Targil 11.

This targil is about identities.

1. Prove that for 2×2 matrixes $\det(A)=\frac{(\mathrm{tr}(A))^2-\mathrm{tr}(A^2)}{2}$ and, more generally, for $n\times n$ matrixes $\det(A)$ can be expressed as a polynomial in $\mathrm{tr}(A),\mathrm{tr}(A^2),...,\mathrm{tr}(A^n)$.

Solution. Assume we brought our matrix A to diagonal, or at least triangular form. The guys on the diagonal $\alpha_1,...,\alpha_n$ are the eigenvalues. The trace, the determinant and all other coefficients of the characteristic polynomial can be easily expressed as elementary symmetric polynomials of $\alpha_1,...,\alpha_n$.
On the diagonal of A^k we have $\alpha_1^k,...,\alpha_n^k$.

So, for instance for 2×2 we get:
$\det A=\alpha_1\alpha_2$
$\mathrm{tr}(A)=\alpha_1+\alpha_2$
$\mathrm{tr}(A^2)=\alpha_1^2+\alpha_2^2$
Hence $\frac{(\mathrm{tr}(A))^2-\mathrm{tr}(A^2)}{2}=\frac{(\alpha_1+\alpha_2)^2-(\alpha_1^2+\alpha_2^2)}{2}=\alpha_1\alpha_2$.

The proof for general case follows directly from the known theorem: if we have any symmetric polynomial of x_1, x_2, … , x_n we can (uniquely) write it as a polynomial in $\sigma_1,\sigma_2,...,\sigma_n$, and as a polynomial in $s_1,s_2,...,s_n$. Here σ_k is a sum of all products of k different x's (a. k. a. elementary symmetric polynomials), and $s_k=\sum_i x_i^k$. Actually, what we need is to show that one of elementary symmetric polynomial is a polynomial in sums of powers.
Indeed, in the story about matrixes, if x_1, x_2, … , x_n are the eigenvalues of A, then is $s_k=\mathrm{tr}(A^k)$, and σ_k are, up to a sign, coefficients of the characteristic polynomial of A.

There are several proves of this claim. My favorite (and the most fitting for this targil) is based on Newton's identities. From Newton's identities we see immediately that σ_k is a polynomial in $\sigma_1, \sigma_2, ..., \sigma_{k-1}, s_1, ..., s_k$, so the claim follows by induction. Newton identities look as follows.

$\sigma_1 - s_1 = 0$

$2\sigma_2 - \sigma_1 s_1 + s_2 = 0$

$3\sigma_3 - \sigma_2 s_1 + \sigma_1 s_2 + s_3 = 0$

...

$n\sigma_n - \sigma_{n-1} s_1 + ... \mp \sigma_1 s_{n-1} \pm s_n = 0$

$\sigma_n s_1 - \sigma_{n-1} s_2 + ... \mp \sigma_1 s_n \pm s_{n+1} = 0$

...

Their proof of the identities is an exercise to the reader – it is nice, and not complicated, so nobody would like it if I would write it downs right away.

2.** Prove that for arbitrary $2n$ matrixes $A_1, A_2, \ldots, A_{2n}$ of size $n \times n$,

$$\sum_{\sigma \in S_{2n}} (-1)^{\mathrm{sgn}(\sigma)} A_{\sigma(1)} A_{\sigma(2)} A_{\sigma(3)} \cdot ... \cdot A_{\sigma(2n)} = 0 .$$

(Here S_{2n} is a group of permutations of $2n$ numbers).

Remark. This is called Amitsur-Levitzky identity; it is named after two mathematicians from the Hebrew university of Jerusalem who found it.

Solution. We shall show the solution based on the trick of ***linearization***.

That is a procedure, given a polynomial, to obtain linear combinations of its values can give us a linear polynomial in several matrixes.

Simple example: if a polynomial is $p(x) = x^2$ we can take

$p(\mathrm{A} + \mathrm{B}) - p(\mathrm{A}) - p(\mathrm{B}) = \mathrm{AB} + \mathrm{BA}$, which is a linear polynomial in A and B.

If $p(x)$ is a polynomial of degree n, we consider:

$$L_p(\mathrm{A}_1, \mathrm{A}_2, ..., \mathrm{A}_n) = \sum_{S \subseteq P(I_n)} (-1)^{n-|S|} p\left(\sum_{i \in S} \mathrm{A}_i \right)$$

Here $I_n = \{1, 2, \ldots, \mathrm{n}\}$ and $P(I_n)$ is a set of its subsets, so we sum over all subsets of indexes. Consider the monomial $\mathrm{A}_{i_1} \mathrm{A}_{i_2} \mathrm{A}_{i_3} ... \mathrm{A}_{i_k}$ in this L_p.

If there are less than n different indexes among $i_1, \ldots, i_k$, then there's an index j which is different from $i_1, \ldots, i_k$. Subsets that produce this monomial can be divided into pairs – two subsets of the same pair differ only by containing j / not containing j. The monomial $A_{i_1}A_{i_2}A_{i_3}...A_{i_k}$ will be produced by the subsets of the same pair with the same coefficient, but of opposite signs (because of the sets in the same pair cardinalities are of different parity), so they cancel out.
Therefore, all terms of degree less than n will die, and from x^n only
$\sum_{\sigma \subseteq S_n} A_{\sigma(1)}A_{\sigma(2)}A_{\sigma(3)}...A_{\sigma(n)}$ will remain.

The same thing can be done for more general polynomials, for instance polynomials with matrix coefficients or polynomial in matrix coefficients, but the idea is the same.

Now, consider Hamilton-Cayley identity: $A^n - (\operatorname{tr} A)A^{n-1} + ... \pm \det A = 0$.
After that, in the same way as in problem 1, we shall rewrite all the coefficients as polynomials in $\operatorname{tr}(A), \operatorname{tr}(A^2), ..., \operatorname{tr}(A^n)$. We shall get and identity of equality to zero of polynomial in $A, \operatorname{tr}(A), \operatorname{tr}(A^2), ..., \operatorname{tr}(A^n)$. Now we shall linearize it.
We shall get an expression in n matrixes: $X_1, X_2, \ldots, X_n$.
A^n will turn into $X_1 \ldots X_n$ and all other terms will contain traces of products of X's.
Then we shall substitute $X_1 = A_1A_2$, $X_2 = A_3A_4$, … , $X_n = A_{2n-1}A_{2n}$.
If we take the sum of linearized identities with permuted $A_1, \ldots, A_{2n}$ multiplied by sign of permutation. The leading term A^n will turn into Amitzur-Levitzky expression; all other terms will vanish.
Indeed, all other terms contain traces of even products of A_k's. Rotating every even product inside trace (for instance replacing tr(*KLMN*) by tr(LMNK)) keeps the value unchanged; however, an even cycle is always an odd permutation, so this terms will come with opposite signs, and thus cancel out. QED.

Remark. It is obvious that in Amitzur-Levitzky we can replace $2n$ by any greater number; however we cannot replace it by a smaller number, so this inequality is tight. Indeed, imagine we write this expression for $2n - 1$ matrixes; now take A_{2k-1} to be a matrix that has 1 at place (k,k) and zeroes elsewhere, and A_{2k} has 1 at place

(k+1, k) and zeroes elsewhere. Out of products of A_1, … , A_{2n-1} only one will be a nonzero matrix, therefore Amitzur-Levitzky expression with $2n - 1$ matrixes can be nonzero.

3. Prove: $\sum_{i=1}^{n} \frac{x_i^m}{\prod_{j\neq i}(x_i - x_j)} = \begin{cases} 0, \text{ for } m = 0,1,...,n-2 \\ 1, \text{ for } m = n-1 \end{cases}$.

Remark. These are called Jacobi identities.

First proof. Let us write the following matrix: first $n - 1$ rows are the same as in Vandermonde (first row consists of ones, second row is x_1, x_2, … , x_n, third row is x_1^2, x_2^2, … , x_n^2 etc.) and the last row is x_1^m, x_2^m, … , x_n^m.
Then if $m < n - 1$ this matrix is degenerate, because two rows coincide; and if m is $n - 1$ this matrix is the Vandermonde, and its determinant is $\prod_{i<j}(x_j - x_i)$.

Now compute the determinant by expansion along the last line. All minors are Vandermondes of order $n - 1$. Divide it by the big Vandermonde, and you get QED.

Second solution. Integrate $\int \frac{z^m dz}{\prod_{i\leq n}(z - x_i)}$ along a very big circle in complex plane.

If $m < n - 1$, integral will be quite close to zero on the large circle; if m = n – 1, the integral will be quite close to $\int \frac{dz}{z} = 2\pi i$.

Now compute the same integral by counting residues; poles are at x_i, and residues are precisely the summands in the LHS of the identity.
(Details are easy exercise. Really.)

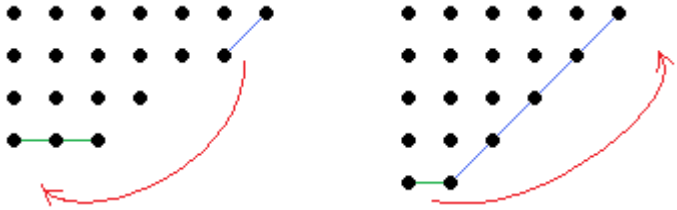
4.* We are given an orthogonal matrix $n\times n$; let d_1 be the determinant of its $k\times k$ upper left corner, and d_2 be the determinant of its of its $(n - k)\times(n - k)$ right bottom corner. Prove that $|d_1| = |d_2|$.

Solution. The original orthogonal matrix will be called A. When we replace its last $n - k$ columns by the corresponding columns of unit matrix, we get matrix B.
Obviously $\det B = d_1$.
Consider matrix $A^{\mathrm{T}}B$. It is easy to see that its first k columns are equal to corresponding columns of the unit matrix, and it last $n - k$ columns are equal to corresponding columns of A^{T}. Therefore its determinant is $\det A^{\mathrm{T}}B = d_2$.
Hence $d_2 = \det A^{\mathrm{T}}B = \det A^{\mathrm{T}} \det B = (\det A)\cdot d_1$.
But A is orthogonal, hence $\det A = \pm 1$, hence $d_2 = \pm d_1$ QED.

5.* Prove: $(1-x)(1-x^2)(1-x^3)\cdot\ldots = 1+\sum_{k=1}^{\infty}(-1)^k\left(x^{\frac{3k^2-k}{2}}+x^{\frac{3k^2+k}{2}}\right)$

Remark. This is Euler's identity. Which in this case is hardly an identifier – too many things are called Euler's formula or Euler's identity ☺.

Solution. We open brackets on the left hand side. Monomial $\pm x^n$ appears, each time we have a partition of n into sum of different natural numbers: $n = k_1 + \ldots + k_m$, where $k_1 < \ldots < k_m$. When m is even, we get this monomial with sign plus, when m is odd, we get it with sign minus. The claim is that almost everything cancels out; that is, we can almost build a bijective correspondence between partitions into odd number of summands and partitions into even number of summands.
To describe the correspondence, we shall represent the partitions by Ferre diagrams. Each number is a row of points, in our case each row will be shorter than the previous (there are also equivalent Young diagrams, that are almost the same, but Young diagrams consist of squares while Ferret diagrams consist of points).



We shall denote the length of the last (and shortest) row by p, and by q length of right diagonal, which consist of right point of the upper row, and of all points which can be reached from that point by a single move of chess bishop (ץר).

So, if $p \le q$, the lower line is distributed between first p rows, one point to each row. But if $p > q$, the right diagonal is taken out and forms a new row below.

It is easy to see that this operation is inverse to itself, and it changes the parity of number of rows; however, in two cases this operation is not well defined.

In the case $p > q$, we usually can put the right diagonal under the right row, because it is smaller, unless $p = q + 1$, and by taking out the right diagonal we reduce the last line. So here a problem can happen when the last row intersects with right diagonal, and they are almost the same size.

In the case when we try to do the opposite, the last row is distributed between the first and largest rows, and there's sufficient number of those because $p \le q$, unless one of these rows is decomposed in the process, and $p = q$.

So, there are precisely two types of diagrams for which our correspondence is not defined: in both cases we have trapezoid (the right diagonal intersects the lowest row), and p equals either q or $q + 1$. The number of points in these cases will be either $\frac{3q^2-q}{2}$ or $\frac{3q^2+q}{2}$. Therefore, precisely for these numbers there is precisely one partition into sum of q numbers which doesn't cancel out.
So, when we open brackets in Euler's identity and cancel things, we shall get zeroes usually, and only for these powers we shall have coefficient of plus or minus 1, where signs are given by the parity of q.

6.** Prove $p\sum_{k=0}^{n}\binom{n}{k}(p-kq)^{k-1}(r+kq)^{n-k}=(p+r)^n$.

Remark. This is called Abel identity; I first heard of this type of things on 2009's SEEMOUS; one of our students on that SEEMOUS (Lior Yanovski) rediscovered a particular case of this identity by combinatorial counting of some trees.
Solution. We shall use two ideas, both of each might be useful olympiad tips.
First idea – if you have some ***natural number*** in your problem (it can be recognized because it might be denoted by n, it might be dimension of the linear space, number of the points, etc.) ***then induction might solve the problem***.
Second idea – use derivative.

So, derive both sides of this identity by r. We get

$$p\sum_{k=0}^{n}\binom{n}{k}(p-kq)^{k-1}(n-k)(r+kq)^{n-k-1}=n(p+r)^{n-1}$$

$$p\sum_{k=0}^{n-1}\frac{n!}{k!(n-k)!}(p-kq)^{k-1}(n-k)(r+kq)^{n-k-1}=n(p+r)^{n-1}$$

$$p\sum_{k=0}^{n-1}\frac{(n-1)!}{k!(n-1-k)!}(p-kq)^{k-1}(r+kq)^{n-k-1}=(p+r)^{n-1}$$

$$p\sum_{k=0}^{n-1}\binom{n-1}{k}(p-kq)^{k-1}(r+kq)^{n-1-k}=(p+r)^{n-1}$$

So, derivative is of Abel's identity of degree n is Abel's identity of degree $n-1$. To make a proof of it, it remains to check Abel's identity for one specific value of r, and to prove it for $n = 0$ (base of induction).

So, let us finish the step of induction first. Take $r = -p$. Then it becomes

$$p\sum_{k=0}^{n}\binom{n}{k}(p-kq)^{k-1}(-p+kq)^{n-k}=0$$

$$p\sum_{k=0}^{n}\binom{n}{k}(p-kq)^{n-1}(-1)^{n-k}=0$$

Consider the following operator "discreet derivative": function $f(x)$ is turned into $\Delta f(x)=f(x+1)-f(x)$. It is really easy to see that if f is a polynomial, then Δf is a polynomial of lower degree. Applying this operator n times results it $\Delta^n f(x)=\sum_{k=0}^{n}(-1)^k f(x+k)$. Therefore if f is a polynomial of degree less then n then $\Delta^n f$ is constant zero. Now apply it to $f(x)=(p+xq)^{n-1}$ and we get what we need in order to finish the step of induction.

It remains to do the base of induction (well, if I'd write it for the jury I would write everything in the reversed order): when $n = 0$,

$$p\sum_{k=0}^{0}\binom{0}{k}(p-kq)^{k-1}(r+kq)^{0-k}=(p+r)^{0}$$

$$p\binom{0}{0}(p-0)^{0-1}(r+0)^{0}=(p+r)^{0}$$

$$p\cdot 1\cdot\frac{1}{p}\cdot 1=1$$

QED.