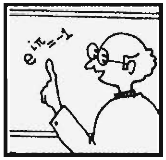

In this section of* Resonance, *we invite readers to pose questions likely to be raised in a classroom situation. We may suggest strategies for dealing with them, or invite responses, or both. "Classroom" is equally a forum for raising broader issues and sharing personal experiences and viewpoints on matters related to teaching and learning science.

Shailesh Shirali
Rishi Valley School
Chittoor District
Rishi Valley 517 352, India.

On Two Problems of Paul Erdös in Combinatorics

That Paul Erdös was a problem poser *par excellence* is well-known. He loved throwing out challenges, and he loved taking his fellow mathematicians along with him on his mathematical journeys. He was not only one of the most prolific mathematicians of all time but also one of the most social. A former student of his wrote(see József Pelikán in Suggested Reading): "The world has just lost one of the greatest, most prolific, most original and most loveworthy mathematicians of all time." It is said that he had the curious ability of knowing the right level of problem with which to challenge those around him. His interests in mathematics encompassed an extremely wide spectrum: number theory, set theory, graph theory, the theory of designs, combinatorial geometry, combinatorial set theory, probability theory, real analysis, ; even elementary geometry! Here are two geometric inequalities stated by him as problems in *The American Mathematical Monthly*: (a) Let P be a point within a triangle ABC, its distances from the sides of the triangle being x, y, z, and its distances from the vertices being u, v, w; then $u+v+w \geq 2(x+y+z)$. (This has long since become a classic; it is now known as the *Erdös-Mordell inequality*.) (b) Let P be a point within a triangle ABC with unit circumradius; then $PA \quad PB \quad PC < 32/27$. (For a pretty solution to (b), see B J Venkatachala, 1996 in Suggested Reading.)

Paul Erdös did more than pose problems; he also solved them, and in the bargain he sometimes created new areas of mathematics. For instance, he pioneered the use of probabilistic methods in combinatorics and number theory, and he was largely responsible for making Ramsey theory a mainstream topic in combinatorics. For more on his life and work, the reader should refer to Erdös, Pelikán and Babai and NBHM Newsletter in Suggested Reading.

In this article we discuss two of Erdös' problems, both fairly elementary in nature. We illustrate, in the first problem, how ideas from probability theory are sometimes used to establish results in combinatorics. In the particular example discussed the use of probabilistic language is by no means essential. However the object here is merely to illustrate the method. The problem, which was first considered by Erdös and Szekeres, is the following.

> *Show that there exists a graph G on* $(n/6)\ 2^{n/2}$ *vertices containing neither a* K_n *nor a* $\overline{K_n}$.

A few definitions are in order. A *graph*[1] G is a set $V = V(G)$ of *vertices* together with a set $E = E(G)$ of *edges*, where E is any subset of the set of all 2-element subsets of V It helps to keep a geometric picture in mind, so we may think of the elements of V as points and the elements of E as line segments or segments of curves joining the points; if a and b are vertices, then we can think of the edge $\{a, b\}$ as 'connecting' a and b.

In a *bipartite graph* the vertex set V can be partitioned as $V = V_1 \cup V_2$, such that all edges of the graph are of the type $\{a, b\}$, with $a \in V_1$ and $b \in V_2$; there are no edges connecting pairs of vertices within V_1 or within V_2. (The notion of bipartite graph is needed in the second problem discussed in this paper. We will also need the notion of a *cycle*. For our purposes, we may think of a cycle as a sequence of distinct edges $e_1, e_2, \ldots, e_r$ along which one can travel outwards from a vertex and eventually return, never passing through the same vertex twice. It is known, and easy to prove, that if a graph on n vertices has no cycles, then it has no more than $n-1$ edges.)

[1] We have only 'undirected graphs' in mind, so this definition suffices. In a directed graph edges are assigned directions. The definition is now more conveniently stated using the language of relations.

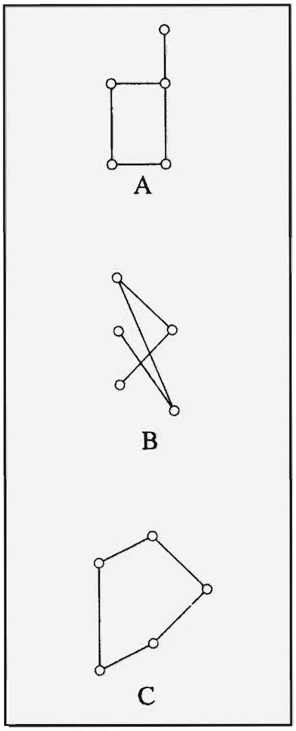

Figure 1. Three graphs.

In *Figure 1* we see depictions of a bipartite graph (B) and a cycle on 5 vertices (C).

Let $G = G(V, E)$ be a graph. For any two vertices $a, b \in V(G)$, if $\{a, b\} \in E(G)$ we say that a, b are *adjacent*; else they are *non-adjacent*. A subset A of V such that any two vertices in A are adjacent to one another is a *clique*; if no two vertices in A are adjacent to one another, then A is a *co-clique*. A clique with n vertices is denoted by K_n, and a co-clique with n vertices is denoted by $\overline{K}_n$.

The size c of the largest clique in G is its *clique number*, and the size $\bar{\mathsf{c}}$ of the largest co-clique in G is its *co-clique number*. If $\mathsf{c} \geq n$, we say that "G contains a K_n", and if $\bar{\mathsf{c}} \geq n$, we say that "G contains a $\overline{K}_n$".

Some introductory words need to be said at this point on the problem considered by Erdös and Szekeres. In 1930, the British logician Frank Plumpton Ramsey proved that in all *sufficiently large* graphs, if c is small then $\bar{\mathsf{c}}$ is large. Formally, given any two positive integers m and n, there is a number R such that any graph on R or more vertices has a clique of size m or a co-clique of size n (i.e., either $\mathsf{c} \geq m$ or $\bar{\mathsf{c}} \geq n$). The least positive integer R for which this statement is true is denoted by $R(m, n)$ and is called a Ramsey number.

The following results are trivial: $R(m, n) = R(n, m)$, $R(1, n) = 1$, $R(2, n) = n$. The result $R(3, 3) = 6$ is widely known and is usually expressed thus in puzzle books: *In any assemblage of 6 persons one can find 3 persons who are all known to one another or who are all unknown to one another.* For the proof we simply focus attention on any one person and then consider the various possibilities.

The Ramsey numbers are notoriously hard to estimate, let alone determine exactly. In 1935, Erdös and G Szekeres established the crude and yet hard-to-beat lower bound $(n/6) \cdot 2^{n/2}$ for the 'diagonal' Ramsey number $R(n, n)$. They also applied Ramsey's theorem to combinatorial geometry and thereby brought Ramsey theory to the mainstream of mathematics.

m\n	3	4	5	6	7	8	9
3	6	9	14	18	23	28	36
4		18	25-27	34-43	47-66	?	?
5			43-52	51-94	76-160	?	?
6				102-169	?	?	?

Table 1.

Before proceeding we reproduce, from V Rödl (see Suggested Reading), a table summarizing current knowledge about some of the smaller Ramsey numbers (thus $25 \leq R(4,5) \leq 27$, and so on).

The question marks seem to stare at us with some disdain!

We now give the proof by Erdös and Szekeres that $R(n,n) \geq (n/6)\ 2^{n/2}$.

Let G denote a graph on N labelled vertices. Suppose that the membership of $E(G)$ is decided by a coin-tossing procedure (with equal chances for heads and tails). Let A, B denote events defined as follows:

$$A : c \geq n, \qquad B : \bar{c} \geq n. \tag{1}$$

We now estimate $\Pr(A)$ and $\Pr(B)$. The number of distinct graphs on N labelled vertices is $2^{\binom{N}{2}}$, and of these the number of graphs in which n specified vertices form a clique is $2^{\binom{N}{2}-\binom{n}{2}}$. Therefore the number of graphs G on N labelled vertices that contain a K_n is at most

$$\binom{N}{n}\ 2^{\binom{N}{2}-\binom{n}{2}}. \tag{2}$$

So we have the following inequality for the probability of A:

$$\Pr(A) \leq \binom{N}{n}\ 2^{-\binom{n}{2}}. \tag{3}$$

By symmetry the same inequality holds for $\Pr(B)$. Therefore we have,

$$\Pr(A \cup B) \leq 2 \cdot \binom{N}{n} \cdot 2^{-\binom{n}{2}}, \Pr(\bar{A} \cap \bar{B}) \geq 1 - 2 \cdot \binom{N}{n} \cdot 2^{-\binom{n}{2}}. \quad (4)$$

This means that if the following relation holds,

$$2 \binom{N}{n} 2^{-\binom{n}{2}} < 1, \quad (5)$$

then there is a non-zero probability that G has neither a K_n nor a $\overline{K_n}$. Equivalently, if (5) holds, then there exists a graph G on N vertices such that both c and $\bar{\text{c}}$ are less than n.

So, for fixed n, we seek an upper bound for N, given that it satisfies (5). Now (5) is equivalent to:

$$N(N-1)(N-2) \quad (N-n+1) < n! \ 2^{\binom{n}{2}-1}. \quad (6)$$

We replace this by a stronger inequality which implies (6). We first replace the quantity on the left by the larger quantity N^n, and then we replace the factor $n!$ on the right by the smaller quantity $(n/3)^n$. (The inequality $n! > (n/3)^n$ is easy to show.) We now obtain the inequalities

$$N^n < \left(\frac{n}{3}\right)^n 2^{\binom{n}{2}-1}, \quad \text{or} \quad N < \frac{n}{3} \ 2^{n/2} \ 2^{-(n+2)/2n}, \quad (7)$$

which imply (6). Since $2^{-(n+2)/2n} > 1/2$ for $n > 2$, a still better inequality is

$$N < \frac{n}{6} \ 2^{n/2}, \quad (8)$$

which implies (7) and therefore (6).So we arrive at the following:*For any positive integer $n > 2$,if $N < (n/6) \ 2^{n/2}$, then there exists a graph G on N vertices that contains neither a K_n nor a $\overline{K_n}$.* This is Erdös' and Szekeres' result. ■

Number crunching may clarify matters somewhat. For $n = 30$, the calculations yield the following: *There exists a graph*

G on 163839 vertices for which both c and $\bar{\mathsf{c}}$ are less than 30. In fact better bounds are possible. We want $N_{\max}$, the largest N for which $\binom{N}{30} < 2^{\binom{30}{2}-1}$, i.e., $\binom{N}{30} < 2^{434}$. Using software packages such as `Derive`, it is possible to compute $N_{\max}$ exactly; it turns out to be 272717. (Note the non-constructive nature of the result; we have no clue as to how such a graph can be constructed!)

The second problem belongs to combinatorial set theory and illustrates Erdös' deft touch. Let the function $\chi(n)$ be defined for positive integers n as follows. For any n-element set S, let each subset of S be assigned a colour. The colouring is said to be 'union-free' if there do not exist distinct subsets A, B and C, all of one colour, such that $A \cup B = C$. Let $\chi(n)$ be the least number of colours needed to accomplish such a colouring of the subsets of S. The problem is to show that $\chi(n)$ lies between $\lceil (n+1)/4 \rceil$ and $\lceil (n+1)/2 \rceil$. (*Notation:* For real numbers z, $\lfloor z \rfloor$ is the greatest integer $\leq z$, and $\lceil z \rceil$ is the least integer $\geq z$; e.g., $\lfloor \pi \rfloor = 3$ and $\lceil \pi \rceil = 4$.)

To show that $\lceil (n+1)/2 \rceil$ colours suffice for a union-free colouring is easy; we use a colour scheme in which the colour of a subset is fixed by its cardinality, as shown in *Table 2.*

It is easy to verify that this does provide a union-free colouring and that the number of colours used is $\lceil (n+1)/2 \rceil$.

For the lower bound it suffices to consider only the 'interval subsets' – subsets of the form $\{i, i+1, i+2, \ldots, j\} =: [i, j]$ where $1 \leq i \leq \lceil n/2 \rceil < j \leq n$. We now show that for a union-free colouring of the interval subsets, the number of colours used cannot be less than $\lceil (n+1)/4 \rceil$.

Colour	Cardinality of Subset
1	0, 1, 3, 7, 15, 31, ...
2	2, 5, 11, 23, 47, ...
3	4, 9, 19, 39, 79, ...
$i+1$	$2i$, $4i+1$, $8i+3$, ...

Table 2.

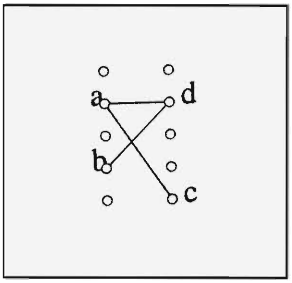

Figure 2. A subgraph.

Represent the interval subsets on a bipartite graph with vertex set $\{1, 2, 3, \quad n\}$, with $[i, j]$ represented by the edge $\{i, j\}$. Observe that the graph has $\lfloor n^2/4 \rfloor$ edges. If any colour class has a cycle with reference to this graph, then there must exist a subgraph of the type shown in *Figure 2*, with $a < b < c < d$. (If the set of vertices in the cycle is C, let a be the vertex in C with least index, d the vertex in C with largest index such that the edge $\{a, d\}$ belongs to the cycle, and b and c the other vertices in C such that the edges $\{b, d\}$ and $\{a, c\}$ belong to the cycle.)

But then $[a, d] = [a, c] \cup [b, d]$ and the colouring is not union-free. So each colour class must be cycle-free and hence can contain at most $n - 1$ edges. Therefore

$$\chi(n) \geq \frac{\lfloor n^2/4 \rfloor}{n-1} \geq \left\lceil \frac{n+1}{4} \right\rceil$$

The proof is by Erdös and Shelah.

Acknowledgement

I thank the referee for making many valuable and helpful suggestions.

Suggested Reading

- **Paul Erdös. *The Art of Counting*. MIT Press, 1973.**
- **József Pelikan and Paul Erdös. 1913–1996. *Journal of the World Federation of National Mathematics Competitions* . Vol. 9. No.2, 1996.**
- **Babai L. In and Out of Hungary: Paul Erdös, His Friends and Times. *Combinatorics: Paul Erdös is Eighty*.Bolyai Society Mathematical Studies, Budapest, Vol.2. pp.7–95,1996.**
- **Obituary of Paul Erdös. *NBHM Newsletter* .Vol. 6, No. 2,Sept. pp.17–18.1996.**
- **B J Venkatachala.Solution to problem 10282. *Am. Math. Monthly*. Vol.102. No.5. p.468.**
- **Rödl V. Some Developments in Ramsey Theory. *Proceedings of International Congress of Mathematicians*. Kyoto, 1990.**