21-373, Algebraic Structures, Department of Mathematical Sciences, Carnegie Mellon University
Fall 2011: (Math Studies Section) Monday, Wednesday, Friday, 10:30 am, Porter Hall 226B.
Luc TARTAR, University Professor of Mathematics, Wean Hall 6212, tartar@cmu.edu

2- Wednesday August 31, 2011.

Paradoxes

Before 1905, when RUSSELL found a *paradox* with the "set of all sets",[1] mathematicians used the term *set* without having defined it, and the paradox was resolved by finding a reasonable definition of the properties which characterize sets, and RUSSELL's idea became a proof that the collection of all sets is not itself a set.

More than fourty years ago, I read an interesting explanation, giving the essence of Russell's paradox: one defines a book as something published, i.e. having an ISBN number (International Standard Book Number),[2] and a catalog as any book just containing a list of books, so that a catalog of catalog is then such a book containing only the ISBN numbers of catalogs; among such catalogs of catalogs, there are those who list their own ISBN number and those who do not list their own ISBN number. One can make the complete list $\mathcal{L}$ of all catalogs of catalogs which do not list their own ISBN number and have no other book in the list, but it is a manuscript and not a book! If one publishes $\mathcal{L}$ into a book $\mathcal{B}$ (i.e. by asking for an ISBN number, and putting it on the cover of the list), then $\mathcal{B}$ will be incomplete, because it should contain the reference to $\mathcal{B}$, since it is a catalog of catalogs which does not list itself; however, if one adds to $\mathcal{L}$ the ISBN number of the book which one is going to publish (while putting it also on the cover of the list), one obtains a book $\mathcal{B}'$ which is not what one wants, since it lists its own reference and it should not: therefore the answer is a manuscript but not a book.

Ancient Greeks discussed of a fake paradox, of a Cretan saying that 'all Cretans are liers': if one assumes that a lier is someone who never says the truth, then it means that there is a Cretan who is truthful, and that the one who speaks then is a lier.[3]

A theory is *consistent* if there is no proposition P such that P and its negation can be proved true by using the axioms of the theory.[4] One should notice that it is still not known if ZF (Zermelo–Fraenkel set theory),[5,6] the standard foundation of modern mathematics, is consistent.[7]

---

[1] Bertrand Arthur William, third earl RUSSELL, Welsh mathematician, 1872–1970. He received the Nobel Prize in Literature in 1950, in recognition of his varied and significant writings in which he champions humanitarian ideals and freedom of thought. He worked in Cambridge, England.

[2] The former *ISBN code* was a $[10, 9]$-code over F_{11} ($\simeq \mathbb{Z}_{11}$) and the verification of the code $c_1 c_2 \cdots c_{10}$ is that one must have $c_1 + 2c_2 + 3c_3 + \ldots + 10c_{10} = 0 \pmod{11}$; the first part of an ISBN codeword was the *group identifier*, which identified a country or a language area, the second part was the *publisher identifier*, which identified a specific publisher in the group, the third part was the *title identifier*, which identified a specific publication of the publisher; the length of the three parts varied, but the total length was 9, and the *check-digit* x_{10} was written X if it was 10. The revised ISBN code uses a 13-digit number, and the verification of the code $c_1 c_2 \cdots c_{13}$ is that one must have $c_1 + 3c_2 + c_3 + 3c_4 + \ldots + c_{11} + 3c_{12} + c_{13} = 0 \pmod{10}$. The ISBN numbers of my first three books, are 978-3-540-35743-8, 978-3-540-71482-8, and 978-3-540-77561-4, so that 978-3-540 seems to identify Springer (Berlin Heidelberg New York), but the ISBN number of my fourth book published by the same publisher is 978-3-642-05194-4.

[3] Ancient Greeks made the mistake in negating the proposition as 'all Cretans are truthful' instead of the correct 'there exists a Cretan who is truthful'; unfortunately, many non-mathematicians still make such silly confusions nowadays.

[4] In such a case, all the propositions in this theory are both true and false, so that such a theory is useless.

[5] Ernst Friedrich Ferdinand ZERMELO, German mathematician, 1871–1953. He worked at Georg-August-Universität, Göttingen, Germany, in Zürich, Switzerland, and at Freiburg im Breisgau, Germany. Zermelo's theorem (conjectured by CANTOR), that every set can be well-ordered, is named after him. The Zermelo–Fraenkel set theory is partly named after him.

[6] Adolf Abraham Halevi FRAENKEL, German-born mathematician, 1891–1965. He worked in Marburg, Germany, and at the Hebrew University in Jerusalem (Israel). The Zermelo–Fraenkel set theory is partly named after him.

[7] ZFC denotes Zermelo–Fraenkel set theory with the axiom of choice.

It was not so easy to imagine that there could exist a consistent theory $\mathcal{T}$ with a proposition P which cannot be proved to be true using only the axioms of $\mathcal{T}$, but the negation of P cannot be proved to be true either using only the axioms of $\mathcal{T}$, and such a proposition is called *undecidable*: in this case one may create two consistent theories, $\mathcal{T}_1$ obtained by adding to the axioms of $\mathcal{T}$ that P is true, and $\mathcal{T}_2$ obtained by adding to the axioms of $\mathcal{T}$ that P is false. GÖDEL had the idea that some paradoxical sentences could be coded into mathematical propositions,[8] and he was able to do it using the integers $\mathbb{N}$, so that his theorem is that in any consistent theory $\mathcal{T}$ which contains $\mathbb{N}$, there is an undecidable proposition.

When HILBERT made a famous list of problems in 1900, he could not imagine for example that CH (the continuum hypothesis) is undecidable:[9] CH, which was conjectured by CANTOR in 1877, states that there is no set whose *cardinality* is strictly between that of $\mathbb{N}$ (denoted $\aleph_0$) and that of $\mathbb{R}$, but the work of GÖDEL in 1940 and that of Paul COHEN in 1963 showed that,[10] assuming that ZF set theory is consistent, CH is undecidable.

A paradox may then result from a lack of correct mathematical definitions, or by using an undecidable proposition, but often it corresponds to a wrong intuition about some mathematical question. After proving that there is a bijection between the interval $I = [0, 1] \subset \mathbb{R}$ and the square $I \times I \subset \mathbb{R}^2$, CANTOR wrote to DEDEKIND 'I see it, but I do not believe it!'.[11]

Before describing more paradoxes, I first recall definitions related to an *order* on a set X: it is a *binary relation* $\mathcal{R}$ on $X \times X$,[12] which is *reflexive* (i.e. for all $a \in X$ one has $a \mathcal{R} a$), *anti-symmetric* (i.e. for all $a, b \in X$, $a \mathcal{R} b$ and $b \mathcal{R} a$ imply $b = a$), and *transitive* (i.e. for all $a, b, c \in X$, $a \mathcal{R} b$ and $b \mathcal{R} c$ imply $a \mathcal{R} c$); it is also called a *partial order* (and sometimes X is called a *poset*, an abbreviation of partially ordered set) by opposition to a *total order* (i.e. such that for all $a, b \in X$, either $a \mathcal{R} b$ or $b \mathcal{R} a$, which maybe both if $b = a$), also called a *linear order*; a *chain* is any subset $A \subset X$ such that the restriction of $\mathcal{R}$ to $A \times A$ is a total order.

$\mathcal{R}$ is usually thought as similar to $a \leq b$ in $\mathbb{R}$, $A \subset B$ for subsets of a set, $a \mid b$ (i.e. a divides b) for positive integers, so that one says that an ordered set X has a *minimum* α if $\alpha \mathcal{R} x$ for all $x \in X$ (and the minimum is unique, although there is not necessarily one),[13] and X has a *maximum* ω if $x \mathcal{R} \omega$ for all $x \in X$ (and the maximum is unique, although there is not necessarily one). A total order on X is called a *well order* if every nonempty subset of X has a minimum.

An element a is *minimal* if $x \mathcal{R} a$ implies $x = a$;[14] an element z is *maximal* if $z \mathcal{R} x$ implies $x = z$. For a subset $A \subset X$, an *upper bound* of A is any $y \in X$ such that $a \mathcal{R} y$ for all $a \in A$, and the *least upper bound* of A, if it exists, is the minimum of all upper bounds; a *lower bound* of A is any $x \in X$ such that $x \mathcal{R} a$ for all $a \in A$, and the *greatest lower bound* of A, if it exists, is the maximum of all lower bounds.

Zermelo's theorem, conjectured by CANTOR, says that every non-empty set can be equipped with a well order.

---

[8] Kurt GÖDEL, Czech-born mathematician, 1906–1978. He worked in Vienna, Austria, and at IAS (Institute for Advanced Study), Princeton, NJ.

[9] David HILBERT, German mathematician, 1862–1943. He worked in Königsberg (then in Germany, now Kaliningrad, Russia) and at Georg-August-Universität, Göttingen, Germany. Hilbert spaces are named after him.

[10] Paul Joseph COHEN, American mathematician, 1934–2007. He received the Fields Medal in 1966 for his fundamental work on the foundations of set theory. He worked at Stanford University, Stanford, CA.

[11] Of course, such a bijection cannot be *continuous*, since its inverse would be continuous (by an argument of *compactness*), and then I minus an *interior* point (which is not *connected*) would be *homeomorphic* to a square minus a point (which is connected).

[12] It means that there is a subset $Y \subset X \times X$ and instead of writing $(a, b) \in Y$ one writes $a \mathcal{R} b$.

[13] One says that *uniqueness* holds for a problem in mathematics if when a_1 and a_2 are two solutions then they must coincide, but there might be no solutions, and *existence* of a solution is a different question. To avoid confusion, one may prefer to say "if a solution exists, it is unique".

[14] In the positive integers with the order $a \mid b$, then 1 is the minimum, but if one removes 1, then there is no minimum and the minimal elements are precisely the *prime numbers* (recalling that 1 is not considered a prime number).

Zorn's lemma,[15] which was actually used by BOCHNER 7 years before Max ZORN,[16] says that if every chain in a non-empty ordered set X has a least upper bound in X, then X has a maximal element.

The *axiom of choice* says that if I is a non-empty index set and for each $i \in I$ one has a non-empty set X_i, then the product $\prod_{i \in I} X_i$ is non-empty, i.e. it is possible to choose an $x_i \in X_i$ for every $i \in I$ and consider the point $x = (x_i, i \in I)$ in the product.

Despite the use of different terms (theorem, lemma, axiom) these three statements are equivalent, but some forms are more natural, and a few questions of maximality occur in algebra, for which one naturally uses "Zorn's lemma"; I do not recall seeing a direct use of Zermelo's theorem in algebra.

The axiom of choice permits to construct surprising sets, and it looks like paradoxes because it contradicts some intuition, which then appears to be misleading. One may want to avoid using the axiom of choice, as the constructivists do, so that such strange constructions cannot be done.

The *Hausdorff–Banach–Tarski paradox* is an improvement of an idea of HAUSDORFF by BANACH and TARSKI,[17,18,19], so that it is sometimes called the Banach–Tarski paradox: it implies that if A is a solid ball of radius 1 in $\mathbb{R}^3$ and B is a solid ball of radius 2 in $\mathbb{R}^3$, then there is an integer N and a *partition* of A as the (disjoint) union of $A_1, \ldots, A_N$ and a *partition* of B as the (disjoint) union of $B_1, \ldots, B_N$ such that for $i = 1, \ldots, N$, the subset B_i is obtained from A_i by a rigid displacement.[20]

The paradox comes from the fact that rigid displacements conserve volume, so that one (mistakenly) thinks that for each $i \in \{1, \ldots, N\}$ the volume of B_i is equal to the volume of A_i, which would imply that the volume of B is equal to the volume of A, and this is obviously not the case.

The resolution of the paradox is that some A_i are *non-measurable* sets, so that it is impossible to define their volume in a consistent way. The construction actually shows that one cannot define a *finitely additive measure* on bounded sets of $\mathbb{R}^3$,[21] with the property that it is invariant by translations and rotations, and that it coincides with the usual (Lebesgue) measure for cubes,[22] or for the *σ-algebra* of Borel sets.[23,24,25]

HAUSDORFF was actually generalizing a previous construction of a non-measurable subset of the circle

[15] Max August ZORN, German-born mathematician, 1906–1993. He worked at UCLA (University of California at Los Angeles), Los Angeles, CA, and at University of Indiana, Bloomington, IN, where I met him in 1980. "Zorn's lemma" is named after him, but it was used 7 years before he did by BOCHNER.

[16] Salomon BOCHNER, Polish-born mathematician, 1899–1982. He worked in München (Munich), Germany, and after 1933 at Princeton University, Princeton, NJ. He used "Zorn's lemma" 7 years before Max ZORN.

[17] Felix HAUSDORFF, German mathematician, 1869–1942. He worked in Leipzig, in Greifswald and in Bonn, Germany. He wrote literary and philosophical work under the pseudonym of Paul MONGRÉ. Hausdorff topologies and Hausdorff measures are named after him.

[18] Stefan BANACH, Polish mathematician, 1892–1945. He worked in Lwów (then in Poland, now Lvov, Ukraine). There is a Stefan Banach International Mathematical Center in Warsaw, Poland. The term Banach space was introduced by FRÉCHET.

[19] Alfred TARSKI (TEITELBAUM), Polish-born mathematician, 1902–1983. He worked in Warsaw, Poland, and at UCB (University of California at Berkeley), Berkeley, CA.

[20] A rigid displacement is a rotation followed by a translation, i.e. a mapping $x \mapsto a + M\,x$ for all $x \in \mathbb{R}^3$, with $a \in \mathbb{R}^3$ and $M \in S\mathbb{O}_3$ (or a translation followed by a rotation, but in general translations and rotations do not commute, since $a + M\,x = M\,(x + b)$ for $b = M^{-1}a = M^T a$).

[21] I.e. a mapping μ from the set bounded subsets of $\mathbb{R}^3$ into $[0, \infty] \subset \mathbb{R} \cup \{\infty\}$, with the property that $\mu(X \cup Y) = \mu(X) + \mu(Y)$ whenever X and Y are disjoint.

[22] Henri Léon LEBESGUE, French mathematician, 1875–1941. He worked in Rennes, in Poitiers, and he held a chair at Collège de France (mathématiques, 1921–1941) in Paris, France. The spaces L^p were named Lebesgue spaces in his honour by F. RIESZ, and the Lebesgue integration theory named after him was discovered two years before him by W.H. YOUNG.

[23] Félix Édouard Justin Émile BOREL, French mathematician, 1871–1956. He worked in Lille and in Paris, France. Borel functions, measures, or sets are named after him.

[24] A σ-algebra $\mathcal{A}$ is a family of subsets (of a set Z) which is stable by complementation and stable by countable unions.

[25] The σ-algebra of Borel sets is the smallest σ-algebra which contains the open sets.

$\mathbb{S}^1$ by VITALI,[26] which implies that there is no *σ-additive measure* on all subsets of $\mathbb{S}^1$ which is invariant by rotation.[27]

What constitutes a proof?

A proof of existence of a solution of a problem may be constructive, in which case it describes an algorithm which constructs a solution, or a sequence which converges to a solution, but if one only knows that a subsequence (of a sequence which one constructs) converges to a solution, it is not considered a constructive proof.

For example, if a real continuous function f from $[0,1]$ satisfies $f(0)\,f(1) < 0$, then there exits $z \in (0,1)$ with $f(z) = 0$, and one constructs a solution in the following way. If for $0 \le a_n < b_n \le 1$ one has $f(a_n)\,f(b_n) < 0$, then one evaluates $f(c_n)$ for $c_n = \frac{a_n+b_n}{2}$, and there are three cases: if $f(c_n) = 0$, then c_n is a solution; if $f(a_n)\,f(c_n) < 0$, one takes $a_{n+1} = a_n$ and $b_{n+1} = c_n$, and one repeats the algorithm, while if $f(a_n)\,f(c_n) > 0$, one takes $a_{n+1} = c_n$ and $b_{n+1} = b_n$, and one repeats the algorithm. Since $b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}$ for all n if one has not the chance to fall on a solution at some step, it implies that a_n and b_n are Cauchy sequences, which both converge to a solution z (because $\mathbb{R}$ is complete).

On the other hand, if a real continuous function f on a connected subset $X \subset \mathbb{R}^N$ for $N \ge 2$ is such that there exist $a, b \in X$ with $f(a) < 0 < f(b)$, then there exists $z \in X$ with $f(z) = 0$ by an argument of connectedness, i.e. $f(X)$ is connected, hence it contains the interval $[f(a), f(b)]$, which contains 0, but there is no precise algorithm behind this proof.

Brouwer's fixed point theorem,[28] that if f is continuous from a (non-empty) *compact* and *convex* set $K \subset \mathbb{R}^N$ into itself has at least one fixed point, i.e. some $z \in K$ satisfies $f(z) = z$, is not constructive for $N \ge 2$,[29] but BROUWER did not like his non-constructive existence proof, and after that he turned to constructivism for the rest of his life.

There is a different question than using non-constructive proofs, which is to be reasonably sure that there is no gap left in a proof which is very long.

At ICM78, the International Congress of Mathematicians of 1978 in Helsinki, Finland, I heard a talk about the classification of *finite simple groups*, by Daniel GORENSTEIN,[30] who was involved in the work. The search for the classification started in 1955, and it was "completed" around 1983: besides a few general families, there are 26 exceptions called *sporadic groups*. However, one may wonder if the "proof" is complete and correct, since it is made up of tens of thousands of pages in about 500 articles written by about 100 authors. Soon after the "completion", specialists started writing simpler proofs in order to reduce the length to something more reasonable.

There was a long standing conjecture, that colouring plane maps of connected countries could always be done with at most four colours, and many years ago there was a "computer proof" of it, because there was too much work involved in checking nearly two thousand cases, and a program was written for a computer to check all these cases.

What is sad about such "proofs" is that no really new mathematical idea has appeared to make the proof easily understandable, and one should remember that doing research in mathematics is about creating new knowledge, but also about discovering simplifying ideas which make those too long "proofs" clearer, if not simple!

---

[26] Giuseppe VITALI, Italian mathematician, 1875–1932. He worked in Modena, in Padova (Padua), and in Bologna, Italy. The department of pure and applied mathematics of Università degli Studi di Modena e Reggio Emilia is named after him.

[27] A σ-additive measure ν defined on a σ-algebra $\mathcal{A}$ is a mapping from $\mathcal{A}$ into $[0,\infty] \subset \mathbb{R} \cup \{\infty\}$ with the property that $\nu\big(\cup_{i\in I} A_i\big) = \sum_i \nu(A_i)$ whenever I is countable, and the A_i are disjoint and belong to $\mathcal{A}$.

[28] Luitzen Egbertus Jan BROUWER, Dutch mathematician, 1881–1966. He worked in Amsterdam, The Netherlands.

[29] For $N = 1$, $K = [a,b]$, and if neither a nor b are fixed points of f, then the function $f(x) - x$ changes sign on the interval $[a,b]$, hence it vanishes at a point.

[30] Daniel GORENSTEIN, American mathematician, 1923–1992. He worked at Clark University, Worcester, MA, at Northeastern University, Boston, MA, and at Rutgers University, Piscataway, NJ.

For doing research, it seems useful to learn enough about what has been done before, but maybe one should not learn too much, because at some point once should realize that a lot of what is published is not really research but development, i.e. using known ideas on various new problems. That one often confuses research and development seems to be a result of the "publish or perish" philosophy, which pushes against discovering simplifying ideas, because writing too much of the same thing is considered good by administrators!

In the early 1980s, I went to a Bourbaki seminar, which met a few times a year at IHP (Institut Henri Poincaré) in Paris, and I heard Jean-Pierre SERRE talk,[31] about a proof by Pierre DELIGNE, of conjectures by WEIL. Jean-Pierre SERRE started by saying that the proof used a tool generalizing an idea of Alexandre GROTHENDIECK, and it involved sheaf theory with p-adic numbers, I think, but since this was not written and only a handful of people would be able to write it down, he assumed that the extension mentioned was correct, and he explained what the proof was after that. Many years after, I heard that a group had worked at constructing the required extension, and it took about 500 pages to write it down!

Additional footnotes: CLARK,[32] FRÉCHET,[33] HARDINGE,[34] HARDY,[35] RIESZ F.,[36] RIESZ M.,[37] RUTGERS,[38] YOUNG L.C.,[39] YOUNG W.H..[40]

---

[31] Jean-Pierre SERRE, French mathematician, born in 1926. He received the Fields Medal in 1954 for his work in algebraic topology. He received the Wolf Prize in 2000 for his many fundamental contributions to topology, algebraic geometry, algebra, and number theory and his inspirational lectures and writing. He received the Abel Prize in 2003 for playing a key role in shaping the modern form of many parts of mathematics, including topology, algebraic geometry and number theory. He held a chair at Collège de France (algebra and geometry, 1956–1994), Paris, France.

[32] Jonas Gilman CLARK, American industrialist, 1815–1900. Clark University, Worcester, MA, is named after him.

[33] Maurice René FRÉCHET, French mathematician, 1878–1973. He worked in Poitiers, in Strasbourg and in Paris, France. Fréchet spaces (which are locally convex, metrizable and complete vector spaces) are named after him.

[34] Sir Charles HARDINGE, 1st baron HARDINGE of Penshurst, English diplomat, 1858–1944. He was Viceroy and Governor-General of India (1910–1916).

[35] Godfrey Harold HARDY, English mathematician, 1877–1947. He worked in Cambridge, in Oxford, England, holding the Savilian chair of geometry (1920–1931), and in Cambridge again, holding the Sadleirian chair of pure mathematics (1931–1942).

[36] Frigyes (Frederic) RIESZ, Hungarian mathematician, 1880–1956. He worked in Kolozsvár (then in Hungary, now Cluj-Napoca, Romania), in Szeged and in Budapest, Hungary. He introduced the spaces L^p in honor of LEBESGUE and the spaces $\mathcal{H}^p$ in honor of HARDY, but no spaces are named after him; the Riesz operators have been introduced by his younger brother Marcel RIESZ.

[37] Marcel RIESZ (younger brother of Frigyes (Frederic) RIESZ), Hungarian-born mathematician, 1886–1969. He worked in Stockholm and in Lund, Sweden. The Riesz operators are named after him.

[38] Henry RUTGERS, American colonel, 1745–1830. Rutgers University, Piscataway, NJ, is named after him.

[39] Laurence Chisholm YOUNG, English-born mathematician, 1905–2000. He worked in Cape Town, South Africa, and at University of Wisconsin, Madison, WI, where I first met him during my first trip to United States, in the spring of 1971. Young measures are named after him, and he introduced them in the Calculus of Variations. I pioneered their use in partial differential equations (from continuum mechanics) in the late 1970s, not knowing at the time that he introduced them, as I heard about them as parametrized measures in seminars on control theory.

[40] William Henry YOUNG, English mathematician, 1863–1942. He worked in Liverpool, England, in Calcutta, India, holding the first Hardinge professorship (1913–1917), in Aberystwyth, Wales, and in Lausanne, Switzerland. He is said to have discovered Lebesgue integration two years before LEBESGUE. There are many results attributed to him which may be joint work with his wife, Grace CHISHOLM-YOUNG, English mathematician, 1868–1944, as they collaborated extensively; their son Laurence is known for his own mathematical results.