The Erdős discrepancy problem

Terence Tao

University of California, Los Angeles

June 15, 2017

- Suppose one has a finite or infinite sequence $f(1), f(2), f(3), \dots \in \{-1, +1\}$ of signs, e.g.

$$+1, -1, -1, +1, +1, -1, -1, \dots$$

- If the sign sequence was distributed "randomly", one would expect any given portion of the sequence (e.g. an initial segment $f(1), f(2), \dots, f(n)$) to have about half of its elements equal to $+1$, and half equal to -1.
- In particular, we expect sums such as $\sum_{i=1}^{n} f(i)$ - which measure the discrepancy of the sequence from this expectation - to be significantly smaller in magnitude than the trivial bound of n.
- But how small can one make such discrepancies? This is the topic of discrepancy theory.

A simple discrepancy problem: given a finite sequence $f(1), f(2), \ldots, f(n) \in \{-1, +1\}$ of signs, how small can one make the largest partial sum

$$\sup_{1 \leq m \leq n} |\sum_{i=1}^{m} f(i)|?$$

- Using tools such as the Chernoff inequality, a randomly chosen sequence will typically have a largest partial sum of $O(\sqrt{n})$.
- But one can do much better simply by using the alternating sequence $f(i) := (-1)^i$:

$$+1, -1, +1, -1, \ldots.$$

The largest partial sum here is of course just 1. This is an example of an sub-random sequence - a sequence with better discrepancy properties than a random sequence.

Now we consider the discrepancy along arbitrary arithmetic progressions. Given a finite sequence $f(1), f(2), \ldots, f(n) \in \{-1, +1\}$ of signs, how small can one make the quantity

$$\sup_{a,r,k \geq 1} |f(a) + f(a+r) + \cdots + f(a+(k-1)r)|$$

(with the convention that $f(i) = 0$ for $i > n$)?

- Again, for a random sequence, this quantity will typically be of size $O(\sqrt{n})$. Can one do better than this?

$$\sup_{a,r,k\geq 1} |f(a) + f(a+r) + \cdots + f(a+(k-1)r)| \quad (*)$$

- van der Waerden's theorem (1927) asserts that if the natural numbers are partitioned into finitely many classes, one of the classes will contain arbitrarily long arithmetic progressions. This implies that the quantity (*) must go to infinity as $n \to \infty$.
- In 1964, Roth shows that the quantity (*) must be as large as $cn^{1/4}$ for some constant $c > 0$. This bound is best possible (Matoušek, Spencer, 1996).

- The Erdős discrepancy problem considers the discrepancy along homogeneous arithmetic progressions:

$$\sup_{a,k\geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)| \quad (**).$$

- In 1932, Erdős asked if this discrepancy was always unbounded for any infinite sign sequence $f(1), f(2), \cdots \in \{-1, +1\}$.
- An equivalent formulation: for any $C \geq 1$, does there exist $n \geq 1$ such that the discrepancy (**) of any finite sign sequence $f(1), \ldots, f(n)$ is at least C?

$$\sup_{a,k\geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)| \geq C \quad (***).$$

What happens for small values of C?

- When $C = 1$, one clearly has (***) as soon as $n \geq 1$.

$$\sup_{a,k\geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)| \geq 2?$$

- The length 11 sequence

$$+1, -1, -1, +1, -1, +1, +1, -1, -1, +1, -1$$

  has discrepancy 1.
- On the other hand, for any $n \geq 12$, one can show by that any length n sequence of signs will have discrepancy at least 2. Sketch of proof: assume for instance that $f(1) = +1$ and that the discrepancy is at most one. This forces $f(2) = -1$, which then forces $f(4) = +1$, which then forces $f(3) = -1$, etc.; continuing in this fashion (much like solving a "Sudoku" puzzle) one eventually deduces that both $f(12) = +1$ and $f(12) = -1$, a contradiction.

$$\sup_{a,k\geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)| \geq 3?$$

- In 2014, Konev and Lisitsa produced a sequence of length 1160 and discrepancy 2. They also showed (using an enormous `3-SAT` computation) that any sequence of signs of length $n > 1160$ would have discrepancy at least 3. Their original proof had a certificate that was 13 gigabytes in size, making it at one point the "longest mathematical proof" ever constructed. (But the proof was later "simplified" to 850 megabytes, and an unrelated computation in 2016 required 200 terabytes of data.)
- Until this computation, it was not known if there existed infinite sequences of discrepancy 2!

$$\sup_{a,k\geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)| \geq 4?$$

- In 2014, Konev and Lisitsa produced a sequence of length $13,900$ and discrepancy 3. However, it is not known if this is the longest such sequence.
- Indeed, until 2015, it was not known if it was possible to have an infinite sequence of discrepancy 3.

Theorem (T. 2015)

If $f : \mathbf{N} \to \{-1, +1\}$ is a sequence, then the sums $f(a) + f(2a) + \cdots + f((k-1)a)$ are unbounded. In other words, the answer to the Erdős discrepancy problem is affirmative.

What finite sequences $f(1), f(2), \ldots, \in \{-1, +1\}$ have small discrepancy

$$\sup_{a,k \geq 1} |f(a) + f(2a) + f(3a) + \cdots + f(ka)|?$$

- If the signs $f(1), f(2), \ldots$ are chosen randomly, then this discrepancy is typically of size $O(\sqrt{n})$. Can one do better?

- A good example of a non-random sign sequence is the Liouville function $\lambda : \mathbf{N} \to \{-1, +1\}$ from analytic number theory.
- $\lambda(n)$ is defined to equal $+1$ when n is the product of an even number of primes, and -1 when it is the product of an odd number of primes. For instance, $\lambda(4) = +1$, $\lambda(30) = -1$.
- It is completely multiplicative, thus $\lambda(a) + \cdots + \lambda(ka)$ has the same magnitude as $\lambda(1) + \cdots + \lambda(k)$.
- It is closely tied to the Riemann zeta function $\xi(s)$. The fact that this function has zeroes on the critical line $\{\frac{1}{2} + it : t \in \mathbf{R}\}$ implies that $\lambda(1) + \cdots + \lambda(k)$ can be at least as large as $\sqrt{k}$ for large k.

Good examples of a low-discrepancy sequences are provided by the Dirichlet characters.

- A Dirichlet character χ is a function $\chi : \mathbf{N} \to \mathbf{C}$ which is periodic with some period q, vanishing on numbers coprime to q, and completely multiplicative (i.e., $\chi(nm) = \chi(n)\chi(m)$ for all n, m). A Dirichlet character is *non-principal* if it has mean zero.
- For instance, the function $\chi_3(n)$ that is equal to $+1$ for $n = 1 \bmod 3$, -1 for $n = 2 \bmod 3$, and 0 for $n = 0 \bmod 3$.

- If χ is a non-principal Dirichlet character, then the partial sums $\chi(1) + \chi(2) + \cdots + \chi(k-1)$ are bounded.
- Using the complete multiplicativity of χ, we conclude that the discrepancy

$$\sup_{a,k \geq 1} |\chi(a) + \chi(2a) + \chi(3a) + \cdots + \chi(ka)|$$

  is also bounded.
- This is not a counterexample to the Erdős discrepancy problem, though, because Dirichlet characters $\chi(n)$ vanish for infinitely many choices of n.

- In 2010, Borwein, Choi, and Coons produced a near-counterexample relating to the Dirichlet character χ_3: the function $f : \mathbf{N} \to \{-1, +1\}$ defined by setting $f(n) = +1$ when $n = 3^k(3m+1)$ for some $k, m \geq 0$, and $f(n) = -1$ if $n = 3^k(3m+2)$.
- The partial sums $f(1) + \cdots + f(k)$ are equal to the number of 1s in the base 3 expansion of k; they do go to infinity, but only as fast as $O(\log k)$ - much slower than a random sequence!
- These examples suggest that the most dangerous sequences $f : \mathbf{N} \to \{-1, +1\}$ are those arising from Dirichlet characters.

In 2010, Timothy Gowers organised a Polymath project to attack the Erdős discrepancy problem. While they did not fully resolve the problem, they made an important reduction:

Theorem (slightly oversimplified)

To solve the Erdős discrepancy problem, it suffices to check sequences $f : \mathbf{N} \to \{-1, +1\}$ that are completely multiplicative.

This reduces the problem to

Simplified discrepancy problem

Let $f : \mathbf{N} \to \{-1, +1\}$ be a completely multiplicative function. Are the partial sums $f(1) + \cdots + f(k)$ unbounded?

Sketch of reduction (omitting some technicalities):

- Using Fourier analysis, one can represent an arbitrary sequence $f : \mathbf{N} \to \{-1, +1\}$ as an average of completely multiplicative functions $g : \mathbf{N} \to S^1$.
- If one forms an ℓ^2 average of the sums $f(d) + f(2d) + \cdots + f(kd)$ with respect to d, one obtains an ℓ^2 average of sums of the form $g(1) + g(2) + \cdots + g(k)$, basically because of Plancherel's theorem.
- Thus, if one can obtain a nontrivial lower bound on the sums $g(1) + \cdots + g(k)$, one of the sums $f(d) + f(2d) + \cdots + f(kd)$ must be large.

How to show that the partial sums $f(1) + \cdots + f(k)$ of a completely multiplicative function $f : \mathbf{N} \to \{-1, +1\}$ are unbounded?

Model problem

Show that the partial sums of the Liouville function $\lambda(1) + \cdots + \lambda(k)$ are unbounded.

As mentioned earlier, this can be established from the properties of the Riemann zeta function. Is there another way to prove this fact?

- When f is a random function, we know that the sums $f(1) + \cdots + f(k)$ grow like $\sqrt{k}$.
- Using the second moment method, one can show the same is true for non-random functions, as long as one can control covariances such as

$$\sum_{n \leq x} f(n) f(n+h)$$

  for fixed h.
- So it becomes natural to study covariances such as $\sum_{n \leq x} \lambda(n)\lambda(n+h)$.

The Chowla conjecture asserts that

$$\sum_{n \leq x} \lambda(n)\lambda(n+h) = o(x)$$

as $x \to \infty$ for any fixed non-zero h. This would be enough to imply the unboundedness of $\lambda(1) + \cdots + \lambda(k)$.
This conjecture remains open (it is similar to, though possibly easier than, the notorious twin prime conjecture). However, we have the following "logarithmically averaged" version:

Theorem (T., 2015)

One has $\sum_{n \leq x} \frac{\lambda(n)\lambda(n+h)}{n} = o(\log x)$ for any non-zero h.

This is still enough to show the unboundedness of $\lambda(1) + \cdots + \lambda(k)$; a more general version of this result can be used to solve the full Erdős discrepancy problem.

- The reason for using logarithmic averages instead of ordinary averages is because one gains approximate dilation invariance: for any function $f : \mathbf{N} \to \{-1, +1\}$ and any small modulus q, one has

$$\sum_{n \leq x} \frac{f(n)}{n} = \sum_{n \leq qx} \frac{f(n/q) q 1_{q|n}}{n}$$
$$\approx \sum_{n \leq x} \frac{f(n/q) q 1_{q|n}}{n}.$$

- Also, the Liouville function has the multiplicativity property $\lambda(pn) = -\lambda(n)$ for any prime p and natural number n.
- These two facts can be used to introduce additional averaging into sums such as $\sum_{n \leq x} \frac{\lambda(n)\lambda(n+1)}{n}$...

For any prime p, one has $\lambda(n) = -\lambda(np)$ and $\lambda(n+1) = -\lambda(np+p)$, hence

$$\begin{aligned}\sum_{n\leq x} \frac{\lambda(n)\lambda(n+1)}{n} &= \sum_{n\leq x} \frac{\lambda(pn)\lambda(pn+p)}{n} \\ &= \sum_{n\leq px} \frac{\lambda(n)\lambda(n+p)p1_{p|n}}{pn} \\ &\approx \sum_{n\leq x} \frac{\lambda(n)\lambda(n+p)p1_{p|n}}{n}.\end{aligned}$$

We can average over p in any finite set $\mathcal{P}$ of primes we choose to conclude

$$\sum_{n\leq x} \frac{\lambda(n)\lambda(n+1)}{n} \approx \frac{1}{|\mathcal{P}|} \sum_{p\in\mathcal{P}} \sum_{n\leq x} \frac{\lambda(n)\lambda(n+p)p1_{p|n}}{n}.$$

We have converted a single average into a double average!

- The most difficult thing about the expression

$$\sum_{p\in\mathcal{P}}\sum_{n\leq x}\frac{\lambda(n)\lambda(n+p)p1_{p|n}}{n}$$

  is the weight $p1_{p|n}$, which is sensitive to the residue class of n modulo various primes.

- But this weight is equal to 1 on the average. Adopting the heuristic $p1_{p|n} \approx 1$, we may expect to approximate this sum by the simpler expression

$$\sum_{p\in\mathcal{P}}\sum_{n\leq x}\frac{\lambda(n)\lambda(n+p)}{n}.$$

  In a breakthrough in 2016, Matomäki and Radziwiłł used methods from multiplicative number theory to obtain new estimates on λ that allow one to control such sums.

- The last remaining difficulty is to justify using the heuristic $p1_{p|n} \approx 1$ to simplify the expression

$$\sum_{p \in \mathcal{P}} \sum_{n \leq x} \frac{\lambda(n)\lambda(n+p)p1_{p|n}}{n}.$$

- Using probabilistic tools such as Hoeffding's inequality, one can use this heuristic unless there is an improbable "conspiracy" between sign patterns such as $(\lambda(n), \lambda(n+1), \ldots, \lambda(n+H))$ (for some medium-sized H) and residue classes n mod p. This "conspiracy" may be formalised using the concept of mutual information from information theory.

- Each set of primes $\mathcal{P}$ would require a different "conspiracy" between the sign pattern $(\lambda(n), \lambda(n+1), \ldots, \lambda(n+H))$ and residue classes n mod p. Using the Shannon entropy inequalities, one can show that such conspiracies cannot happen for many disjoint families of primes $\mathcal{P}$.
- Using the pigeonhole principle, one can then choose a family $\mathcal{P}$ for which there is no "conspiracy", and for which the argument works.

Thanks for listening!