The history of mathematical impossibility

Jesper Lützen, June 2019

When I studied mathematics in Aarhus in the 1970s the curriculum was very Bourbakist. For example Mathematics 1 began with an introduction to general topological spaces. I found it all very beautiful and interesting but it left me wondering how and why these abstract structures had been developed in the first place. In order to satisfy my curiosity I began taking courses at the Department of History of the Exact Sciences and I decided to write my thesis for the master's degree on the history of the concept of functions. For my PhD thesis I continued with an investigation of the prehistory of the theory of generalized functions (distributions). In both cases, I was particularly interested in uncovering the forces that drove the concept formation. I found out that often they stemmed from problems arising in connection with application of mathematics, in particular to physics. For example from the time d'Alembert formulated the wave equation in 1747, there has been attempts to extend the type of functions that can serve as solutions. Such attempts were crucial for the development of the concept of function and the development of the theory of distributions.

My next major research project was centered not around a concept but around a person namely Joseph Liouville, the leading French mathematician between Cauchy and Hermite. This gave me a very different experience. First because it forced me to deal with all the aspects of mathematics that Liouville worked on (which was pretty much all of early 19th century mathematics) and secondly because I could draw on an enormous unpublished Nachlass. His Liouville's 240 notebooks and his other manuscripts allowed me to follow the creative process leading to some of his famous results for example on Sturm-Liouville theory and on integration in finite terms. Moreover the note books also contained ingenious ideas and theories that he never published. For example his notes revealed a theory of the equilibrium shapes of rotating fluid planets as well as a spectral theory of a certain type of integral operator. These series of notes both anticipate later developments with about 40 years.

After a long period working on the history of mechanics and in particular Heinrich Hertz's famous *Prinzipien der Mechanik* I have for the last couple of years dealt with mathematical impossibility. From a logical point of view impossibility theorems do not constitute a well-defined class of theorems. Indeed, an impossibility theorem can always be reformulated into a positive universal statement. For example, we can formulate Fermat's last theorem in the positive form: For all natural numbers $n>2$ and all integers $a,b,c,$ we have

$a^n + b^n \neq c^n$. However, it is more usual to formulate the result as the impossibility of finding integer solutions a,b,c to the equation $a^n + b^n = c^n$ when $n>2$. The reason for the negative formulation is that our experience with Pythagorean triples suggests that the equation should have solutions. In general, impossibility theorems in mathematics usually say that a problem that one would expect to have a solution in fact has none.

Such theorems are as old as deductive mathematics itself, going back to the ancient Greeks. Already the Pythagoreans proved the incommensurability of the side and the diagonal of the square: It is impossible to find a line segment that measures both the side and the diagonal of a square a whole number of times. Yet, until well into the 19th century, impossibility theorems were often considered less interesting than positive statements. This changed with Hilbert who insisted that an impossibility proof would count as a kind of solution of a problem. This allowed him to formulate the (later disproved) meta-theorem that all mathematical problems are solvable in the sense that either one can find a solution or one can prove that the problem is impossible. “In mathematics there is no ignorabimus”, as he famously phrased it.

One of the things that called Hilbert’s attention to impossibility theorems was the recent (1882) proof of the transcendence of π, which implies the impossibility of constructing the quadrature of the circle by ruler and compass. This was the last of the classical Greek construction problems to be solved in the negative, the duplication of the cube and the trisection of the angle having been proved impossible in 1837. It is conspicuous that the Greeks came to suspect that the latter two of the three problems were unsolvable by ruler and compass but they do not seem to have felt the need for, and probably not even the possibility of, a proof of the impossibility. Indeed, the method of indirect proof that allowed them to prove the incommensurability, was not so immediately applicable in the case where one tries to argue that a construction rather than a line segment does not exist.

However, in the 17th century when Descartes and Fermat began connecting geometry to the new algebra (analytic geometry) they also opened the road for an impossibility proof of the classical problems. In two recent papers I have analyzed some of the impossibility arguments for the classical problems given in the 17th century. Descartes loosely argued that problems that can be constructed by ruler and compass will lead to quadratic equations; but since the duplication of the cube and the trisection of the angle lead to cubic equations they can not be constructed by such means. The argument contains a good idea and was apparently accepted by many of his contemporaries, but it is in fact far from satisfactory. The most conspicuous problem is that the translation of “solvable by ruler and compass” into algebra was unprecise. This translation was clarified by Gauss in 1799.

Another problem with Descartes’ impossibility “proof” is the argument that a certain cubic equation cannot be reduced to the solution of a (or a number of) quadratic equations. However, this vagueness was cleared up using algebraic methods that were developed in connection with the proof of another impossibility result: the impossibility of solving the general quintic by radicals. The methods developed by Lagrange, Abel and Galois allowed Wantzel in 1837 to prove the impossibility of constructing the duplication of the cube and the trisection of the angle by ruler and compass.

Wantzel’s proof is mentioned in modern histories of mathematics as the final solution of the two classical problems, but Wantzel’s paper was hardly noticed until well into the 20th century. This struck me as odd, so I wrote a paper trying to understand this lack of recognition. The oversight was probably a result of several factors: 1. Many mathematicians in the beginning of the 19th century thought that the problems had been solved in the 17th century. 2. Many mathematicians did not value such impossibility results highly. 3. Wantzel’s proof was unnecessarily complicated. In fact, I may be one of the few mathematicians who have worked my way through it. At least no one seems to have noticed a minor mistake in the proof until Robin Hartshorne pointed it out in 2000. The first completely rigorous proof of the two impossibility results seems to be due to Julius Petersen (1871 and 1879). This remarkable Danish mathematician has been studied in great detail by Bjarne Toft in Odense and I have been fortunate to take part in some of the research.

About the same time as Wantzel, the young Liouville was busy using the new Galois-theoretic methods to prove that one cannot express certain integrals (e.g. elliptic integrals) in finite terms. This was the first kind of impossibility theorem whose history I studied. My latest paper on the history of mathematical impossibility is concerned with fields of applied mathematics that I have never studied before: social choice theory and the theory of elections. The theorem in question is Arrow's impossibility theorem stating that it is impossible to construct election procedures satisfying a small number of desirable properties. Its history turned out to be dramatic including priority disputes and the actors' emotional reaction to the surprising and unpleasant result: "my stomach revolted in something akin to physical sickness", as Duncan Black formulated it.

Having published papers on the history of special impossibility theorems, I now plan to write a semi-popular book on the role of impossibility in the development of mathematics from the Babylonians to the 20th century. I have even written seven pages of it! My only fear is that towards the end of the book I have to write about Gödel's theorem and Turing's results that in a sense disproved Hilbert's optimistic rejection of ignorabimus. I find it very hard to grasp these results and their proofs and thus I am afraid that my historical analyses of them will be either vague, naïve or wrong. Here I hope that my logical colleagues can help me.

It is remarkable that in mathematics we can use the methods of mathematics itself to establish impossibilities. This probably makes mathematics the only science that can deal with its own limits using its own methods. This deserves to be communicated to a broader audience.