Second stage of Israeli students competition, 2011.

1. In each vertex of a connected simple graph a number is written. The following action is repeated infinitely many times: all numbers are replaced simultaneously by the average of their neighbors. Consider the sequence of numbers which appear at a specific vertex of the graph. Assume that one of those sequences does not converge. Prove that the graph is bipartite (which means that its vertices can be painted in black and white so that neighbors are always of opposite colors).

Solution. Let N be a number of vertexes. Choose some numbering for the set of vertexes. Then any set of numbers written by the vertexes corresponds to a vector in $\mathbb{R}^N$. "Averaging" is a linear operator. It consists of subsequent application of two linear operators: first summing of neighbors, then dividing by degrees. In terms of matrixes we get: A = DG, where A is the matrix of averaging, D is the diagonal matrix of reciprocal degrees, and G is the matrix of zeroes and ones in which one appears in square (i, j) iff vertexes i and j are connected in the graph, a.k.a. "the graph matrix". We shall also consider matrix R which is "a square root" of D: it is also a diagonal matrix, but the numbers on diagonal are square roots of respective numbers of D. Obviously, $R^2 = D$, and R is an invertible matrix.
Applying the averaging procedure k times is multiplying by $A^k = DGDG\ldots DG$.
This matrix is conjugated to $RGDG\ldots GR = B^k$, where B = RGR, so A^k and B^k have the same eigenvalues with the same geometric and algebraic multiplicities. It is easier to analyze eigenvalues of B^k, since both B and B^k are symmetric matrixes. Therefore, the eigenvalues are real, and geometric multiplicities are equal to algebraic multiplicities, so B has diagonal form with real numbers $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_N$ on the diagonal, and B^k also has a diagonal form with $\lambda_1^k, \lambda_2^k, \lambda_3^k, ..., \lambda_N^k$ on the diagonal, with the same eigenbasis. Notice that since eigenvalues are real, the eigenbasis can also be chosen to be real. So, A^k also has a diagonal form with $\lambda_1^k, \lambda_2^k, \lambda_3^k, ..., \lambda_N^k$ on the diagonal, and real eigenbasis.
Choose a real eigenbasis for A: it will consist of vectors $a_1, a_2, \ldots, a_N$. It is also an eigenbasis of A^k. Now we shall show that $|\lambda_j| \le 1$ for every j. Otherwise, if we start with a_j, then after k averagings we get $A^k a_j = \lambda_j^k a_j$, and absolute values at all nonzero coordinates keep growing with every step, but that is impossible: maximal absolute value cannot grow during averaging.

Therefore, $-1 \le \lambda_1, \lambda_2, \lambda_3, ..., \lambda_N \le 1$. Then we ask whether -1 is an eigenvalue of A.
Consider both possible answers:

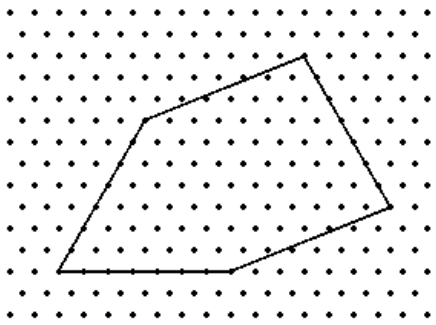
No Then we can easily prove that for any initial vector v, the vectors $A^k v$ will converge. Indeed, if $v = \sum_j v_j a_j$, then $A^k v = \sum_j \left(\lambda_j^k v_j\right) a_j$. Coefficient which correspond to eigenvalues equal to 1 remain the same, all other coefficients converge to zero. Therefore $A^k v$ converges, and all its coordinates converge.

Yes Then consider a nonzero eigenvector. Each number is minus average of its neighbors. Assume that a number m has maximal absolute value. WLOG it is positive (otherwise, multiply all numbers by -1). Then all its neighbors are at least $-m$, and to get m as minus their average they have to be exactly $-m$. F or the same reason all neighbors of $-m$ are m. Since graph is connected, then by induction with this argument we conclude that all numbers in a graph are either m or $-m$, and neighbors are always of opposite signs, so there is "a chess coloring" for the graph.

To conclude: if averagings don't converge, then we have are in the "yes" case (there is a -1 eigenvalue), then there is a chess coloring.

2. Is it possible to find a planar strictly convex equilateral pentagon, all vertices of which are in $\mathbb{Z}^3$ (integer three-dimensional points)?
Remark. A polygon is called equilateral if all its sides are of the same length. It is possible for a polygon to be equilateral but not regular.

Solution. Consider a plane $x + y + z = 0$. Within this plane, we have a triangular lattice (spanned by integer linear combinations of the vectors formed by sides of an equilateral triangle).
On that lattice, we build an equilateral pentagon (see the picture). It is based on the fact that triangle with sides 3, 5, 7 has an angle of 120º (that happens because $3^2 + 3 \cdot 5 + 5^2 = 7^2$).



3. There is an urn with 5 balls: 2 blue and 3 white. Every minute, a random ball is chosen from the urn and returned with another ball of the same color. What is the

limit of the probability that more than a half of the balls are blue, as the time goes to infinity?

Solution. Denote that at any time *i*, there are m_i blue balls and n_i white balls inside the urn. We will compute the probability that at time *t* there are exactly m_t, n_t blue and white balls in the urn, given that there were m_0 and n_0 time 0 (in our case, m_0=2 and n_0=3). Denote also m' = m_t-m_0, n' = n_t-n_0.
For this to happen, we must draw exactly m' blue balls and n' white balls in the first *t* steps. Denote by *I* the set of steps up to *t* on which a blue ball was drawn. It is clear that *I* may be any subset of {0,…,t-1} of size m', and that for each such set, the probability that it is represents the sequence of draws is exactly

$P_I = \prod_{i\in I} \frac{m_i}{m_i+n_i} \prod_{j\notin j} \frac{m_j}{m_j+n_j}$, as the probability of drawing a blue ball at a single step *i* is $\frac{m_i}{m_i+n_i}$, and so on.

We observe that the numerator in the left product comprise all integer numbers sequentially from m_0 to m_t-1, as the number of blue balls increases by exactly one between two blue draws, and similarly the numerators on the right product are all integers from n_0 to n_t-1. Furthermore, all the denominators together are simply the integers m_0+n_0 to m_t+n_t-1, as m_i+n_i = m_0+n_0+i. Thus after rearrangement, we obtain

$$P_I = \frac{(m_t-1)!}{(m_0-1)!}\cdot\frac{(n_t-1)!}{(n_0-1)!}\Big/\frac{(m_t+n_t-1)!}{(m_0+n_0-1)!}$$

In particular, this probability is independent of *I*, and thus to obtain the total probability of reaching m_t and n_t we need only multiply by the number of choices for *I*, which is $\binom{m'+n'}{m'} = \frac{(m'+n')!}{m'!n'!}$. With more rearrangement, we get

$$P_{m_t,n_t} = \frac{(m'+n')!}{m'!n'!}P_I = \frac{(m_t-1)!}{m'!(m_0-1)!}\cdot\frac{(n_t-1)!}{n'!(n_0-1)!}\cdot\frac{(m'+n')!(m_0+n_0-1)!}{(m_t+n_t-1)!} = \frac{\binom{m_t-1}{m_0-1}\binom{n_t-1}{n_0-1}}{\binom{m_t+n_t-1}{m_0+n_0-1}}$$

This probability can now be considered in a new way: consider an ordered set of size S = m_t+n_t-1, for example {1, 2, …, m_t+n_t-1}. From this set, we choose uniformly a random subset of size m_0+n_0-1. We will denote the subset's elements by $\{x_1, x_2, ..., x_{m_0+n_0-1}\}$, and assume that the sequence of *x*'s is increasing. Consider the element x_{m_0}. It can be immediately computed that the probability of $x_{m_0} = m_t$ is exactly the same expression as (*), i.e. the same as P_{m_t,n_t}. We are interested in the probability that $\frac{m_t}{m_t+n_t} < \frac{1}{2}$, which is therefore equal to the probability that $\frac{x_{m_0}}{m_t+n_t} < \frac{1}{2}$.

As *t* goes to infinity, the limit of the last distribution can be easily computed: it is similar to choosing m_0+n_0-1 points uniformly distributed on the interval [0,1], sorting them as $\{x_1, x_2, ..., x_{m_0+n_0-1}\}$, and then considering only x_{m_0}. It is easy to see that the probability density of the random variable x_{m_0} is $\frac{x^{m_0-1}(1-x)^{n_0-1}}{\mathrm{B}(m_0,n_0)}$, where $B(m_0,n_0) = \frac{(m_0-1)!(n_0-1)!}{(m_0+n_0-1)!}$ is the Beta function.

Finally, we are interested in the probability that $x_{m_0} < \frac{1}{2}$, which is simply:

$$\frac{1}{\mathrm{B}(m_0,n_0)}\int_0^{1/2} x^{m_0-1}(1-x)^{n_0-1}\,dx = \frac{1}{\mathrm{B}(2,3)}\int_0^{1/2} x(1-x)^2\,dx = \frac{4!}{1!2!}\int_0^{1/2}\left(x-2x^2+x^3\right)dx =$$

$$= 12\left(\frac{1}{2}\cdot\frac{1}{2^2} - 2\cdot\frac{1}{3}\cdot\frac{1}{2^3} + \frac{1}{4}\cdot\frac{1}{2^4}\right) = \boxed{\frac{11}{16}}$$

4. We have a hyperbola and two distinct points A and B on it. For any point X on the same hyperbola, we define 3 numbers:
α = the distance from X to the straight line which is tangent to the hyperbola at A.
β = the distance from X to the straight line which is tangent to the hyperbola at B.
γ = the distance from X to the straight line AB
Prove that $\frac{\alpha\beta}{\gamma^2}$ doesn't depend on the choice of X.

Solution. The solution works for any conic (ellipse, parabola, or hyperbola) so from now on we shall talk of conics. We shall denote the tangents to the hyperbola A and B by t_a and t_b respectively. The distance from point (x, y) to t_a can be written as $|l_a(x, y)|$ where $l_a(x, y)$ is a linear function: $kx + my + n$. Linear function l_b is chosen similarly for the line t_b. The third linear function l is such that $|l(x, y)|$ = distance from the line AB to the point (x, y).

Equations of curves of order at most two form a six-dimensional linear space

$$Q = \{\ q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0\ \}$$

Inside that space, the equations of curves that pass through A form a sub-space Q_1; it is strictly smaller (since some curves of order 2 don't pass through A), and is defined by one linear condition – substituting coordinates of A to $q(x, y)$ specifies one linear condition on the coefficients; therefore, equations of order 2 of curves containing A form a five-dimensional space. For similar reasons, since some curves of order 2 contain A but not B, equations of second degree satisfied by A and B form a linear space Q_2 of dimension four.

Inside Q_2, consider such equations, that when we reduce them to t_a we get multiple root at A. In other words, if K is a non-zero vector parallel to t_a, we substitute the coordinates of A + sK to the polynomial q(x, y), we get a polynomial $q_a(s)$ of second degree in t; for all polynomials in Q_2, this polynomial has a root at zero (since the curve goes through A); and we define a subspace Q_3 in Q_2 by the condition that $q_a(t)$ should have a multiple root at zero; thus it should be of form $q_a(s) = hs^2$.

Finally, consider the subspace Q_4 in Q_3 of curves which, when reduced to t_b, have double root at B (similarly to the previous condition, but with B instead of A).

It is easy to find examples of curves of degree two in Q_2 but not in Q_3 (for example the product of two linear equations, one of line AB and another of a line parallel to AB, so it has two distinct roots on the line t_a) so dim $Q_3 = 3$. It is also easy to find an example of something in Q_3 but not in Q_4: for instance $l(x, y) \cdot l_a(x, y)$.

Therefore dim $Q_4 = 2$.

Now we shall show three examples of equations in Q_4.

The first example is the equation of the original conic. The other two obvious examples of curves are $l_b(x, y) \cdot l_a(x, y) = 0$ and $(l(x, y))^2 = 0$.

But dim $Q_4 = 2$, and all three examples define different curves, so neither two of them are linearly dependent. Therefore the equation of the conic can be expressed as a linear combination of the other two. That means it can be written as:

$\lambda \cdot l_a(x, y) \cdot l_b(x, y) + \mu \cdot (l(x, y))^2 = 0$

Where λ, μ are some fixed real numbers. Therefore, for any X on the conic we get $\pm\lambda\alpha\beta \pm \mu\gamma^2 = 0$. Since X doesn't coincide with A or B we are allowed to divide:

$$\frac{\alpha\beta}{\gamma^2} = \left|\frac{\mu}{\lambda}\right|$$

The right hand side obviously doesn't depend on the choice of X.

5. Compute $\int_0^1 \frac{\ln(1+x)}{1+x^2}dx$.

Solution. Perform change of variables: $\arctan x = y$, then $\frac{dx}{1+x^2} = dy$

$$\int_0^1 \frac{\ln(1+x)}{1+x^2}dx = \int_0^{\pi/4} \ln(1+\tan y)dy = \int_0^{\pi/4} \ln\left(\frac{\sin y+\cos y}{\cos}\right)dy =$$

$$= \int_0^{\pi/4} \ln(\sin y+\cos y)dy - \int_0^{\pi/4} \ln(\cos y)dy$$

Recall that $\sin y+\cos y = \sqrt{2}\left(\cos\frac{\pi}{4}\sin y+\sin\frac{\pi}{4}\cos y\right) = \sqrt{2}\sin\left(\frac{\pi}{4}+y\right)$. Therefore

$$\int_0^1 \frac{\ln(1+x)}{1+x^2}dx = \int_0^{\pi/4} \ln(\sin y+\cos y)dy - \int_0^{\pi/4} \ln(\cos y)dy =$$

$$= \int_0^{\pi/4} \ln\sqrt{2}dy + \int_0^{\pi/4} \ln\left(\sin\left(\frac{\pi}{4}+y\right)\right)dy - \int_0^{\pi/4} \ln(\cos y)dy =$$

(here $z = \frac{\pi}{2} - y$)

$$= \int_0^{\pi/4} \ln\sqrt{2}dy + \int_{\pi/4}^{\pi/2} \ln(\sin(y))dy - \int_0^{\pi/4} \ln(\cos y)dy =$$

$$= \int_0^{\pi/4} \ln\sqrt{2}dy + \int_0^{\pi/4} \ln(\cos(z))dz - \int_0^{\pi/4} \ln(\cos y)dy = \frac{\pi}{4}\ln\sqrt{2}$$