THE PI MU EPSILON 100TH ANNIVERSARY PROBLEMS: PART III

STEVEN J. MILLER*

As 2014 marks the 100^{th} anniversary of Pi Mu Epsilon, I thought it would be fun to celebrate with 100 problems related to important mathematics milestones of the past century. The problems and notes below are meant to provide a brief tour through some of the most exciting and influential moments in recent mathematics. As editor I have been fortunate to have so many people contribute (especially James Andrews and Avery Carr, who assisted greatly in Parts I and II); for each year a contributor has written a description of the event and proposed a problem for the reader's enjoyment. No list can be complete, and of course there are far too many items to celebrate. This list must painfully miss many people's favorites.

As the goal is to introduce students to some of the history of mathematics, accessibility counted far more than importance in breaking ties, and thus the list below is populated with many problems that are more recreational. Many others are well known and extensively studied in the literature; however, as the goal is to introduce people to what can be done in and with mathematics, I've decided to include many of these as exercises since attacking them is a great way to learn. We have tried to include some background text before each problem framing it, and references for further reading. This has led to a very long document, so for space issues we split it into four parts (based on the congruence of the year modulo 4). That said: Enjoy!

1915

General Relativity and The Absolute Differential Calculus

Ricci-Curbastro (1853 - 1925) developed a branch of Mathematics known as the *Absolute Differential Calculus* in his studies of geometrical quantities and physical laws that are invariant under general coordinate transformations. The concept of a *tensor* first appeared in Ricci's work although a restricted form of tensors had been previously introduced in *Vector Analysis*. In 1901, Ricci and his student T. Levi-Civita, published a complete account of the methods of absolute differential calculus and their applications [7]. Their work was a natural extension of the mathematics of curved surfaces introduced by Gauss and developed by Riemann and others, and of the Vector Analysis of Gibbs and Heaviside.

Einstein's *Special Theory of Relativity* deals with the study of the dynamics of matter and light in frames of reference that move uniformly with respect to each other – the so-called inertial frames. Those quantities that are invariant under the (Lorentz) transformation from one frame to another are of fundamental importance. They include the invariant interval between two events $(ct)^2 - x^2$, the energy-momentum invariant $E^2 - (pc)^2$, and the frequency-wave number invariant $\omega^2 - (kc)^2$. Here c is the invariant speed of light in free space. The Special Theory is formulated in a gravity-free universe. Ten years after Einstein completed his Special Theory he published his crowning achievement, *The General Theory of Relativity* [4, 5]. This is a theory of space-time and dynamics in the presence of gravity. The essential mathematical methods used in the General Theory are Differential Geometry, and the Absolute Differential Calculus (that Einstein referred to as *Tensor Analysis*). Einstein devoted more than five years to mastering the necessary mathematical techniques.

*Williams College, Editor

He corresponded with Levi-Civita, asking for his advice on applications of Tensor Analysis.

A tensor is a set of functions, fixed in a coordinate system that transforms under a change of the coordinate system according to definite rules. Each tensor component in a given coordinate system is a linear, homogeneous function of the components in another system. If there are two tensors with components that are equal when both are written in one coordinate system, then they are equal in all coordinate systems; these tensors are invariant under a transformation of the coordinates [8]. Physical laws are true in their mathematical forms for all observers in their own frames of reference (coordinate systems) and therefore the laws are necessarily formulated in terms of tensors.

Einstein's belief that matter generates a curvature of space-time led him to the notion that space-time is Riemannian; it is locally flat. The entire curved surface can be approximated by tiling with flat frames. Einstein assumed that in such locally flat regions, in which there is no appreciable gradient in the gravitational field, a freely-falling observer experiences all physical aspects of Special Relativity; the effects of gravity are thereby locally removed! (This assumption is known as the Principle of Equivalence).

In Special Relativity, the energy-momentum invariant is of fundamental importance. It involves not only the concepts of energy E and momentum p but also that of mass m: $E^2 - (pc)^2 = (mc^2)^2$, where m is the rest mass of the object.

Einstein therefore proposed that in General Relativity, it is mass/energy that is responsible for the curvature. He therefore introduced the stress-energy tensor, well known in Physics, to be the quantity related to the curvature. He proposed that the relationship between them is the simplest possible: they are proportional to each other [**?**]: Curvature tensor =k·Stress − Energy tensor, where the constant k is chosen so that the equation agrees with Newton's Law of Gravity for the motion of low - velocity objects in weak gravitational fields ($k = 8\pi G/c^4$, with G Newton's constant).

Centennial Problem 1915. *Proposed by Frank W. K. Firk, Yale University.*

In 1907, Einstein [3] combined his Principle of Equivalence with the Theory of Special Relativity (1905) and predicted that clocks run at different rates in a gravitational potential, and light rays bend in a gravitational field. This work predated his introduction of the Theory of General Relativity (1915). In General Relativity, objects falling in a gravitational field are not being acted upon by a gravitational force (in the Newtonian sense). Rather, they are moving along geodesics (straight lines) in the warped space-time that surrounds massive objects. The observed deflection of light, grazing the sun, is a test of the Principle of Equivalence. Tests of General Relativity are, at the present time, an active part of research in Physics and Astronomy. The problem below is related to one of these tests; for a review of early tests of Gravitation Theory see Will [6].

The famous Schwarzschild line element, in the region of a spherical mass M (obtained as an exact solution of the Einstein field equations) is, in polar coordinates

$$ds^2 \;=\; c^2(1 - 2GM/rc^2)dt^2 - (1 - 2GM/rc^2)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

If $\chi = 2GM/rc^2 \ll 1$, the coefficient $(1-\chi)^{-1}$, of dr^2 in the Schwarzschild line element can be replaced by the leading term of its binomial expansion to give the "weak field" line element

$$ds_W^2 \;=\; (1-\chi)(cdt)^2 - (1+\chi)dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

At the surface of the Sun, the value of χ is $4.2 \cdot 8^{-6}$, so that the weak field approximation is valid in all gravitational phenomena in our solar system).

Consider a beam of light traveling radially in the weak field of a mass M, then

$$ds_W^2 \;=\; 0 \text{ (a light} - \text{like interval)}, \quad \text{and} \;\; d\theta^2 + \sin^2\theta d\phi^2 \;=\; 0,$$

giving

$$0 \;=\; (1-\chi)(cdt)^2 - (1+\chi)dr^2.$$

The "velocity" of the light $v_L = dr/dt$, as determined by observers far from the gravitational influence of M, is therefore

$$v_L \;=\; c\sqrt{(1-\chi)/(1+\chi)} \;\neq\; c$$

if $\chi \neq 0$. Note observers in free fall near M always measure the speed of light to be c.

Expanding the term $\sqrt{(1-\chi)/(1+\chi)}$ to first order in χ, we obtain

$$v_L(r) \;\approx\; c(1 - 2GM/rc^2 + \cdots),$$

so that $v_L(r) < c$ in the presence of a mass M according to observers far removed from M.

In Geometrical Optics, the refractive index n of a material is defined as $n := c/v_{\text{medium}}$, where v_{medium} is the speed of light in the medium. We introduce the concept of the *refractive index of space-time* $n_G(r)$ at a point r in the gravitational field of a mass M:

$$n_G(r) \;:=\; c/v_L(r) \;\approx\; 1 + 2GM/rc^2.$$

The value of $n_G(r)$ increases as r decreases. This effect can be interpreted as an increase in the "density" of space-time as M is approached.

As a plane wave of light approaches a spherical mass, those parts of the wave front nearest the mass are slowed down more than those parts farthest from the mass. The speed of the wave front is no longer constant along its surface, and therefore the normal to the surface must be deflected. The deflection of a plane wave of light by a spherical mass M of radius R, as it travels through space-time, can be calculated in the weak field approximation. Show that in the weak field approximation the total deflection $\Delta\alpha$ equals $4GM/Rc^2$. This is Einstein's famous prediction on the bending of light in a gravitational field.

REFERENCES

[1] P. A. M. Dirac, "General Theory of Relativity", Wiley, New York, 1975.

[2] A. Einstein, "On the Electrodynamics of Moving Bodies", Annalen der Physik **17** (1905), 891–921. http://www.fourmilab.ch/etexts/einstein/specrel/www/. For more of Einstein's papers from this time period, see http://www.loc.gov/rr/scitech/SciRefGuides/einstein.html.

[3] A. Einstein, Jahrbuch Rad. **4**, 410 (1907). http://www.relativitycalculator.com/pdfs/Einstein_1907_Comprehensive_Essay_PartsI_II_III.pdf.

[4] A. Einstein, "The Foundation of the General Theory of Relativity", Annalen der Physik (1916). http://web.archive.org/web/20060831163721/http://www.alberteinstein.info/gallery/pdf/CP6Doc30_English_pp146-200.pdf. Also http://www.marxists.org/reference/archive/einstein/works/1910s/relative/relativity.pdf for "Relativity: The Special and General Theory", a book by Einstein in 1920 collecting the two papers.

[5] A. Einstein, "The Meaning of Relativity", MJF Books, New York, 5th Ed., 1956.

[6] D. F. Lawdon, "An Introduction to Tensor Calculus, Relativity and Cosmology", Wiley, New York, 3rd Ed. (1982).
[7] G. Ricci-Curbastro and T. Levi-Civita, "Methods of Differential Calculus and their Applications", Math. Ann. **54** (1901), 125–201.
[8] C. M. Will, in "General Relativity", Eds. S. W. Hawking and W. Israel, Chapter 2, Cambridge University Press, Cambridge, 1979.

1919

Brun's Theorem

One of the most tantalizing conjectures about prime numbers is that there are infinitely many pairs of primes differing by 2 (or, more generally, given any even integer $2m$ there are infinitely many pairs of primes differing by $2m$). Though still open, there has been remarkable progress on this problem over the past 100 years, culminating in Yitang Zhang's groundbreaking work in 2013 proving that there is *some* even number such that infinitely many pairs of primes differ by this. This result has been improved and generalized by many, especially Maynard, Tao and the Polymath8 project.

One of the earliest results in the field is due to Brun, who proved the sum of the reciprocals of the twin primes converged (if that sum diverged there would have to be infinitely many twin primes; sadly its convergence can't resolve the question of the infinitude of such primes). The value of this sum,

$$\sum_{p:p,p+2 \text{ prime}} \left(\frac{1}{p}+\frac{1}{p+2}\right) = \left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\cdots,$$

is called Brun's constant in his honor. It equals approximately 1.9021605, and the search for a good approximation to it led Thomas Nicely of Lynchburg College to discovering a floating point error in Intel's Pentium processor (which led hundreds of millions of dollars of loss for Intel, demonstrating the power of pure mathematics!).

Centennial Problem 1919. *Proposed by Steven J. Miller, Williams College.*

Let $\mathbb{N}_{\text{twin}}$ be the set of all integers whose only prime factors are twin primes. Thus 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21 and 25 are all in $\mathbb{N}_{\text{twin}}$ while 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 23 and 24 are not. Does

$$\mathcal{S} := \sum_{n\in\mathbb{N}_{\text{twin}}} \frac{1}{n}$$

converge or diverge? If it converges approximate the sum.

REFERENCES

[1] V. Brun, "La série 1/5 + 1/7 + 1/11 + 1/13 + 1/17 + 1/19 + 1/29 + 1/31 + 1/41 + 1/43 + 1/59 + 1/61 + ..., oú les dénominateurs sont nombres premiers jumeaux est convergente ou finie", Bulletin des Sciences Mathématiques **43** (1919), 100–104, 124–128. Available online at http://gallica.bnf.fr/ark:/12148/bpt6k486270d.
[2] J. Maynard, "Small gaps between primes", http://arxiv.org/abs/1311.4600.
[3] T. Nicely, "The pentium bug", available online at http://www.trnicely.net/pentbug/pentbug.html.
[4] T. Nicely, "Enumeration to 10^{14} of the Twin Primes and Brun's Constant", Virginia J. Sci. **46** (1996), 195–204.
[5] Polymath8, "Bounded gaps between primes", http://michaelnielsen.org/polymath1/index.php?title=Bounded_gaps_between_primes.

1923

The Circle Method

In 1920 Hardy and Littlewood published their first joint work on the Circle Method, where they used it to attack Waring's Problem (writing all positive integers as a sum of at most a given number of k-powers); see the problem from that year for more details about the Circle Method. In 1923 they continued attacking related additive problems. In *Some problems of 'Partitio Numerorum': III. On the expression of a number as a sum of primes*, they connected the distribution of zeros of Dirichlet L-functions to the writing odd numbers as the sum of three primes (if no Dirichlet L-function has a zero with real part $3/4$ or larger, then all sufficiently large odd numbers are the sum of three primes). Their paper attacks many additive questions, including how many twin prime pairs there should be up to x, as well as generalizations to admissible tuples of primes. Not all tuples of primes can appear; for example, the only triple of primes with the difference between adjacent primes equal to 2 is (3, 5, 7); this is because if the neighbor difference is 2 then exactly one of the three numbers must be a multiple of 3, and thus the only way all can be prime is if the prime 3 is one of the numbers. Hardy and Littlewood conjectured that as long as there are no congruence obstructions to a k-tuple of primes, then there are infinitely many such k-tuple. Further, they give formulas for how many such k-tuples exist when the primes are all at most x. The main term of their answer is the product of a multiplicative factor that depends on the $k-1$ neighbor differences, which vanishes if there is a congruence obstruction, and $x/\log^{k+1} x$.

Centennial Problem 1923. *Proposed by Steven J. Miller, Williams College.*

Spectacular recent work by Green and Tao prove that the primes contain arbitrarily long arithmetic progressions; this of course is a trivial consequence of the Hardy-Littlewood k-tuple conjectures, which are far beyond our ability to prove. Consider the significantly easier problem of whether or not there are infinitely many triples of primes with the same difference; in other words, can you prove there are infinitely many triples of primes of the form $(p, p + 2m_p, p + 4m_p)$ (so the difference between neighboring primes can depend on the first prime in the sequence)? Note we are *not* asking that two different triples have the same difference between primes; thus (11, 17, 23) (with a common difference of 6) and (29, 41, 53) (with a common difference of 12) would count. While this follows immediately from the Green-Tao Theorem, can you prove this elementarily? More generally, consider any sequence of integers such that the number of terms of $\{1, \ldots, x\}$ in our sequence is $f(x)$. Find a function $g(x)$ such that if $f(x) > g(x)$ then there will be infinitely many triples of terms in our sequence where each triple is of the form $(n, n + m_n, n + 2m_n)$ (note different triples can have different n_m's). Find a function $h(x)$ such that if $f(x) < h(x)$ then there is a choice for our sequence of integers such that we do not have infinitely many such triples. Of course one could take $g(x) = x$ and $h(x) = 2$; the question is to find the best values of these. For the primes, if x is large then the number of primes at most x is about $x/\log x$; what do your results say about the primes? If you can't find a function which will yield infinitely many triples, can you at least insure the existence of one, or find a special sequence so that there are no triples?

REFERENCES

[1] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", Annals of Mathematics **167** (2008), 481–547. arXiv:math.NT/0404188.

[2] G. H. Hardy and J. E. Littlewood, "Some problems of 'Partitio Numerorum': III. On the expression of a number as a sum of primes", Acta Math. **44** (1923), 1–70.

1927

William Lowell Putnam Mathematical Competition

Many a problem solver will be aware of the William Lowell Putnam Mathematical Competition, a North American undergraduate contest administered by the Mathematical Association of America, and founded in 1927 by Elizabeth Lowell Putnam in honor of her late husband who firmly believed in the virtues of more academic rivalry between universities. Among the many unwritten traditions of the Putnam exam is that every exam should have at least one problem that uses the year number as part of a problem statement or its solution. So, it is also a very fitting reversal of logic, or a meta-twist, that the Putnam exam itself is the subject of this section titled "1927".

Joe Gallian has written a fabulous overview of the Putnam exam's history, milestones, statistics, and trivia [2]. Offered every year since 1938 (except in 1943–45), the Putnam exam's roots include a math competition also sponsored by Elizabeth Lowell Putnam and held in 1933 between ten Harvard students and ten West Point cadets; the cadets both won the team contest and had the top individual score. Earlier Putnam exams featured problems in areas closer to the introductory technical undergraduate curriculum such as calculus, differential equations, or geometry; in more recent years, a very recognizable blend of topics including also linear algebra, some abstract algebra, combinatorics, number theory (or even an occasional advanced topic on harder questions) characterizes each year's 12 problems. Five most successful contestants each year are named Putnam Fellows, one of whom is also awarded a fellowship for graduate study at Harvard; eight persons so far have been a Putnam Fellow the maximum possible four times. Other substantial monetary team and individual prizes are given, and an Elizabeth Lowell Putnam prize may be awarded to one female contestant. The original intent to boost team spirit and provide an avenue for students to fight for their institution's glory in an academic subject helps understand the peculiar ranking system, in which every participating institution must designate a three-person team in advance, and the team ranking is obtained by adding the team members' individual ranks (rather than their scores). Since higher scores are obtained by many fewer students, a university whose all three team members solve seven problems will usually rank markedly higher than the one where two brilliant team members solve nine problems and the third solves three.

The Putnam exam has been called "the hardest math test in the world" [1, 3], and not without reason: the median score has budged above 1 point out of 120 in only four years since 1999 and then never above 3, while fully 62.6% of 2006 entrants scored 0. A score of 1 point should be worn as a badge of honor, since a student must make substantial progress toward an actual solution to receive any points for a problem; checking small examples or stating some immediate conclusions typically doesn't make the cut. Each of the 12 problems is graded on a scale from 0 to 10 points, with the only scores allowed being 0,1,2,8,9,10; thus, the grader must decide whether the problem is essentially solved or not. A submission that solves one of the two main cases, or one that contains the structure of a full solution but has a serious flaw might get 1 or 2 points. On the other hand, a submission that contains all ideas of a full solution but neglects to check a straightforward sub-case might get 8 or 9 points; the full mark of 10 points is reserved for essentially perfect solutions. The first round of grading currently occurs in December at Santa Clara University; imagine several dozens of mathematicians in a few rooms tackling, one paper at a time over the span of four days, boxes containing the collective output of over 4,000 competitors from over 500 colleges on 12 problems, and you will have a pretty good picture.

Undergraduate students solve problems in several competitions around the world, such as the annual Schweitzer competition in Hungary, the Jarník competition in central Europe, the famous competitions at Moscow's and Kiev's Mech-Mats, or the International Mathematics Competition for University Students [5], an annual contest held in Europe that has also seen participation from several American universities.

Opinions differ on the extent to which the Putnam exam or other competitions mimic the mathematical research experience or are somehow reflective of the student's research aptitude; see several Putnam Fellows' perspectives in [1]. The Putnam exam was, of course, never designed for such use. Five Fellows (Feynman, Milnor, Mumford, Quillen, and Wilson) have been subsequently recognized with a Fields Medal or a Nobel Prize, and many dozens more have become distinguished mathematicians at top universities and research institutes. Notable Putnam competitors include many Abel Prize winners, MacArthur Fellows, AMS and MAA presidents, members of the National Academy of Sciences as well as many winners of the Morgan Prize for undergraduate research. Many others have chosen entirely different careers, and obviously many top-notch mathematicians have never taken or particularly enjoyed problem mathematics. (Also, while many of the Putnam bourgeoisie were successful in the high-school IMOs, the two contests retain very distinct mathematical profiles.) Ask mathematics graduate program admissions chairs or hedge fund managers and many will tell you that, while neither a prerequisite or a guarantee of success, a candidate's good showing on the Putnam exam gets their attention. Putnam problems test a specific kind of ingenuity over technical mastery and are sometimes seen as occupying a universe of their own, but here as in Hamming [4] we must remember that Putnam problems "were not on the stone tablets that Moses brought down from Mt. Sinai" – they are composed by a committee of working mathematicians designated by the MAA, and so their evolution over time perhaps reflects our collective style and taste.

What makes a good Putnam problem then? Bruce Reznick [7] has written about the process of writing for the Putnam exam in charm and detail that neither the length nor the writer's ability allow here. André Weil [8], paraphrasing the English poet Housman who had used an example of a fox-terrier hunting for a rat to explain why he can't define poetry, famously quipped: "When I smell number-theory I think I know it, and when I smell something else I think I know it too". (He then proceeded to argue that analytic number theory is *not* number theory, but this is clearly a subject for another article.) Putnam takers and experienced problem-solvers will similarly spot a juicy problem. It will be accessible but not trivial, and challenging but not impossibly so, it will relate to important mathematics but every time with an unexpected twist, it will make you smile and, in Reznick's words, leave the exam whistling it in your mind like a catchy tune.

I propose the following Putnam problem for your whistling enjoyment. It appeared as problem A3 in the 2013 exam and requires nothing beyond first-year college mathematics, but it was borne out of thinking about measure theory. Don't just solve it and bask in your glory. Make yourself a hot beverage, relate the solution to your other mathematical experiences, continue the story that it tells; you will have new problems of your own in no time.

Centennial Problem 1927. *Proposed by Djordje Milićević, Bryn Mawr College.*

Suppose that the real numbers $a_0, a_1, \ldots, a_n$ and x, with $0 < x < 1$, satisfy

$$\frac{a_0}{1-x} + \frac{a_1}{1-x^2} + \cdots + \frac{a_n}{1-x^{n+1}} \ = \ 0.$$

Prove that there exists a real number y with $0 < y < 1$ such that

$$a_0 + a_1 y + \cdots + a_n y^n \ = \ 0.$$

REFERENCES

[1] G. L. Alexanderson, "How Putnam fellows view the competition", Focus, December 2004, 14–15, http://www.maa.org/sites/default/files/pdf/pubs/dec04.pdf

[2] J. A. Gallian, "The Putnam competition from 1938–2013", available at http://www.d.umn.edu/~jgallian/putnam.pdf. This is an updated version of *The First sixty-six years of the Putnam competition*, American Mathematical Monthly **111** (2004), 691–699.

[3] L. Grossman, "Crunching the numbers", Time magazine, December 16, 2002, http://content.time.com/time/magazine/article/0,9171,400000,00.html.

[4] R. W. Hamming, "The Unreasonable Effectiveness of Mathematics", The American Mathematical Monthly **86**, no. 2 (1980), 81–90.

[5] International Mathematics Competition for University Students, http://www.imc-math.org.uk.

[6] K. S. Kedlaya, B. Poonen, R. Vakil, "The William Lowell Putnam Mathematical Competition 1985–2000: Problems, Solutions, and Commentary", Mathematical Association of America, 2002.

[7] B. Reznick, "Some thoughts on writing for the Putnam", available at http://www.math.uiuc.edu/~reznick/putnam.pdf. This is an updated version of "Some Thoughts on Writing for the Putnam", in Mathematical Thinking and Problem Solving, edited by A. Schoenfeld, Erlbaum, 1993, 19–29; a revised version appeared in "The William Lowell Putnam Mathematical Competition: 1985–2000 Problems, Solutions and Commentary", by Kiran S. Kedlaya, Bjorn Poonen and Ravi Vakil, MAA, 2002, pp. 311–320.

[8] A. Weil, "Essais historiques sur la theorie des nombres", Monographie de l'Enseignement mathematique, 1975.

1931

The Ergodic Theorem

A mathematical, discrete dynamical system consists of a set of states X and a map or transformation T from X onto X, which we assume invertible. For example, for the case of a pendulum a state would be a point consisting of the position of the pendulum bob and the bob's momentum. There is an equation based on the laws of mechanics whose solution gives us what the state would be after a certain amount of time has elapsed; this gives us the transformation on the set of states.

One of the first questions concerning a dynamical system is the eventual behavior of points. Given a system in state y, let $T(y)$ be the state of the system one unit of time later. Thus if the system starts off in state x, it next moves to $T(x)$, then to $T(T(x))$ (which for notational ease we denote by $T^2(x)$, though some authors prefer $(T \circ T)(x)$ or $T^{\circ 2}(x)$), and more generally after n units of time it is at $T^n(x)$. If A is a set of states and $\mathbb{I}_A$ is the characteristic or indicator function of A, $\sum_{i=1}^n \mathbb{I}_A(T^i(x))$ counts the number of visits of x to A up to time n. The time average of visits of x to A is the limit, if it exists, as $n \to \infty$ of $\frac{1}{n}\sum_{i=1}^n \mathbb{I}_A(T^i(x))$.

A consequence of Gibbs and Boltzman's investigations in statistical mechanics was the *Ergodic Hypothesis*. A version of the ergodic hypothesis states that the time average of a system should equal the space average, which is the relative size of A in X. This space average is given by a measure or probability defined on a class of subsets of X, called the measurable sets. We denote the measure of a set A by $m(A)$, and we study systems where $m(X)$ is 1 (thus we may view $m(A)$ as the probability we are in A). In addition we assume that the measure m is preserved by the transformation T, i.e., the measure of the set of points in any set A remains the same as they evolve through time, so $m(A) = m(T(A))$.

In 1931 von Neumann [2], and shortly after though published a bit earlier Birkhoff [1], proved that time averages exist and equal the space averages for measure-preserving systems satisfying a condition called ergodicity. Ergodicity means that the only invariant sets of T are the empty set $\emptyset$ and the entire space X, or differ from one of these by a set of measure zero. In other words, if A is invariant then A differs from either $\emptyset$ or X by a set of measure zero. A set E is said to be invariant if $T(x)$ is in E if and only if x is in E.

If E is invariant and x starts in E, all its iterates stay in E and no point outside E visits E. That means that if T were not ergodic there would exist sets E and E^c, both of positive measure, where the dynamics of T on E will be totally unrelated to the dynamics of T on E^c, in other words, one could decompose T into simpler systems. The remarkable fact is that when T is ergodic one obtains a time average which is a limit that exists and equals what is expected.

We can now state the Birkhoff Ergodic Theorem: for all measurable sets A there exists a set of measure zero N so that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_A(T^i(x)) \;=\; m(A) \text{ for all } x \text{ outside } N.$$

This convergence is called almost everywhere pointwise convergence. An important part of the theorem is that the limit exists. In fact, once we know the limit exists, using standard results from analysis it is possible to show the limit must equal the measure of A. An immediate consequence is that for all sets A of positive measure, every point of X outside a set of measure zero visits the set A, and furthermore, the visits are with the right frequency. Thus we know a lot about the orbit of almost every point.

This theorem has had a strong influence in analysis and has many consequences. For example, it can be used to prove Weyl's uniform distribution property[1] and the Law of Large Numbers[2] in probability. The ergodic theorem is in fact a bit more general: one can replace $\mathbb{I}_A$ by any Lebesgue integrable function f, and then $m(A)$ is replaced by the integral of f. The theorem proved by von Neumann is similar to Birkhoff's but differs in two respects. In von Neumann's theorem the convergence is not pointwise but in the norm of the Hilbert space where the functions are, and also the theorem is in the context of unitary operators on a Hilbert space. For an introduction and proof the reader may consult [5]. Further historical details and current developments can be found in [4].

Centennial Problem 1931. *Proposed by Cesar E. Silva, Williams College.*

In 1988 Bourgain [3] proved that for every function f whose square is integrable, if $p(x)$ is a polynomial with integer coefficients then the time average along polynomial times exists outside a set of measure zero. In other words $\frac{1}{n}\sum_{i=1}^{n} f(T^{p(i)}(x))$

---

[1]If α is irrational, then the set $\{n\alpha \bmod 1\}_{n=1}^{\infty}$ is equidistributed in $[0,1]$

[2]Let $X_1, \dots, X_n$ be independent random variables drawn from a common, nice distribution with mean μ. If we observe values $x_1, \dots, x_n$ then as $n \to \infty$ the sample average $(x_1 + \cdots + x_n)/n$ converges (in some sense) to μ.

converges for almost all x. When all powers of T are ergodic it follows that this limit equals the integral that is expected. It is reasonable to ask what happens when the function f is merely integrable, even in the case of the squares: $p(i) = i^2$. It was shown recently by Buczolich and Mauldin [6] that the theorem for the squares fails when f is only assumed to be integrable. This proof has been extended recently by P. LaVictoire. It would be interesting to find simpler proofs of all of these results.

REFERENCES

[1] G. Birkhoff, "Proof of the ergodic theorem", Proc. Nat. Acad. Sci, USA **17** (1931), 656–660.

[2] J. von Neumann, "Proof of the quasi-ergodic hypothesis", Proc. Nat. Acad. Sci., USA **18** (1932), 70–82.

[3] J. Bourgain, "On the maximal ergodic theorem for certain subsets of the integers", Israel J. Math. **61** (1988), 39–72.

[4] V. Bergelson, "Some historical comments and modern questions around the Ergodic Theorem", Dynamics of Complex Systems 1-11, Research Institute for Math. Sciences, Kyoto, 2004.

[5] C. E. Silva, "Invitation to ergodic theory", Student Math. Library **42**, AMS, Providence, RI, 2008.

[6] Z. Buczolich and R. Daniel Mauldin, "Divergent square averages", Ann. of Math. (2) **171** (2010), 1479–1530.

1935

Hilbert's Seventh Problem

Our collection of one hundred problems celebrating 100 years is inspired, as are so many other collections, by the set of problems David Hilbert proposed at his keynote address at the International Congress of mathematicians in Paris in 1900 (see [2]). These problems were meant to chart important directions of research for the new century. Solution to any of these brings instant fame and membership in "The Honors Class" [3].

Hilbert seventh problem is the following: Let α and β be two algebraic numbers (this means they are solutions to polynomials of finite degree and integer coefficients; a transcendental number is a number that is not algebraic). Assume further that β is irrational. Then α^β is transcendental whenever α is neither 0 nor 1.

Problems along these lines have a long history. They have been studied by many mathematicians over the years. For example, in 1748 Euler proposed that if $\alpha \notin \{0, 1\}$ is a non-zero rational number and β is an irrational, algebraic number, then α^β is irrational. Notice this is a weak version of Hilbert's seventh problem.

In 1934 and 1935 Gelfond and Schneider independently resolved Hilbert's problem, successfully proving that under the assumptions above, α^β is transcendental.

Centennial Problem 1935. *Proposed by Jesse Freeman and Steven J. Miller, Williams College.*

The following problems are designed to give the reader an intuition of the development of the study of transcendental numbers and of the power of Gelfond and Schneider's result.

1. **Euler:** For $\alpha \in \mathbb{C}$, let $B_\alpha = \{\beta \in \mathbb{C} : \alpha^\beta \in \mathbb{Q}\}$. Show that for a general algebraic, but irrational γ, we have

$$B_\gamma \subsetneq \bigcap_{\alpha \in \mathbb{Q}} B_\alpha;$$

denote the union on the right by B. You may assume Gelfond and Schneider's result. Describe B, and investigate the algebraic structure of B_γ for a general γ.

2. **Cantor:** Using the fundamental theorem of algebra, prove that the set of algebraic numbers is countable. As the set of real numbers is uncountable, this implies that almost all real numbers are transcendental. Interestingly, while this argument proves that almost all numbers are transcendental it doesn't explicitly give us any!
3. **Liouville:** Liouville gave a method to construct transcendental numbers. The key ingredient is the following. Suppose α is an algebraic number of degree $d > 1$. Then there exists a positive constant $c(\alpha)$ such that for any rational number a/b,

$$\left|\alpha - \frac{a}{b}\right| \ > \ \frac{c(\alpha)}{b^d}.$$

We say $\alpha \in \mathbb{R}$ is a **Liouville number** if for every positive integer n there are integers a and b with $b > 1$ such that

$$0 \ < \ \left|\alpha - \frac{a}{b}\right| \ < \ \frac{1}{b^n}.$$

The result above implies that all Liouville numbers are transcendental; however, not all transcendental numbers are Liouville numbers. Show that the set of Liouville numbers in the interval $[-1, 1]$ has measure 0.
4. **Gelfond/Schneider/Hilbert:** Using Gelfond and Schneider's result, show that if the ratio of two angles in an isosceles triangle is algebraic and irrational that the ratio between the sides opposite those angles is transcendental.

REFERENCES

[1] M. Filaseta, "The Gelfond-Schneider Theorem and Some Related Results", accessed on-line February 2014. http://www.math.sc.edu/~filaseta/gradcourses/Math785/Math785Notes8.pdf.

[2] Wikipedia, "Hilbert's problems", http://en.wikipedia.org/wiki/Hilbert's_problems.

[3] B. Yandell, "The Honors Class: Hilbert's Problems and Their Solvers", A K Peters/CRC Press, 2001.

1939

The Power of Positive Thinking

A student doing a homework problem has an enormous advantage over a researcher: almost always the problem is solvable! This is especially true in undergraduate and beginning graduate classes, as these assignments are meant to reinforce lessons and help students learn techniques and how to attack problems in the field. It is hard to understate how important this is in successfully attacking a problem; psychologically it's a huge boost to know a solution exists.

There are many anecdotes and studies of people who were unaware of the difficulty of a problem, and blissfully unaware proceeded to make great progress. The following story and its variants have been circulating for years, and are the subject of this year's entry. We'll meet the protagonist, George B. Dantzig, again in the 1947 entry. The quote below is from an interview of him which appeared in the College Mathematics Journal in 1986 (available online; see [1]). He was asked why his PhD was on a statistics topic when he had taken so few stats courses.

> *It happened because during my first year at Berkeley I arrived late one day at one of Neyman's classes. On the blackboard there were two*

> *problems that I assumed had been assigned for homework. I copied them down. A few days later I apologized to Neyman for taking so long to do the homework – the problems seemed to be a little harder to do than usual. I asked him if he still wanted it. He told me to throw it on his desk. I did so reluctantly because his desk was covered with such a heap of papers that I feared my homework would be lost there forever. About six weeks later, one Sunday morning about eight o'clock, Anne and I were awakened by someone banging on our front door. It was Neyman. He rushed in with papers in hand, all excited: "I've just written an introduction to one of your papers. Read it so I can send it out right away for publication." For a minute I had no idea what he was talking about. To make a long story short, the problems on the blackboard that I had solved thinking they were homework were in fact two famous unsolved problems in statistics. That was the first inkling I had that there was anything special about them.*

Later in the interview he discusses how the story found its way into sermons.

> *The origin of that minister's sermon can be traced to another Lutheran minister, the Reverend Schuler of the Crystal Cathedral in Los Angeles. Several years ago he and I happened to have adjacent seats on an airplane. He told me his ideas about thinking positively, and I told him my story about the homework problems and my thesis. A few months later I received a letter from him asking permission to include my story in a book he was writing on the power of positive thinking. Schuler's published version was a bit garbled and exaggerated but essentially correct. The moral of his sermon was this: If I had known that the problems were not homework but were in fact two famous unsolved problems in statistics, I probably would not have thought positively, would have become discouraged, and would never have solved them.*

Centennial Problem 1939. *Proposed by Steven J. Miller, Williams College.*

Find the statements of the two problems Dantzig solved, read papers, and believe in yourself when confronted with challenges in the future.

REFERENCES

[1] D. J. Albers and C. Reid, "An Interview of George B. Dantzig: The Father of Linear Programming", College Mathematics Journal **17** (1986), no. 4, 293–314. Available online: http://www.jstor.org/stable/2686279.

1943

Breaking Enigma

One group of mathematicians played a crucial role in the Allied victory in World War II: the codebreakers. The German army encrypted its communications with Enigma machines, typerwriter-like devices that produce a fiendishly complicated code. The Polish Cipher Bureau developed the strategies to break the Enigma code in the early 1930's, but the largest codebreaking operation was British, headquartered at Bletchley Park, a Victorian manor northwest of London. The top-secret Bletchley Park project, codenamed "Ultra," is legendary. It employed mathematicians, linguists, chess masters, academics, composers, and puzzle experts. Recruiters once asked the *Daily Telegraph* to organize a crossword competition, and then secretly offered jobs to the winners. One of the leaders of Ultra was Alan Turing, the mathematician and pioneer of theoretical computer science.

Mathematically speaking, the Enigma machine generates a permutation $\tau \in S_{26}$ of the 26 letters of the alphabet. The permutation τ changes with every keystroke. Typing one letter sends electrical current through scrambling mechanisms–a plugboard, then a set of rotors, then a reflector, then back through the rotors and the plugboard–causing a different letter to light up. It also turns the rotors so that the next letter will be scrambled differently.

The scramblers are wired as follows: the plugboard has one plug for each letter, and 10 pairs of letters wired together. It defines a permutation π, which is a product of 10 two-cycles. The rotors are rotating wheels with a circle of 26 brass pins on one side and 26 electrical contacts on the other. The wiring from contacts to pins gives a fixed permutation ρ; depending on the position of the rotor, this permutation is conjugated by a power of $\alpha = (1\,2\,3\ldots 26)$. The reflector simply has 26 electrical contacts, connected in pairs by 13 wires. It gives a fixed permutation σ, a product of 13 two-cycles.

All together, the permutation τ is:

$$\pi^{-1}(\alpha^{-i_1}\rho_1\alpha^{i_1})^{-1}(\alpha^{-i_2}\rho_2\alpha^{i_2})^{-1}(\alpha^{-i_3}\rho_3\alpha^{i_3})^{-1}\sigma(\alpha^{-i_3}\rho_3\alpha^{i_3})(\alpha^{-i_2}\rho_2\alpha^{i_2})(\alpha^{-i_1}\rho_1\alpha^{i_1})\pi,$$

where the i_1, i_2, i_3, which represent the positions of rotors 1, 2, and 3, are varying. Since each permutation τ is a conjugate of σ, τ is also a product of 13 two-cycles, and $\tau^{-1} = \tau$, so a message can be encrypted and decrypted by machines with the same settings. The operator could choose 10 pairs of letters to connect in the plugboard, three out of five exchangeable rotors in any order, and 26 initial positions for each rotor: a total of $150,738,274,937,250$ initial settings for the machine.

This vast number of settings makes the Enigma code almost unbreakable, but it does have weaknesses. Since τ is a product of 13 two-cycles, no letter is ever encoded as itself. A codebreaker can look for common words and phrases in the encrypted text, and rule them out if any letters match. German messages also had various common formats which made them easier to guess. Furthermore, the Allied spies captured parts of enigma machines, decrypted messages, and information about initial settings. All this was just enough to break the code. By 1943, British Intelligence was able to decrypt most Enigma codes without knowing the initial settings of the machine. This capability was kept utterly secret–the Nazis never knew. Winston Churchill later told King George VI, "It was thanks to Ultra that we won the war."

Centennial Problem 1943. *Proposed by Ian Whitehead, University of Minnesota.*

Here, in honor of the Bletchley Park crossword contest, is a cryptography-themed cryptic crossword (see Figure 1), jointly written with Joey McGarvey. As in all cryptic crosswords, each clue contains a regular definition and a pun/anagram/wordplay hint. You must figure out how to parse the clue.

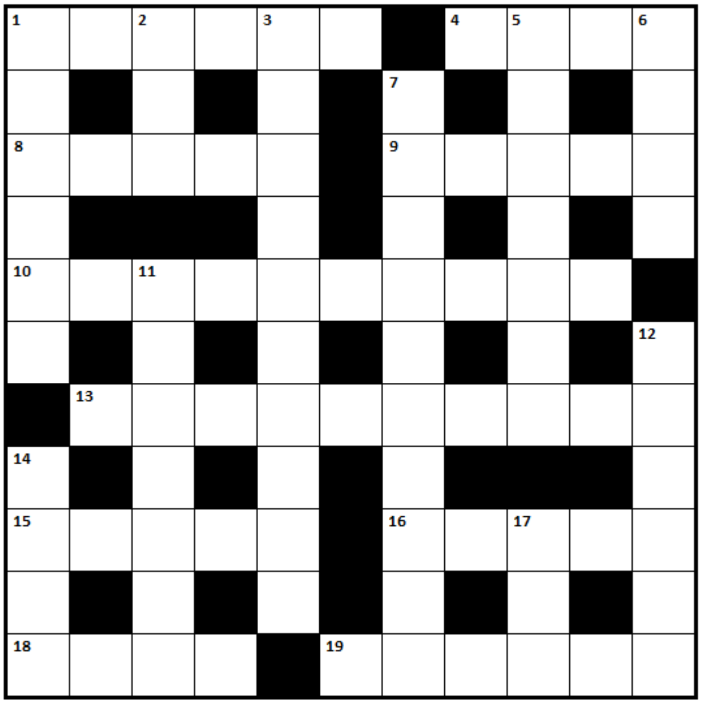
REFERENCES

[1] M. Cozzens and S. J. Miller, "The Mathematics of Encryption: An Elementary Introduction", AMS Mathematical World series **29**, Providence, RI, 2013, 332 pages.

[2] F. W. Hinsley and A. Stripp, "Codebreakers: The Inside Story of Bletchley Park", Oxford University Press, 1997.

[3] A. R. Miller, "The Cryptographic Mathematics of Enigma", NSA Pamphlet, 2001. http://ed-thelen.org/comp-hist/NSA-Comb.html.

[4] S. Singh, "The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography", Fourth Estate, 1999.



Across:

1 Film spooling backwards, noir, shot through the middle (6)
4 Therefore, um, leave (4)
8 Rambler has a stroll (5)
9 She sends girl to Wonderland (5)
10 Mixed-up, stuttering computer conversation (6,4)
13 Coded message is a hit; excerpt decoded (10)
15 Exploder, England's first decoder (5)
16 Zambian love god is subject of fundamental theorem of algebra (5)
18 Leonard's monocle (4)
19 One trades kernels for rings (6)

Down:

1 Responds to broken crates (6)
2 He receives short haircut (3)
3 Density of primes up to 30, reinterpreted in E. Noether (3,2,5)
5 Put out again regarding matter of dispute (7)
6 Without boundary, Peru, Ecuador absorbed by Ontario (4)
7 What primes can't be: stunned, eating 100 doughnuts (10)
11 Savant ran around in Massachusetts (4,3)
12 Calm state—as is (6)
14 Abraham Lincoln's first mathematician (4)
17 Code hidden in caesar salad (3)

FIG. 1. *Cryptography themed crossword.*

[5] W. Trappe and L. Washington, "Introduction to Cryptography and Coding Theory", Pearson Prentice Hall, 2006.

[6] J. Wilcox, "Solving the Enigma - History of the Cryptanalytic Bombe", NSA Pamphlet, 2001. http://www.nsa.gov/about/_files/cryptologic_heritage/publications/wwii/solving_enigma.pdf.

1947

The Simplex Method

We often encounter important problems where, while it is easy to write down an algorithm to find the solution, the run-time is so long that the method is worthless for practical purposes. One of the first examples of this is factorization: given an integer N, write it as a product of primes. It is ridiculously simple to write down a solution; we give one below without any attempts to improve its efficiency.

- `Step 1:` Initialize *Factors*(N) to be the empty set; as the name suggests, we'll store N's factors here. Let $M = N$ and $n = 2$ and continue to `Step 2`.
- `Step 2:` If n divides M then add n to *Factors*(N), replace M with M/n, and continue to `Step 3`. If n does not divide M then let $n = n + 1$; if $n = M$ then add n to *Factors*(N) and go to `Step 4`, else repeat this step.
- `Step 3:` If $M > 1$ then set $n = 2$ and repeat `Step 2`, else go to `Step 4`.
- `Step 4:` Print *Factors*(N) and stop.

This algorithm is painfully slow, and requires us to check all numbers up to N as potential divisors. While there are many improvements we can make to this algorithm, it will still be too slow to be practical for large N. The first is that once we find an n dividing M we should see how many times n goes into M; this would save us from having to return to $n = 2$ each time we restart `Step 2`. Next, we can notice that any prime factor of N is at most $\sqrt{N}$, and hence once $n > \sqrt{N}$ we know M is prime. Finally, if we are able to store the earlier prime numbers we need only check n that are prime. Even if we do all of these, however, we still have to check all primes at most $\sqrt{N}$. The Prime Number Theorem tells us that the number of primes at most x is approximately $x/\log x$; thus if N is around 10^{406} we would need to check about $2 \cdot 10^{200}$ numbers! This is *well* beyond what computers can do.

The lesson from the above is that, while factorization is easy to do in principle, in practice the 'natural' approach is too slow to be useful. It is a major open problem to find a fast way to factor numbers; if such an algorithm existed then encryption schemes such as RSA (described in the 1977 entry) would not be secure. Interestingly, while we cannot quickly factor a number, we *can* quickly tell if a number is prime (see the Primes in P entry from 2002).

Our topic for this year concerns a different problem where we are more fortunate, and a fast algorithm is available. Linear Programming is a beautiful subject, and a natural outgrowth of linear algebra. In linear algebra we try and solve systems of linear equations, such as $A\overrightarrow{x} = \overrightarrow{b}$. In linear programming we have a constraint matrix A and are now looking for a solution to $A\overrightarrow{x} = \overrightarrow{b}$ that maximizes the profit $\overrightarrow{c} \cdot \overrightarrow{x}$. Initially one allows inequalities in the linear system of constraints; however, by introducing additional variables we can replace all the inequalities with equalities. We also require each component of $\overrightarrow{x}$ to be non-negative (it is a nice exercise to show we may always do this, though we may need to introduce some additional variables); doing so allows us to put our linear programming problem into a standard, canonical form.

For example, one of the earliest successes in the subject concerns the Diet Problem. Here the entries of $\overrightarrow{x}$ are constrained to be non-negative, with x_k equaling the amount of product k consumed. Each food provides a different amount of essential vitamins and minerals, and we wish to find the cheapest diet that will keep us alive while ensuring that we get the minimum daily recommended allowance of each nutrient. See [2] for a humorous recounting of the meeting between linear programming and the Diet problem.

One of the first theorems proved in the subject concerns the candidates for our solution. We say $\overrightarrow{x}$ is feasible if it solves $A\overrightarrow{x} = \overrightarrow{b}$. It turns out that the space of feasible solutions has many nice properties. We call a solution $\overrightarrow{x}$ of the constraints a basic solution if the columns corresponding to the non-zero entries of $\overrightarrow{x}$ are linearly independent. It turns out that if there is an optimal solution to our problem, then that optimal solution is a basic solution. Moreover, there are only finitely many basic solutions. Thus we need *only* check all the basic solutions to find the optimal solution!

The problem is that if we try to search the space of basic solutions in the same manner we did to factor N, we'll never finish for large, real world problems. Let A be an $m \times n$ matrix. Thus our vector $\overrightarrow{x}$ has n components. We assume $n > m$, as otherwise the system is over-determined and there is at most one solution. If every subset of at most m columns of A is linearly independent, then the number of basic solutions is at most $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m}$. For n, m large this is approximately $n^m/m!$; to get some feel for how quickly this grows, if $n = 10000$ and $m = 100$ then the number of candidates exceeds 10^{241}, which is well beyond our factorization nightmare.

While it is nice that there are only finitely many candidates, typically there are too many to make checking each practical. We need an efficient way to navigate the space. Amazingly, there is such an approach. It is called *The Simplex Algorithm*, and was introduced by George Dantzing (whom we met in the 1939 entry) in 1947. His procedure, and other people's generalizations, allow us to solve many real world problems in reasonable amounts of time on everyday laptops, and transforms linear programming from theoretically interesting to practical.

Centennial Problem 1947. *Proposed by Steven J. Miller, Williams College.*

Building on the success of the Simplex Algorithm, it is natural to consider other generalizations of linear programming and ask if they too can be solved efficiently. The first natural candidate is to replace the word linear with quadratic. Unfortunately, while quadratic objective functions can often be handled, to date we still require the constraints to be linear.

To see why, we first consider another generalization. Instead of requiring the solution vector $\overrightarrow{x}$ to have non-negative real entries, let us require it to have non-negative integral entries. This is an extremely important special case of optimization problems. In the special case when the integers are either 0 or 1, we can interpret the components as binary indicator variables. Do we have a plane leave from Albany to Charlotte at 2:45pm? Do we show The Lego Movie on our biggest screen at 10:30am? If we're trying to solve the Traveling Salesman Problem (what is the route of least distance through a given set of cities), is the fifth leg of our trip from Boston to Rochester? These examples should hopefully convey the importance of solving binary integer programming problems.

Prove that if we could modify the simplex method to handle problems with quadratic constraints then we could solve all integer programming problems! For those familiar with the P versus NP problem (see the description at the Clay Mathematics Institute's list of Millenial Problems [1]), this would prove P equals NP.

REFERENCES

[1] Clay Mathematics Institute, "Millenial Problems".
http://www.claymath.org/millennium-problems.

[2] G. Dantzig, "The Diet Problem", Interfaces **20** (1990), no. 4, The Practice of Mathematical Programming, pp. 43–47. Published by: INFORMS.
http://dl.dropboxusercontent.com/u/5317066/1990-dantzig-dietproblem.pdf.

[3] J. Franklin, "Mathematical Methods of Economics: Linear and Nonlinear Programming, Fixed-Point Theorem", Springer-Verlag, New York, 1980.

1951

Tennenbaum's proof of the irrationality of $\sqrt{2}$

One of the rights of passage to being a mathematician, as opposed to simply a calculator and number cruncher, is the proof of the irrationality of $\sqrt{2}$ (which means we cannot write $\sqrt{2}$ as a/b for two integers a and b). There are now many proofs of this fact. Perhaps the most famous runs roughly as follows: (1) We may assume a and b are relatively prime. (2) Cross multiply and square, which yields $2b^2 = a^2$. (3) Since 2 divides the right hand side, 2 divides the left hand side (this is non-trivial and must be proved!). After a little work we find $2|a$, and thus we can write $a = 2x$ for an integer x. (4) We then find $2b^2 = 4x^2$, so $b^2 = 2x^2$, and a similar argument now gives $b = 2y$ for y an integer. This is a contradiction, as we assumed a and b were relatively prime.

There are, however, other proofs. Sometime in the 1950s Tennenbaum came up with the following geometric gem. Again we proceed by contradiction, and assume $\sqrt{2} = a/b$ (so $a^2 = 2b^2$). We may assume b is the smallest integer where we have such a relation. Consider a square with sides of length a, and draw squares of length b in the upper left and lower right corners. By assumption, the area of the two squares of length b equals that of the large square of length a (as $a^2 = 2b^2$). If you draw the picture, you see the two squares miss two small squares of length $a - b$, and double count a square of length $b - 2a$. Thus the double counted region must have the same area as the two missing squares, or $(b-2a)^2 = 2(a-b)^2$. Thus $\sqrt{2} = (b-2a)/(a-b)$, and a little work shows $a - b < b$. Thus we've found a smaller denominator that works, contradiction!

Centennial Problem 1951. *Proposed by Steven J. Miller, Williams College.*

Tennenbaum's construction is beautiful, and gives the irrationality of $\sqrt{2}$. Miller and Montague used similar geometric arguments to get the irrationality of $\sqrt{3}$, $\sqrt{5}$, $\sqrt{6}$ and $\sqrt{10}$. Can you geometrically prove the irrationality of $\sqrt{7}$, or $\sqrt[3]{2}$?

REFERENCES

[1] S. J. Miller and D. Montague, "Rational irrationality proofs", Mathematics Magazine **85** (2012), no. 2, 110–114. Available online: http://arxiv.org/pdf/0909.4913.pdf. For a discussion of the history of Tennenbaum's proof, see http://web.williams.edu/Mathematics/sjmiller/public_html/math/papers/Neode_Tennenbaum.PDF.

1955

Furstenberg's topological proof of the infinitude of primes

Every so often a proof comes along that acts as a looking glass, providing a rare glimpse at the beauty and simplicity of mathematics. In 1955 one such opportunity presented itself, in the form of Hillel Furstenberg's paper, "On the infinitude of primes". In a brilliant proof by contradiction, Furstenberg used topological language with the basic properties of arithmetic sequences and open and closed sets to show the existence of infinitely many prime numbers. The essence of the proof was highlighted by Idris D. Mercer in 2009 when he provided a variant of Furstenberg's proof without the use of topology.

Furstenberg begins his proof by constructing a topology using the set of all arithmetic sequences (from $-\infty$ to $+\infty$) as a basis. It is easily verified that this construction satisfies all necessary conditions for it to be considered a topology. If a set in this topology is open, then it is either empty or a union of arithmetic sequences. Similarly, every open set must also be closed since its compliment is a union of arithmetic sequences and therefore open. As the finite union of closed sets is closed, it follows that a finite union of arithmetic sequences is closed. Consider the set $A = \cup A_p$ where A_p is the set of all integer multiples of p where p runs over all prime numbers. The only integers not in A are -1 and 1. Since the set $\{-1, 1\}$ is clearly not open, it's compliment cannot be a finite union of closed sets. This implies that there are infinitely many primes.

Centennial Problem 1955. *Proposed by James M. Andrews, University of Memphis.*

Consider the topology generated by the base $B_{a,b} := \{an + b : n \in \mathbf{Z}\}$ where $a \in \mathbb{N}$ and b is a non-negative integer. Is the topology Hausdorff? Is it regular? Is it normal? (If you haven't taken a topology class, here is a quick explanation of the terminology. A space $\mathbb{X}$ is a *Hausdorff space* if and only if whenever $x, y \in \mathbb{X}$ are distinct points, there are disjoint open sets $U, V \subset \mathbb{X}$ where $x \in U$ and $y \in V$. A space $\mathbb{X}$ is a *regular space* if and only if whenever $A \subset \mathbb{X}$ is closed and $x \in X \backslash A$, there are disjoint open sets $U, V \subset \mathbb{X}$ where $x \in U$ and $A \subset V$. A space $\mathbb{X}$ is *normal* if and only if whenever $A, B \subset \mathbb{X}$ are disjoint closed sets, there are disjoint open sets $U, V \subset \mathbb{X}$ where $A \subset U$ and $B \subset V$.

REFERENCES

[1] H. Furstenberg, "On the infinitude of primes", American Mathematical Monthly **62** (1955), no. 5, 353.
http://thales.doa.fmph.uniba.sk/sleziak/vyuka/2010/semtc/clanky/FurstenbergPrimes.pdf

[2] I. D. Mercer, "On Furstenberg's Proof of the Infinitude of Primes", American Mathematical Monthly **116**, 355-356. Available online at http://www.idmercer.com/primes.pdf.

1959

100th Anniversary of Riemann's Paper on the Zeta Function

Though Riemann only wrote one paper on the function which now bears his name [4], he made it count and packed in enough mathematics to keep researchers busy for a hundred years and more. Initially defined for the real part of s exceeding 1 by

$$\zeta(s) \ = \ \sum_{n=1}^{\infty} \frac{1}{n^s},$$

by the Fundamental Theorem of Arithmetic (every positive integer $n \geq 2$ can be written uniquely as a product of prime powers, up to reordering the factors) and the Geometric Series Formula, in this region it also equals $\prod_{p \text{ prime}}(1 - 1/p^s)^{-1}$; this is called the Euler product of the zeta function. The function can be meromorphically continued to the entire complex plane, with the only pole a simple one at $s = 1$.

The reason this function occupies such a central role in number theory is that it relates the prime numbers to the integers, and the hope is that we can pass from knowledge of the integers to knowledge of the primes. We discussed one of these connections in the 1948 entry, where we translated results on the sums of $1/n$ and $1/n^2$ to proofs of the infinitude of primes.

Riemann pursued another connection in his paper. He showed how knowledge of the zeros of $\zeta(s)$ yields information on the number of primes. His idea was to use complex analysis. There one of the fundamental quantities to study is the logarithmic derivative of a function f ($\frac{d}{ds} \log f = f'(s)/f(s)$), as integrals of $f'(s)/f(s)$ are related to sums of the residues of the function at its zeros and poles. While the logarithmic derivative of the sum formulation of $\zeta(s)$ is not pleasant, the situation is very different for the Euler product. The logarithm of the product becomes a sum of logarithms, and the resulting integrals are tractable. The problem now reduces to a contour integral of the logarithmic derivative, which yields sums over the zeros and poles; thus it is important to know the locations of these. Up to lower order terms, when we integrate $\zeta'(s)/\zeta(s)$ times x^s/s along the line $\mathrm{Re}(s) = 2$, we find

$$\sum_{p\le x} \log p \;=\; x - \sum_{\substack{\rho=\sigma+i\gamma\\ \zeta(\rho)=0}} \frac{x^\rho}{\rho}.$$

Some care is needed in writing down the sum so that it converges (this is typically done by taking the zeros in complex conjugate pairs). When we write the zeros of $\zeta(s)$ we always take σ and γ to be real numbers.

As remarked above the continuation of $\zeta(s)$ has only one pole, which is at $s = 1$ with residue 1; this is responsible for the $x = x^1/1$ term. The remaining terms come from the zeros of $\zeta(s)$. One can show that these zeros have real part at most 1 without too much trouble; the key step in the proof of the prime number theorem (which says $\sum_{p\le x} 1 \approx x/\log x$, which by partial summation is the same as $\sum_{p<x}\log p \approx x$) is to extend this zero-free region to show that if $\zeta(\sigma + i\gamma) = 0$ then $\sigma < 1$. Riemann conjectured this is the case; in fact, he conjectured more. He wrote in [4]: *and it is very probable that all roots are real. Certainly one would wish for a stricter proof here; I have meanwhile temporarily put aside the search for this after some fleeting futile attempts, as it appears unnecessary for the next objective of my investigation.* This is now known as the Riemann Hypothesis. It makes appearances both on Hilbert's list [6] and the Clay list of millenium problems [1], and is considered by many to be the most important open conjecture in mathematics.

It is interesting to note that the complex analytic proofs of the Prime Number Theorem start by considering the weighted sum $\sum_{p<x}\log p$, and then pass from this to the unweighted sum $\sum_{p<x} 1$ by partial summation (see the 1948 entry for remarks on proofs that avoid complex analysis). The reason it is easier to understand the weighted sums is that when we take the logarithmic derivative of the Euler product we get $-\sum_p \frac{d}{ds}\log(1-p^{-s})$, and the derivative gives us $\log p\cdot p^{-s}/(1-p^{-s})$; we then expand the denominator by using the geometric series formula, and find the main term is $\log p/p^s$. See [5] for a more fleshed out sketch of this proof, or [2, 3] for full details.

Centennial Problem 1959. *Proposed by Steven J. Miller, Williams College.*

The purpose of this problem is to prove that $\zeta(s) \neq 0$ whenever $s = \sigma + i\gamma$ with $\sigma, \gamma \in \mathbb{R}$ and $\sigma \ge 1$. First use the series expansion to prove $\zeta(\sigma) \neq 0$ if $\sigma > 0$, and then use the product expansion to show that $\zeta(\sigma + it) \neq 0$ if $\sigma > 0$.

We are left with the case of $\sigma = 1$. This was originally independently proved by Hadamard and de la Vallée-Poussin in 1896; fill in the details of Mertens' elegant proof from a few years later by proving the following statements (where we write $\Re z$ to denote the real part of a complex number z):

1. $3 + 4\cos\theta + \cos 2\theta \ge 0$ (*Hint:* Consider $(\cos\theta + 1)^2$).
2. For $s = \sigma + it$, $\log\zeta(s) = \sum_p \sum_{k=1}^\infty \frac{p^{-k\sigma}}{k} e^{-itk\log p}$.
3. $\Re\log\zeta(s) = \sum_p \sum_{k=1}^\infty \frac{p^{-k\sigma}}{k}\cos\left(t\log p^k\right)$.
4. $3\log\zeta(\sigma) + 4\Re\log\zeta(\sigma+it) + \Re\log\zeta(\sigma+2it) \ge 0$.
5. $\zeta(\sigma)^3|\zeta(\sigma+it)^4\zeta(\sigma+2it)| \ge 1$.
6. If $\zeta(1 + it) = 0$, then as σ decreases to 1 from above, $|\zeta(\sigma + it)| < A(\sigma - 1)$ for some A.
7. As $\zeta(\sigma) \sim (\sigma-1)^{-1}$ ($\zeta(s)$ has a simple pole of residue 1 at $s = 1$) and $\zeta(\sigma + 2it)$ is bounded as $\sigma \to 1$ (the only pole of $\zeta(s)$ is at $s = 1$), the above

implies that if $\zeta(1+it)=0$ then as $\sigma \to 1$, $\zeta(\sigma)^3|\zeta(\sigma+it)^4\zeta(\sigma+2it)| \to 0$. As the product must be at least 1, this proves $\zeta(1+it) \neq 0$.

The key to Mertens' proof is the positivity of our trigonometric expression.

While there are partial results towards the Riemann Hypothesis, it is still unproved whether or not there is a $c<1$ such that all zeros of $\zeta(s)$ have real part at most c (the Riemann Hypothesis is equivalent to being able to take $c=1/2$). The best results are zero free regions where how far to the left of the line $\mathrm{Re}(s)=1$ we can go tends to zero rapidly with the height t, giving regions where $\zeta(\sigma+it)\neq 0$ if $\sigma > 1 - A(\log|t|)^{-r_1}(\log\log|t|)^{-r_2})$ for some positive constants A, r_1, r_2.

REFERENCES

[1] Clay Mathematics Institute, "Millenial Problems". http://www.claymath.org/millennium-problems.
[2] H. Davenport, "Multiplicative Number Theory", 2nd edition, revised by H. Montgomery, Graduate Texts in Mathematics, Vol. 74, Springer-Verlag, New York, 1980.
[3] H. Iwaniec and E. Kowalski, "Analytic Number Theory", AMS Colloquium Publications, Vol. 53, AMS, Providence, RI, 2004.
[4] G. F. B. Riemann, "Über die Anzahl der Primzahlen unter einer gegebenen Grösse", Monatsber. Königl. Preuss. Akad. Wiss. Berlin, Nov. 1859, 671–680. See H. M. Edwards, "Riemann's Zeta Function", Academic Press, New York, 1974 for an English translation, or go to http://www.claymath.org/millennium/Riemann_Hypothesis/1859_manuscript/EZeta.pdf.
[5] S. J. Miller and R. Takloo-Bighash, "An Invitation to Modern Number Theory", Princeton University Press, Princeton, NJ, 2006.
[6] Wikipedia, "Hilbert's problems", http://en.wikipedia.org/wiki/Hilbert's_problems.

1963

Continuum Hypothesis

One of the first mathematical concepts learned is that of comparison. This collection has one element, this has two and so on. While there is no dearth of picture books on the market with different sets for different numbers (for example, one dinosaur, two planes, three cars [1]), not surprisingly these books only show small finite sets. What about infinite sets? Can we look at two different sets and determine if one is 'larger' than the other?

One of the most important results in set theory is that there are different orders of infinity. The smallest is the size of the integers $\mathbb{Z}$, which also equals the size of the rationals $\mathbb{Q}$ or algebraic numbers $\mathcal{A}$; we denote this size by $\aleph_0$. There is no largest infinity, as given any set S its power set (which is the set of all subsets of S) has strictly larger cardinality. We often write 2^S for the power set of S. The real number line, often called the continuum, can be viewed as the power set of the integers. To see this, we may regard a subset of the positive integers as an infinite string of 0s and 1s, giving the binary expansion of a number in $[0,1]$. For more on infinities, especially the difference in the cardinalities of the integers and the reals, see the problems from 1918 and 1935.

If X is a set we write $|X|$ for its cardinality; if X is finite then its cardinality is simply the number of elements. The notation for the power set of S is highly suggestive. For finite sets S, we have $|2^S| = 2^{|S|}$. Let's investigate what happens for different S. There is only one set with no elements, the empty set $\emptyset$. It has cardinality zero, and its power set is simply $\{\emptyset\}$ (the set containing the empty set), which has cardinality 1. All sets of just one element are equivalent to $\{\emptyset\}$; the powerset here has two elements: the empty set and the set containing the empty set. Thus $2^{\{\emptyset\}} = \{\emptyset, \{\emptyset\}\}$. Arguing along these lines we see that if S has n elements (for n a non-negative integer) then 2^S has $2^{|S|} = 2^n$ elements, which justifies our notation.

Interestingly, for finite sets the only time $|2^S| = |S| + 1$ is when $|S|$ is the empty set or the set containing the empty set. In all other cases there is a set T such that $|S| < |T| < |2^S|$. It is natural to ask if this is true for infinite sets. In particular, denoting the size of the integers by $\aleph_0$, how should we denote the size of the continuum $\mathbb{R}$ (i.e., the size of the real numbers)? Generalizing our notation from the power set of finite sets, it seems reasonable to write $|\mathbb{R}| = 2^{\aleph_0} = 2^{|\mathbb{Z}|}$; however, this is just notation. What is the difference in size between $2^{|\mathbb{Z}|}$ and $|\mathbb{Z}|$? Using Cantor's diagonalization argument (see the entry from 1918) we have $|2^{\mathbb{Z}}| > |\mathbb{Z}|$; how much can they differ by? In particular, for finite sets we saw that if S has at least two elements then we can always find a set of size strictly between $|S|$ and $2^{|S|}$; does this hold for infinite sets as well? Is $2^{\aleph_0}$ the next smallest infinity and thus merits being called $\aleph_1$, or is there an infinity between the two? Could there be infinitely many infinities between?

Cantor's Continuum Hypothesis states that there is no set of cardinality strictly between that of the integers and that of the reals. Hilbert made the resolution of this conjecture the first problem in his celebrated list [10] (see the entry for 1935 for more on Hilbert's problems). Interestingly, this question is independent of the standard axioms of set theory. What this means is that if the standard axioms are consistent, then so too is the model obtained by adding the Continuum Hypothesis, or by adding its negation. Typically one uses either ZFC, the Zermelo-Fraenkel axioms with the additional assumption of the Axiom of Choice, or ZF without Choice. Kurt Gödel [6] proved in 1940 that the Continuum Hypothesis cannot be disproved if one assumes ZF (or ZFC) is consistent. Almost 20 years later, in 1963 Paul Cohen [2, 3] introduced the concept of forcing (see [4]) and proved that one cannot prove the Continuum Hypothesis under the same assumptions. Thus the Continuum Hypothesis is independent of the standard axioms of set theory, and one can construct models where it holds or fails: take your pick! See [**?**, 8] for remembrances of Paul Cohen (on a personal note, the editor is proud to be one of his mathematical grandsons).

Centennial Problem 1963. *Proposed by Steven J. Miller, Williams College.*

Cardinality is merely one way of measuring the complexity of a set. Another option is to use the notion of dimension. Given a set S and a positive real number r, let rS denote the dilation of S by r; for example, if S is a unit square then rS is an $r \times r$ square. A good working definition is to say a set S is of dimension d if the 'measure' (we are being deliberately vague as to what this means!) of rS is r^d times the measure of S. Note that for a square we have $2S$ has measure 4 times that of S, and thus its dimension would be 2. Similarly a cube would have dimension 3.

Interestingly, not all sets have integral dimension! A terrific example is the Cantor set. It is defined as follows. Let C_0 equal the unit interval, and let C_1 equal the unit interval minus the middle third, so $C_1 = [0, 1/3] \cup [2/3, 1]$. We continue, and define C_{n+1} to be C_n minus the middle third of each of the 2^n subintervals of length $1/3^n$ making up C_n. The Cantor set C is $\cap_n C_n$. Clearly the length of C is zero, as the length is less than that of C_n for each n, and the length of C_n is $(2/3)^n$. *Prove that the dimension of* C *is* $\log_3 2$*, which is approximately .63. This proves that there exists a set whose dimension is strictly between sets of dimension 0 and 1.*

The Cantor set is one of the first encountered fractals; for more on fractals see the entry for 1978. For another interesting fractal, look at an equilateral triangle and subdivide into four equilateral triangles. If you continue this process you get the Sierpinski triangle, which you should prove has dimension of $\log_2 3 \approx 1.585$, which is strictly between 1 and 2. Interestingly, one can also obtain the Sierpinski triangle by looking at Pascal's triangle modulo 2; we sketch the idea and leave it to you to

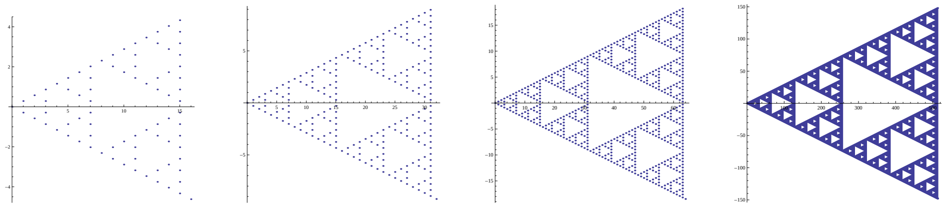


FIG. 2. *Pascal's triangle modulo 2. For ease of plotting in Mathematica, instead of filling in the cell we put a dot only if we should take the cell. The three triangles correspond to taking 31, 63 and 511 rows (for displaying purposes, it is easier for the program to include the first entry of the next row).*

make it rigorous. For each non-negative integer k draw the first $n = 2^k - 1$ rows of Pascal's triangle, where row m is the row that corresponds to the coefficients arising from expanding $(x+y)^m$. Thus the entries of this row are $\binom{m}{0} = 1$, $\binom{m}{1} = m$, $\dots$, $\binom{m}{\ell} = \frac{m!}{\ell!(m-\ell)!}$, $\dots$, $\binom{m}{m} = 1$. Adjust the scale so that the three vertices of this finite version of Pascal's triangle give an equilateral triangle of height 1. We divide the corresponding triangle into small equilateral triangular cells centered about the different lattice points corresponding to the locations of the elements of Pascal's triangle. Keep all the points in a cell if the entry in Pascal's triangle is odd, and delete the cell if the entry is even. Taking the limit as $k \to \infty$ yields the Sierpinski triangle; see Figure 2 for some finite approximations to this limit.

Thus the notion of dimension is more involved than you might expect. We proved that there is a set of dimension strictly between that of 0 and 1, and one of dimension between 1 and 2. More generally, if you have two sets of positive dimension d_1 and d_2, can you always construct a set whose dimension is strictly between these two?

REFERENCES

[1] SCHOLASTIC EDITORIAL, "Learn With Lego: Numbers: Counting", Cartwheel Books 2007.

[2] P. J. COHEN, "The Independence of the Continuum Hypothesis", Proceedings of the National Academy of Sciences of the United States of America **50** (1963), no. 6, 1143–1148. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC221287/.

[3] P. J. COHEN, "The Independence of the Continuum Hypothesis, II", Proceedings of the National Academy of Sciences of the United States of America **51** (1964), no. 1, 105–110. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC300611/.

[4] T. Y. CHOW, "A beginner's guide to forcing", Communicating Mathematics: A Conference in Honor of Joseph A. Gallian's 65th Birthday, Contemporary Mathematics **479**, 25–40. http://arxiv.org/abs/0712.1320.

[5] L. GILLMAN, "Two Classical Surprises Concerning the Axiom of Choice and the Continuum Hypothesis", American Mathematical Monthly **109** (2002), 544–553. http://www.maa.org/sites/default/files/pdf/upload_library/22/Ford/Gillman544-553.pdf.

[6] K. GÖDEL, "The Consistency of the Continuum-Hypothesis", Princeton University Press 1940.

[7] P. SARNAK, "Remembering Paul Cohen", MAA Focus **27** (2007), no. 9, 21–22. http://web.math.princeton.edu/sarnak/RememberingPaulCohen.pdf.

[8] P. SARNAK (coordinating editor), "Remembering Paul Cohen", Notices of the AMS **57** (2010), no. 7, 824–838.

[9] WIKIPEDIA, "Continuum Hypothesis". http://en.wikipedia.org/wiki/Continuum_hypothesis.

[10] WIKIPEDIA, "Hilbert's problems", http://en.wikipedia.org/wiki/Hilbert's_problems.

1967

The Langlands Program

Robert Langlands's 1967 handwritten letter to the master number theorist André

Weil begins modestly: "Your opinion of these questions would be appreciated.... I hope you will treat them with the tolerance they require at this stage." But Langlands's letter was in fact a tour de force, a manifesto that would shape the next half-century (and more) of number theory.

The main characters in Langlands's drama are automorphic forms: functions on a topological space which are invariant under a discrete group of symmetries (the actual definition is much longer and much more particular). There are two crucial supporting characters: first, Galois representations, or homomorphisms $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(\mathbb{C})$, where $\bar{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$. This Galois group is one of the richest objects in algebraic number theory, and describing its representations is a complicated problem. Second, L-functions, of which the simplest example is the Riemann zeta function:

$$\zeta(s) \;=\; \sum_{n=1}^{\infty} \frac{1}{n^s} \;=\; \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{0.1}$$

Every L-function can be written as a product over primes in this way, and extended to all $s \in \mathbb{C}$, with a symmetry in $s \mapsto 1-s$. L-functions encode all kinds of data, from the distribution of prime numbers to point counts on algebraic varieties.

The Langlands Program, very roughly, is to show that wherever we see a Galois representation or an L-function in number theory, there is an automorphic form lurking behind it. Here is one example: suppose that we have an L-function:

$$L(s) \;=\; \prod_{p \text{ prime}} \left(1 - \frac{\alpha_p}{p^s}\right)^{-1} \left(1 - \frac{\beta_p}{p^s}\right)^{-1} \tag{0.2}$$

where $\alpha_p, \beta_p \in \mathbb{C}$ with $\alpha_p\beta_p = 1$. This kind of L-function might come from an elliptic curve $y^2 = x^3 + ax + b$, a representation $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{C})$, a modular form, or a more mysterious Maass form. (In the case of an elliptic curve, α_p and β_p are functions of the number of points on the curve mod p.) Langlands conjectures that not only do these L-functions have automorphic forms behind them, but so too do the "symmetric power L-functions"

$$L_r(s) \;=\; \prod_{p \text{ prime}} \prod_{i=0}^{r} \left(1 - \frac{\alpha_p^i \beta_p^{r-i}}{p^s}\right)^{-1}. \tag{0.3}$$

Just the convergence of the symmetric powers implies two famous conjectures: the Ramanujan conjecture, that all α_p, β_p are on the unit circle, and the Sato-Tate conjecture, that they are equidistributed on the circle.

The Langlands Program encompasses a vast range of conjectures and theorems, more than one person could ever prove. For example, class field theory is just the simplest case of the Langlands Program. Andrew Wiles's proof of Fermat's Last Theorem? Just part of the next simplest case. There have been huge breakthroughs on the Langlands Program since 1967, but we'll be working on it for a long time to come.

Centennial Problem 1967. *Proposed by Ian Whitehead, University of Minnesota.*

In this problem we'll show that Langlands' conjecture for symmetric power L-functions implies the Ramanujan conjecture. Consider one factor of the product of symmetric power L-functions $L_0(s)L_2(s)L_4(s)\cdots L_{2m}(s)$:

$$\prod_{r=0}^{m} \prod_{i=0}^{2r} (1 - \alpha^i \beta^{2r-i} x)^{-1}. \tag{0.4}$$

Here we've substituted α, β for α_p, β_p, and x for p^{-s}. Assume that $\alpha\beta = 1$ and $\alpha+\beta \in \mathbb{R}$. Prove that this expands as a power series in x with positive real coefficients (hint: you could take a logarithm first). This fact, together with Langlands's conjecture, implies that the series converges for $x < 1/p$, regardless of m. Conclude that $|\alpha| = |\beta| = 1$.

REFERENCES

[1] D. Bump, "Automorphic Forms and Representations", Cambridge Studies in Advanced Mathematics, 1998.

[2] R. Langlands, "Letter to André Weil", Institute for Advanced Study. http://publications.ias.edu/rpl/paper/43.

[3] R. Langlands, "Problems in the Theory of Automorphic Forms", Lecture Notes in Mathematics 170, 1970.

1971

Society for American Baseball Research

Founded in Cooperstown, New York by Bob Davids in 1971, the Society for American Baseball Research (http://www.sabr.org/) has many objectives. One of them is to encourage and aid the application of mathematics and statistics to the analysis of baseball. One consequence of their efforts is the alphabet soup of acronyms for new metrics to measure player performance (VORP, WAR, OPS, ...). This is an extremely important task, not just for baseball but for other fields as well. It is important to know *what* to measure. For example, originally walks were viewed as errors by the pitcher and not a positive event by the batter. This led to an enormous undervaluation of walks, which is remedied now with on-base percentage. As the annual revenues in Major League Baseball (MLB) are measured in the *billions*, there is a lot at stake and a team that has a better understanding of which statistics truly matter can assemble a better team for less, which can translate to World Series rings and greater revenue. Most teams now have sabermetricians helping with the analysis, player selection and strategy. *Moneyball*, by Michael Lewis, is an excellent popular account of how the Oakland A's applied these principles and, with a budget a third to a fourth of other teams, fielded competitive teams reaching the playoffs.

Centennial Problem 1971. *Proposed by Steven J. Miller, Williams College.*

Only seven times in MLB history has a team had four consecutive batters hit home runs: the Milwaukee Braves in 1961, the Cleveland Indians in 1963, the Minnesota Twins in 1964, the Los Angeles Dodgers in 2006, the Boston Red Sox in 2007, the Chicago White Sox in 2008, and the Arizona Diamondbacks in 2010. Estimate the probability that some team during the season performs this feat. *Hint:* a simple model is to calculate the average home run frequency of players, and essentially raise that to the fourth power. What is wrong with this argument? There are *many* articles with arguments and calculations of this problem; what makes this problem interesting is trying to account for all the different factors.

1975

Szemerédi's Theorem

Let us call a set $A \subset \mathbb{N}$ AP-rich (correspondingly, GP-rich) if it contains arbitrarily long arithmetic progressions (correspondingly arbitrarily long geometric progressions). In his 1975 paper Szemerédi [1] has shown that any set $A \subset \mathbb{N}$ having positive upper density $\overline{d}(A) := \limsup_{N\to\infty} \frac{1}{N}|A \cap \{1, \dots, N\}| > 0$ is AP-rich, thereby confirming the Erdős-Turán conjecture formulated in [5]. It is not too hard to show that Szemerédi's theorem implies an ostensibly stronger theorem which states that if for some sequence of intervals $I_n = \{a_n, \dots, b_n\} \subset \mathbb{N}$ with $b_n - a_n \to \infty$, one has $\overline{d}_{(I_n)}(A) := \limsup_{N\to\infty} |A \cap I_n|/|I_n| > 0$, then A is AP-rich. It is customary to view Szemerédi theorem as a density version of van der Waerden's result [6], which states that for any finite coloring $\mathbb{N} = \cup_{i=1}^{r} C_i$, one of the C_i is AP-rich. Now, it is also true that one of the C_i is GP-rich (consider the restriction of our r-coloring to the set $\{2^n : n \in \mathbb{N}\}$ and apply van der Waerden's theorem).

This naturally leads to the question whether sets of positive upper density are GP-rich. The answer turns out to be **NO!**. Consider for example the set R of square-free numbers. IT clearly cannot contain length 3 geometric progressions $\{x, xq, xq^2\}$ with $q > 1$. At the same time, it is known that $\lim_{N\to\infty} \frac{1}{N}|R \cap \{1, \dots, N\}| = 6/\pi^2$. So, if one believes in the idea that any partition result of Ramsey theory has a density version (see [3] for a discussion of some of the principles of Ramsey theory), a new notion of largeness, geared towards the multiplicative structure of $\mathbb{N}$, should be looked for.

Let $(a_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in $\mathbb{N}$, let $(p_n)_{n\in\mathbb{N}}$ be any ordering of the set of primes $P = \{2, 3, 5, 7, \dots\}$ and let, for any $j \in \mathbb{N}$, $(N_n^{(j)})_{n\in\mathbb{N}}$ be an increasing sequence of natural numbers. Let $F_n = \{a_n p_1^{i_1} \cdots p_n^{i_n}, 0 \le i_\ell \le N_n^{(j)}\}$. Finally, for a set $A \subset \mathbb{N}$ let us define the upper multiplicative density with respect to the family $(F_n)_{n\in\mathbb{N}}$ as $\overline{d}_\times(A; (F_n)) := \limsup_{N\to\infty} |A \cap F_n|/|F_n|$. The main reason we have chosen the sequence $(F_n)_{n\in\mathbb{N}}$ to appear in this definition is that the quantity $\overline{d}_\times(A; (F_n))$ is invariant with respect to multiplication and division: for any $k \in \mathbb{N}$ one has

$$\overline{d}_\times(A; (F_n)) \;=\; \overline{d}_\times(kA; (F_n)) \;=\; \overline{d}_\times(A/k; (F_n))$$

(where $kA = \{ka : a \in A\}$ and $A/k = \{b : kb \in A\}$). The best way of thinking of the sets F_n, $n \in \mathbb{N}$, is to view them as multiplicative counterparts of the family of intervals I_n, $n \in \mathbb{N}$, which appear above in the definition of the "generalized" upper density $\overline{d}_{(I_n)}$.

While the additive semigroup of natural numbers $(\mathbb{N}, +)$ have a single generator (and $k \in \mathbb{N}$ is a multiple of 1), the multiplicative semigroups $(\mathbb{N}, \times)$ has infinitely many generators (the primes). This accounts for the somewhat cumbersome definition of $\overline{d}_\times(\cdot; (F_n))$. Let us call a set $A \subset \mathbb{N}$ multiplicatively large if for some sequence $(F_n)_{n\in\mathbb{N}}$ as defined above, $\overline{d}_\times(A; (F_n)) > 0$, and let us call A additively large if for some sequence of intervals I_n with $|I_n| \to \infty$, $\overline{d}_{(I_n)}(A) > 0$. It is not hard to see that these two notions of largeness do not overlap (see the discussion in [3]). The following result, obtained in [3], may be viewed as a multiplicative analogue of Szemerédi's theorem.

Theorem: Any multiplicatively large set is GP-rich.

Returning for a moment to van der Waerden's theorem, let us observe that one can actually derive from it the fact that for any finite partition $\mathbb{N} = \sum_{i=1}^{r} C_i$, one of C_i is simultaneously AP-rich and GP-rich. It turns out, surprisingly, that the notion of multiplicative largeness provides for a density version of this result as well.

Theorem [3]: Any multiplicatively large set is AP-rich.

Note that the set R of square-free number, while being free of geometric progressions, has quite a bit of multiplicative structure: for any distinct primes $p_1, \dots, p_k$, the product $p_1 \cdots p_k$ belongs to R. It turns out that any set of positive upper logarithmic density (a notion slightly stronger than that of upper density) contains an infinite divisibility chain, i.e., a sequence $(x_n)_{n\in\mathbb{N}}$ such that, for any n, x_n divides x_{n+1}. See [4] for the details.

Assume now that S is a syndetic set in $(\mathbb{N}, +)$, that is a set with the property that finitely many of its shifts cover $\mathbb{N}$. Equivalently, S is syndetic if it has bounded gaps. This property is quite a bit stronger than that of positive upper logarithmic density, and one may expect that syndetic sets, in addition to having a divisibility chain, have some additional multiplicative structure.

Centennial Problem 1975. *Proposed by Vitaly Bergelson, The Ohio State University.*

Assume that S is a syndetic set in $\mathbb{N}$. Is S GP-rich?

See [1, 2] for discussion and some equivalent forms of this conjecture. As a matter of fact, it is not even known whether any syndetic set contains a pair of the form $\{x, xq^2\}$ with $q > 1$.

REFERENCES

[1] M. Beiglböck, V. Bergelson, N. Hindman and D. Strauss, "Multiplicative structures in additively large sets," J. Combin. Theory Ser. A **113** (2006), 1219–1242.

[2] M. Beiglböck, V. Bergelson, N. Hindman and D. Strauss, "Some new results in multiplicative and additive Ramsey theory," Trans. Amer. Math. Soc. **360** (2008), 819–847.

[3] V. Bergelson, "Multiplicatively large sets and ergodic Ramsey theory," Israel J. Math. **148** (2005), 23–40.

[4] H. Davenport and P. Erdős, "On sequences of positive integers", Acta Arith. **2** (1936), 147–151.

[5] P. Erdős and P. Turán, "On Some Sequences of Integers," J. London Math. Soc. **11** (1936), 261–264.

[6] B. L. van der Waerden, "Beweis einer Baudetschen Vermutung," Nieuw. Arch. Wisk. **15** (1927), 212–216.

[7] E. Szemerédi, "On sets of integers containing no k elements in arithmetic progression", Acta Arith. **27** (1975), 299–345.
http://www.scholarpedia.org/article/Szemer%C3%A9di's_Theorem.

1979

TeX

This entry honors two fundamental contributions of computer science to the mathematical and scientific communities: first, Donald Knuth's creation of the TeX typesetting system, released in 1978; and second, off-by-one errors, which is why this entry is listed under 1979.[3]

Donald Knuth is perhaps best known for his monumental, encyclopedic, and stunningly readable series *The Art of Computer Programming*. Begun in 1962 while he was a graduate student at Caltech, the project continues to this day, with volume 4A published in 2011 and several remaining volumes in preparation. While preparing a second edition of Volume 2, Knuth was dismayed with the quality of the typesetting done by the publisher. Realizing that digital typesetting boiled down to 0s and 1s—is this pixel black or not?—Knuth saw this as a problem amenable to computer science and set out to design his own system.

[3]For the purists, though, Knuth was honored with the National Medal of Science in 1979.

Then came another classic computer science insight: Knuth estimated he could have the system ready in six months. Instead, it was almost ten years before the “first” TEX was released. This was called version 3. (You may have heard that computer scientists start counting at 0, but this is apparently incorrect: they start at 3.) The next version was 3.1, which was followed by version 3.14. The current version is 3.14159265. You get the idea.

TEX is now used extensively. Essentially every contemporary paper in mathematics and computer science is typeset using some system based on TEX, including this very document! TEX has several features that distinguish it from other digital typesetting and publishing systems, but we only discuss two points of interest here.

First, TEX is designed as a programming language. The user writes a program that describes both the content and layout of the document. The program is then interpreted by TEX to produce the desired document. This design choice means that TEX is extraordinarily flexible and customizable. The price is that TEX and related systems can be hard for beginners to pick up from scratch. Fortunately there are many templates available online, and by looking at the code and compiled documents you can learn over 90% of what you need fairly quickly, and then search the web or ask experts for the rest. For example, the editor of these problems maintains TEX templates (for papers and presentations) online:

http://web.williams.edu/Mathematics/sjmiller/public_html/math/handouts/latex.htm

(the website also has a link to a YouTube video going through how to write simple articles using the above TEX template).

Second, TEX uses sophisticated algorithms to layout text on the page. Consider the problem of breaking a paragraph of text into justified lines. Each line must begin at the left margin and end at the right margin, and there cannot be too much nor to little space between words. Line breaks are allowed only between words and, if necessary, inside a word at a known hyphenation point.

How would you solve this problem? The solution used in most digital typesetting systems and word processors is a greedy strategy. We consider the words of the paragraph one at a time, adding them to the current line. When the current line is full, it is added to the page, and we begin adding words to the next line. This approach is fast—it considers each word only once—but it can lead to unappealing results because it never changes its mind about lines that have already been added to the page. For example, the greedy algorithm may put vastly different amounts of space between words on different lines, which looks strikingly bad.

Centennial Problem 1979. *Proposed by James Wilcox, University of Washington.*

Instead, TEX tries harder to optimize for a good-looking paragraph. To do so, it uses a notion called *badness*, which is computed using rather complex rules that are designed to penalize ugly layouts. For example, we wish to penalize paragraphs that contain lines with too much or too little space between words. Given a definition of badness, the problem is now to minimize badness over all possible sets of line breaks. A naïve implementation of this approach would consider all exponentially many choices, but it is possible to do better. Give a quadratic-time algorithm for finding the optimal set of line breaks. For a detailed discussion of TEX’s line breaking algorithm, see [1].

REFERENCES

[1] D. E. Knuth and M. F. Plass “Breaking paragraphs into lines”, Software: Experience and Practice, Vol. 11, Iss. 11, November 1981.

[2] D. E. Knuth, "The Art of Computer Programming, Volume 1: Fundamental Algorithms", 3rd ed., Addison-Wesley, 1997.
[3] D. E. Knuth, "The Art of Computer Programming", http://www-cs-faculty.stanford.edu/~uno/taocp.html.
[4] The TeX User Group, "History of TeX", http://www.tug.org/whatis.html.

1983

Julia Robinson

The year 1983 marks the service of Julia Robinson as the first woman to hold office as president of the American Mathematical Society. Born on December 8, 1919, Robinson shared a passion for mathematics at an early age with her sister and notable biographer, Constance Reid. Reid wrote extensively about Robinson in several publications, including her award winning Mathematical Association of America article, "The Autobiography of Julia Robinson". In the article's introduction, Reid expresses some personal remarks about her sister. "She herself, in the normal course of events, would never have considered recounting the story of her own life. As far as she was concerned, what she had done mathematically was all that was significant." Indeed, significant is a fitting word when speaking of the magnitude of Robinson's mathematical ability, especially with her collaborative involvement that ultimately led to the resolution of Hilbert's Tenth Problem.

In the early years of the 20th century, the German mathematician David Hilbert published a set of twenty-three open problems that challenged the foundation of mathematical formalism. One of the underlying themes of the list was the question of decidability. Decision problems that originate from these questions are summarized in the following way: Given a mathematical problem that falls into a certain class of problems, is there a general algorithm that can solve every problem in that class? In the case of Hilbert's Tenth Problem, Hilbert proposed the question for the general solvability of Diophantine equations (a Diophantine equation is an equation of the form $p(x_0, x_1, \ldots, x_n) = 0$ where p is a polynomial with integer coefficients and only integer solutions are considered). A more formal statement of Hilbert's Tenth Problem is the following: Is there a general algorithm that can determine the solvability of an arbitrary Diophantine equation?

Over the span of several decades, Robinson and several collaborators, including Martin Davis and Hillary Putnam, formulated a necessary condition to answer Hilbert's question in the negative. In particular, they proved that if there is at least one Diophantine relation of exponential growth, then no such solvability algorithm exists. A general family of Diophantine equations is given by $p(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0$, where $a_1, \ldots, a_n$ are parameter relations which hold if and only if the equation has a solution in the remaining variables $x_1, \ldots, x_m$. Furthermore, the form of p can be addition, multiplication, and exponentiation constructed from the variables $x_1, \ldots, x_m$. Robinson, Putnam, and Davis established that if a Diophantine equation $Q(a, b, c, x_1, \ldots, x_m) = 0$ has a solution in integers $x_1, \ldots, x_m$ if and only if $a = b^c$, then the necessary condition of finding a Diophantine relation of exponential growth is met. Furthermore, Robinson showed that it is equally sufficient to have an equation $G(a, b, x_1, \ldots, x_m)$ which defines the relation $J(a, b)$ such that for any a and b, $J(a, b)$ implies that $a < b^b$ and for any k, there exist a and b such that $J(a, b)$ and $a > b^k$. In 1970, a young Russian mathematician named Yuri Matiyasevich discovered such an equation that ultimately ended the search and answered Hilbert's Tenth Problem in the negative.

With her notable achievements aside, Julia Robinson was a remarkable mathematician and pioneer for women in modern mathematics. Though she passed away in 1984, her work still forms a foundation for further questions to be explored. It is such questions that find their way into the imaginations of the mathematically curious and stretch the boundary of what is decidable.

Centennial Problem 1983. *Proposed by Avery T. Carr, Emporia State University.*

Decision problems are found in many areas of mathematics, including Combinatorics and Graph Theory. A discrete graph G is a set of n vertices (much like points in a projective plane) and m edges (similar to straight lines in a plane) such that there is an irreflexive relation between the vertices and edges. The order of G is the number of vertices of G. For a simple discrete graph, every edge has exactly two vertices, one at each end, called end vertices. A cycle in G is a path traversing a subset of vertices, each only once, and closing with the starting vertex of the path. Hamiltonian cycles of G are cycles that span every vertex of G. Weighted edges are edges with numerical values assigned to them. Summing the weighted edges of an arbitrary cycle gives the total weight of the cycle. An open unresolved question is: Given an arbitrary edge-weighted discrete graph G with sufficiently large order, is there a general algorithm to find the Hamiltonian cycle (provided at least one exists) of G with the minimal weight? This is an equivalent abstract version of the famed Traveling Salesman Problem, and a positive solution would have notable implications in several industries, including logistics and computer engineering.

REFERENCES

[1] Y. Matiyasevich, "My Collaboration with JULIA ROBINSON", http://logic.pdmi.ras.ru/~yumat/personaljournal/collaborationjulia/index.html.

[2] C. Reid, "The Autobiography of Julia Robinson", The College Mathematics Journal **17** (1986), no. 1, 3–21. http://www.ams.org/notices/199612/reid.pdf.

[3] C. Reid, "Being Julia Robinson's Sister", Notices of the American Mathematical Society **43** (1996), no. 12, 1486–1492. http://www.ams.org/notices/199612/reid.pdf.

[4] Wikipedia, "Hilbert's tenth problem", http://en.wikipedia.org/wiki/Hilbert's_tenth_problem.

[5] Wikipedia, "Julia Robinson", http://en.wikipedia.org/wiki/Julia_Robinson.

[6] Wolfram Mathworld, "Hilbert's Problems", http://mathworld.wolfram.com/HilbertsProblems.html.

[7] Wolfram Mathworld, "Diophantine Equation", http://mathworld.wolfram.com/DiophantineEquation.html.

[8] Wolfram Mathworld, "Traveling Salesman Problem", http://mathworld.wolfram.com/TravelingSalesmanProblem.html.

[9] C. Wood, "Julia Robinson and Hilbert's Tenth Problem", Film Review, Notices of the American Mathematical Society (2008), 573–575. http://www.ams.org/notices/200805/tx080500573p.pdf.

1987

Primes, the zeta function, randomness, and physics

The Centennial Problem 1942 featured the Riemann zeta function, $\zeta(s)$, and its intimate connection with primes. For $\mathrm{Re}(s) > 1$, $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, and simultaneously $\zeta(s) = \prod_{p \text{ prime}}(1 - p^{-s})^{-1}$. The equality of the series and the product is a simple consequence of the unique factorization of integers into prime powers. But it does lead to a potentially much more fruitful observation. The zeta function can be continued analytically to the entire complex plane with the exception of a first order pole at $s = 1$, and aside from so-called trivial zeros at negative even integers, has all its other (called non-trivial) zeros inside the critical strip $0 < \mathrm{Re}(s) < 1$. Those non-trivial zeros determine the primes through explicit formulas that date back to Riemann. Thus the very discrete primes, the fundamental building blocks of arithmetic, dance to the tune of zeros of an analytic function! This unusual connection of continuous and discrete has fascinated mathematicians for a long time.

The Riemann Hypothesis (RH), one of the most famous unsolved problems in mathematics, says that all the non-trivial zeros lie on the critical line, $\mathrm{Re}(s) = \frac{1}{2}$. Conrey's 2003 article provides a nice survey of the main approaches that have been taken to it, and there are many other sources of information on the Web. The conventional analytic methods have not produced a proof so far, and there is general opinion among experts in the area that other approaches are needed. The one that has stimulated the largest effort in the last few decades is founded on the Hilbert and Pólya conjecture, which says that the RH is true because there is a positive operator whose eigenvalues correspond in a canonical way to the non-trivial zeros. This conjecture leads to potential connections with extensive work on physics on random matrices. It gained credibility as a result of a theoretical advance by Montgomery in the early 1970s, followed by the production of extensive empirical data by Odlyzko in 1987. While this line of attack has produced many intriguing results, it is not clear whether it will lead to a proof of the RH.

Many approaches to the RH, such as that through random matrix theory, require substantial background. However, it is still conceivable that simpler attacks will succeed. If we don't insist for direct attacks on the RH, the range of interesting questions related to the RH that are amenable to experimentation and simple proofs grows substantially.

Centennial Problem 1987. *Proposed by Andrew Odlyzko, University of Minnesota.*

The sieve of Eratosthenes is the standard method for producing all primes up to a given bound. In order to understand the extent to which conjectures such as the RH or the twin prime conjecture are outcomes of general sieving procedures, Hawkins proposed a probabilistic version of the sieve of Eratosthenes. We start with a list of integers > 1, and perform an infinite series of passes through the list, each pass producing what is called a *Hawkins prime.* On the first pass, we identify 2 as the first element in the list, and we call it the first *Hawkins prime,* $p_1 = 2$. We then go through the remainder of the list, and cross out each element with probability $1/p_1 = 1/2$. On the k-th pass, for $k > 1$, we look for the first element in the list that is $> p_{k-1}$ and has not been crossed out, declare it to be p_k, and cross out every subsequent integer with probability $1/p_k$. Thus it might happen that on the first pass, 3 will be crossed out, but not 4, in which case we will obtain the counter-intuitive result that $p_2 = 4$. (The Hawkins sieve loses many natural properties of the ordinary primes, and for finite sequence of integers $2 < q_3 < q_4 < \cdots < q_k$ there will be a positive probability that $p_j = q_j$ for $2 < j \leq k$.)

The Hawkins sieve is easy to simulate on a computer, and, of more interest, to analyze rigorously. It turns out that many properties of primes, both proven and conjectured, can be shown to hold for Hawkins primes, but only with probability one. Can one come up with other random sieves that can be analyzed rigorously, even though probabilistically, and which produce numbers that are closer to ordinary primes?

REFERENCES

[1] J. B. Conrey, "The Riemann Hypothesis," Notices Amer. Math. Soc. **30** 2003, 341–353.
[2] J. Lorch and G. Ökten, "Primes and probability: The Hawkins random sieve," Mathematics Magazine **80** 2007, 112–119.
[3] A. Odlyzko, "On the distribution of spacings between zeros of the zeta function," Math. Comp. **48** 1987, 273–308.

1991

arXiv

In many lists of the most influential people of all time, Gutenberg frequently not only makes the cut, but is often towards the top. The reason is that the ability to mass produce printed works allowed information to be greatly disseminated, which fosters collaborations and creates a powerful feedback loop. This year's problem honors the founding of the arXiv, http://arxiv.org/. Begun by Paul Ginsparg in August 1991 as a repository for physics papers, there are now well over half a million articles in Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics; each day over a hundred papers are added in math alone!

The arXiv is generously supported by Cornell University, the Simons Foundation (see https://www.simonsfoundation.org/ for their homepage; their magazine https://www.simonsfoundation.org/quanta/ is a wonderful source of interesting articles), and member institutions. It allows researchers and interested parties all over the world immediate access to research and expository papers. As the time from when a paper is submitted to when it appears in press at a journal is usually measured in **years**, the importance of the arXiv becomes apparent, as it allows people to share their ideas in real time. See http://arxiv.org/help/stats for some interesting graphs on the growth of the use of the arXiv over the years (both in posting and downloading articles).

Centennial Problem 1991. *Proposed by Steven J. Miller, Williams College.*

Every day you can check for new posts to the arXiv. Spend ten minutes a day, every day for a month, skimming the titles of papers, the names and affiliations of the authors, and the abstracts. For me, there's exponential (or worse!) decay between these categories and one more, actually clicking and opening the paper! That said, it is excellent advice to keep abreast of what is happening in your field. Get to know whom the players are, and what they study. Get a sense of what topics are popular. Years ago Jeff Lagarias gave me this advice; I've mostly followed it and frequently this is how I learn about developments in my field.

For the especially brave: go to the *general math* category, find a paper claiming a proof of the Riemann Hypothesis, the Twin Prime Conjecture, or Goldbach's problem. Find the mistake in the proof, and if you're brave communicate with the author.

Spend ten minutes a day for a month reading the titles, abstracts, and names of authors who have posted a paper

REFERENCES

[1] http://arxiv.org/.

[2] P. Ginsparg, "Winners and Losers in the Global Research Village, P. Ginsparg", Invited contribution for Conference held at UNESCO HQ, Paris, 19–23 Feb 1996, during session Scientist's View of Electronic Publishing and Issues Raised, Wed 21 Feb 1996. Available online at http://www.cs.cornell.edu/~ginsparg/physics/blurb/pg96unesco.html.

[3] P. Ginsparg, "arXiv.org blurb", http://www.cs.cornell.edu/~ginsparg/physics/blurb/.

1995

Fermat's Last Theorem

Fermat's Last Theorem (FLT) states: *Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.* Or, as most of us know, there are no solutions in non-zero integers to $x^n + y^n = z^n$ if $n > 2$. This year marked the publication of papers by Andrew Wiles and Richard Taylor and Andrew Wiles finally proving this claim. Their work is the outgrowth of numerous other people, which connected Fermat's problem to the theory of elliptic curves. Thus, while Fermat's result has held mathematicians' interest for centuries, the method of proof was at least as important as the result, as it yielded important results in active areas of research.

Centennial Problem 1995. *Proposed by proposed by the students in Frank Morgan's 'The Big Questions' class at Williams College, Fall 2008).*

The status of Fermat's Last Theorem for rational exponents is known; what about real exponents? Are there integral solutions to $x^r + y^r = z^r$ for r real? If yes, can you give a nice example?

REFERENCES

[1] K. Devlin, F. Gouvêa and A. Granville, "Fermat's Last Theorem, a Theorem at Last", FOCUS, August 1993, 3–5, online at http://www.dms.umontreal.ca/~andrew/PDF/FLTatlast.pdf.

[2] F. Gouvêa, "A Marvelous Proof", The American Mathematical Monthly **101** (1994), 203–222, online at http://mathdl.maa.org/images/upload_library/22/Ford/Gouvea203-222.pdf.

[3] F. Morgan, "Fermat's Last Theorem for Fractional and Irrational Exponents", College Math. J. **41** (2010), 182–185.

[4] R. Taylor and A. Wiles, "Ring-theoretic properties of certain Hecke algebras", Ann. Math. **141** (1995), 553–572, available online at http://www.jstor.org/discover/10.2307/2118560?uid=3739696&uid=2&uid=4&uid=3739256&sid=21102444307757.

[5] A. Wiles, "Modular elliptic curves and Fermat's Last Theorem", Annals of Mathematics **142** (1995), 443–551, available online at http://www.cs.berkeley.edu/~anindya/fermat.pdf.

1999

Baire's Category Theorem

A very important result in analysis, Baire's Category Theorem was published by the French mathematician René-Louis Baire in 1899 in his doctoral thesis titled "Sur les fonctions de variables réelles". As explained in [1], [2], and [3], Baire's Category Theorem has numerous applications to various branches of analysis; in particular, it is the main ingredient in the proof of three fundamental theorems in Functional Analysis: the Open Mapping Theorem, the Closed Graph Theorem, and the Uniform Boundedness Principle.

A few definitions are necessary in order to state this important theorem. A subset A of a topological space is called nowhere dense if its closure $\overline{A}$ has empty interior. A subset A of a topological space is called a set of the first Baire category (or "meagre") if it can be written as the countable union of nowhere dense sets; otherwise A is called a set of the second Baire category. The first version of Baire's Category Theorem is: In a complete metric space X, the set X is of the second Baire category.

An equivalent version of the Baire's Category Theorem is obtained by using the concept of Baire space: a topological space X is a Baire space if it has the property that any intersection of open dense sets is dense. The second version of Baire's Category Theorem is: Any complete metric space is a Baire space.

Centennial Problem 1999. *Proposed by Mihai Stoiciu, Williams College.*

Let $\mathbb{C}[x]$ be the vector space of polynomials in one variable with complex coefficients and let $\|\cdot\| : \mathbb{C}[x] \to [0, \infty)$ be a norm on $\mathbb{C}[x]$ (for example, $\|P\| =$

$\sqrt{\int_0^1 |P(x)|^2\,dx}$, for $P \in \mathbb{C}[x]$). Use Baire's Category Theorem to prove that the metric induced by $\|\cdot\|$ on $\mathbb{C}[x]$ is not complete (or, equivalently, $(\mathbb{C}[x], \|\cdot\|)$ is not a Banach space).

REFERENCES

[1] G. B. Folland, "Real analysis. Modern techniques and their applications." Second edition. Pure and Applied Mathematics (New York). A Wiley-Interscience Publication. *John Wiley & Sons, Inc., New York*, 1999.

[2] S. H. Jones, "Applications of the Baire category theorem". Real Anal. Exchange **23** (1997/98), no. 2, 363-394.

[3] T. Tao, "The Baire category theorem and its Banach space consequences", available online at http://terrytao.wordpress.com/2009/02/01/245b-notes-9-the-baire-category-theorem-and-its-banach-space-consequences/.

2003

Poincaré Conjecture

In 2003 the first of the million dollar Clay millennial prizes (see the problem from 1925 for more on these problems) fell as Perelman proved the Poincaré Conjecture, which says that a compact, simply connected 3-dimensional manifold must be a sphere. To understand what this means, it's worthwhile exploring the 2-dimensional equivalent. Consider a compact two-dimensional surface without a boundary. If every closed loop on it can be contracted, staying on the surface, to a point, then the surface is topologically homeomorphic to the two-dimensional sphere. Thus a donut is not the same as a basketball, as a ring around the donut cannot be contracted to a point. Thus the Poincaré conjecture says that this condition detects spheres in 3-dimensions as well.

Centennial Problem 2003. *Proposed by Frank Morgan, Williams College.*

Show that a nice simply connected 4-dimensional manifold need not be a 4-sphere. (One also needs some assumption on the 2-homotopy.)

REFERENCES

[1] J. Milnor, "Poincare Conjecture", http://www.claymath.org/millennium/Poincare_Conjecture/ (see in particular the problem description posted there by John Milnor).

2007

Flatland

The year 2007 marks the latest attempt to capture the classic book "Flatland" on film, and it is one of the most successful. "Flatland: the Movie" is a creation of three filmmakers from the University of Texas, Austin who discovered that they had all read and enjoyed the book when they were young students. Seth Caplan is the producer, Jeffrey Travis the director, and Dano Johnson the chef animator and their film has become a hit among teachers in middle and high school geometry classes.

Although some of the social satire of the original book has been modified, the central part of the story remains. The precocious hexagonal grandson in the classic has been replaced by an equally precocious granddaughter, Hex, with the voice of Kristin Bell (who voiced Princess Anna in the recent smash "Frozen"). She is a good spokesperson for mathematics in the extras at the end of the DVD, along with Martin Sheen (of "West Wing").

The major innovation in the film is a mysterious artifact left in the two-dimensional world of Flatland by a visitor from the third dimension, namely a cube that rotates about a point so that the Flatlanders can see all of its various cross-sections. The challenge is for the onlookers is to imagine what kind of object could produce all of those slices, and it is only when A Square and Hex are taken up into the third dimension by the three-dimensional visitor Spherius that they can begin to appreciate geometric phenomena in a dimension higher than their own.

Edwin Abbott Abbott definitely wanted to challenge his readers to imagine the analogous situation if we were confronted by phenomena originating in a fourth dimension of space. The film concludes with views of a four-dimensional cube, a "hypercube", that is projected into our space as it rotates in various ways in the fourth dimension.

Centennial Problem 2007. *Proposed by Thomas Banchoff, Brown University.*

For starters, describe all of the kinds of slices that occur when a cube in three-space is sliced through it center by a various planes. In particular when the plane is parallel to a face of the cube or to an edge of the cube or corner first. The real challenge is to generalize this exercise to the fourth dimension–what are the three-dimensional slices through the origin of a hypercube?

It may be helpful to think in terms of 2-dimensional coordinate geometry, where a two-dimensional square has four vertices (1,1), (-1,1), (-1,-1) and (1,-1). What happens when we slice by the y-axis? What about the line $y = x$?

In three dimensions, consider the cube with eight vertices (1,1,1), (-1,1,1), (-1,-1,1), (1,-1,1) and the four vertices with the third coordinate changed to -1. What are the slices by the plane perpendicular to (1,0,0), or the plane perpendicular to (1,1,0), or, most interestingly, the plane perpendicular to (1,1,1), the longest diagonal?

The analogue is clear. A four-dimensional cube has sixteen vertices, each with four coordinates that are 1 or -1. The most symmetric slicing hyperplanes are the ones perpendicular to (1,0,0,0), or (1,1,0,0), or (1,1,1,0), and, most interestingly, (1,1,1,1), the long diagonal.

Martin Gardner, who is being celebrated this year on the hundredth anniversary of his birth, treated the hypercube more than once in his columns. He raised the question of which of the central slices of the three-dimensional cube has the greatest area (the answer might be surprising). The analogous question asks which central slice of the hypercube has the greatest volume.

A harder problem, also quite fascinating, asks for the structure of the central slice of the five-dimensional cube by a four-dimensional hyperplane perpendicular to its long diagonal.

REFERENCES

[1] E. A. Abbot with an introduction by T. Banchoff, "Flatland: A romance of many dimensions", Princeton Science Series, 2005. See the LaTeX version from Ives van der Flaas, September 20, 2011. Available online: https://github.com/Ivesvdf/flatland/blob/master/oneside_a4.pdf?raw=true.

[2] E. A. Abbott with T. Banchoff and the filmmakers of Flatland, "Flatland: The Movie, Princeton University Press, 2008

[3] E. A. Abbot, W. F. Lindgren and T. Banchoff, "Flatland: An Edition with Notes and Commentary", Cambridge University Press, 2009.

[4] Banchoff and Strauss Productions, "The Hypercube: Projections and Slicing," 1978. Available online: http://www.math.brown.edu/~banchoff/video/hypercube.mp4.

[5] T. Banchoff, "Additional notes on Flatland", 2014. Available online: http://www.math.brown.edu/~banchoff/HexCentralSlices/HexCentralSlices4308.html.

[6] Flatland Homepage, "Flatland the Movie". http://www.flatlandthemovie.com/.

2011

100th Anniversary of Egorov's theorem

"On voit sans peine que ce théorème est susceptible d'un grand nombre d'applications" ("One sees with no effort that this theorem is prone to a great number of applications"). This is the proud, yet unostentatious, end of Dmitri F. Egorov's paper on the fundamental result in measure theory commonly know as *Egorov's theorem*. (It would be unfair to reduce Egorov's contribution to Mathematics to this theorem, for he is one of the most influential intellectuals of his generation. If we accept to measure someone's 'mathematical influence' by the number of academic 'descendants' recorded by the Mathematics Genealogy Project, then Egorov is the most prominent mathematician among the graduates of 1901. As of September 10th 2013, he has 4587 descendants, far ahead of the second most descendant-prolific mathematician of the 1901 cohort (I. Schur, 2115 descendants)). His theorem allows us –for a small price– to recoup the familiar notion of *uniform* convergence for a pointwise convergent sequence of functions over an interval, assuming only their *measurability*. The price we pay is to restrict the functions to a subset of the interval whose complement has arbitrarily small measure. Egorov has to share the paternity of his theorem with Carlo Severini, a lesser known mathematician who actually preceded Egorov by few months. Severini published his theorem in April 1910 in a journal of rather limited diffusion, and his contribution was not acknowledged outside Italy until 1924 when the famous L. Tonelli published a short note crediting him with the initial discovery. Severini's and Egorov's proof are practically identical but discovered independently.

Let us formulate the celebrated theorem precisely. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of real-valued measurable functions defined over $[a,b]\subset\mathbb{R}$. Assume that $\lim_{n\to\infty} f_n(x) = f(x)$ for almost every $x\in[a,b]$, where the function f is automatically measurable. Severini-Egorov's theorem states that for every $\varepsilon>0$ there exists a measurable subset $E\subset[a,b]$ such that $|E|<\varepsilon$ and (f_n) converges to f *uniformly* on $[a,b]\smallsetminus E$. Here $|\cdot|$ denotes the Lebesgue measure.

Centennial Problem 2011. *Proposed by Francesco Cellarosi, University of Illinois Urbana-Champaign.*

Severini-Egorov's theorem has been assumed to be true (e.g. page 79 of Hardy-Rogosinski) when (f_n) is replaced by a family of functions that depends on a real parameter and tends continuously to some measurable function. This is false, however, as shown by J. D. Weston in 1958. The problem we propose is to re-discover Weston's counterexample. *Hint*: Find a family of functions $(f_h)_{0\le h<1}$ on $[0,1)$ such that (i) for each $h\in[0,1)$, $f_h(x)=0$ except possibly at a single point x where $f_h(x)=1$; (ii) for each $x\in[0,1)$, $f_h(x)\to 0$ as $h\to 0$; (iii) the convergence is not uniform on any set of positive measure. *Warning: If you are reluctant to use the Axiom of Choice to construct a useful non-measurable set, you might be out of luck here!*

REFERENCES

[1] D. T. Egoroff, "Sur les suites des fonctions mesurables", Comptes rendus hebdomadaires des séances de l'Académie des sciences **152** (1911), 244–246 [in French].

[2] G. H. Hardy and W. W. Rogosinski, "Fourier series", 1st edition, Cambridge tracts in mathematics and mathematical physics, no. 38, 1944.

[3] C. Severini, "Sopra gli sviluppi in serie di funzioni ortogonali", Atti della Accademia Gioenia di scienze naturali in Catania, Series V **III** (1910), 1–7 [in Italian].

[4] J. D. Weston, "A counter-example concerning Egoroff's theorem", Journal of the London Mathematical Society **34** (1959), 139–140.