# Amicable numbers

Brown University
Number Theory Seminar
April 6, 2015

Carl Pomerance, Dartmouth College
(U. Georgia, emeritus)

Let $s(n)$ be the sum of the proper divisors of $n$. That is, $s(n) = \sigma(n) - n$, where $\sigma(n)$ is the sum of all of $n$'s natural divisors.

Two different numbers $n, n'$ form an *amicable pair* if $s(n) = n'$ and $s(n') = n$. This concept goes back to Pythagoras who found the pair 220 and 284.

The condition is easily seen to be equivalent to

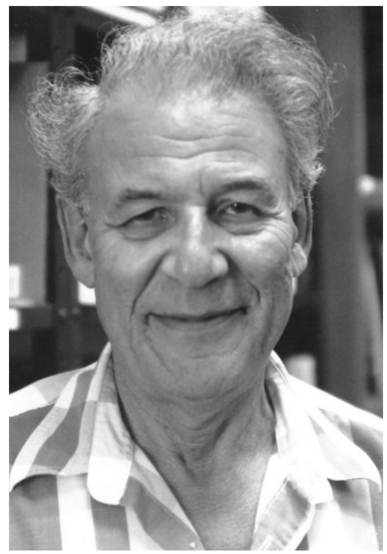
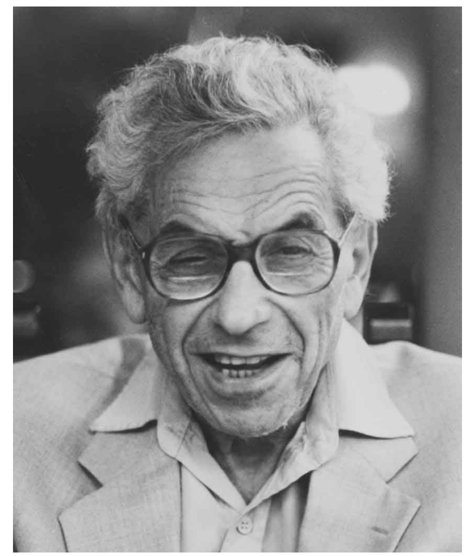
$$n + n' = \sigma(n) = \sigma(n').$$

We currently know about 12 million different amicable pairs.

Erdős has a heuristic that there should be infinitely many amicable pairs, in fact, more than $x^{1-\epsilon}$ of them up to $x$, for each fixed $\epsilon > 0$ and $x$ sufficiently large in terms of $\epsilon$:

There is already a well-known and widely believed heuristic of Erdős that there are infinitely many numbers $N$ with $\sigma(n) = N$ having more than $N^{1-\epsilon}$ solutions $n$. (This follows from the Elliott–Halberstam conjecture, and is proved for $\epsilon = 0.3$.) So among these solutions $n$, it should not be that unusual to have two of them with $n + n' = N$, in fact there ought to be about $N^{1-2\epsilon}$ such pairs.

But if $n + n' = \sigma(n) = \sigma(n')$ and $n \neq n'$, we have seen that $n, n'$ form an amicable pair.

Nevertheless, we have not proved that there are infinitely many amicable numbers.

Can we prove that amicable numbers are rare among the natural numbers?

This quest was begun by Kanold in 1954, who showed that the number of integers $n \leq x$ that belong to an amicable pair is at most $.204x$ for all sufficiently large values of $x$.

To fix notation, let $\mathcal{A}(x)$ denote the number of integers $n \leq x$ that belong to some amicable pair. Here's what's happened since Kanold:

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Rieger (1973): $\mathcal{A}(x) \le x/(\log\log\log\log x)^{1/2}$, $x$ large.

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Rieger (1973): $\mathcal{A}(x) \le x/(\log\log\log\log x)^{1/2}$, $x$ large.

Erdős & Rieger (1975): $\mathcal{A}(x) = O(x/\log\log\log x)$.

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Rieger (1973): $\mathcal{A}(x) \le x/(\log\log\log\log x)^{1/2}$, $x$ large.

Erdős & Rieger (1975): $\mathcal{A}(x) = O(x/\log\log\log x)$.

P (1977): $\mathcal{A}(x) \le x/\exp((\log\log\log x)^{1/2})$, $x$ large.

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Rieger (1973): $\mathcal{A}(x) \le x/(\log\log\log\log x)^{1/2}$, $x$ large.

Erdős & Rieger (1975): $\mathcal{A}(x) = O(x/\log\log\log x)$.

P (1977): $\mathcal{A}(x) \le x/\exp((\log\log\log x)^{1/2})$, $x$ large.

P (1981): $\mathcal{A}(x) \le x/\exp((\log x)^{1/3})$, $x$ large.

Erdős (1955): $\mathcal{A}(x) = o(x)$ as $x \to \infty$. Said his method would give $\mathcal{A}(x) = O(x/\log\log\log x)$.

Rieger (1973): $\mathcal{A}(x) \le x/(\log\log\log\log x)^{1/2}$, $x$ large.

Erdős & Rieger (1975): $\mathcal{A}(x) = O(x/\log\log\log x)$.

P (1977): $\mathcal{A}(x) \le x/\exp((\log\log\log x)^{1/2})$, $x$ large.

P (1981): $\mathcal{A}(x) \le x/\exp((\log x)^{1/3})$, $x$ large.

P (2014): $\mathcal{A}(x) \le x/\exp((\log x)^{1/2})$, $x$ large.

Here's a quick proof that $\mathcal{A}(x) = o(x)$ as $x \to \infty$, using a seemingly unrelated result of Erdős from the 1930s:

Say a number $a$ is "pnd" (stands for "primitive non-deficient") if $s(a) \ge a$, yet for each proper divisor $b$ of $a$, $s(b) < b$. It is clear that $s(n) \ge n$ if and only if $n$ is divisible by some pnd. Now Erdős showed that

$$\sum_{a \text{ is pnd}} \frac{1}{a} < \infty. \tag{1}$$

Say $n, m$ form an amicable pair with $n < m$. Then $s(n) = m > n = s(m)$, so that $n$ is divisible by some pnd and $m$ is not. By (1), but for a negligible set of numbers $n$, we may assume that some pnd $a = O(1)$ divides $n$.

For a fixed number $a$, the primes $p \equiv -1 \pmod a$ have density $1/\varphi(a)$ in the primes, so almost all numbers $n$ are divisible by such a prime $p$ to exactly the first power. Then $a \mid p+1 \mid \sigma(n)$, so that $a \mid \sigma(n) - n = s(n) = m$.

But $m$ is not divisible by any pnd, so we have a contradiction. This shows that amicable numbers lie in the two negligible sets discarded along the way.

The proof shows that the set $\mathcal{S}$ of integers $n$ with $s(n) > n$ and $s(s(n)) \le s(n)$ has density 0. In fact, using finer information on pnd's, one can show that the number of members of $\mathcal{S}$ in $[1, x]$ is at most $x/\exp((\log\log\log x)^{1/2})$ for large $x$ (my 1977 theorem), and at least $x/(\log\log x)^{1+\epsilon}$.

To make further progress on bounding $\mathcal{A}(x)$, we need to use more properties of amicable numbers.

As typical with an Erdős-style argument, one divides the problem into a number of cases, some being routine, some not. But there is an overarching strategy which is sometimes lost in the details.

Here's the strategy. We have $n, n'$ an amicable pair. Write

$$n = pk, \quad n' = p'k',$$

where $p, p'$ are the largest primes in $n, n'$, respectively. Assume that $p > p',\ p \nmid k,\ p' \nmid k'$. (The cases $p = p',\ p \mid k,\ p' \mid k'$ are easily handled.)

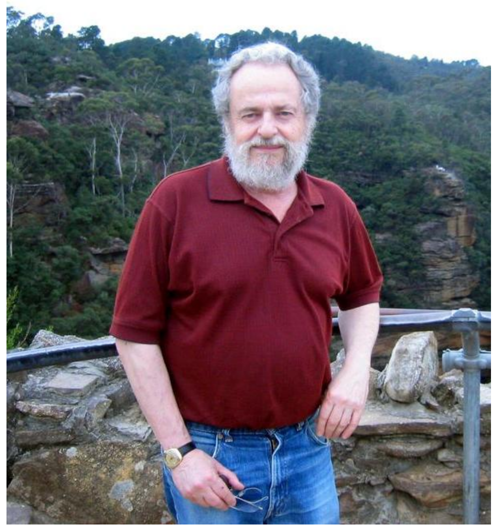
We may assume that $k, k'$ are largely squarefree, their smooth parts are not too large, and they both have some size. That is, $p, p'$ don't overly dominate. (For $n \le x$, we have $p, p' < x^{3/4}$, approximately.)

Since $k$ is large, we may assume that $r$, the largest prime factor of $\sigma(k)$ is large, say $r > L$ (where $L$ is to be chosen). (Prove directly or use a recent paper of Banks, Friedlander, P, & Shparlinski. Lamzouri and Yamada have also considered smooth values of multiplicative functions.)

We have $r \mid q+1$ for some prime $q \parallel k$. Since $r \mid \sigma(k) \mid \sigma(n) = \sigma(n')$, we have $r \mid q'+1$ for some prime $q' \parallel n'$, and $q' \neq q$. Note that

$$q' \mid n' = s(n) = s(pk) = (p+1)s(k) + \sigma(k).$$

Thus, with $k, q'$ given, $p$ is constrained to a residue class modulo $q'$.



We thus count as follows:

$$\sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \sum_{\substack{k<x \\ q|k}} \sum_{\substack{q'<x \\ r|q'+1}} \sum_{\substack{p\le x/k \\ p>q' \\ p\equiv a_{k,q'} \pmod{q'}}} 1.$$

We thus count as follows:

$$\sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \sum_{\substack{k<x \\ q|k}} \sum_{\substack{q'<x \\ r|q'+1}} \sum_{\substack{p \le x/k \\ p>q' \\ p \equiv a_{k,q'} \pmod{q'}}} 1.$$

And we're laughing.

$$\sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \sum_{\substack{k<x \\ q|m}} \sum_{\substack{q'<x \\ r|q'+1}} \sum_{\substack{p\le x/k \\ p>q' \\ p\equiv a_{k,q'} \pmod{q'}}} 1.$$

The inner sum can be replaced with $\frac{x}{kq'}$:

$$x \sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \sum_{\substack{k<x \\ q|k}} \sum_{\substack{q'<x \\ r|q'+1}} \frac{1}{kq'}.$$

The new inner sum can be replaced with $\frac{\log x}{kr}$:

$$x \log x \sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \sum_{\substack{k<x \\ q|k}} \frac{1}{kr}.$$

The new inner sum can be replaced with $\frac{\log x}{qr}$:

$$x(\log x)^2 \sum_{r>L} \sum_{\substack{q<x \\ r|q+1}} \frac{1}{qr}$$

The new inner sum can be replaced with $\frac{\log x}{r^2}$:

$$x(\log x)^3 \sum_{r>L} \frac{1}{r^2}.$$

And this sum is smaller than $1/L$, so we have the estimate $x(\log x)^3/L$. By choosing $L$ as large as possible so that the various assumptions may be justified, we have our result. And in fact we can choose $L$ a tad larger than $\exp(\sqrt{\log x})$. So, we have

**P** (**2014**): For all large $x$, $\mathcal{A}(x) \le x/\exp(\sqrt{\log x})$.

My 1981 result with $x/\exp((\log x)^{1/3})$ did not get such a good lower bound on $m, m'$ so that it was difficult to show that $\sigma(m), \sigma(m')$ had large prime factors.

Beyond showing there are few amicable numbers, it should be true that there are few *sociable* numbers.

Definition: Say a number $n$ with $s_k(n) = n$ for some $k$ is sociable.

That is, sociable numbers are the numbers involved in a cycle in the dynamical system introduced by Pythagoras.

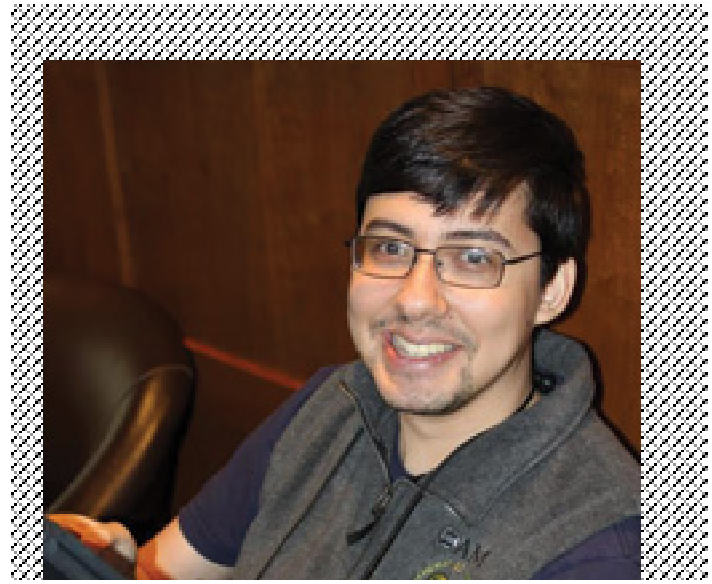
Here is what we know about the distribution of sociable numbers.

From a 1976 Erdős result (*if* $n < s(n)$, *then almost surely the sequence continues to increase for* $k$ *terms*), one can show that the sociable numbers that belong to a cycle of any fixed length have asymptotic density 0. (This is how we argued that $\mathcal{A}(x) = o(x)$ earlier in this talk.)

In Kobayashi, Pollack, & P (2009), we showed that

- The even sociable numbers have asymptotic density 0.

- The odd sociable numbers $n$ with $n > s(n)$ have asymptotic density 0.

This would leave the odd sociables $n$ with $n < s(n)$. The odd numbers $n$ with $n < s(n)$ have an asymptotic density of about 1/500, so we're talking about a fairly sparse set to begin with. But the problem of showing the sociable numbers in this set have density 0 is still open.

Here are some other solved and unsolved problems in connection with the function $s(n)$.

Erdős showed in 1973 that there is a positive proportion of even numbers that are not in the form $s(n)$. (It is easy to see as a consequence of attacks on Goldbach's conjecture that almost all odd numbers are in the form $s(n)$.)

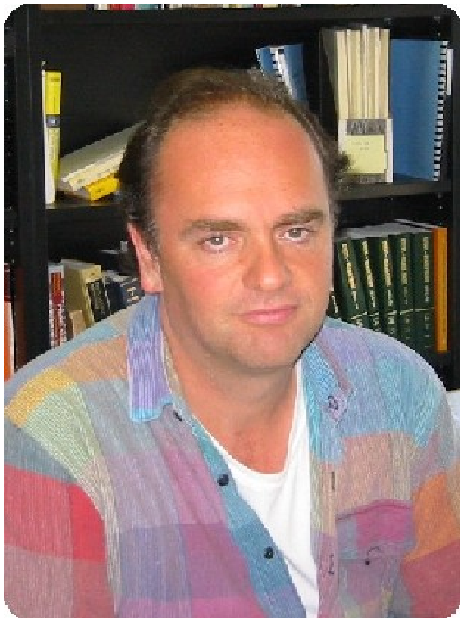
Luca, P (2014): *Every residue class* $a \bmod b$ *contains a positive proportion of numbers of the form* $s(n)$.

**Conjecture (Erdős, Granville, P, & Spiro – 1990):**
*If $\mathcal{S} \subset \mathbb{N}$ has density 0, then $s^{-1}(\mathcal{S})$ has density 0.*

A consequence: If $n > s(n)$, then almost surely the sequence $n, s(n), s(s(n)), \dots,$ continues to decrease for another $k$ steps. Here's why. Let $\mathcal{S}_k$ be the set of $n$ which decrease for $k$ steps and then do not increase at the next step. Then $s(\mathcal{S}_k) \subset \mathcal{S}_{k-1}$, so that $\mathcal{S}_k \subset s^{-1}(\mathcal{S}_{k-1})$. We know (Erdős, Granville, P, & Spiro) that $\mathcal{S}_1$ has density 0, so the conjecture implies that each $\mathcal{S}_k$ has density 0.

Note that there is a set $\mathcal{S}$ of density 0 such that $s(\mathcal{S})$ has density $\frac{1}{2}$, namely the set $\{pq : p, q \text{ prime}\}$.

The conjecture is that $s$ cannot map a positive proportion of the integers to a set of density 0.



There are a number of other attractive problems here.

It is easy to describe all solutions $n$ to $\varphi(n)/n = r$, where $r$ is a given rational number. For example, if $r = 1/2$, then $n$ is a power of 2, and if $r = 1/3$, then $n$ has radical 6.

What about solving $\sigma(n)/n = r$?

Remark: No one knows a solution to $\sigma(n)/n = 5/3$, but if such a solution $n$ should exist, then $5n$ would be an odd perfect number.

Idea of the proof that a positive proportion of even numbers are values of $s(n) = \sigma(n) - n$ (Luca & P, 2014):

Consider even numbers $n$ with several constraints:

- $n$ is deficient (means that $s(n) < n$);

- $n = pqrk \in [\frac{1}{2}x, x]$ with $p > q > r > k$ and $p, q, r$ primes;

- $k \le x^{1/60}, \quad r \in [x^{1/15}, x^{1/12}], \quad q \in [x^{7/20}, x^{11/30}]$;

- $n$ is “normal”.

If $n$ satisfies these conditions, then $s(n) \leq x$ is even.

Let $r(s)$ denote the number of representations of $s$ as $s(n)$ from such numbers $n$.

We have $\sum_s r(s) \gg x$.

The trick then is to show that $\sum_s r(s)^2 \ll x$.

For this, the sieve is useful.

We know less about the function $s_\varphi(n) = n - \varphi(n)$.

As with $s(n)$, what's known about Goldbach's conjecture implies that almost all odd numbers can be represented in the form $s_\varphi(n)$. And the proof with Luca that every residue class has a positive proportion of its members in the form $s(n)$ should go over for $s_\varphi(n)$.

But the Erdős argument that $s$ misses a positive proportion of even numbers just doesn't work for $s_\varphi$.

There are probabilistic heuristics that this should hold, but some new idea would be needed to prove it.

I would like to close this talk with a heuristic argument for the distribution of values of $s(n)$ developed recently with Pollack.

For some reason there has been some particularly colorful nomenclature surrounding numbers missing from the range of $s$. In his 1972 Yale Ph.D. thesis, Alanen called such numbers *untouchable*.

Ibn Tahir al-Baghdadi (980–1037) referred to $n$ in the equation $s(n) = m$ as the *begetter* and $m$ as the *begotten*. He goes on to write:

*“As to the begetter and the begotten among the numbers, so the sum of the aliquot parts is the begotten of this number, which itself is the begetter of its aliquot parts. Now,* 5 *among the odd numbers and* 2 *among the even numbers have no begetter, since there is no number such that the sum of its aliquot parts be* 5 *or* 2*. Hence, they stand among the numbers like a bastard among the people.”*

We will simply refer to these numbers unjudgmentally as *nonaliquot*.

The sequence of nonaliquot numbers:

$$2,\ 5,\ 52,\ 88,\ 96,\ 120,\ 124,\ 146,\ 162,\ 188, \dots\ .$$

Odd numbers here are scarce, and if we believe a slightly stronger form of Goldbach's conjecture, 5 is the only one. Provably, the odd nonaliquot numbers have density 0.

We've seen that a positive proportion of evens are nonaliquot and a positive proportion of evens are aliquot, that is, missing from this list.

In his 1976 Ph.D. thesis, te Riele gave a heuristic argument for the density of nonaliquot numbers. He views $s(n)$ as a random function, and if $M(x)$ is the number of integers $n$ with $s(n) \le x$ and $s(n)$ even, then the chance that a random even $m \le x$ is nonaliquot should be $(1-2/x)^{M(x)}$.

This heuristic is refined in a recent paper with Pollack: For each even $y$-smooth integer $a$, we consider those $n$ with $y$-smooth part $a$. For $a, y$ fairly small, the $y$-smooth part of $s(n)$ is also usually $a$, so we consider $s$ as random on this subset. This leads to a conjectured density of nonaliquots of

$$\lim_{y\to\infty} \frac{1}{\log y} \sum_{\substack{a \le y \\ 2|a}} \frac{1}{a} e^{-a/s(a)} \approx 0.1718.$$

Reality check: at $10^{10}$, the density of nonaliquots is $\approx 0.1682$.

# Thank You