemple : l'arithmétique)
à des manipulations quasiment mécaniques à
partir des axiomes de cette théorie.

On espère pouvoir démontrer de cette manière
la non-contradiction de l'arithmétique, et plus
généralement de toutes les mathématiques.

Extrait d'un texte qu'on reverra plus loin :

Die Entwicklung der Mathematik in der Richtung zu größerer Exaktheit hat bekanntlich dazu geführt, daß weite Gebiete von ihr formalisiert wurden, in der Art, daß das Beweisen nach einigen wenigen mechanischen Regeln vollzogen werden kann.

Le développement des mathématiques en direction d'une plus grande
exactitude a conduit, comme on le sait, à ce que de larges domaines
en soient formalisés de telle manière que les preuves puissent y être
ramenées à un petit nombre de règles mécaniques.

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 » ;
von Neumann introduit une nouvelle vision des ordinaux.

Notre formulation de l'axiome de remplacement page suivante est largement aussi imprécise que celle qu'on a donnée pour l'axiome de compréhension (cet axiome de remplacement est encore, en fait, un *schéma d'axiome*). La « vraie définition », qu'on trouvera par exemple chez J. L. Krivine, « *Théorie des ensembles* », passe par la notion de relation fonctionnelle R.

Pour en dire un petit peu plus : il faut avoir un énoncé R(x, y) écrit « selon les règles », et tel que pour tout x il y ait au plus un y satisfaisant R(x, y) (c'est le caractère *fonctionnel* de la relation R ; l'unicité de y doit résulter de l'application des axiomes de la théorie et des règles générales de logique). Autrement dit, l'énoncé :

$$\left(\mathrm{R}(x, y_1) \text{ ET } \mathrm{R}(x, y_2)\right) \Rightarrow y_1 = y_2$$

doit être *démontrable*.

Pour simplifier, nous parlons page suivante d'une « fonction » f définie sur le domaine B tout entier ; la relation fonctionnelle qui doit remplacer cette fonction f dans la définition « correcte » n'est pas nécessairement définie sur le domaine B entier.

Remplacement :

Si $x \mapsto f(x)$ est une application bien définie du domaine $\mathfrak{B}$ dans lui-même, et si un objet M du domaine est un ENSEMBLE, il existe un ENSEMBLE N, objet de $\mathfrak{B}$ dont les ÉLÉMENTS sont précisément les *images* par f des ÉLÉMENTS de M :

$$\mathrm{N} = \{f(x) \,:\, x \in \mathrm{M}\},$$

deux exemples élémentaires de remplacement :

M = $\mathscr{P}(\varnothing)$ = $\{\varnothing\}$;

f(x) = a pour tout x ;

N = {a} (singleton)

M = $\mathscr{P}\mathscr{P}(\varnothing) = \{\varnothing, \{\varnothing\}\}$;

$f(\varnothing)$ = a et

f(x) = b pour tout x $\neq \varnothing$;

N = {a, b} (paire)

désormais en théorie axiomatique on peut tout ramener à un ENSEMBLE :

le couple, ensemble produit = ensemble des couples

$\langle x, y\rangle = \{x, \{x, y\}\}$, $\qquad \mathrm{X} \times \mathrm{Y}$

identifiée à

application f de X dans Y $\equiv$ son graphe G, qui est un sous-ENSEMBLE de $\mathrm{X} \times \mathrm{Y}$

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

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

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

L'ordinal représente l'entier :

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

$$0 \quad 1 = \{0\} \quad 2 = \{0, 1\} \quad 3 = \{0, 1, 2\} \quad \ldots$$

L'ordre des ordinaux :

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

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

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

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

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

Dans le point de vue de von Neumann, les ordinaux sont des ensembles, et les cardinaux correspondant à ce point de vue sont des ordinaux particuliers : un *cardinal* κ est un ordinal qui en tant qu'ensemble, ne peut être mis en bijection avec aucun ordinal $\alpha < \kappa$.

Ainsi, ω est en même temps le premier ordinal infini et le premier cardinal infini, que Cantor avait noté $\aleph_0$ (aleph-zéro ; pour Cantor, les cardinaux ne sont pas des ordinaux, ce qui explique ce « double langage »). Les cardinaux (après von Neumann) forment une sous-collection des ordinaux, et cette collection des cardinaux est elle aussi bien ordonnée par inclusion.

Le premier cardinal plus grand que $\aleph_0$ est noté $\aleph_1$. *L'hypothèse du continu* revient à demander si $\aleph_1$ est le cardinal du continu (en tout cas, le continu, étant strictement supérieur à $\aleph_0$, est supérieur ou égal à $\aleph_1$).

Et après...

Kurt Gödel vers 1925

Kurt Gödel (1906–1978)

Les théorèmes d'incomplétude

(1931)

Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I[1]).

Von **Kurt Gödel** in Wien. (1931)

(fragment de texte qu'on a déjà vu)

1.

Die Entwicklung der Mathematik in der Richtung zu größerer Exaktheit hat bekanntlich dazu geführt, daß weite Gebiete von ihr formalisiert wurden, in der Art, daß das Beweisen nach einigen wenigen mechanischen Regeln vollzogen werden kann. Die umfassendsten derzeit aufgestellten formalen Systeme sind das System der Principia Mathematica (PM)[2]) einerseits, das Zermelo-Fraenkelsche (von J. v. Neumann weiter ausgebildete) Axiomensystem der Mengenlehre[3]) andererseits. Diese beiden Systeme sind so weit, daß alle heute in der Mathematik angewendeten Beweismethoden in ihnen formalisiert, d. h. auf einige wenige Axiome und Schlußregeln zurückgeführt sind. Es liegt daher die Vermutung nahe, daß diese Axiome und Schlußregeln dazu ausreichen, alle mathematischen Fragen, die sich in den betreffenden Systemen überhaupt formal ausdrücken lassen, auch zu entscheiden. Im folgenden wird gezeigt, daß dies nicht der Fall ist, sondern daß es in den beiden angeführten Systemen sogar relativ einfache Probleme aus der Theorie der gewöhnlichen ganzen Zahlen gibt[4]), die sich aus den Axiomen nicht

page précédente :

« On n'est pas loin alors de présumer que ces axiomes et règles de déduction puissent parvenir à décider toute question mathématique que l'on peut exprimer formellement dans les systèmes logiques sus-mentionnés. Il sera montré dans la suite que ce n'est pas le cas. Au contraire, on montrera qu'il existe, dans ces deux systèmes, des questions relativement simples portant sur la théorie usuelle des nombres entiers, qui ne peuvent pas être décidées à partir des axiomes. »

Principia Mathematica, Zermelo – Fraenkel

Dans cet article, Gödel mentionne le *paradoxe du menteur* ; on va en rappeler des versions page suivante :

Pseudologos le Crétois est un menteur pathologique ;

il n'ouvre la bouche que pour proférer des mensonges.

Un jour, Pseudologos a déclaré :

« je suis un menteur »

?

Que peut-on en penser ?

Une forme plus classique de ce paradoxe :

Pseudologos le Crétois déclare :
« tous les Crétois sont des menteurs »

Au début de son article, Gödel réussit à expliquer, en une seule page (la page qui suit), l'idée de fond de son théorème **:** il s'agit de formaliser une forme du paradoxe du menteur, en produisant un certain énoncé E *qui dit de lui-même* qu'il n'est pas démontrable.

Alors, si on peut démontrer cet énoncé E, c'est qu'il est faux ! D'un autre côté, si les axiomes sont *vrais* et si les règles logiques de déduction sont correctes, ce qui est *démontrable* est *vrai*. On ne pourra donc pas démontrer cet énoncé E.

L'article de Gödel met en lumière une différence fondamentale entre ce qui est *démontrable* à partir d'un système d'axiomes et ce qui est *vrai*. Il ruine les espoirs de réaliser les versions radicales du *programme de Hilbert*.

bemerken, daß sich der Begriff „Klassenzeichen" sowie die ordnende Relation R im System PM definieren lassen. Sei α ein beliebiges Klassenzeichen; mit $[\alpha; n]$ bezeichnen wir diejenige Formel, welche aus dem Klassenzeichen α dadurch entsteht, daß man die freie Variable durch das Zeichen für die natürliche Zahl n ersetzt. Auch die Tripel-Relation $x = [y; z]$ erweist sich als innerhalb PM definierbar. Nun definieren wir eine Klasse K natürlicher Zahlen folgendermaßen:

(article de Gödel, p. 175)

$$n \varepsilon K \equiv \overline{Bew}\,[R(n); n]^{11a)} \qquad (1)$$

(wobei $Bew\, x$ bedeutet: x ist eine beweisbare Formel). Da die Begriffe, welche im Definiens vorkommen, sämtlich in PM definierbar sind, so auch der daraus zusammengesetzte Begriff K, d. h. es gibt ein Klassenzeichen S[12]), so daß die Formel $[S; n]$ inhaltlich gedeutet besagt, daß die natürliche Zahl n zu K gehört. S ist als Klassenzeichen mit einem bestimmten $R(q)$ identisch, d. h. es gilt

(traduction à suivre)

$$S = R(q)$$

für eine bestimmte natürliche Zahl q. Wir zeigen nun, daß der Satz $[R(q); q]$[13]) in PM unentscheidbar ist. Denn angenommen der Satz $[R(q); q]$ wäre beweisbar, dann wäre er auch richtig, d. h. aber nach dem obigen q würde zu K gehören, d. h. nach (1) es würde $\overline{Bew}\,[R(q); q]$ gelten, im Widerspruch mit der Annahme. Wäre dagegen die Negation von $[R(q); q]$ beweisbar, so würde $\overline{n \varepsilon K}$, d. h. $Bew\,[R(q); q]$ gelten. $[R(q); q]$ wäre also zugleich mit seiner Negation beweisbar, was wiederum unmöglich ist.

Traduction du passage signalé en bleu page précédente (avec des notes de traduction telles que ([1]), visibles page suivante) :

« Nous montrons maintenant que l'énoncé [R(q) ; q] ([1]) est indécidable dans PM ([2]). Car en admettant que l'énoncé [R(q) ; q] soit démontrable, il serait alors vrai ; mais alors, par ce qui précède, l'entier q appartiendrait à K ([3]), donc d'après (1) l'énoncé NON *Bew* [R(q) ; q] ([4]) serait satisfait, en contradiction avec l'hypothèse. Si au contraire la négation de [R(q) ; q] était démontrable, on aurait NON ($n\ \varepsilon$ K) ([5]), et *Bew* [R(q) ; q] serait satisfait. Alors [R(q) ; q] serait démontrable en même temps que sa négation, ce qui à nouveau est impossible ([6]). »

Notes du « traducteur » :

([1]) On a établi une liste R(n) de tous les énoncés R à une variable libre x, et [R(n) ; m] désigne l'énoncé obtenu en remplaçant dans R(n) la variable libre x par l'entier m. Pour l'entier q, voir la note ([3]) ci-dessous.

([2]) « PM » est pour *Principia Mathematica.*

([3]) K désigne l'ensemble des entiers n tels que [R(n) ; n] ne soit pas démontrable. Cet ensemble peut être défini par un énoncé R à une variable libre, on suppose que q est son « numéro » dans la liste R(n) : ainsi, on a $n \,\varepsilon$ K si et seulement si [R(q) ; n] est vrai ; c'est ce qu'on vient d'utiliser, juste avant la note ([3]).

([4]) « *Bew* » pour « *beweisbar* », démontrable.

([5]) Il y a probablement une faute de frappe : au lieu de « n », il faut je pense lire « q ».

([6]) Dans un système non-contradictoire, on ne peut *démontrer* un énoncé et son contraire.

Reprenons calmement.

Les axiomes sont réputés *vrais*, et
les règles de déduction correctes.
Un énoncé qui est déduit des axiomes au moyen
des règles de déduction est dit *démontrable*,

on a donc : démontrable $\Rightarrow$ vrai.

Dans un système *non contradictoire*,

E démontrable $\Rightarrow$ « NON E » n'est pas démontrable

On se place en arithmétique, les « valeurs » considérées sont des entiers naturels. On s'intéresse aux énoncés $\mathrm{E} = \mathrm{E}(\underline{x})$ à une variable libre $\underline{x}$, par exemple

$$\mathrm{E}_0 \equiv \text{« 5 divise } \underline{x} \text{ » :} \qquad\qquad \exists y\,(\underline{x} = 5\,.\,y)$$

on note $[\mathrm{E}\,;\,k]$ l'énoncé « clos » obtenu en remplaçant la variable libre x par l'entier k ;

cet énoncé clos est une *affirmation*, vraie ou fausse, démontrable ou pas : $[\mathrm{E}_0\,;\,20]$ est « 5 divise 20 », qui est à la fois vraie et démontrable.

Dans la preuve de Gödel, les énoncés sont associés injectivement à des *entiers,* et Gödel introduit des fonctions f *calculables* qui miment les principes de démonstration :

si on a $n = f(m)$, l'énoncé de numéro n peut être déduit de l'énoncé m par un certain axiome ou une certaine règle logique qui correspond à cette fonction f particulière. De cette façon, on peut concevoir qu'on pourra obtenir, en composant ces fonctions, une nouvelle fonction calculable capable d'indiquer si l'énoncé de numéro n est *démontrable*.

On considère une liste $(\mathrm{R}_n)_{n \geqslant 0}$ de tous les énoncés à une variable libre, puis on pose

$$\mathrm{K} \;=\; \{n \;:\; \textsc{non}\big(\mathrm{Dem}\,[\mathrm{R}_n\,;\,n]\big)\,\}\;;$$

(ensemble des entiers naturels n tels que l'énoncé clos $\mathrm{R}_n\,(n)$ ne soit pas démontrable)

il existe $\mathrm{S} = \mathrm{S}(x)$ énoncé à une variable libre tel que

$$n \in \mathrm{K} \;\;\Leftrightarrow\;\; [\mathrm{S}\,;\,n]$$

et il existe q tel que $\;\mathrm{S} = \mathrm{R}_q\;;$

l'énoncé

$$[\mathrm{R}_q\,;\,q] \;\equiv\; [\mathrm{S}\,;\,q] \;\Leftrightarrow\; q \in \mathrm{K} \;\equiv\; \textsc{non}\big(\mathrm{Dem}\,[\mathrm{R}_q\,;\,q]\big)$$

est indécidable !

À partir du premier théorème d'incomplétude, Gödel déduit un *second théorème d'incomplétude* **:**

si T est une théorie suffisamment riche pour pouvoir formaliser l'arithmétique, alors, il est impossible de démontrer la *consistance* de T à partir des axiomes de T.

Cela s'applique en particulier à la théorie des ensembles **:** *on ne peut pas démontrer la non-contradiction* du système d'axiomes ZF à partir des axiomes de ZF (la même remarque s'appliquerait aussi bien au système Z de Zermelo).

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 « 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.

Le langage adapté au modèle : les *énoncés* de théorie axiomatique des ensembles.

On dispose de variables libres, en quantité potentiellement illimitée, notées $x, y, z, \ldots$ ou $x_0, x_1, \ldots$. Il n'y a qu'un seul type de variables : elles sont toutes appelées à représenter (à être remplacées par) des objets $a, b, c, \ldots$ du modèle.

Sans entrer dans les détails les plus fins, les énoncés sont obtenus ainsi, par récurrence sur le nombre d'opérations de construction effectuées :

• $x = y$, $x = a$, $a = b$, $x \in y$, $a \in y$, $x \in b$, $a \in b$ sont des énoncés ;

• si $\mathrm{E}, \mathrm{E}_1, \mathrm{E}_2$ sont des énoncés déjà construits, alors NON E, E_1 OU E_2, E_1 ET E_2, $\forall x\, \mathrm{E}$ et $\exists x\, \mathrm{E}$ sont aussi des énoncés.

On se permet des abréviations : $A \Rightarrow B$ au lieu de (NON A) OU B, puis $A \Leftrightarrow B$ au lieu de $(A \Rightarrow B)$ ET $(B \Rightarrow A)$

Les axiomes de la théorie des ensembles sont des énoncés particuliers, par exemple l'axiome d'extensionnalité :

$$\forall x\, \forall y\, [\forall z\, (z \in x \Leftrightarrow z \in y) \Rightarrow x = y]$$

Modèles

une collection de
de « choses » si
vous voulez !

Modèle de la théorie des ensembles :

au sens naïf, usuel, de tous vos jours mathématiques . . .

On peut considérer qu'un modèle est un ensemble $\mathcal{M}$ non vide qui a une structure de *graphe orienté*.

Les arcs du graphe modélisent l'« appartenance » :
l'arc noté (a, b) ou bien a $\rightarrow$ b existe dans le graphe *M*
quand on veut exprimer a $\in$ b au sens du modèle ;

le graphe est un *modèle de la théorie des ensembles* si, muni de cette notion « d'appartenance », ce graphe vérifie les axiomes de la théorie des ensembles : on verra des exemples d'application des axiomes pages suivantes.

Modèle :

d'après les axiomes de ZF, il **doit** exister un objet b du modèle tel qu'on n'ait *jamais* $a \in b$ **;** cet objet b noté **0** *représente* l'ensemble vide.

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.**

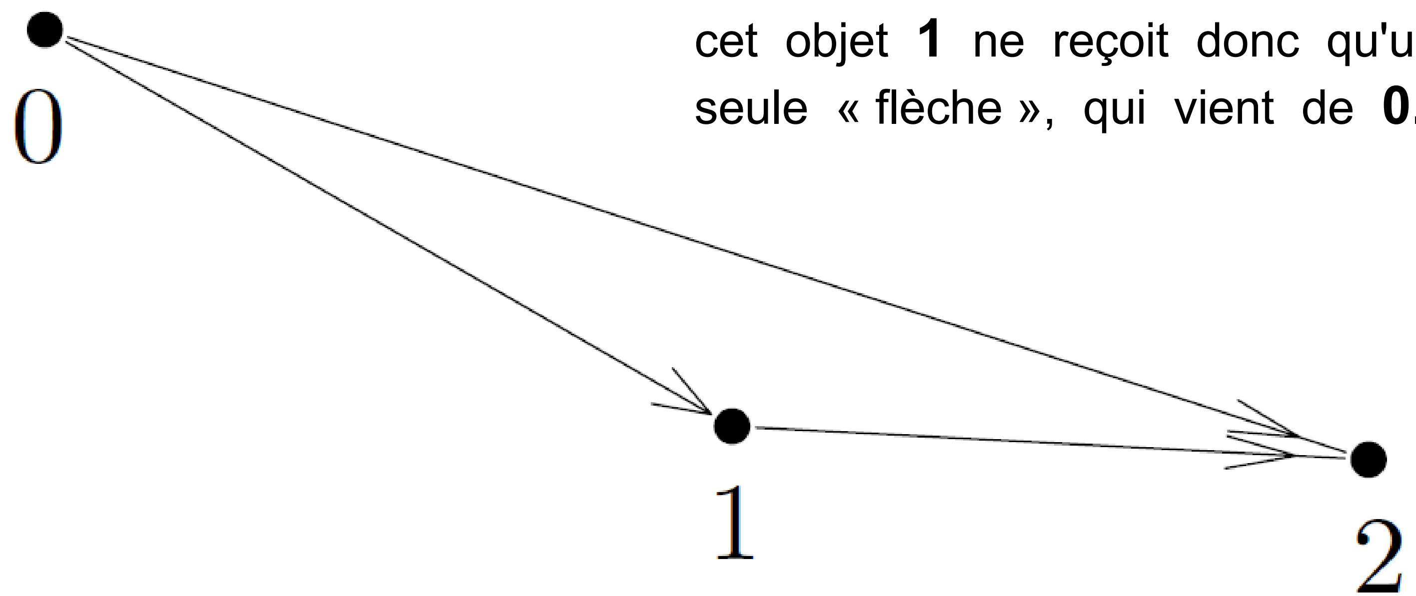
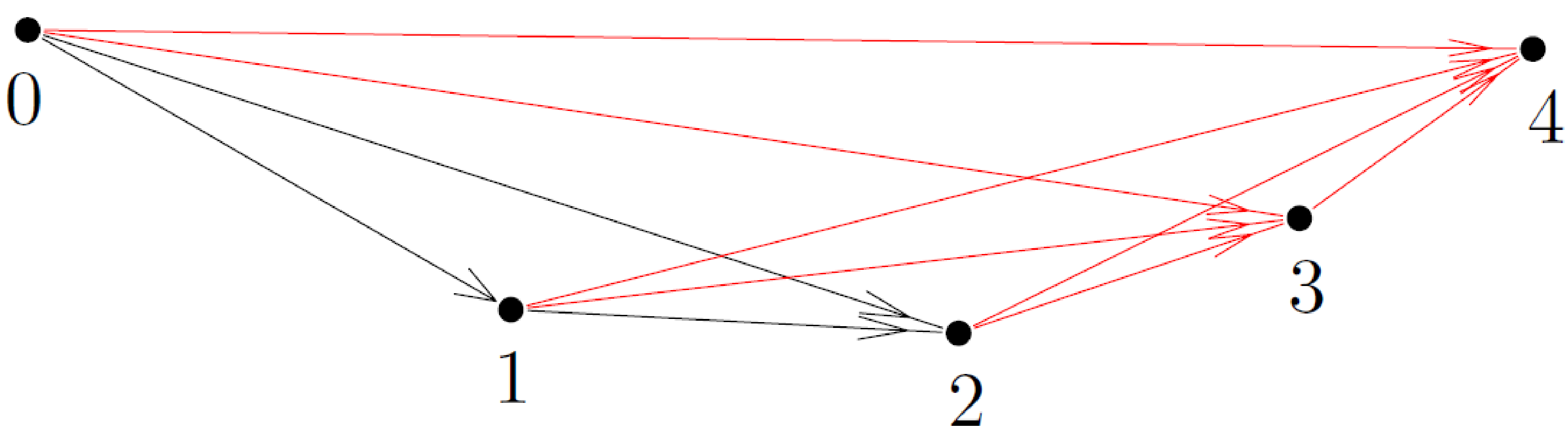


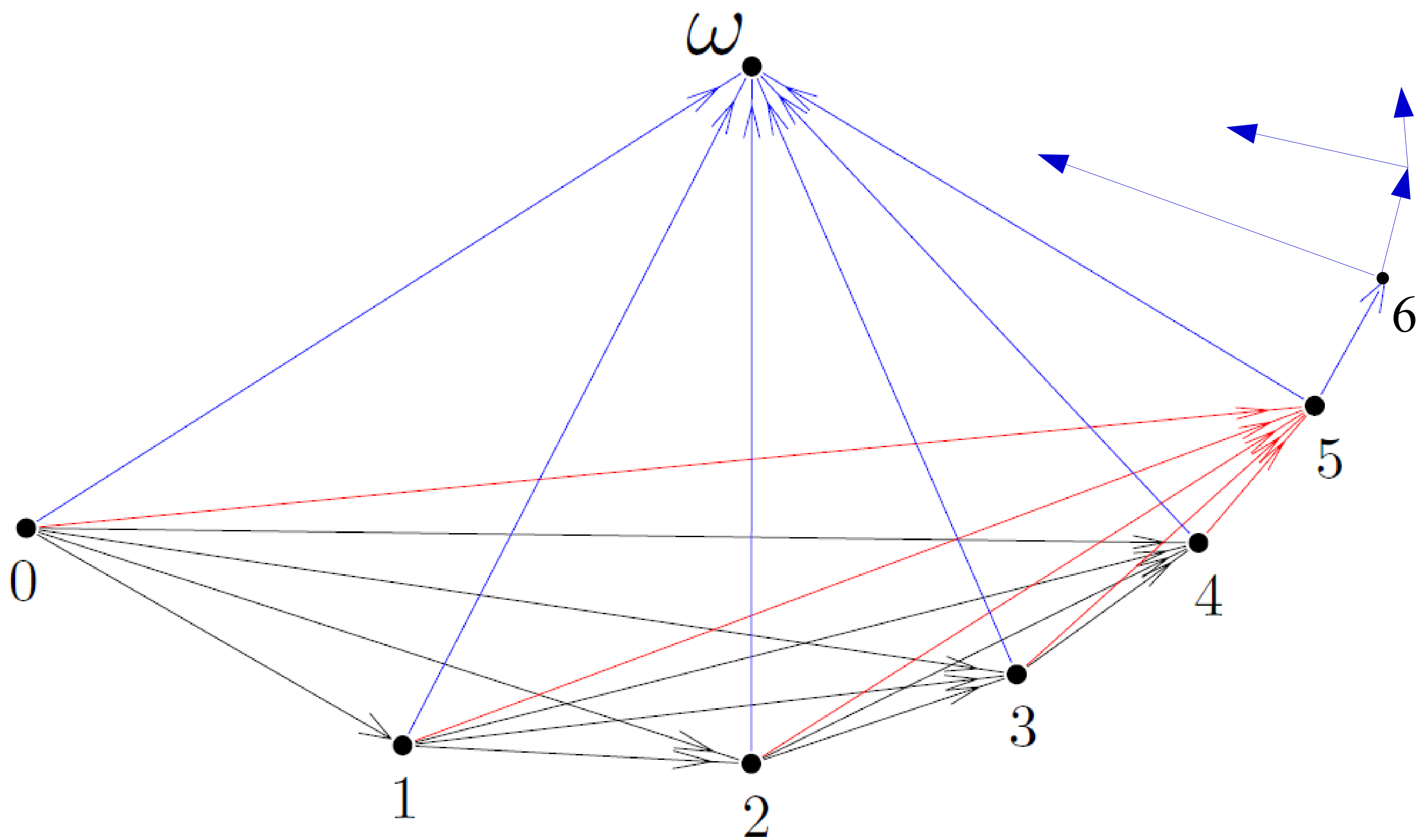
schéma pour les premiers entiers de von Neumann :
aucune flèche vers **0** (**0** modélise l'ensemble vide)

Modèle :

un schéma avec deux entiers de plus, 3 et 4 :



Modèle : une figuration « fléchée » de ω



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

Extensionnalité

Axiome de la somme $\longrightarrow$ si X est un ensemble, il existe un ensemble Y dont les éléments sont tous les *éléments des éléments* de X :

$$a \in \mathrm{Y} \Leftrightarrow \exists b,\ a \in b \in \mathrm{X}$$

ou bien : $$\mathrm{Y} = \bigcup_{b \in \mathrm{X}} b$$

Ensembles des parties

Axiome de l'infini

Remplacement

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

Modèle :

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 $\gamma+1$ (β est un *successeur*)

Modèle : à l'intérieur d'un modèle se recréent toutes les notions mathématiques, mais réinterprétées au sens du modèle ; on pourrait dire que les objets du modèle vivent dans leur monde et ignorent, ne voient pas le « monde extérieur ».

On ne peut pas « fabriquer » un modèle de la théorie des ensembles (Gödel).

Mais on peut « bricoler » un modèle existant pour construire un modèle différent, avec des propriétés additionnelles.

Certains des résultats mentionnés à la page suivante sont obtenus de cette façon : si on suppose qu'on dispose d'un modèle M, on va définir, éventuellement par construction ordinale, une certaine collection M' d'objets de M dont l'existence est garantie par les axiomes. Si cette sous-collection M' est bien choisie, elle vérifiera d'une part les axiomes de ZF (elle sera donc un *nouveau modèle* de ZF) et d'autre part, une propriété supplémentaire telle que l'axiome du choix par exemple.

Si une telle construction est possible, elle prouvera qu'on ne peut pas contredire l'axiome du choix à partir des axiomes de ZF. C'est un résultat de *consistance relative* : on peut supposer que l'axiome du choix est vrai, sans introduire de contradiction s'il n'y en avait pas déjà avant.

Quelques résultats.

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)

Le *paradoxe de Skolem* consiste en l'existence d'un modèle M de ZF qui est *dénombrable*, alors que ce modèle M doit contenir par exemple un objet qui représente l'ensemble **R** des réels, dont on a prouvé la non-dénombrabilité ! Ce « paradoxe » peut s'expliquer par une opposition *intérieur/extérieur*.

Dans le film « The Truman Show », le héros, interprété par Jim Carrey, ne se rend pas compte (pendant un long moment) qu'il est enfermé dans une (énorme) bulle observée par le monde extérieur. On peut ainsi imaginer que les bijections visibles de l'extérieur de M, qui montrent à l'observateur extérieur que le modèle M est dénombrable, ne soient pas visibles pour l'être qui est enfermé dans le modèle. Pour ce que cet être confiné peut voir, il n'existe pas de bijection entre **N** et « son » ensemble de nombres réels **;** pour lui c'est sûr **:** l'ensemble **R** n'est pas dénombrable. Le pauvre !

En présence d'un modèle dénombrable, on serait en droit d'affirmer que le modèle n'est pas une *image des mathématiques*, mais une *image de ce qu'on peut démontrer* en mathématiques.

Donnons une idée abusivement simpliste de la preuve du théorème de Löwenheim–Skolem. On suppose qu'on dispose d'un premier modèle M de la théorie. On ne peut écrire qu'une quantité dénombrable d'énoncés syntaxiquement valides, on en fait donc une liste (E_n). Pour chaque énoncé E_n qui est vrai dans M, on sélectionne dans M un objet a_n qui satisfait cet énoncé et on le place dans (on l'ajoute à) une collection M_0 qui sera, à la fin du traitement de toute la liste, le nouveau modèle dénombrable.

En réalité, les choses ne sont pas aussi simples, parce que l'adjonction des objets a_n dans M_0 a « créé » de nouveaux énoncés valides dans M_0 qu'il faut considérer aussi ! Mais on peut le faire... (on refait une liste d'énoncés, cette fois des énoncés *avec paramètres* dans M_0, on produit ainsi un M_1 contenant M_0, et on continue en une suite *dénombrable* d'étapes...)

Avec Löwenheim–Skolem, on a *simplifié* un premier modèle M donné **;** l'introduction *d'entiers non standard* procède *en sens inverse*, par la construction d'un modèle M**'** *plus complexe* que M.

Grâce à l'axiome du choix, on peut produire un *ultrafiltre* $\mathcal{U}$ sur **N**, qui permet d'introduire l'expression « pour n assez loin selon $\mathcal{U}$ » dotée de la propriété suivante **:** si (a_n) est une suite ***bornée*** d'entiers, il existe une valeur b telle que $a_n = b$ « pour n assez loin selon $\mathcal{U}$ » **:** l'ultrafiltre agit comme un super-extracteur de sous-suites convergentes, en fournissant un sous-ensemble infini U formé d'indices n tels que $a_n = b$ pour tout $n \in$ U (et $\mathcal{U}$ est l'ensemble de ces U, voir page suivante).

De la même façon qu'on a défini un réel comme classe d'équivalence de suites de Cauchy de rationnels, on obtiendra un entier non-standard ainsi **:** à chaque suite (a_n) quelconque d'entiers (non nécessairement bornée) du modèle M, on associe l'entier non standard **a** formé de toutes les suites (b_n) telles que $b_n = a_n$ « pour n assez loin selon $\mathcal{U}$ ».

Comme on a plongé les rationnels dans les réels, on « plonge » un entier ordinaire m dans les entiers non-standard en lui associant la classe de la suite constante définie par $a_n = m$ pour tout n.

Puisque « vrai » et « faux » peuvent s'interpréter comme deux valeurs entières 1 et 0, on peut dire aussi qu'une propriété P(n) dépendant d'un entier n est vraie « pour n assez loin selon $\mathcal{U}$ ». On peut alors définir un ordre $\mathbf{a} \leq \mathbf{b}$ sur les entiers non-standard en disant que : $a_n \leq b_n$ est vrai « pour n assez loin selon $\mathcal{U}$ ».

Considérons la suite non-bornée $a_n = n$; elle définit un entier non-standard $\mathbf{a}_0$ dont on voit facilement que $m < \mathbf{a}_0$ pour tout entier m standard : « l'entier » $\mathbf{a}_0$ est *infiniment grand*. Il possède un prédécesseur $\mathbf{a}_0 - 1$, qui est la classe de la suite $(a_n - 1) = (n - 1)$.

Un *ultrafiltre* $\mathcal{U}$ (non trivial) est une famille de sous-ensembles *infinis* de **N** telle que : pour tout sous-ensemble A de **N**, ou bien A ou bien son complémentaire A^c est élément de $\mathcal{U}$, et pour tous U, V $\in \mathcal{U}$, l'intersection $U \cap V$ est élément de $\mathcal{U}$.

Il y a bien pire **:** l'observateur extérieur voit que l'entier infiniment grand $\mathbf{a}_0$ de la page précédente possède *une infinité de prédécesseurs.* Pourtant, dans le modèle non-standard M* (que nous ne définirons pas ici **;** disons seulement qu'il contient les entiers non-standard) cet élément $\mathbf{a}_0$ est strictement plus petit que l'ordinal ω du modèle M* **:** $\mathbf{a}_0$ est un *entier fini au sens du modèle* M*. Chaque entier standard m a un représentant dans le modèle M*, mais l'ensemble « extérieur » **N** formé de tous les entiers standard n'a pas de représentant dans M*.

Les entiers non-standard vérifient les bons axiomes **:** enfermés dans un modèle non-standard, vous n'y verrez que du feu. De là à nous demander si nous ne vivons pas nous-mêmes dans un modèle non-standard, à notre insu **. . .**

Pour tout entier standard m on peut considérer $\mathbf{a}_0 - m$, qui provient de la suite (b_n) définie par $b_n = n - m$ si $n \geq m$ et $b_n = 0$ sinon **;** ces entiers non-standard $\mathbf{a}_0 - m$, en quantité infinie, sont des prédécesseurs successifs de $\mathbf{a}_0$.

Paul Cohen **:**

il invente (en 1963) la méthode dite de *forcing*
qui produit un modèle de ZFC où CH est fausse !

Compte tenu du résultat de cohérence relative de CH
dû à Gödel, qu'on a mentionné précédemment,
on sait maintenant que **:**

l'hypothèse du continu est

indépendante

des axiomes de ZFC

(ZFC = les axiomes usuels de la théorie des ensembles **:**

Zermelo–Fraenkel ZF avec axiome du choix AC ajouté)

Paul Cohen (1934–2007)

Arrivé à ce point, en ma ¿qualité ? de Monsieur Tout le Monde des mathématiques, je suis bien perplexe : doit-on penser qu'avec l'hypothèse du continu, Cantor s'est posé une question qui au fond n'a pas de sens, la formation de l'ensemble des parties d'un ensemble étant finalement une opération trop indécise. Ou bien, Gödel nous ayant appris qu'un système d'axiomes ne peut pas capturer toute la « vérité », doit-on se demander s'il faut considérer que l'hypothèse du continu est *vraie* (ou pas) et l'ajouter (ou son contraire) comme nouvel axiome ?

Laissant de côté ces questions existentielles, je vais essayer d'en dire un tout petit peu plus sur la méthode de Paul Cohen.

Pour moi qui ne suis pas logicien, cette méthode de Cohen apparaît comme un « tour de force » d'une complexité folle, et je suis très loin de vraiment comprendre ce qu'il s'y passe **:** par un regard extérieur au modèle M, on envisage un objet G qui est paradoxal pour le monde de M (comme était paradoxal le modèle dénombrable de Löwenheim–Skolem), et on réussit à introduire l'objet G « de force » dans un nouveau modèle M**'** de ZF qui est *plus riche* que M. Cet objet G va servir à produire une quantité de sous-ensembles de $\mathscr{P}(\mathbf{N})$ (des sous-ensembles au sens de M**'**), de « beaucoup » de cardinalités infinies différentes (toujours au sens du nouveau modèle M**'**), alors que l'hypothèse du continu n'en prévoyait que *deux*, le *dénombrable* et le *continu.*

La méthode de forcing a engendré une quantité de résultats en théorie axiomatique des ensembles ; je mentionne le suivant, qui fera plaisir aux étudiants qui ont souffert avec la notion de mesurabilité : Robert Solovay (né en 1938) a démontré que sous l'hypothèse ZF+AC+CI (pour l'axiome CI, voir page suivante), on peut construire un modèle de ZF où

tous les sous-ensembles de la droite réelle sont Lebesgue-mesurables.

Ce modèle ne peut pas satisfaire l'axiome du choix usuel, puisque celui-ci implique l'existence d'ensembles non-mesurables, mais on peut garder une forme dénombrable d'axiome du choix, qui est celle qu'on utilise très souvent en Analyse, « l'axiome du choix dépendant ».

On s'est posé précédemment la question d'ajouter l'hypothèse du continu comme nouvel axiome. Les spécialistes de théorie axiomatique des ensembles ont souvent joué à ce jeu d'adjonction d'axiomes. Un de ces axiomes qu'on a pu ajouter à ZF est *l'existence d'un cardinal inaccessible* (en abrégé axiome CI). Pour rester vague, un cardinal inaccessible est un cardinal si grand qu'on ne peut pas l'atteindre par les opérations ensemblistes définies en n'utilisant que des cardinaux inférieurs. On a vu avec Gödel qu'on ne peut pas démontrer la consistance de ZF avec les seuls outils de ZF, mais si on ajoute l'axiome CI, la conjonction ZF+AC+CI implique la consistance de ZF. Et comme d'hab, on ne peut pas démontrer la consistance de ZF+AC+CI avec les moyens de ZF+AC+CI, ça ne finit jamais...

L'*axiome du choix dépendant* (en abrégé ACD), formulé en langage naïf, est le suivant : si a est un ensemble et R une relation ayant la propriété que pour tout $x \in a$, il existe $y \in a$ tel que $x \,\mathrm{R}\, y$, alors, pour tout $x_0 \in a$, il existe une fonction f définie sur **N** telle que $f(0) = x_0$ et telle que pour tout entier n on ait $f(n) \,\mathrm{R}\, f(n+1)$ (au lieu de dire « fonction définie sur **N** », on dit plus souvent : il existe *une suite* (x_n) d'éléments de a telle que $x_n \,\mathrm{R}\, x_{n+1}$ pour tout n).

Vous allez dire : c'est bien évident, « yaka » choisir un élément après l'autre ; justement, c'est ça un axiome qu'on utilise tous les jours !

À côté de son application à la démonstration d'autres résultats nouveaux, la méthode de forcing a aussi fait l'objet de présentations sous des formes diverses. Dans les pages suivantes, on va tenter d'expliquer un peu plus précisément le ressort de la preuve de Cohen, en indiquant un petit nombre d'éléments de la version donnée par J. L. Krivine dans son livre « *Théorie des Ensembles* » déjà mentionné ; je n'oserai pas prétendre que les quelques pages qui suivent sont lisibles !

Contrairement à ce qu'on peut voir dans ce livre de Krivine, on choisira ici de ne pas employer directement, pour désigner les objets d'un modèle de ZF, les mots du langage habituel de théorie des ensembles tels que « *ensemble, éléments, contenir, couple, produit* », etc. mais de leur adjoindre systématiquement un préfixe : M-ensemble, M-élément, etc. Ce parti-pris serait vite insupportable s'il fallait le maintenir plus longtemps.

Partant d'un premier modèle $\mathcal{U}$ de ZFC, Löwenheim et Skolem ont concocté un nouveau modèle $\mathcal{M}$ qui est un ensemble dénombrable, au sens de tout le monde, le sens qu'on dit « naïf » : c'est une collection $\mathcal{M}$ que l'on peut énumérer sous la forme $\mathcal{M} = \{m_0, m_1, \ldots\}$. Pour rester dans des limites raisonnables, on va tricher, en omettant d'indiquer des propriétés de $\mathcal{U}$ et de $\mathcal{M}$ dont on a besoin pour démontrer certaines des affirmations qui seront faites plus loin sans aucune preuve. Pour ne pas dire les « choses » exactement comme Hilbert, nous dirons que $\mathcal{M}$ est une collection *d'objets*, notés $a, b, c, \ldots$; ce modèle est livré avec sa relation pour coder la notion d'appartenance, relation qui sera notée $a \sqsubset b$ (noter la petite différence avec le signe usuel). Comme on l'a dit, on peut reproduire dans le modèle $\mathcal{M}$ toutes les notions mathématiques, on les fera précéder du préfixe « $\mathcal{M}$- », ce qui aura peut-être le mérite de la précision mais ce qui sera à coup sûr rapidement très désagréable : ainsi, les objets de $\mathcal{M}$ seront appelés *$\mathcal{M}$-ensembles* dans la suite, on dira que l'objet a de $\mathcal{M}$ est un *$\mathcal{M}$-élément de b* si $a \sqsubseteq b$ et on dira aussi dans ce cas que b *$\mathcal{M}$-contient* a ; on dira que b est un *$\mathcal{M}$-sous-ensemble* de c si tous les $\mathcal{M}$-éléments de b sont aussi $\mathcal{M}$-éléments de c, on notera dans ce cas $b \sqsubset c$.

Les axiomes de ZF impliquent l'existence dans le modèle $\mathcal{M}$ de certains objets, ainsi que la possibilité de certaines constructions ; nous allons mentionner ce dont nous aurons besoin, sans le justifier ici à partir des axiomes. En premier lieu, il existe un objet qui n'a aucun $\mathcal{M}$-élément, on le notera $\mathbf{0}$, son $\mathcal{M}$-ensemble de parties noté $\mathbf{1}$ a exactement un $\mathcal{M}$-élément, à savoir $\mathbf{0}$. Un *$\mathcal{M}$-couple* associé à deux objets a, b sera noté $(\!(a, b)\!)$ avec de bizarres parenthèses. Si c, d sont deux objets, le *$\mathcal{M}$-produit* $c \cdot d$ est le

$\mathcal{M}$-ensemble dont les $\mathcal{M}$-éléments sont tous les $\mathcal{M}$-couples $\langle\!\langle a, b\rangle\!\rangle$ avec $a \sqsubseteq c$ et $b \sqsubseteq d$.

Un *$\mathcal{M}$-graphe de fonction* est un objet $g \sqsubset c \cdot d$, pour un certain c et un certain d, tel que, lorsque $\langle\!\langle a, b_1\rangle\!\rangle \sqsubseteq g$ et $\langle\!\langle a, b_2\rangle\!\rangle \sqsubseteq g$, alors $b_1 = b_2$. On dira qu'un $\mathcal{M}$-graphe g_2 *prolonge* un $\mathcal{M}$-graphe g_1 lorsque $g_1 \sqsubset g_2$. Si g est un $\mathcal{M}$-graphe, il sera commode d'utiliser la notation de fonction $b = g(a)$ pour traduire le fait que $\langle\!\langle a, b\rangle\!\rangle \sqsubseteq g$; il sera commode aussi de dire « $g(a)$ est défini » pour traduire le fait qu'il existe b tel que $\langle\!\langle a, b\rangle\!\rangle \sqsubseteq g$; le *$\mathcal{M}$-domaine* de g est le $\mathcal{M}$-ensemble $c_0 \sqsubset c$ qui $\mathcal{M}$-contient *précisément tous les a* tels que $g(a)$ soit défini : on dira que g est le *$\mathcal{M}$-graphe d'une fonction de c_0 dans d*, on pourra parler de la « $\mathcal{M}$-fonction » $a \mapsto g(a)$ de c_0 dans d ; la $\mathcal{M}$-image de g est le $\mathcal{M}$-ensemble dont les $\mathcal{M}$-éléments sont tous les $g(a)$, pour $a \sqsubseteq c_0$. Si $c_1 \sqsubset c_0$, on définit la *restriction* de g à c_1 :

c'est la $\mathcal{M}$-intersection de g avec $c_1 \cdot d$. On définira sans peine les notions de $\mathcal{M}$-fonction injective, surjective, bijective de c_0 vers d. Il est clair que la notion de *prolongement* des $\mathcal{M}$-graphes de fonction correspond à la notion habituelle de prolongement des $\mathcal{M}$-fonctions associées.

On appellera *$\mathcal{M}$-cardinal* d'un objet a de $\mathcal{M}$ le plus petit $\mathcal{M}$-ordinal κ tel qu'il existe une $\mathcal{M}$-bijection de κ sur a. Il existe dans $\mathcal{M}$ un objet ω qui représente l'ensemble $\mathbb{N}$ des entiers naturels, son $\mathcal{M}$-cardinal est $\aleph_0$ par définition. Un autre objet $\mathbf{r}$ de $\mathcal{M}$ représente l'ensemble $\mathbb{R}$ des nombres réels ; au sens de $\mathcal{M}$, l'objet $\mathbf{r}$ a la $\mathcal{M}$-puissance du continu, mais vu de l'extérieur, l'ensemble des $\mathcal{M}$-éléments de $\mathbf{r}$ est dénombrable, puisque $\mathcal{M}$ tout entier est dénombrable ! Par la preuve de Cantor, on sait que le $\mathcal{M}$-cardinal de $\mathbf{r}$ est $> \aleph_0$, il est donc supérieur ou égal à $\aleph_1$, qui est défini comme le plus petit $\mathcal{M}$-cardinal $> \aleph_0$ (de même, le

$\mathcal{M}$-cardinal $\aleph_2$ est le plus petit $\mathcal{M}$-cardinal $> \aleph_1$). Il existe aussi dans $\mathcal{M}$ un objet $\mathbf{p}$ dont les $\mathcal{M}$-éléments sont tous les $\mathcal{M}$-sous-ensembles de $\mathbf{r}$; on sait, par Cantor à nouveau, que le $\mathcal{M}$-cardinal de l'objet $\mathbf{p}$ est strictement supérieur à celui de $\mathbf{r}$, le $\mathcal{M}$-cardinal de $\mathbf{p}$ est donc $\geqslant \aleph_2$: il faudra s'en souvenir. On rappelle qu'un $\mathcal{M}$-ordinal α est *$\mathcal{M}$-fini* quand tous les $\mathcal{M}$-ordinaux β *non nuls* tels que $\beta < \alpha$ sont des successeurs, de la forme $\beta = \gamma + 1$; un $\mathcal{M}$-ensemble F est dit *$\mathcal{M}$-fini* quand il existe une $\mathcal{M}$-bijection de F sur un $\mathcal{M}$-ordinal $\mathcal{M}$-fini. L'objet ω n'est pas $\mathcal{M}$-fini : il a été introduit comme tel par l'axiome de l'infini.

Tout cela étant posé, on en vient aux définitions spécifiques de la « preuve » qui suit. Le $\mathcal{M}$-produit $\omega \cdot \mathbf{p}$ sera utilisé comme un outil de la preuve, on le notera $\mathbf{q}$. Il existe un $\mathcal{M}$-ensemble X dont les $\mathcal{M}$-éléments sont tous les $\mathcal{M}$-graphes de fonction g d'un $\mathcal{M}$-sous-ensemble $\mathcal{M}$-fini quelconque $\mathrm{F} \sqsubset \mathbf{q}$, à valeurs dans la $\mathcal{M}$-paire $\{\mathbf{0}, \mathbf{1}\}$ (l'objet $\mathcal{M}$-fini F est le $\mathcal{M}$-domaine du $\mathcal{M}$-graphe g).

Définition. On dit qu'un $\mathcal{M}$-sous-ensemble Y de l'objet X défini ci-dessus est *dense* si pour tout $\mathcal{M}$-graphe $g_1 \sqsubseteq \mathrm{X}$, il existe un $\mathcal{M}$-graphe g_2 qui prolonge g_1 et qui est $\mathcal{M}$-élément de Y, autrement dit :

$$\forall g_1 \sqsubseteq \mathrm{X},\ \exists g_2 \sqsubseteq \mathrm{Y},\ g_1 \sqsubset g_2.$$

Adoptons le point de vue extérieur : puisque $\mathcal{M}$ est dénombrable, on peut dresser une liste $\mathrm{Y}_0, \mathrm{Y}_1, \ldots$ de tous les $\mathcal{M}$-sous-ensembles *denses* de X, et on peut procéder à l'opération suivante : on choisit d'abord un $\mathcal{M}$-graphe $\gamma_0 \sqsubseteq \mathrm{Y}_0$, quelconque ; ensuite, par récurrence, on raisonne ainsi : supposons que $\gamma_0, \ldots, \gamma_j$ soient déjà choisis, avec $\gamma_i \sqsubseteq \mathrm{Y}_i$ et γ_i prolongeant γ_{i-1} pour $0 < i \leqslant j$; puisque Y_{j+1} est dense, on peut trouver un $\mathcal{M}$-graphe $\gamma_{j+1} \sqsubseteq \mathrm{Y}_{j+1}$

qui prolonge $\gamma_j \sqsubseteq \mathrm{Y}_j \sqsubset \mathrm{X}$. *L'ensemble paradoxal* Γ est l'ensemble formé de tous les objets γ_i $(i \in \mathbb{N})$ qu'on vient d'introduire par récurrence. Par construction, on a le fait suivant :

$(\mathbf{P})$. *Pour tout* $\mathcal{M}$*-ensemble dense* $\mathrm{Y} \sqsubset \mathrm{X}$, *il existe un objet* $\gamma \in \Gamma$ *tel que* $\gamma \sqsubseteq \mathrm{Y}$ (en effet, on sait que $\mathrm{Y} = \mathrm{Y}_j$ pour un certain $j \in \mathbb{N}$, donc $\gamma = \gamma_j \sqsubseteq \mathrm{Y}_j = \mathrm{Y}$ et $\gamma \in \Gamma$).

Si $q \sqsubseteq \mathbf{q}$, considérons le $\mathcal{M}$-ensemble Y_q dont les $\mathcal{M}$-éléments sont tous les $\mathcal{M}$-graphes $g \sqsubseteq \mathrm{X}$ tels que $g(q)$ soit défini ; on va montrer que Y_q est dense. Pour tout $g_1 \sqsubseteq \mathrm{X}$, ou bien on a $g_1 \sqsubseteq \mathrm{Y}_q$, auquel cas on prend simplement $g_2 = g_1$, ou bien on peut prolonger g_1 en g_2 en « ajoutant » à la famille des $\mathcal{M}$-éléments de g_1 le $\mathcal{M}$-couple $(\!(q, \mathbf{0})\!)$ — ou bien $(\!(q, \mathbf{1})\!)$: peu importe, il s'agit seulement de prolonger g_1 ; l'objet g_2 est la $\mathcal{M}$-réunion de g_1 et du $\mathcal{M}$-singleton $\{(\!(q, \mathbf{0})\!)\}$ — ; dans les deux cas, on a trouvé une extension $g_2 \sqsubseteq \mathrm{Y}_q$, ce qui prouve que Y_q est dense. Par la propriété paradoxale (**P**) de Γ, il existe un $\mathcal{M}$-graphe $\gamma \in \Gamma$ tel que $\gamma \sqsubseteq \mathrm{Y}_q$. Comme les $\mathcal{M}$-graphes dans Γ se prolongent les uns les autres (par construction), on voit que $v(q) = \gamma(q)$ ne dépend pas de $\gamma \in \Gamma$ tel que $\gamma \sqsubseteq \mathrm{Y}_q$. L'observateur extérieur voit ainsi une fonction v qui est définie pour tout $\mathcal{M}$-élément $q = (\!(n, p)\!)$ de $\mathbf{q}$, une fonction v telle que $v(n, p) = \mathbf{0}$ ou $\mathbf{1}$. Par ailleurs, il n'existe certainement pas un $\mathcal{M}$-ensemble dont la collection des $\mathcal{M}$-éléments soit Γ, ni un $\mathcal{M}$-graphe qui représente la fonction v : l'ensemble Γ et la fonction v ne sont pas « du même monde » que le monde du modèle $\mathcal{M}$...

La méthode très délicate de Cohen produit un nouveau modèle $\mathcal{M}^*$ de ZF qui, pour le dire très vite, contient à la fois $\mathcal{M}$ et un objet $\mathfrak{G}$ dont la famille des $\mathcal{M}^*$-éléments est Γ. Si b est un objet de $\mathcal{M}$, c'est aussi un objet de $\mathcal{M}^*$;

si a est un $\mathcal{M}$-élément de b, c'est encore un $\mathcal{M}^*$-élément de b, et la famille des $\mathcal{M}^*$-éléments de b coïncide avec la famille des $\mathcal{M}$-élément de b. Il est cependant possible (c'est même fait pour) que la famille des $\mathcal{M}^*$-sous-ensembles de $b \in \mathcal{M}$ contienne plus d'objets « qu'avant ». À côté de l'objet $\mathcal{P}^{\mathcal{M}}(b)$ de $\mathcal{M}$ qui contient comme $\mathcal{M}^*$-éléments (ou $\mathcal{M}$-éléments, on a dit que c'est pareil dans ce cas) tous les $c \in \mathcal{M}$ tels que $c \sqsubset b$, il y a maintenant un objet $\mathcal{P}^{\mathcal{M}^*}(b)$ qui peut $\mathcal{M}^*$-contenir, en plus des $\mathcal{M}$-éléments de $\mathcal{P}^{\mathcal{M}}(b)$, des $c^* \in \mathcal{M}^*$ « nouveaux » dont tous les $\mathcal{M}^*$-éléments sont des $\mathcal{M}^*$-éléments de b, ce qu'on pourrait noter $c^* \sqsubset^* b$. C'est ce qui se passe avec l'objet $\mathfrak{G} \notin \mathcal{M}$ qui représente Γ : les $\mathcal{M}^*$-éléments de $\mathfrak{G}$ sont des $\gamma \sqsubseteq \mathrm{X}$ (où X, objet de $\mathcal{M}$, a été défini plus haut) ; ainsi, on a $\mathfrak{G} \sqsubset^* \mathrm{X}$.

En revanche, les $\mathcal{M}$-ordinaux α restent des $\mathcal{M}^*$-ordinaux, et de plus, tous les $\mathcal{M}^*$-ordinaux étaient déjà des

$\mathcal{M}$-ordinaux. Un $\mathcal{M}$-cardinal κ est un $\mathcal{M}$-ordinal qui ne peut être mis en $\mathcal{M}$-bijection avec aucun $\mathcal{M}$-ordinal $\alpha < \kappa$. On vient de dire que κ reste un ordinal dans $\mathcal{M}^*$, mais il serait possible que le modèle plus riche $\mathcal{M}^*$ fasse apparaître une $\mathcal{M}^*$-bijection, qui n'existait pas dans $\mathcal{M}$, entre κ et un $\alpha < \kappa$. C'est une propriété essentielle du modèle $\mathcal{M}^*$ que cela ne se produise pas : les $\mathcal{M}$-cardinaux restent des $\mathcal{M}^*$-cardinaux, mais c'est loin d'être évident.

Désignons par $\mathbf{N}$ l'ensemble (naïf) des $\mathcal{M}$-éléments de ω, et pour chaque $p \sqsubseteq \mathbf{p}$, considérons le sous-ensemble (naïf) de $\mathbf{N}$ défini par

$$\mathrm{D}_p = \{n \sqsubseteq \omega \ : \ v(n,p) = \mathbf{0}\} \subset \mathbf{N}.$$

Comme Γ (de même que la fonction v) est représenté par un objet de $\mathcal{M}^*$, on saura qu'il existe un objet d_p dans ce nouveau modèle tel que D_p soit la famille des $\mathcal{M}^*$-éléments de d_p, et de plus, on pourra démontrer qu'il existe un $\mathcal{M}^*$-graphe g^* dont les $\mathcal{M}^*$-éléments sont tous les $\mathcal{M}^*$-couples $(\!(p, d_p)\!)$, pour $p \sqsubseteq \mathbf{p}$; dit plus librement, $p \mapsto d_p$ est une $\mathcal{M}^*$-fonction. On verra plus loin que les d_p (ou les D_p) sont deux à deux distincts ; on aura donc une $\mathcal{M}^*$-bijection entre $\mathbf{p}$ et la famille des $(d_p)_{p \sqsubseteq \mathbf{p}}$, bijection correspondant au $\mathcal{M}^*$-graphe g^*. On a dit qu'on peut montrer que le $\mathcal{M}$-ensemble $\mathbf{p}$, de $\mathcal{M}$-cardinalité $\geqslant \aleph_2$, garde dans le nouveau modèle $\mathcal{M}^*$ une $\mathcal{M}^*$-cardinalité $\geqslant \aleph_2$. On aura ainsi produit une famille de $\mathcal{M}^*$-sous-ensembles de $\omega \equiv \mathbb{N}$ dont la $\mathcal{M}^*$-cardinalité est celle de $\mathbf{p}$, donc, supérieure ou égale à $\aleph_2$. *L'hypothèse du continu est fausse dans* $\mathcal{M}^*$: en effet, d'après ce qui précède, la $\mathcal{M}^*$-cardinalité de $\mathcal{P}(\omega)$ dans $\mathcal{M}^*$ est au moins égale à $\aleph_2$; par l'axiome du choix dans $\mathcal{M}^*$ (et le théorème de Zermelo), on peut construire une $\mathcal{M}^*$-injection j_2 de $\aleph_2$ dans $\mathcal{P}(\omega)$, qui admet deux restrictions, j_0 à $\aleph_0 \sqsubset \aleph_2$ et j_1 à $\aleph_1 \sqsubset \aleph_2$. Les $\mathcal{M}^*$-images de j_0, j_1 et j_2 donneront *trois* cardinalités infinies distinctes pour les sous-ensembles de $\mathcal{P}(\mathbb{N})$, en contradiction avec l'hypothèse du continu (la technique de forcing pourrait aussi bien produire une *infinité* de cardinalités distinctes).

On va montrer pour finir que les D_p sont deux à deux distincts : si $p_1 \neq p_2$ sont deux $\mathcal{M}$-éléments de $\mathbf{p}$, on posera $q_{1,n} = (\!(n, p_1)\!) \sqsubset \mathbf{q}$ et $q_{2,n} = (\!(n, p_2)\!) \sqsubset \mathbf{q}$, pour $n \sqsubset \omega$; considérons le $\mathcal{M}$-ensemble Y dont les $\mathcal{M}$-éléments $g \sqsubset \mathrm{X}$ vérifient ceci : il existe $n \sqsubset \omega$ tel que $g(q_{1,n})$ et $g(q_{2,n})$ sont définis mais $g(q_{1,n}) \neq g(q_{2,n})$. On va voir que Y est dense : si $g \sqsubset \mathrm{X}$, son $\mathcal{M}$-domaine est $\mathcal{M}$-fini et par conséquent, le $\mathcal{M}$-ensemble F, qui $\mathcal{M}$-contient tous les $n \sqsubset \omega$ tels que $g_1(q_{1,n})$ ou $g_2(q_{2,n})$ soit défini, est $\mathcal{M}$-fini également ; puisque l'objet ω n'est pas $\mathcal{M}$-fini, il existe $n \sqsubset \omega$ tel que n ne soit pas $\mathcal{M}$-élément de F, c'est-à-dire que ni $g_1(q_{1,n})$ ni $g_1(q_{2,n})$ ne soient définis ; on peut alors construire un prolongement g_2 de g_1 tel que $(\!(q_{1,n}, \mathbf{0})\!)$ et $(\!(q_{2,n}, \mathbf{1})\!)$ soient $\mathcal{M}$-éléments de g_2. Ainsi, on aura $g_2(q_{1,n}) = \mathbf{0} \neq \mathbf{1} = g_2(q_{2,n})$ donc $g_2 \sqsubset \mathrm{Y}$, on a prouvé que Y est dense. Puisque Y est dense, il existe par la propriété $(\mathbf{P})$ un $\gamma \in \Gamma$ tel que $\gamma \sqsubset \mathrm{Y}$; par la définition de l'objet Y, il existe $n \sqsubset \omega$ tel que $\gamma(q_{1,n})$ et $\gamma(q_{2,n})$ soient définis mais $\gamma(q_{1,n}) \neq \gamma(q_{2,n})$. Il en résulte que $v(n, p_1) = \gamma(q_{1,n}) \neq \gamma(q_{2,n}) = v(n, p_2)$: par conséquent, *un et un seul* de $v(n, p_1)$ et $v(n, p_2)$ est égal à $\mathbf{0}$, l'objet n est élément d' *un et un seul parmi* D_{p_1} et D_{p_2}, donc $d_{p_1} \neq d_{p_2}$. You know what : I'm happy qu'ça soit fini !

Je suis prêt à parier 10 cents d'euro que moins de un pour cent des mathématiciens « normaux » a essayé de lire la preuve de Paul Cohen ; j'ai l'impression que le monde de la logique mathématique leur est presque totalement étranger. J'irai même jusqu'à avancer que ces « vrais mathématiciens » se fichent complètement des axiomes de ZF : ils pratiquent les opérations qui leur sont habituelles, dans leur domaine particulier, sans imaginer devoir jamais demander la permission aux axiomes de la théorie des ensembles. Peut-être, tout au plus, vivent-ils dans un certain sentiment de sécurité, le sentiment d'avoir la garantie que cette théorie justifierait, si on le lui demandait, la validité de leurs constructions mathématiques. Pour terminer sur une provocation gratuite, on pourrait dire que les mathématiciens sont à l'abri sous le parapluie de ZFC comme les Français sont à l'abri sous leur *parapluie nucléaire* : bien heureusement, il n'a jamais servi.

Voilà, c'est fini . . .

Présentation visible sur internet à l'adresse :

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

La présentation « d'origine » :
https://webusers.imj-prg.fr/~bernard.maurey/articles/MlV2021/MlV-2021.pdf

accompagnée d'un texte contenant une liste de références :
https://webusers.imj-prg.fr/~bernard.maurey/articles/MlV2021/Cantor-CH.pdf

la présente présentation « augmentée » :
https://webusers.imj-prg.fr/~bernard.maurey/articles/MlV2021/Cantor-et-HC.pdf

Trois références utilisées (pillées) ici, en particulier pour leurs images :

Heinz-Dieter Ebbinghaus, in cooperation with Volker Peckhaus,
Ernst Zermelo. An Approach to His Life and Work. Second Edition. Springer, 2015.

Walter Purkert & Hans Joachim Ilgauds, *Georg Cantor: 1845–1918*.
Vita Mathematica 1, Birkhäuser Verlag, 1987.

Jean van Heijenoort, *From Frege to Gödel.*
A Source Book in Mathematical Logic, 1879–1931.
Harvard University Press, Cambridge, Mass., 1967.