Anurag Bishnoi
Delft Institute of Applied Mathematics
TU Delft
a.bishnoi@tudelft.nl

The Solution

Finite geometry paves the way for breakthroughs in Ramsey theory

This has been a year of breakthroughs in Ramsey theory, with remarkable progress on longstanding fundamental research problems. In this article, Anurag Bishnoi will focus on one particular work closely aligned with his own research. Drawing on a blend of algebra, geometry, and probability theory, Sam Mattheus (VUB, Belgium) and Jacques Verstraete (UCSD, USA) have successfully determined the correct asymptotics of a specific Ramsey number [10]. Furthermore, they have paved a new path for future developments in the area that relies on constructions from finite geometry.

Our story takes us back to the 1960s, when a Hungarian sociologist named Sándor Szalai made an intriguing observation. In every classroom with a minimum of twenty students, he always found either a clique of four students (each of them being friends with one another) or an anti-clique of four students (none of them being friends with each other). After sharing this sociological phenomenon with his mathematician friends, which included the renowned Paul Erdős, he came to a fascinating realization: he had stumbled upon the already known mathematical concept of Ramsey numbers.

The *Ramsey number*, denoted by $R(s,t)$, is the smallest n such that any group of n people, you're guaranteed to find either a clique of size s (where everyone knows everyone) or an anti-clique of size t (where nobody knows anyone). We can express this concept in the language of graph theory, where we represent people as vertices, and their friendships are illustrated by edges connecting these vertices. In this framework, $R(s,t)$ is the smallest number n for which every graph on n vertices either has a clique of size s or an anti-clique (commonly referred to as an *independent set*) of size t. Take a look at Figure 1 for a five-vertex graph illustrating $R(3,3) > 5$. I also encourage you to prove that $R(3,3)$ is equal to 6. You can either just check all the

$$2^{\binom{6}{2}} = 32768$$

graphs on 6 vertices, or find a direct argument without going through so many cases.

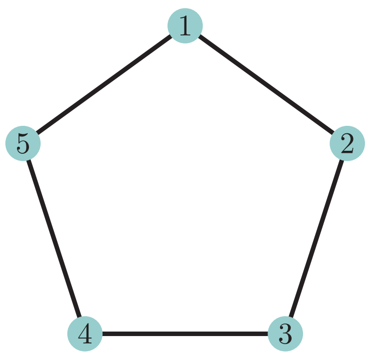

Figure 1 A 5-vertex graph with no cliques or independent sets of size 3.

The sociologist's observation can be reformulated as the inequality $R(4,4) \le 20$. In fact, it is known that $R(4,4)$ is exactly equal to 18, but even determining $R(5,5)$ is a major open problem in mathematics: we only know that $43 \le R(5,5) \le 48$. Note that there are 2^{1128} possible graphs on 48 vertices, and thus it is impossible to use a brute-force approach to prove this upper bound. Given the disheartening state of affairs, a natural question arises: can we at least understand the behaviour of the function $R(s,t)$ when either s,t, or both, tend to infinity? In other words, *what are the asymptotic properties of the Ramsey numbers?*

A classic inductive argument of Erdős and Szekeres [8] shows that

$$R(s,t) \le \binom{s+t-2}{s-1},$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the binomial coefficient. For $s = t$, we then have:

$$R(t,t) \le \binom{2t-2}{t-1} \le 4^t.$$

For any fixed s, we have:

$$R(s,t) \le \binom{t+s-2}{s-1} \le C_s t^{s-1},$$

where C_s is a constant that only depends on s. In other words, the 'diagonal' Ramsey number $R(t,t)$ is at most an exponential function in t, while the 'off-diagonal'

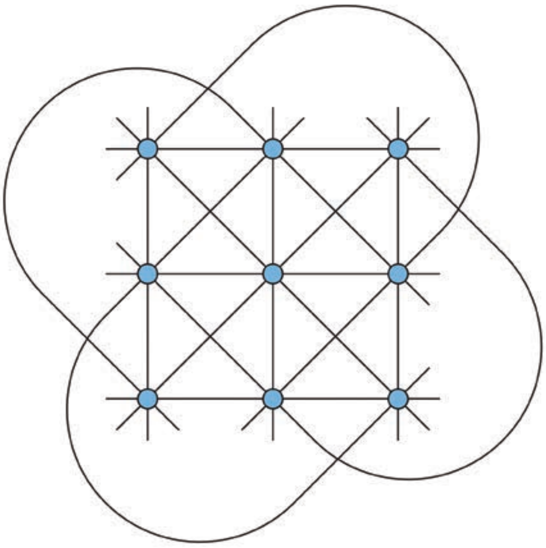

Figure 2 A finite affine plane.

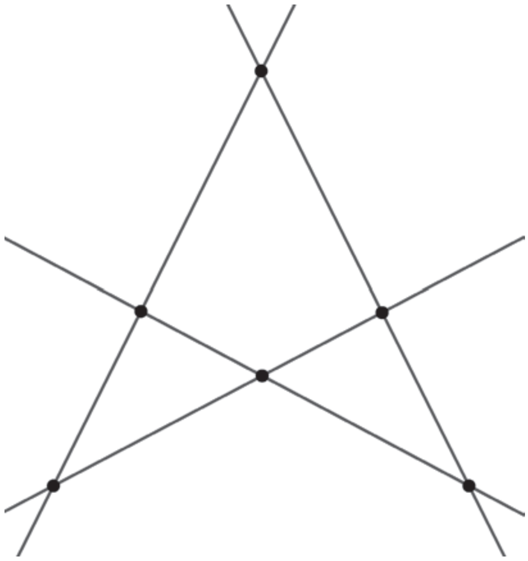

Figure 3 Four lines in general position with six intersection points.

Ramsey number $R(s,t)$ is at most a polynomial function in t of degree $s-1$. The challenge is then to determine how close these upper bounds are to the true behaviour of these functions. In the diagonal case, Erdős introduced a beautiful argument [7] to show that $R(t,t) > \sqrt{2}^t$. He picked graphs on these many vertices by tossing a fair coin to decide whether each pair of vertices is an edge or not. He then showed that with positive probability the graph will not have any cliques or independent sets of size t. This argument paved the way for the development of two mathematical fields: *Probabilistic Combinatorics* [2] and the theory of *Random Graphs* [9]. In fact, new developments in probabilistic combinatorics are often motivated by the problems on Ramsey numbers. Despite years of research, we haven't been able to improve the base of the exponent in the lower bound on $R(t,t)$ from $\sqrt{2}$ to something closer to the upper bound which has base 4. However, the upper bound has recently been improved to 3.99^t [6], which was the other major breakthrough this year.

The situation in *the off-diagonal case* is somewhat more promising. Through the development of intricate probabilistic arguments over several decades [1, 5, 13] it has been shown that

$$\left(\frac{1}{4}-o(1)\right)\frac{t^2}{\log t} \leq R(3,t) \leq (1+o(1))\frac{t^2}{\log t}.$$

Therefore, we know this Ramsey number up-to a factor of $4+o(1)$, where $o(1)$ is a function whose limit is 0 when t approaches infinity. However, for $s=4$, these probabilistic arguments can only show a lower bound of the form

$$R(4,t) \geq c\frac{t^{5/2}}{\log^2 t},$$

which is far away from the best-known upper bound of

$$R(4,t) \leq C\frac{t^3}{\log^2 t}.$$

The lower bound is widely believed to be the best one can hope for by purely probabilistic techniques. Given the decades long absence of any alternate techniques, some mathematicians started believing that the true value of $R(4,t)$ must be closer to the lower bound. World renowned mathematician and field medalist Timothy Gowers has openly admitted to working on improving the upper bound on $R(4,t)$ (see his tweet https://twitter.com/wtgowers/status/1666724170783244288). However, Sam Mattheus and Jacques Verstraet [10] have shattered any such beliefs by proving

$$R(4,t) \geq c'\frac{t^3}{\log^4 t},$$

thus matching the best upper bound up-to a factor of $\log^2 t$ and a constant.

The *key insight* in this work is to abandon the realm of pure randomness and start with a finite geometric construction. This idea is from an earlier work of Mubayi and Verstraete [11], where they proved that finding an 'optimally pseudorandom' graph with no cliques of size s suffices to determine the asymptotics of $R(s,t)$. In this context pseudorandomness is related to certain edge distributions in a deterministic graph that make it look like a random graph, and optimal pseudorandomness is about being as close as possible to the random graph. Mubayi and Verstraete had then reduced the problem on Ramsey numbers to a problem on the existence of these clique-free optimally pseudorandom graphs, which is perhaps more tractable. However, the current best-known construction for those, which is also based on finite geometry [4], falls short of improving the probabilistic bounds on the Ramsey numbers.

Luckily, Mubayi and Verstraete also proved a much more general result: if an n vertex graph has at most m^t independent sets of size t, then we can sample a set of n/m vertices on which there is no independent set of size t. To be able to use this result, Mattheus and Verstraete found a construction of graphs with no cliques of size 4 and 'few' independent sets of size t. At the heart of their construction lie the so-called *Hermitian curves* over finite fields. These curves naturally emerge in various contexts, most notably within the realms of algebraic geometry and finite geometry [3]. The latter is the field of expertise for Sam Mattheus, who recently finished his PhD at Vrije Universiteit Brussel and then visited Jacques Verstraete as a postdoc. In finite geometry, we study finite collections of 'points' and 'lines' that satisfy some geometric axioms inspired from classical geometry. For example, in Figure 2 (created by David Eppstein) you can see a collection of 9 points and 12 lines

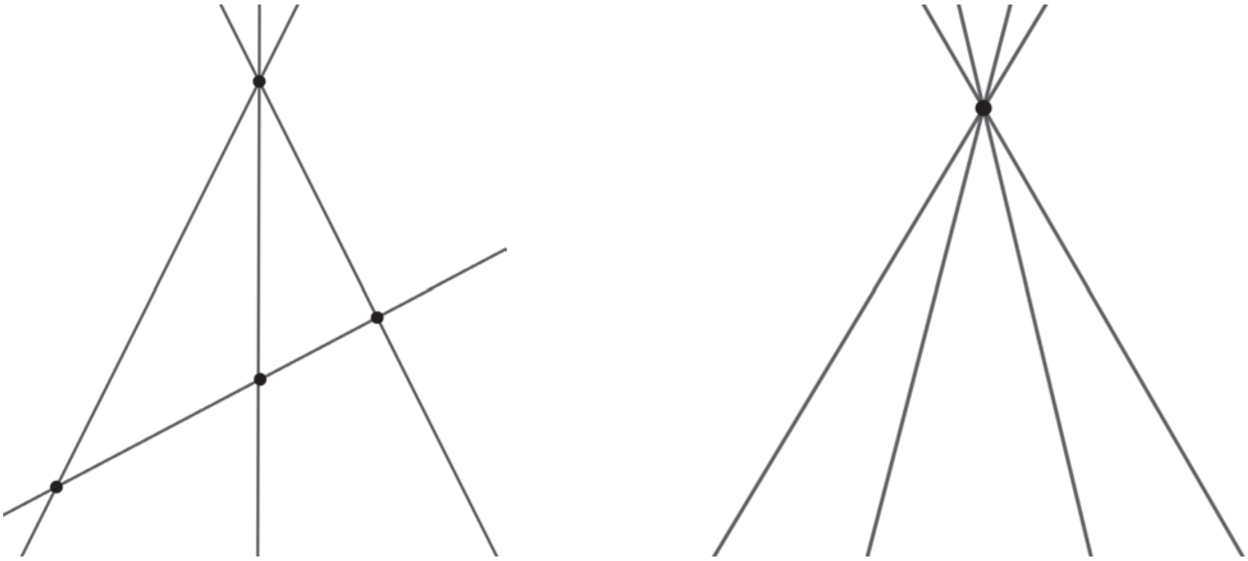

Figure 4 The two kinds of cliques of size 4.

(both the straight and the curved lines in the picture). Each line contains exactly three points on it, and through each point there are exactly four lines. Moreover, these points and lines satisfy the following properties:

1. Through any two points there is a unique line.
2. For any line l and a point P outside l, there is a unique line l' through P that is completely disjoint from the l.

Both of these properties are inspired from the incidence axioms of the real Euclidean plane $\mathbb{R}^2$, and this finite geometrical structure is an example of an affine plane.

In their affine formulation, Hermitian curves can be defined as the set of all points $\{(x,y) \in \mathbb{F}_{q^2}: x^{q+1} = y^q + y\}$, where q is a prime power and $\mathbb{F}_{q^2}$ is the finite field of size q^2. In a paper on unital groups from 1972, O'Nan showed [12] that it is impossible to find four lines in general position (no three concurrent) whose pairwise intersection points all lie on the Hermitian curve (see Figure 3). While this result has been well-known in the finite geometry community for years, nobody imagined that it could be useful for Ramsey theory! In fact, this property of Hermitian curves, along with the fact that they have about q^3 points, turn out to be the key properties that we need to lower bound the Ramsey number $R(4,t)$.

Mattheus and Verstraete studied the graph G defined on the lines in the plane with two lines adjacent in G if their intersection point does lie on the Hermitian curve. The result of O'Nan shows that the only cliques of size 4 in G are three lines passing through a fixed point of the curve along with a fourth line meeting all of these lines inside the curve, or four lines passing through a fixed point (see Figure 4). By cleverly tweaking the graph G through a random process, Mattheus and Verstraete managed to obtain a new graph G' on approximately q^4 vertices (sine that is roughly the total number of lines in the finite plane) by deleting some edges such that G' has no cliques of size 4 and only a few independent sets of size slightly larger than q. Subsequently, they took a random subgraph of G' on about q^3 vertices, by deleting each vertex with a certain probability, to destroy all independent sets of size $t \approx q \log^2 q$. They concluded that with positive probability this gives rise to a graph required for their lower bound on $R(4,t)$.

This breakthrough has given a *fresh perspective to Ramsey theory*. For instance, one can now explore other algebraic curves or finite-geometric objects to asymptotically determine the off-diagonal Ramsey numbers $R(s,t)$ for any fixed $s \geq 5$. The work of Mattheus and Verstraete have shown us that perhaps the key to solving such Ramsey problems lies in finite geometry.

References

1 M. Ajtai, J. Komlos and E. Szemeredi, A note on Ramsey numbers, *J. Combin. Theory Ser.* A 29 (1980), 354–360.

2 N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley, 2015, fourth edition.

3 S. Ball, *Finite Geometry and Combinatorial Applications*, Cambridge University Press, 2015.

4 A. Bishnoi, F. Ihringer and V. Pepe, A construction for clique-free pseudorandom graphs, *Combinatorica* 40 (2020), 307–314.

5 T. Bohman and P. Keevash, The early evolution of the H-free process, *Invent. Math.* 181 (2010), 291–336.

6 M. Campos, S. Griths, R. Morris and J. Sahasrabudhe, An exponential improvement for diagonal Ramsey (2023), arXiv:2303.09521.

7 P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* 53 (1947), 292–294.

8 P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* 2 (1935), 463–470.

9 A. Frieze and M. Karonski, *Introduction to Random Graphs*, Cambridge University Press, 2016.

10 S. Mattheus and J. Verstraete, The asymptotics of $r(4,t)$ (2023), arXiv:2306.04007.

11 D. Mubayi and J. Verstraete, A note on pseudorandom Ramsey graphs, *J. Eur. Math. Soc.* (2023).

12 M.E. O'Nan, Automorphisms of unitary block designs, *J. Algebra* 20 (1972), 495–511

13 G. Fiz Pontiveros, S. Griths and R. Morris, *The Triangle-free Process and the Ramsey Number* $R(3,k)$, Mem. Amer. Math. Soc., Vol. 263, AMS, 2020, 125 pp.