Continuity Beyond the Real Numbers

Vlad-Ștefan Constantinescu

Introduction

In high school, continuity is defined for functions of one real variable: if $x_k \to x$, then $f(x_k) \to f(x)$. I have not seen many people talk about continuity beyond the kind studied in class, although it can be very useful in olympiad problems.

In many olympiad problems the variable is not a number but a more complicated object: a matrix, a function, a geometric configuration. The continuity we define in Section 2, with respect to the "largest deviation" distance, is new relative to the grade 11 curriculum. The purpose of this text is practical: to show when we can formalize the concept of passing to the limit for objects that are not real numbers as well.

We start from three competition problems whose solutions share the same structure: the required property is first proved on a convenient class of objects, and then extended to all objects by approximation. For the first problem this is exactly the route of the official solution; the other two also have solutions with no limits at all, which we point out from the start, but we keep them as training examples. Building on them, we set up the general framework (Section 2), apply it to other competition problems (Section 3), show where it fails (Section 4) and end with exercises. The high school theory we use (limits, continuity, dense sets, the Darboux property, matrices and eigenvalues, symmetric polynomials, integrals) can be found in my books [1, 2, 3, 4] and in their translations [5, 6, 7, 8, 9, 10, 11, 12].

1. Three problems

1.1 Problem 1. Romanian National Olympiad (ONM), 2022, grade 11, problem 4

Statement. Let $A, B \in \mathcal{M}_n(\mathbb{C})$ be such that $A^2 + B^2 = 2AB$. Prove that

$$\det(A - xI_n) = \det(B - xI_n),$$

for every $x \in \mathbb{C}$.

Source: [13]. The solution below follows Solution 1 of the marking scheme; in Step 3 we also give the version with a limit.

Solution. *Step 1: the case of invertible A, B.* Multiplying the relation by A^{-1} on the left and by B^{-1} on the right, we get $AB^{-1} + A^{-1}B = 2I_n$. With $X = A^{-1}B$ we have $B = AX$, so $AB^{-1} = AX^{-1}A^{-1}$, and the relation becomes $AX^{-1}A^{-1} = 2I_n - X$. Hence the matrices $2I_n - X$ and X^{-1} are similar and have the same eigenvalues, with the same multiplicities. Let $Y = I_n - X$: the matrices $I_n + Y$ and $(I_n - Y)^{-1}$ have the same eigenvalues. The eigenvalues of $(I_n - Y)^{-1}$ are the numbers $\frac{1}{1-\beta}$, with β an eigenvalue of Y. If α is an eigenvalue of Y, then $1 + \alpha$

is an eigenvalue of $I_n + Y$, so $1 + \alpha = \frac{1}{1-\beta}$ for some eigenvalue β of Y; in particular $1 + \alpha \neq 0$ and $\beta = \frac{\alpha}{1+\alpha}$. Applying this observation repeatedly, by induction, all the numbers $\frac{\alpha}{1+k\alpha}$, $k \in \mathbb{N}$, are eigenvalues of Y. If $\alpha \neq 0$, they are pairwise distinct, and Y would have infinitely many eigenvalues, which is absurd. Hence all eigenvalues of Y are 0, all eigenvalues of $X = I_n - Y$ are 1, $\det X = 1$ and $\det B = \det A \cdot \det X = \det A$.

Step 2: the notions we use. We now pass from invertible pairs to all pairs, by means of a limit. First we define, from scratch, every notion we need.

Definition 1. A *sequence of pairs of matrices* is a rule that assigns to every natural number $k \geq 1$ a pair (P_k, Q_k) of matrices in $\mathcal{M}_n(\mathbb{C})$. A pair is described by $2n^2$ numbers: the entries of P_k and those of Q_k. The sequence *converges* to the pair (P, Q), and we write $(P_k, Q_k) \to (P, Q)$, if each of these $2n^2$ numbers converges to the number in the same position in P, respectively in Q.

Definition 2. Let $\mathcal{H}$ be a class of pairs of matrices and $\mathcal{H}^*$ a subclass of it. We say that $\mathcal{H}^*$ is *dense* in $\mathcal{H}$ if every pair in $\mathcal{H}$ is the limit of a sequence of pairs in $\mathcal{H}^*$.

Definition 3. A *quantity* on $\mathcal{H}$ is a rule Φ that assigns a complex number to every pair in $\mathcal{H}$. It is *continuous* on $\mathcal{H}$ if, whenever $(P_k, Q_k) \to (P, Q)$ with all the pairs in $\mathcal{H}$, we have $\Phi(P_k, Q_k) \to \Phi(P, Q)$.

Let

$$\mathcal{H} = \{(P, Q) : \ P^2 + Q^2 = 2PQ\}$$

be the class of pairs that satisfy the hypothesis, and let $\mathcal{H}^* \subseteq \mathcal{H}$ be the class of those in which P and Q are invertible. Step 1 says that $\det P = \det Q$ for every $(P, Q) \in \mathcal{H}^*$.

The hypothesis is preserved when we shift both matrices by the same multiple of I_n: for $(P, Q) \in \mathcal{H}$ and $c \in \mathbb{C}$, expanding the brackets,

$$(P - cI_n)^2 + (Q - cI_n)^2 = P^2 + Q^2 - 2c(P + Q) + 2c^2 I_n = 2(P - cI_n)(Q - cI_n),$$

so $(P - cI_n, Q - cI_n) \in \mathcal{H}$.

Lemma 1. *The class $\mathcal{H}^*$ is dense in $\mathcal{H}$, that is, every pair in $\mathcal{H}$ is the limit of a sequence of pairs in $\mathcal{H}^*$.*

Proof. Let $(P, Q) \in \mathcal{H}$. We construct a sequence of pairs in $\mathcal{H}^*$ that converges to (P, Q), in three steps.

(i) Choosing the numbers c_k. Let V be the set of eigenvalues of P and of Q. The eigenvalues of P are the roots of the polynomial $\det(P - cI_n)$, which has degree n in c, so there are at most n of them; the same holds for Q. Hence V has at most $2n$ elements. Each element of V equals $\frac{1}{k}$ for at most one k, so only finitely many indices k have $\frac{1}{k} \in V$. Discarding them, we obtain a sequence $c_k \to 0$ with $c_k \notin V$ for every k.

(ii) The pairs $(P - c_k I_n, Q - c_k I_n)$ lie in $\mathcal{H}^*$. Since c_k is not an eigenvalue of P, we have $\det(P - c_k I_n) \neq 0$, so $P - c_k I_n$ is invertible; the same holds for $Q - c_k I_n$. By the computation above, with $c = c_k$, the pair satisfies the relation that defines $\mathcal{H}$. Hence it lies in $\mathcal{H}^*$.

(iii) Convergence. Off the diagonal, the matrix $P - c_k I_n$ has the same entries as P: these form constant sequences, which converge to the entries of P. On the diagonal it has $p_{ii} - c_k$, and $p_{ii} - c_k \to p_{ii}$ because $c_k \to 0$. The same holds for $Q - c_k I_n$. All $2n^2$ sequences of numbers converge, so $(P - c_k I_n, Q - c_k I_n) \to (P, Q)$ (Definition 1). □

Step 3: the extension. Let $\Phi(P, Q) = \det P - \det Q$.

(i) Φ is continuous on $\mathcal{H}$. Let $(P_k, Q_k) \to (P, Q)$. By the definition of the determinant,

$$\det P_k = \sum_{\sigma} \varepsilon(\sigma)\, p^{(k)}_{1\sigma(1)}\, p^{(k)}_{2\sigma(2)} \cdots p^{(k)}_{n\sigma(n)},$$

where the sum runs over the $n!$ permutations σ of the set $\{1, \dots, n\}$, and $\varepsilon(\sigma) = \pm 1$ is the sign of the permutation. Each factor $p^{(k)}_{i\sigma(i)}$ is a sequence of numbers that converges to $p_{i\sigma(i)}$ (Definition 1). A product of n convergent sequences converges to the product of the limits, and a finite sum of convergent sequences converges to the sum of the limits. Hence $\det P_k \to \det P$; likewise $\det Q_k \to \det Q$, and their difference converges to the difference of the limits: $\Phi(P_k, Q_k) \to \Phi(P, Q)$.

(ii) Φ vanishes on $\mathcal{H}^*$. This is exactly Step 1: for $(P, Q) \in \mathcal{H}^*$, $\det P = \det Q$.

(iii) Φ vanishes on all of $\mathcal{H}$. Let $(P, Q) \in \mathcal{H}$ and let (P_k, Q_k) be the sequence from Lemma 1. By (ii), $\Phi(P_k, Q_k) = 0$ for every k, so the sequence $\Phi(P_k, Q_k)$ is constantly equal to 0 and has limit 0. By (i), this same limit is $\Phi(P, Q)$. Since a sequence of numbers has only one limit, $\Phi(P, Q) = 0$. (In the language of Section 2, this is Theorem 1 a) applied to the class $\mathcal{H}$ and the dense class $\mathcal{H}^*$.)

(iv) Conclusion. Let $x \in \mathbb{C}$. The pair (A, B) lies in $\mathcal{H}$, so, by the computation in Step 2 with $c = x$, the pair $(A - xI_n, B - xI_n)$ also lies in $\mathcal{H}$. By (iii), $\Phi(A - xI_n, B - xI_n) = 0$, that is, $\det(A - xI_n) = \det(B - xI_n)$. □

> *Remark.* The marking scheme concludes algebraically: $\det(A - xI_n) - \det(B - xI_n)$ is a polynomial in x that vanishes at every point x that is not an eigenvalue, so it is the zero polynomial. The marking scheme also contains a second, purely algebraic solution, based on ranks. The key point of the solution above is that the approximation takes place *inside the hypothesis class*: the approximating pairs also satisfy $P^2 + Q^2 = 2PQ$, so we know something about them. It would not be enough for the invertible pairs to be dense in the set of all pairs. In Section 4 we see a problem in which precisely this condition is missing.

1.2 Problem 2. Romanian Olympiad, county round (OJM), 2026, grade 11, problem 3

Statement. Let $n \geq 2$. Find all matrices $M \in \mathcal{M}_n(\mathbb{C})$ with the following property: if $A, B \in \mathcal{M}_n(\mathbb{C})$ and $AB = M$, then $BA = M$.

Answer. $M = \lambda I_n$ with $\lambda \in \mathbb{C}^*$.

Source: [14, 15]. Let us say from the start that the official marking scheme uses no limit at all: after Step 1 below, it works directly with the invertible matrices $I_n + E_{1j}$. We keep the problem as a training example.

Personal note. I took part in OJM 2026, and the solution below is the first one I found during the contest. I did not write up this one, judging that an argument with limits might not be accepted as complete.

Solution. We denote by $\mathrm{GL}_n(\mathbb{C})$ the set of invertible matrices. The solution passes to the limit along a sequence of matrices. The notions are those of Problem 1, now written for a single matrix; we give them again, from scratch.

> **Definition 4.** A *sequence of matrices* is a rule that assigns to every natural number $k \geq 1$ a matrix $X_k \in \mathcal{M}_n(\mathbb{C})$. We denote the entries of X_k by $x^{(k)}_{ij}$. For each fixed position (i, j) we thus obtain an ordinary sequence of complex numbers, $\left(x^{(k)}_{ij}\right)_{k \geq 1}$.

Definition 5. The sequence of matrices (X_k) *converges* to the matrix $X = (x_{ij})$, and we write $X_k \to X$, if for each position (i, j) we have $x_{ij}^{(k)} \to x_{ij}$, that is, $\left|x_{ij}^{(k)} - x_{ij}\right| \to 0$. Equivalently, since there are finitely many positions, the largest deviation $\max_{i,j}\left|x_{ij}^{(k)} - x_{ij}\right|$ tends to 0.

Lemma 2. *Let $M \in \mathcal{M}_n(\mathbb{C})$ be fixed. The quantity $X \mapsto MX - XM$ is continuous, that is: if $X_k \to X$, then $MX_k - X_kM \to MX - XM$. In particular, if $MX_k - X_kM = O_n$ for every k, then $MX - XM = O_n$.*

Proof. The entry in position (i, j) of the matrix $MX_k - X_kM$ is

$$\sum_{l=1}^{n} m_{il}\, x_{lj}^{(k)} - \sum_{l=1}^{n} x_{il}^{(k)}\, m_{lj},$$

a linear combination, with fixed coefficients, of finitely many entries of X_k. By the rules for limits of sequences of numbers, it tends to the corresponding entry of $MX - XM$. If all the matrices $MX_k - X_kM$ are zero, each of these sequences of numbers is constantly equal to 0, so its limit is 0. □

Step 1: M commutes with every invertible matrix. For $C \in \mathrm{GL}_n(\mathbb{C})$, take $A = MC$, $B = C^{-1}$: $AB = M$, so $BA = C^{-1}MC = M$, that is, $MC = CM$.

Step 2: approximating an arbitrary matrix.

Lemma 3. *The set $\mathrm{GL}_n(\mathbb{C})$ is dense in $\mathcal{M}_n(\mathbb{C})$, that is, for every $X \in \mathcal{M}_n(\mathbb{C})$ there is a sequence of invertible matrices X_k with $X_k \to X$.*

Proof. The function $p(t) = \det(X + tI_n)$ is a polynomial in t of degree n with leading coefficient 1, so it has at most n roots. Each root equals $\frac{1}{k}$ for at most one k; discarding these finitely many indices, we obtain a sequence $t_k \to 0$ with $p(t_k) \neq 0$ for every k. The matrices $X_k = X + t_kI_n$ have determinant $p(t_k) \neq 0$, so they are invertible. Off the diagonal, the entries of X_k are those of X (constant sequences); on the diagonal they are $x_{ii} + t_k \to x_{ii}$. Hence $X_k \to X$ (Definition 5). □

Let $X \in \mathcal{M}_n(\mathbb{C})$ be arbitrary and let X_k be the sequence of invertible matrices from Lemma 3, with $X_k \to X$. By Step 1, applied to the invertible matrix X_k, we have $MX_k - X_kM = O_n$ for every k. By Lemma 2, $MX_k - X_kM \to MX - XM$; a sequence constantly equal to O_n has limit O_n, so $MX - XM = O_n$, that is, $MX = XM$. Since X was arbitrary, M commutes with every matrix.

Step 3: M is scalar. Let E_{ij} be the matrix with 1 in position (i, j) and 0 elsewhere. Column j of ME_{ij} is column i of M, the other columns being zero; column j of $E_{ij}M$ has the entry m_{jj} in position i and 0 elsewhere. From $ME_{ij} = E_{ij}M$, with $j \neq i$, it follows that $m_{ki} = 0$ for $k \neq i$ and $m_{ii} = m_{jj}$. Hence $M = \lambda I_n$.

Step 4: verification. For $\lambda = 0$ the property fails: $A = E_{12}$, $B = E_{11}$ give $AB = O_n$ and $BA = E_{12} \neq O_n$ (this is where $n \geq 2$ is used). For $\lambda \neq 0$, $AB = \lambda I_n$ implies $\det A \det B = \lambda^n \neq 0$, so $B = \lambda A^{-1}$ and $BA = \lambda I_n$. □

Remark. Here $MX - XM$ does not change when X is replaced by $X + tI_n$, so the limit can be avoided. The version with a limit, however, shows the general mechanism, which also works when no such simplification is available.

1.3 Problem 3. OJM 2004, grade 12, problem 2

Statement. Let $f : [0,1] \to \mathbb{R}$ be a continuous function such that

$$\int_0^1 f(x)g(x)\,dx = \int_0^1 f(x)\,dx \cdot \int_0^1 g(x)\,dx$$

for every continuous non-differentiable function $g : [0,1] \to \mathbb{R}$. Prove that f is constant.

Source: [16]. Let us say from the start that the limit can be avoided completely (see the Remark after the solution), and that the official solution uses a localized test function, with no limit at all.

Personal note. I came across this problem while working on the ROHS difficulty scale (a scale from 0 to 50 that I am building for OJM and ONM problems, separately for each grade). I estimate my solution below at 20 to 25 on the ROHS scale; the official solution is easier.

Solution. Here too the solution passes to the limit, this time along a sequence of functions. We first define the notions (see also [2]).

Definition 6. A *sequence of continuous functions* is a rule that assigns to every natural number $k \geq 1$ a continuous function $g_k : [0,1] \to \mathbb{R}$. For each fixed $x \in [0,1]$ we obtain an ordinary sequence of numbers, $\big(g_k(x)\big)_{k\geq 1}$.

Definition 7. For continuous $g_1, g_2 : [0,1] \to \mathbb{R}$, we call the *distance* between them the largest deviation between their values,

$$d(g_1, g_2) = \max_{x\in[0,1]} |g_1(x) - g_2(x)|$$

(the maximum exists, since it is the maximum of a continuous function on a closed interval). The sequence (g_k) *converges uniformly* to g, and we write $g_k \to g$, if $d(g_k, g) \to 0$.

It is not enough that $g_k(x) \to g(x)$ for each x separately; we require the deviation to become small *simultaneously* at all points. Section 2.2 shows why this requirement is necessary.

Let $c = \int_0^1 f$, $\varphi = f - c$ and $L(g) = \int_0^1 \varphi g$, for continuous g. For continuous non-differentiable g, $L(g) = \int_0^1 fg - c\int_0^1 g = 0$.

Step 1: L behaves well under limits.

Lemma 4. *The quantity L is continuous with respect to uniform convergence, that is: if $g_k \to g$ uniformly, then $L(g_k) \to L(g)$.*

Proof. For continuous g_1, g_2, we have $|\varphi(x)\big(g_1(x) - g_2(x)\big)| \leq |\varphi(x)|\, d(g_1, g_2)$ for every x, so, integrating,

$$|L(g_1) - L(g_2)| \leq \int_0^1 |\varphi| \cdot d(g_1, g_2).$$

With $g_1 = g_k$ and $g_2 = g$, the right-hand side tends to 0. □

Step 2: approximating an arbitrary continuous function.

Lemma 5. *The continuous non-differentiable functions form a dense class in the class of continuous functions on $[0,1]$, that is, every continuous $g : [0,1] \to \mathbb{R}$ is the uniform limit of a sequence of continuous non-differentiable functions.*

Proof. Let $\tau(x) = |x - \frac{1}{2}|$, which is continuous, not differentiable at $\frac{1}{2}$, and satisfies $\max|\tau| = \frac{1}{2}$. The function $g + \varepsilon\tau$ is not differentiable at $\frac{1}{2}$ for all real values of ε except at most one: if $g + \varepsilon_1\tau$ and $g + \varepsilon_2\tau$, with $\varepsilon_1 \neq \varepsilon_2$, were both differentiable at $\frac{1}{2}$, then $(\varepsilon_1 - \varepsilon_2)\tau$ would also be differentiable at $\frac{1}{2}$, which is false. We choose $\varepsilon_k \to 0$ avoiding this value; the functions $g_k = g + \varepsilon_k\tau$ are continuous and non-differentiable, and $d(g_k, g) = \frac{|\varepsilon_k|}{2} \to 0$. □

Step 3: conclusion. Let g be an arbitrary continuous function and let g_k be the sequence of continuous non-differentiable functions from Lemma 5, with $g_k \to g$ uniformly. Each g_k is non-differentiable, so $L(g_k) = 0$: the sequence $L(g_k)$ is constantly equal to 0. By Lemma 4, $L(g_k) \to L(g)$, so $L(g) = 0$. For $g = \varphi$ we get $\int_0^1 \varphi^2 = 0$. The function φ^2 is continuous and nonnegative; if $\varphi(x_0) \neq 0$, then $\varphi^2 > \frac{1}{2}\varphi^2(x_0)$ on an interval around x_0, and the integral would be positive. Hence $\varphi \equiv 0$ and $f \equiv c$. □

> *Remark* (L vanishes, in two lines). If g is continuous and differentiable at every point of $[0, 1]$, then $g+\tau$ is not differentiable at $\frac{1}{2}$, so $L(g) = L(g+\tau)-L(\tau) = 0$. Every continuous function is either non-differentiable somewhere (and then the hypothesis directly gives $L(g) = 0$) or differentiable everywhere (and then we have just seen that $L(g) = 0$). Hence L is identically zero without any limit, and Lemmas 4 and 5 are not needed. I have nevertheless kept the solution by approximation, because it shows the mechanism in its cleanest form.

1.4 From three problems to a generalization

These three problems suggest the following generalization. All three solutions consist of the same three steps.

1. The property is proved on a particular class of objects: the pairs of invertible matrices in $\mathcal{H}$; the invertible matrices; the non-differentiable functions.

2. One shows that every object can be approximated arbitrarily well by objects of that class (Lemmas 1, 3 and 5).

3. The quantity that expresses the property ($\det P - \det Q$, $MX - XM$ and $L(g)$, respectively) behaves well under limits (Step 3 of Problem 1, Lemmas 2 and 4), so the property passes from the approximating objects to the approximated object.

Moreover, the definitions of convergence have the same form. A pair of matrices is described by its $2n^2$ entries, a matrix by the numbers x_{ij}, a function by the numbers $g(x)$, and in all cases a sequence converges when *the largest deviation between corresponding numbers* tends to 0 (Definitions 1, 5 and 7). The only difference is that a pair of matrices or a matrix is described by finitely many numbers, while a function is described by infinitely many.

So we generalize as follows: every object (matrix, function, triangle, line) is described by a list of numbers, its parameters; the distance between two objects is the largest deviation between corresponding parameters; a quantity associated with the objects is continuous if it changes little when the parameters change little. In the next section we develop this framework and show that the three steps above become a single theorem.

2. The theory

2.1 Objects described by parameters

Definition 8. Let I be a nonempty set (the index set). An *object with parameters indexed by I* is a function $X : I \to \mathbb{C}$; the numbers $X(i)$, $i \in I$, are called the *parameters* of X. In many applications the parameters are real.

In other words, an object is given by the list of its parameters, one for each index $i \in I$, and we identify it with this list. The list may be finite or infinite (for functions it has one entry for each number in $[0,1]$).

Example. a) *Matrices.* A matrix $A \in \mathcal{M}_n(\mathbb{C})$ is the object with $I = \{1,\dots,n\}^2$. The function is

$$A : \{1,\dots,n\}^2 \to \mathbb{C}, \qquad A(i,j) = a_{ij},$$

so the matrix has n^2 parameters, its entries. With the distance of Section 2.2, two matrices are close when all their corresponding entries are close. For example, for

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2.1 \\ 3 & 3.95 \end{pmatrix},$$

the functions are $A(1,1) = 1$, $A(1,2) = 2$, $A(2,1) = 3$, $A(2,2) = 4$ and $B(1,1) = 1$, $B(1,2) = 2.1$, $B(2,1) = 3$, $B(2,2) = 3.95$. The deviations are 0, 0.1, 0 and 0.05, so the distance between A and B is 0.1. Similarly, a *pair* of matrices (P,Q) as in Problem 1 is the object with $I = \{1,2\} \times \{1,\dots,n\}^2$ and $2n^2$ parameters; the function is $(1,i,j) \mapsto p_{ij}$, $(2,i,j) \mapsto q_{ij}$.

b) *Triangles.* A triangle ABC in the plane is the object with $I = \{1,\dots,6\}$, the parameters being the coordinates of the vertices. The function is $X : \{1,\dots,6\} \to \mathbb{R}$,

$$X(1) = x_A, \quad X(2) = y_A, \quad X(3) = x_B, \quad X(4) = y_B, \quad X(5) = x_C, \quad X(6) = y_C.$$

Not every function of this kind is a triangle: we require A, B, C to be non-collinear, so the class of triangles is a part of the set of all lists of six numbers. Two triangles are at distance at most δ when each vertex of one is obtained from the corresponding vertex of the other by moving each coordinate by at most δ. For example, the triangle with vertices $(0,0)$, $(4,0)$, $(0,3)$ is the function with values $0,0,4,0,0,3$, and the one with vertices $(0,0.01)$, $(4,0)$, $(0,2.98)$ is the function with values $0,0.01,4,0,0,2.98$. The deviations are 0, 0.01, 0, 0, 0, 0.02, so the two triangles are at distance 0.02.

c) *Polynomials.* A polynomial $a_0 + a_1X + \cdots + a_nX^n$ of degree at most n is the object with $I = \{0,\dots,n\}$, the parameters being the coefficients. The function is

$$\{0,\dots,n\} \to \mathbb{C}, \qquad k \mapsto a_k.$$

For example, for $n = 2$, the polynomial $X^2 + 1$ is the function $0 \mapsto 1$, $1 \mapsto 0$, $2 \mapsto 1$, and $X^2 + 0.01X + 1$ is the function $0 \mapsto 1$, $1 \mapsto 0.01$, $2 \mapsto 1$; they differ only at index 1, so they are at distance 0.01. If two polynomials of degree at most n are at distance δ, then for $|x| \le 1$ their values differ by at most $(n+1)\delta$, because each of the $n+1$ terms differs by at most δ.

d) *Functions.* A function $g : [0,1] \to \mathbb{R}$ is the object with $I = [0,1]$. Here the function of Definition 8 is g itself: the parameter with index x is the value $g(x)$, one for each number x in $[0,1]$, so there are infinitely many parameters. The distance between g

and h is the largest value of $|g(x) - h(x)|$; it is at most δ exactly when the graph of h lies in the band of width 2δ around the graph of g. For example, $g(x) = x$ and $h(x) = x + 0.01\sin(10x)$ are at distance at most 0.01. This is the object of Problem 3.

e) *Lines.* Every line has an equation $ax + by + c = 0$, with $(a, b) \neq (0, 0)$. Dividing it by $r = \sqrt{a^2 + b^2}$ and choosing the angle θ with $\cos\theta = \frac{a}{r}$, $\sin\theta = \frac{b}{r}$, it becomes

$$x\cos\theta + y\sin\theta = p, \qquad p = -\frac{c}{r},$$

where $(\cos\theta, \sin\theta)$ is a unit vector perpendicular to the line, and $|p|$ is the distance from the origin to the line. This form also covers vertical lines: $x = 1$ is obtained for $\theta = 0$, $p = 1$. So a line is the object with two parameters (θ, p); the function is $X : \{1, 2\} \to \mathbb{R}$, $X(1) = \theta$, $X(2) = p$. The same line has several pairs: (θ, p), $(\theta + 2\pi, p)$ and $(\theta + \pi, -p)$ give the same set of points. We work with the pair, that is, with the equation, and not with the set of points.

f) *Configurations.* In Exercise 4 a geometric configuration (a trapezoid and two points on its sides) will be the object with six parameters a, b, c, η, s, t; the function is $X : \{1, \dots, 6\} \to \mathbb{R}$, with $X(1) = a$, $X(2) = b$, $X(3) = c$, $X(4) = \eta$, $X(5) = s$, $X(6) = t$. In all cases, "close objects" means "all parameters are close".

2.2 Distance and limit

Definition 9. The distance between two objects X, Y with parameters indexed by the same set I is

$$d(X, Y) = \sup_{i \in I} |X(i) - Y(i)| \in [0, +\infty],$$

that is, the largest deviation between corresponding parameters. A sequence of objects $(X_k)_{k \geq 1}$ *converges* to the object X, and we write $X_k \to X$, if $d(X_k, X) \to 0$ (in particular, $d(X_k, X)$ is finite from some index on).

When I is finite, the supremum is a maximum, hence a number. When I is infinite, the supremum may be $+\infty$ (for example, for an unbounded function and the zero function); we allow this value. For continuous functions on $[0, 1]$ the supremum is a maximum, so we recover Definition 7.

Proposition 1. *If I is finite, then $X_k \to X$ if and only if $X_k(i) \to X(i)$ for each $i \in I$.*

Proof. For each i, $|X_k(i) - X(i)| \leq d(X_k, X)$, so convergence of the objects implies convergence of each parameter. Conversely, if each of the finitely many parameters converges, then the maximum of the $|I|$ deviations tends to 0. □

So for matrices, triangles or polynomials, the limit of a sequence of objects is computed parameter by parameter. For functions, where $I = [0, 1]$ is infinite, the definition asks for more: the deviation must tend to 0 *simultaneously* at all points. This is uniform convergence; the remark below shows why it is the right choice.

Remark (why not parameter by parameter). With infinitely many parameters, convergence of each parameter (*pointwise convergence*: $g_k(x) \to g(x)$ for each x) is too weak. For $k \geq 2$,

let

$$g_k(x) = \begin{cases} k^2 x, & 0 \le x \le \frac{1}{k}, \\ 2k - k^2 x, & \frac{1}{k} \le x \le \frac{2}{k}, \\ 0, & \frac{2}{k} \le x \le 1. \end{cases}$$

For each fixed x, $g_k(x) \to 0$: $g_k(0) = 0$, and for $x > 0$ we have $g_k(x) = 0$ as soon as $\frac{2}{k} < x$. Yet $\int_0^1 g_k = 1$ for every k, while the integral of the limit is 0. So the integral, the quantity we care about, would not be continuous with respect to pointwise convergence. With respect to uniform convergence, it is continuous (Lemma 4).

2.3 Continuous quantities

Definition 10. Let $\mathcal{O}$ be a class of objects with parameters indexed by I, and let $\Phi : \mathcal{O} \to \mathbb{C}$ be a function (a *quantity* associated with the objects). We say that Φ is *continuous at the object* $X \in \mathcal{O}$ if, for every sequence (X_k) of objects in $\mathcal{O}$ with $X_k \to X$, we have $\Phi(X_k) \to \Phi(X)$. The quantity Φ is *continuous on* $\mathcal{O}$ if it is continuous at every object of $\mathcal{O}$.

This is exactly the high school definition recalled in the introduction: if the parameters of X_k approach the parameters of X, then $\Phi(X_k)$ approaches $\Phi(X)$. For a function of one variable (I with a single element) we recover the usual definition. (As for functions of one variable, one can show that this definition is equivalent to the formulation with ε and δ; we will not need this equivalence.)

Proposition 2. *a) Each parameter $X \mapsto X(i)$ is a continuous quantity.*

b) Sums, products and quotients (with denominator nonzero on $\mathcal{O}$) of continuous quantities are continuous, and if Φ is continuous, then so is $|\Phi|$. If Φ is continuous and real-valued, all its values lie in an interval J, and $\omega : J \to \mathbb{R}$ is continuous, then $\omega \circ \Phi$ is continuous.

c) Consequently, every polynomial expression in finitely many parameters is a continuous quantity.

Proof. a) $|X_k(i) - X(i)| \le d(X_k, X) \to 0$. b) Let $X_k \to X$ in $\mathcal{O}$. Then $\Phi(X_k) \to \Phi(X)$ and $\Psi(X_k) \to \Psi(X)$ for the continuous quantities Φ, Ψ, and the rules for limits of sequences give the conclusion for sums, products and quotients. For the absolute value, $\big||\Phi(X_k)| - |\Phi(X)|\big| \le |\Phi(X_k) - \Phi(X)|$. For composition, the numbers $\Phi(X_k)$ lie in J and tend to $\Phi(X) \in J$, and ω is continuous at $\Phi(X)$, so $\omega(\Phi(X_k)) \to \omega(\Phi(X))$ (the sequential characterization of continuity for functions of one variable). c) Follows from a) and b). □

The version of b) with an arbitrary interval J matters: it covers $\sqrt{\cdot}$ (defined on $[0, +\infty)$), as well as arccos and arcsin (defined on $[-1, 1]$).

Proposition 3 (stability of strict inequalities)**.** *Let $\Phi : \mathcal{O} \to \mathbb{R}$ be continuous at $X \in \mathcal{O}$, with $\Phi(X) > 0$. Then there is $\delta > 0$ such that $\Phi(Y) > 0$ for every $Y \in \mathcal{O}$ with $d(Y, X) < \delta$.*

Proof. Suppose the contrary: for each $k \ge 1$ there is $Y_k \in \mathcal{O}$ with $d(Y_k, X) < \frac{1}{k}$ and $\Phi(Y_k) \le 0$. Then $Y_k \to X$, so $\Phi(Y_k) \to \Phi(X)$; non-strict inequalities are preserved under limits, so $\Phi(X) \le 0$, a contradiction. □

Applied to $-\Phi$ and to $|\Phi|$, the proposition says the same thing for $\Phi(X) < 0$ and for $\Phi(X) \neq 0$, respectively. Intuitively: a strict inequality between continuous quantities remains true under small perturbations of the parameters.

2.4 Dense classes and the extension principle

Definition 11. A subclass $\mathcal{D} \subseteq \mathcal{O}$ is *dense* in $\mathcal{O}$ if every object in $\mathcal{O}$ is the limit of a sequence of objects in $\mathcal{D}$.

Theorem 1 (extension principle)**.** *Let $\Phi, \Psi : \mathcal{O} \to \mathbb{C}$ be continuous quantities and let $\mathcal{D}$ be dense in $\mathcal{O}$.*

a) If $\Phi = \Psi$ on $\mathcal{D}$, then $\Phi = \Psi$ on $\mathcal{O}$.

b) If Φ, Ψ are real-valued and $\Phi \leq \Psi$ on $\mathcal{D}$, then $\Phi \leq \Psi$ on $\mathcal{O}$.

Proof. Let $X \in \mathcal{O}$ and $X_k \in \mathcal{D}$ with $X_k \to X$. By Definition 10, $\Phi(X_k) \to \Phi(X)$ and $\Psi(X_k) \to \Psi(X)$. a) $\Phi(X_k) = \Psi(X_k)$ for every k; passing to the limit, $\Phi(X) = \Psi(X)$. b) Similarly, the non-strict inequality is preserved under limits. □

For functions of one real variable, part a) is the "density principle" of [1].

The hypothesis class. In Theorem 1, $\mathcal{O}$ is the class on which we know that Φ and Ψ are defined and continuous, and $\mathcal{D}$ must be dense in $\mathcal{O}$ itself. If the problem gives a hypothesis that defines a smaller class (for example, the pairs (A, B) with $A^2 + B^2 = 2AB$), then what matters is density in that class, not in the space of all objects: the approximating objects must satisfy the hypothesis as well. In Problem 1 the shift by xI_n preserves the hypothesis; in Problem 2 the hypothesis concerns the fixed matrix M, and the matrices X being approximated need not satisfy anything. In Section 4.1 we see an olympiad problem in which this condition is missing and the conclusion no longer extends.

The strategy. For a problem that asks for a property of all objects:

1. write the property as an equality (or a non-strict inequality) between continuous quantities;
2. find a dense class on which the property is easy to prove;
3. apply Theorem 1.

A strict inequality is not preserved under limits: if $\Phi < \Psi$ on $\mathcal{D}$, one only gets $\Phi \leq \Psi$ on $\mathcal{O}$.

Checklist. Before we can say "by continuity", the following must be made explicit:

(a) the objects and their parameters (what varies and what stays fixed);

(b) that the quantities used are defined and continuous on the *entire* class $\mathcal{O}$ (denominators do not vanish, square roots and arccos receive arguments in their domain);

(c) a continuous quantity that expresses the conclusion, and a class that is dense in the very class on which the hypothesis holds;

(d) that from an equality (or inequality) between quantities one can get back to the required property.

2.5 The problems revisited

Problem 1. The object is the pair (P, Q), with $2n^2$ parameters, the entries of the two matrices; by Proposition 1, the convergence in Definition 1 is exactly the limit in Definition 9. The quantity $\Phi(P, Q) = \det P - \det Q$ is polynomial in the parameters, hence continuous (Proposition 2), and Lemma 1 says that $\mathcal{H}^*$ is dense in $\mathcal{H}$. Theorem 1 a), applied to the class $\mathcal{H}$, is Step 3.

Problem 2. By Proposition 1, the limit of a sequence of matrices in the sense of Definition 5 (entry by entry) is exactly the limit of Definition 9. Lemma 3 says that $\mathrm{GL}_n(\mathbb{C})$ is dense in $\mathcal{M}_n(\mathbb{C})$. The entries of $X \mapsto MX - XM$ are continuous quantities (Proposition 2) that vanish on $\mathrm{GL}_n(\mathbb{C})$; Theorem 1 gives Step 2 of the solution.

The same mechanism yields classical results. For example, for any $A, B \in \mathcal{M}_n(\mathbb{C})$, the matrices AB and BA have the same characteristic polynomial: if A is invertible, $\det(tI_n - BA) = \det\big(A^{-1}(tI_n - AB)A\big) = \det(tI_n - AB)$; with B and t fixed, both sides are polynomial in the entries of A, hence continuous, and they agree on the dense class $\mathrm{GL}_n(\mathbb{C})$.

Problem 3. For continuous functions, the limit of Definition 9 is the uniform convergence of Definition 7. Lemma 4 says exactly that L is a continuous quantity on the class of continuous functions on $[0, 1]$, Lemma 5 says that the non-differentiable functions form a dense class, and Theorem 1 once again gives the conclusion.

2.6 Triangles

Let $\mathcal{O}$ be the class of nondegenerate triangles, each given by the six coordinates of its vertices. By Proposition 1, a sequence of triangles $A_kB_kC_k$ converges to ABC if and only if $A_k \to A$, $B_k \to B$, $C_k \to C$ (coordinate by coordinate). We use this class in the exercise on scalene triangles at the end of the text.

2.7 Continuous families and the Darboux property

In many problems the objects do not come as a sequence, but as a family depending on a real number: the tangents to a circle, the lines rotating around a point, the points moving along segments.

Definition 12. Let $J \subseteq \mathbb{R}$ be an interval and $\mathcal{O}$ a class of objects with parameters indexed by a finite set I. A *continuous family* of objects is a function $t \mapsto X_t$, $J \to \mathcal{O}$, for which each parameter $t \mapsto X_t(i)$ is a continuous function of t.

By Proposition 1, if $t_k \to t$ in J, then $X_{t_k} \to X_t$.

Proposition 4. *Let $(X_t)_{t\in J}$ be a continuous family and $\Phi : \mathcal{O} \to \mathbb{R}$ a continuous quantity. Then the function $t \mapsto \Phi(X_t)$ is continuous on J. In particular, if $\alpha, \beta \in J$ and $\Phi(X_\alpha) \le 0 \le \Phi(X_\beta)$, there is t between α and β with $\Phi(X_t) = 0$.*

Proof. If $t_k \to t$, then $X_{t_k} \to X_t$, so $\Phi(X_{t_k}) \to \Phi(X_t)$ (Definition 10). Therefore the function $t \mapsto \Phi(X_t)$ is continuous on the interval J and has the Darboux property [1]. □

3. Applications

3.1 Problem 4. IMC 2004, day 2, problem 4

Statement. For $n \geq 1$, let M be an $n \times n$ complex matrix whose distinct eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_k$, with multiplicities $m_1, m_2, \ldots, m_k$. Consider the linear operator L_M defined by $L_M(X) = MX + XM^T$ for every $n \times n$ complex matrix X. Find the eigenvalues of L_M and their multiplicities.

Source: [23]. The official solution first treats the case in which the eigenvalues of M are distinct and all the sums $\lambda_r + \lambda_s$ are different, then observes that every M is a limit of matrices of this type and concludes "by the continuity of the eigenvalues". The claim is correct, but the eigenvalues have no natural order, so, taken one at a time, they are not quantities in the sense of Section 2. We show how to make the argument rigorous.

Answer. The eigenvalues are $2\lambda_r$, with multiplicity m_r^2 $(1 \leq r \leq k)$, and $\lambda_r + \lambda_s$, with multiplicity $2m_r m_s$ $(1 \leq r < s \leq k)$; if some of these numbers coincide, their multiplicities add up.

Solution. Denote by $\mu_1, \ldots, \mu_n$ the eigenvalues of M, listed with multiplicities (each λ_r appears m_r times). Let $p_M(t) = \det(t\,\mathrm{Id} - L_M)$ be the characteristic polynomial of L_M, that is, the determinant of the $n^2 \times n^2$ matrix of $X \mapsto tX - L_M(X)$ in the basis (E_{ij}), and let

$$q_M(t) = \prod_{i=1}^{n} \prod_{j=1}^{n} (t - \mu_i - \mu_j).$$

Grouping the factors of q_M according to the values λ_r, λ_s taken by μ_i, μ_j, we obtain

$$q_M(t) = \prod_{r=1}^{k} (t - 2\lambda_r)^{m_r^2} \prod_{1 \leq r < s \leq k} (t - \lambda_r - \lambda_s)^{2m_r m_s}.$$

The answer is therefore equivalent to the identity $p_M(t) = q_M(t)$, for every $M \in \mathcal{M}_n(\mathbb{C})$ and every $t \in \mathbb{C}$. We fix t; the objects are the matrices M, with n^2 parameters.

Step 1: $M \mapsto p_M(t)$ is continuous. The matrix of L_M in the basis (E_{ij}) has entries that are linear in the entries of M, so $p_M(t)$ is a polynomial expression in the entries of M (Proposition 2).

Step 2: $M \mapsto q_M(t)$ is continuous. This is the delicate part, the "continuity of the eigenvalues". We do not follow the eigenvalues one at a time, but only a symmetric expression in them. Permuting the numbers $\mu_1, \ldots, \mu_n$ only permutes the factors of $q_M(t)$, so $q_M(t)$ is a symmetric polynomial in $\mu_1, \ldots, \mu_n$ (with coefficients that depend polynomially on t). By the fundamental theorem of symmetric polynomials [4], it can be written as a polynomial in $e_1, \ldots, e_n$, the elementary symmetric polynomials of $\mu_1, \ldots, \mu_n$. By Vieta's formulas, $(-1)^k e_k$ is the coefficient of t^{n-k} in $\det(tI_n - M)$, hence a polynomial expression in the entries of M. Therefore $q_M(t)$ is a polynomial expression in the entries of M, hence a continuous quantity.

Step 3: equality on a dense class. Let $\mathcal{D}$ be the class of matrices with n distinct eigenvalues. Such a matrix has a basis of eigenvectors $v_1, \ldots, v_n$, with $Mv_i = \mu_i v_i$ [3]. Then

$$L_M(v_i v_j^T) = (Mv_i)v_j^T + v_i(Mv_j)^T = (\mu_i + \mu_j)\, v_i v_j^T.$$

The n^2 matrices $v_i v_j^T$ are linearly independent: if $\sum c_{ij} v_i v_j^T = O_n$, then $VCV^T = O_n$, where V is the matrix with columns $v_1, \ldots, v_n$ (which is invertible) and $C = (c_{ij})$, so $C = O_n$. They therefore form a basis of $\mathcal{M}_n(\mathbb{C})$ consisting of eigenvectors of L_M; in this basis the matrix of L_M

is diagonal, with diagonal entries $\mu_i + \mu_j$, and the characteristic polynomial does not depend on the basis. Hence $p_M(t) = q_M(t)$ on $\mathcal{D}$. (We do not need the sums $\mu_i + \mu_j$ to be distinct.)

Step 4: $\mathcal{D}$ is dense. Every complex matrix is similar to an upper triangular matrix [3]: $M = PTP^{-1}$, and the diagonal of T contains the eigenvalues $t_{11}, \dots, t_{nn}$. For real ε, the matrix $T_\varepsilon = T + \varepsilon \operatorname{diag}(1, 2, \dots, n)$ has on its diagonal the numbers $t_{ii} + i\varepsilon$, which are pairwise distinct except for at most $\frac{n(n-1)}{2}$ values of ε (each equality $t_{ii} + i\varepsilon = t_{jj} + j\varepsilon$, with $i \neq j$, determines ε). We choose $\varepsilon_k \to 0$ avoiding these values. The matrices $M_k = PT_{\varepsilon_k}P^{-1} = M + \varepsilon_k P \operatorname{diag}(1, \dots, n)P^{-1}$ have n distinct eigenvalues, and $M_k \to M$.

Step 5: the extension. With t fixed, $M \mapsto p_M(t)$ and $M \mapsto q_M(t)$ are continuous on $\mathcal{M}_n(\mathbb{C})$ and agree on the dense class $\mathcal{D}$; by Theorem 1 a), they agree on $\mathcal{M}_n(\mathbb{C})$. □

> *Remark* (a route without limits). Triangularization also gives a direct solution. With $X = PYP^T$ we have $L_M(X) = P\,(TY + YT^T)\,P^T$, so L_M has the same characteristic polynomial as L_T. A direct computation gives
>
> $$L_T(E_{ij}) = (t_{ii} + t_{jj})E_{ij} + \sum_{l<i} t_{li}E_{lj} + \sum_{l<j} t_{lj}E_{il},$$
>
> so the matrix of L_T in the basis (E_{ij}), ordered lexicographically by (i, j), is triangular, with diagonal entries $t_{ii} + t_{jj}$. Since we used triangularization in Step 4 anyway, continuity does not shorten the solution here. Its role is to show that the sentence in the official solution can be made rigorous with the tools of Section 2.

3.2 Problem 5. The adjugate of a 3×3 matrix

Statement. Let $X, Y, Z \in \mathcal{M}_3(\mathbb{C})$. We denote by A^* the adjugate of the matrix A.

(a) Prove the identity

$$(X + Y + Z)^* + X^* + Y^* + Z^* = (X + Y)^* + (Y + Z)^* + (Z + X)^*.$$

(b) Prove the equivalence

$$(X + Y + Z)^* = 3(X^* + Y^* + Z^*) \iff (X - Y)^* + (Y - Z)^* + (Z - X)^* = O_3.$$

Source: Gazeta Matematică [19].

Solution. *Idea.* For an invertible matrix, the adjugate is $(\det A)A^{-1}$, and the Cayley-Hamilton theorem writes it as a polynomial of degree two in A. The formula obtained in this way extends by continuity to all matrices, and with it both parts become computations with an expression "of degree two".

Step 1: the formula for invertible matrices. For $A \in \mathcal{M}_3(\mathbb{C})$ let $s_1(A) = \operatorname{tr} A$, $s_2(A) = \frac{1}{2}\big((\operatorname{tr} A)^2 - \operatorname{tr}(A^2)\big)$ and $s_3(A) = \det A$. The Cayley-Hamilton theorem gives $A^3 - s_1(A)A^2 + s_2(A)A - s_3(A)I_3 = O_3$, that is,

$$A\big(A^2 - s_1(A)A + s_2(A)I_3\big) = (\det A)\, I_3.$$

If A is invertible, we multiply on the left by A^{-1} and, since $A^* = (\det A)A^{-1}$, we obtain

$$A^* = A^2 - s_1(A)A + s_2(A)I_3. \tag{F}$$

Step 2: the formula for every matrix. The objects are the matrices $A \in \mathcal{M}_3(\mathbb{C})$, with 9 parameters. Each entry of the matrix $\Phi(A) = A^* - A^2 + s_1(A)A - s_2(A)I_3$ is a polynomial

expression in the entries of A (the entries of A^* are minors of order 2), so it is a continuous quantity (Proposition 2). By Step 1, Φ vanishes on the invertible matrices, which form a dense class (Lemma 3, for $n = 3$). By Theorem 1 a), applied to each entry, $\Phi = O_3$ on all of $\mathcal{M}_3(\mathbb{C})$: formula (F) holds for every matrix.

Step 3: an expression of degree two. We write $Q(A) = A^2 - s_1(A)A + s_2(A)I_3$ and, for $U, V \in \mathcal{M}_3(\mathbb{C})$,

$$B(U,V) = UV + VU - (\operatorname{tr} U)V - (\operatorname{tr} V)U + \big(\operatorname{tr} U \operatorname{tr} V - \operatorname{tr}(UV)\big)I_3.$$

Expanding, $Q(U+V) = Q(U) + Q(V) + B(U,V)$; moreover, B is symmetric and linear in each argument, and $Q(-U) = Q(U)$. By Step 2, $A^* = Q(A)$ for every A.

Step 4: conclusion. a) Applying the relation of Step 3 twice,

$$(X+Y+Z)^* = Q(X) + Q(Y) + Q(Z) + B(X,Y) + B(Y,Z) + B(Z,X),$$

while $(X+Y)^* + (Y+Z)^* + (Z+X)^* = 2\big(Q(X)+Q(Y)+Q(Z)\big) + B(X,Y) + B(Y,Z) + B(Z,X)$. So the difference between the two sides of the identity is 0.

b) Since $Q(X-Y) = Q(X) + Q(Y) - B(X,Y)$,

$$(X-Y)^* + (Y-Z)^* + (Z-X)^* = 2\big(Q(X)+Q(Y)+Q(Z)\big) - B(X,Y) - B(Y,Z) - B(Z,X),$$

while $(X+Y+Z)^* - 3(X^*+Y^*+Z^*) = B(X,Y)+B(Y,Z)+B(Z,X) - 2\big(Q(X)+Q(Y)+Q(Z)\big)$. The two expressions are opposite, so one of them is O_3 exactly when the other one is. □

> *Remark.* The limit can be avoided: the entries of A^* are minors of order 2, hence homogeneous polynomials of degree two in the entries of A, and a) and b) are identities that hold for any such map. The solution above shows, however, how formula (F) is obtained for every matrix from the invertible case, exactly as in Problems 1 and 2.

3.3 Problem 6. ONM 2009, grade 12, problem 4

Statement. Find all continuous bijective functions $f : [0,1] \to [0,1]$ such that

$$\int_0^1 g\big(f(x)\big)\,dx = \int_0^1 g(x)\,dx$$

for every continuous function $g : [0,1] \to \mathbb{R}$.

Source: [20].

Solution. *Idea.* If g could be the indicator of an interval $[a,b]$, the hypothesis would say that f maps a set of length $b-a$ into $[a,b]$. The indicator is not continuous, so we squeeze it between two continuous "trapezoid" functions and pass to the limit.

Step 1: the preimage of an interval. A continuous bijective function on $[0,1]$ is strictly monotonic, and its inverse is also continuous and strictly monotonic. For $0 \le a < b \le 1$, the set $f^{-1}([a,b])$ is therefore an interval $[c,d]$, and $\int_0^1 \mathbf{1}_{[a,b]}\big(f(x)\big)\,dx = d - c$.

Step 2: the squeeze. For $\varepsilon > 0$, let g_ε be the continuous function equal to 1 on $[a,b]$, to 0 outside $[a-\varepsilon, b+\varepsilon]$, and linear in between (restricted to $[0,1]$). Since $g_\varepsilon \ge \mathbf{1}_{[a,b]}$,

$$d - c \le \int_0^1 g_\varepsilon\big(f(x)\big)\,dx = \int_0^1 g_\varepsilon(x)\,dx \le b - a + \varepsilon.$$

For $0 < \varepsilon < \frac{b-a}{2}$, let h_ε be the continuous function equal to 1 on $[a+\varepsilon, b-\varepsilon]$, to 0 outside $[a,b]$, and linear in between. Since $h_\varepsilon \le \mathbf{1}_{[a,b]}$,

$$d - c \ge \int_0^1 h_\varepsilon\big(f(x)\big)\,dx = \int_0^1 h_\varepsilon(x)\,dx \ge b - a - 2\varepsilon.$$

Letting $\varepsilon \to 0$, we get $d - c = b - a$: the preimage of every interval has the same length as the interval.

Step 3: conclusion. If f is increasing, $f^{-1}([0,y]) = [0, f^{-1}(y)]$ has length y, so $f^{-1}(y) = y$ and $f(x) = x$. If f is decreasing, $f^{-1}([0,y]) = [f^{-1}(y), 1]$ has length y, so $f^{-1}(y) = 1 - y$ and $f(x) = 1 - x$. Both functions satisfy the hypothesis (for the second one, by the change of variable $x \mapsto 1 - x$). □

> *Remark.* The limit in Step 2 is an ordinary one, in a single number ε: the trapezoids do not approach the indicator in the "largest deviation" distance (the distance stays equal to 1). The next problem starts from here and genuinely needs the continuity of Section 2.

3.4 Problem 7. A variant with powers

Statement. Let $f : [0,1] \to [0,1]$ be a continuous bijective function such that

$$\int_0^1 f(x)^k \, dx = \frac{1}{k+1} \quad \text{for every } k \in \mathbb{N}.$$

Prove that $f(x) = x$ for every x, or $f(x) = 1 - x$ for every x.

A new problem, built from Problem 6.

Solution. *Idea.* The hypothesis says that the equality of Problem 6 holds for $g(x) = x^k$, hence for every polynomial. Polynomials are dense among continuous functions, and the quantity expressing the equality is continuous; we extend it to all continuous functions and apply Problem 6.

Step 1: the objects and the quantity. The objects are the continuous functions $g : [0,1] \to \mathbb{R}$, with the "largest deviation" distance $d(g,h) = \max_{[0,1]} |g - h|$, as in Problem 3 (example d) in Section 2). Let

$$\Phi(g) = \int_0^1 g\big(f(x)\big) \, dx - \int_0^1 g(x) \, dx.$$

Since f takes values in $[0,1]$, we have $|g(f(x)) - h(f(x))| \le d(g,h)$ for every x, so

$$|\Phi(g) - \Phi(h)| \le \int_0^1 |g(f(x)) - h(f(x))| \, dx + \int_0^1 |g(x) - h(x)| \, dx \le 2\, d(g,h),$$

and Φ is a continuous quantity.

Step 2: the dense class. For $g(x) = x^k$, the hypothesis says exactly that $\Phi(g) = \int_0^1 f(x)^k \, dx - \frac{1}{k+1} = 0$. Since Φ is linear, $\Phi(P) = 0$ for every polynomial P. By the Weierstrass approximation theorem (which we use without proof), every continuous function on $[0,1]$ is the limit of a sequence of polynomials in the distance d: the polynomials form a dense class.

Step 3: the extension. By Theorem 1 a), $\Phi(g) = 0$ for every continuous function g. So f satisfies the hypothesis of Problem 6, hence $f(x) = x$ or $f(x) = 1 - x$; both also satisfy the hypothesis here, because $\int_0^1 x^k \, dx = \int_0^1 (1-x)^k \, dx = \frac{1}{k+1}$. □

4. When it fails

Each of the examples below violates one of the checks in the checklist of Section 2.4.

4.1 The dense class must lie in the hypothesis class: OJM 2025, grade 11, problem 2

Statement. Let $n \in \mathbb{N}$, $n \geq 3$. We say that a matrix $A \in \mathcal{M}_n(\mathbb{C})$ has property $(\mathcal{P})$ if $\det(A + X_{ij}) = \det(A + X_{ji})$ for all $i, j \in \{1, 2, \ldots, n\}$, where $X_{ij} \in \mathcal{M}_n(\mathbb{C})$ is the matrix with 1 in position (i, j) and 0 elsewhere.

(a) Prove that if $A \in \mathcal{M}_n(\mathbb{C})$ has property $(\mathcal{P})$ and $\det(A) \neq 0$, then $A = A^T$.

(b) Give an example of a matrix $A \in \mathcal{M}_n(\mathbb{C})$ that has property $(\mathcal{P})$ but $A \neq A^T$.

Source: [17]. The solution below is the one in the marking scheme.

Solution. a) Expanding along row i, $\det(A + X_{ij}) = \det A + \delta_{ij}$, where δ_{ij} is the cofactor of the entry in position (i, j). Property $(\mathcal{P})$ therefore means $\delta_{ij} = \delta_{ji}$, that is, the adjugate A^* is symmetric. For $d = \det A \neq 0$ we have $AA^* = dI_n = (dI_n)^T = (A^*A)^T = A^T(A^*)^T = A^T A^*$, and $\det A^* = d^{n-1} \neq 0$, so $A = A^T$. b) The matrix $A = X_{11} + X_{12}$: for $n \geq 3$, each $A + X_{ij}$ has at least one zero row, so all the determinants are 0.

The lesson. Invertible matrices are dense in $\mathcal{M}_n(\mathbb{C})$ (Lemma 3), the quantity $A \mapsto A - A^T$ is continuous and, by a), it vanishes on the invertible matrices with property $(\mathcal{P})$. Yet the conclusion does not extend: the matrix in b) has $(\mathcal{P})$ and is not symmetric. The explanation is check (c). The class on which the hypothesis holds is $\mathcal{O} = \{A : \ A \text{ has } (\mathcal{P})\}$, and its invertible members are *not* dense in $\mathcal{O}$: they are symmetric, and a limit of symmetric matrices is symmetric, so $X_{11} + X_{12} \in \mathcal{O}$ is not a limit of them. The perturbation $A \mapsto A + tI_n$, which worked in Problems 1 and 2, does not stay inside $\mathcal{O}$ here: for $A = X_{11} + X_{12}$ and $t \notin \{0, -1\}$, the matrix $A + tI_n$ is invertible (upper triangular, with determinant $(1 + t)t^{n-1}$) and not symmetric, so, by a), it does not have $(\mathcal{P})$.

4.2 Quantities that are not continuous: rank and similarity

Let $A = E_{12}$ and $B = E_{11}$, the pair from Step 4 of Problem 2: $AB = O_n$ and $BA = E_{12}$. For invertible A, $\operatorname{rank}(AB) = \operatorname{rank} B = \operatorname{rank}(BA)$, and invertible matrices are dense; yet here the ranks are 0 and 1. Rank is not a continuous quantity, because it does not have the Darboux property: on the continuous family $t \mapsto tI_n$, $t \in [0, 1]$, it takes the value 0 for $t = 0$ and the value n for $t \neq 0$, but no value in between (for example $\frac{1}{2}$), which, by Proposition 4, cannot happen for a continuous quantity. Check (b) fails. Similarly, for invertible A, $BA = A^{-1}(AB)A$ is similar to AB, but for the pair above $AB = O_n$ and $BA \neq O_n$ are not similar (the only matrix similar to O_n is O_n), even though they have the same characteristic polynomial, t^n (Section 2). Similarity cannot be written as an equality between continuous quantities.

4.3 The choice of distance matters: the length of a graph

For $k \geq 1$, let $g_k : [0, 1] \to \mathbb{R}$ be the function that, on each interval $\left[\frac{j}{k}, \frac{j+1}{k}\right]$, increases with slope 2 on the first half and is constant on the second half, with $g_k(0) = 0$; thus $g_k\left(\frac{j}{k}\right) = \frac{j}{k}$. Then $|g_k(x) - x| \leq \frac{1}{2k}$, so g_k converges uniformly to $g(x) = x$. However, on each interval the graph of g_k consists of a segment of length $\frac{\sqrt{5}}{2k}$ and one of length $\frac{1}{2k}$, so its total length is $\frac{1+\sqrt{5}}{2}$ for every k, while the graph of the limit has length $\sqrt{2}$. The length of the graph is not a continuous quantity with respect to the uniform distance (check (b)). The "largest deviation" distance controls values, not slopes.

4.4 The warnings, collected

- *A strict inequality becomes non-strict.* On $\mathcal{O} = [0, +\infty)$, with $\mathcal{D} = (0, +\infty)$ dense, $\Phi(x) = x$ and $\Psi(x) = 0$ satisfy $\Phi > \Psi$ on $\mathcal{D}$, but $\Phi(0) = \Psi(0)$.
- *Pointwise convergence does not preserve the integral*: the spike example in Section 2.2.
- *Concurrency can be lost in the limit.* The lines $y = 0$, $\frac{1}{k}x - y - 1 = 0$ and $\frac{1}{k}x + y - 1 = 0$ all pass through $(k, 0)$. Their coefficient triples tend to $(0, 1, 0)$, $(0, -1, -1)$ and $(0, 1, -1)$, that is, to the parallel lines $y = 0$, $y = -1$ and $y = 1$, which have no common point. This is why, when we express concurrency as the vanishing of the determinant of the coefficients, the first two lines must not be parallel (check (d), Exercise 4).

5. Exercises

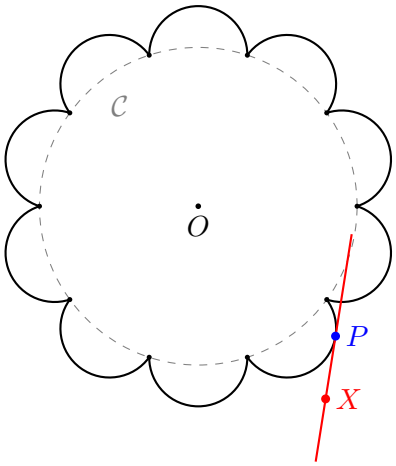
1. *(IMC 2004, day 1, problem 3 a) [18].)* Let $n \geq 2$ and let S_n be the set of sums $x_1 + \cdots + x_n$ with $0 \leq x_1, \ldots, x_n \leq \frac{\pi}{2}$ and $\sin x_1 + \cdots + \sin x_n = 1$. Prove that S_n is an interval.

Hint: with $y_k = \sin x_k$, S_n is the set of values of $\sum \arcsin y_k$ on $Y = \{y \in [0,1]^n : \ y_1 + \cdots + y_n = 1\}$. For $y, y' \in Y$, the points $(1-t)y + ty'$, $t \in [0,1]$, lie in Y and form a continuous family; the quantity $\sum \arcsin y_k$ is continuous (Proposition 2 b), with arcsin *continuous on $[0,1]$), so Proposition 4 gives all intermediate values. The official solution uses the notion of connectedness.*

2. Prove the following identities, first for invertible matrices and then by means of Theorem 1: a) $\det(I_n + AB) = \det(I_n + BA)$; b) $(AB)^* = B^*A^*$; c) $(A^*)^* = (\det A)^{n-2}A$, for $n \geq 3$; d) the Cayley-Hamilton theorem, $p_M(M) = O_n$, where p_M is the characteristic polynomial of M.

Hint: in b) the objects are the pairs (A, B), with $2n^2$ parameters, and they are approximated by the pairs $(A + t_k I_n, B + t_k I_n)$ in which both matrices are invertible. In c), for invertible A, $A^* = (\det A)A^{-1}$. In d), the dense class is the one from Step 4 of Problem 4: a matrix with n distinct eigenvalues can be written as $P\Lambda P^{-1}$ with Λ diagonal, and $p_M(P\Lambda P^{-1}) = P\, p_M(\Lambda)\, P^{-1} = O_n$.

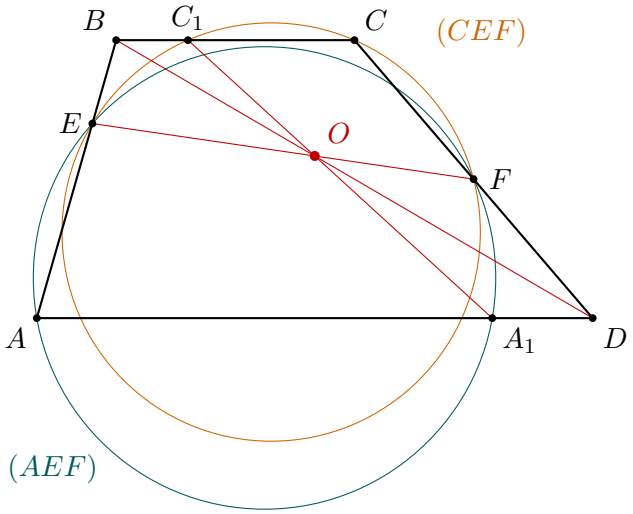
3. *(My own problem, part (a) of problem OJM.17 in [1].)* Let $n \geq 4$ and let $A_1A_2 \ldots A_n$ be a regular polygon inscribed in the circle $\mathcal{C}$ with center O and radius R (the planet). On each side A_iA_{i+1} (with $A_{n+1} = A_1$) we construct, outside $\mathcal{C}$, the semicircle with diameter A_iA_{i+1}. We call a *tangent* any line tangent to one of these semicircles, including at its endpoints (at an endpoint, the tangent is the perpendicular to the side). Prove that through every point X with $OX \geq R$ there passes a tangent.



The case $n = 10$: the tangent at the point P of a petal passes through the point X.

4. *(RMM Shortlist 2017, G1 [24].)* Let $ABCD$ be a trapezoid with $AD \parallel BC$, and let E and F be points in the interior of the sides $[AB]$ and $[DC]$, respectively. The circumcircle of triangle AEF meets the line AD again at A_1, and the circumcircle of triangle CEF meets the line BC again at C_1 (if the circle is tangent to the line, then $A_1 = A$, respectively $C_1 = C$). Prove that the lines BD, EF and A_1C_1 are concurrent.

Assume the problem has been solved in the case $EF \nparallel AD$ (a solution can be found in [25, pages 39 to 40]). How do you finish the problem, that is, how do you handle the case $EF \parallel AD$?



Case 1 ($EF \nparallel AD$): the lines BD, EF and A_1C_1 pass through the point O.

5. *(Putnam 2003, A2 [22].)* Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be nonnegative real numbers. Prove that

$$(a_1 a_2 \cdots a_n)^{1/n} + (b_1 b_2 \cdots b_n)^{1/n} \leq \left[(a_1 + b_1)(a_2 + b_2) \cdots (a_n + b_n)\right]^{1/n}.$$

Hint: the objects are the pairs of n-tuples with nonnegative entries, with $2n$ parameters, and both sides are continuous quantities (the n-th root is continuous on $[0, \infty)$, Proposition 2 b). The pairs with all $a_i > 0$ form a dense class (replace a_i by $a_i + \frac{1}{k}$). On this class, if some b_i is 0 the inequality is immediate; otherwise, dividing by $(a_1 \cdots a_n)^{1/n}$ and setting $y_i = \ln \frac{b_i}{a_i}$, it becomes Jensen's inequality for the convex function $y \mapsto \ln(1 + e^y)$. Theorem 1 b) finishes the proof.

6. *(Tournament of Towns, 1990/91, autumn round, basic version, grades 10 to 11, problem 4 [21].)* There are 20 points in the plane, no three of them collinear: 10 blue and 10 red. Prove that there is a line with 5 blue points and 5 red points on each side.

7. Prove that the scalene triangles (those whose side lengths are pairwise different) form a dense class in the class of nondegenerate triangles.

Hint: let ABC be nondegenerate, w a unit vector along the line AB, and $C_t = C + tw$. The point C_t stays at the same nonzero distance from the line AB, so ABC_t is nondegenerate. The condition $C_tA = C_tB$ means that C_t lies on the perpendicular bisector of AB, which meets the line $\{C + tw\}$ (perpendicular to it) at a single point; the condition $C_tA = AB$ means that C_t lies on a circle, which meets a line in at most two points; the same goes for $C_tB = AB$. At most five values of t are excluded.

Conclusion

We started from the high school notion of continuity, defined for functions of one real variable, and extended it to objects that are not numbers: matrices, functions, lines, geometric configurations. The goal was to see when we can formalize the idea of passing to the limit for such objects as well. The answer is the one in Section 2.4: we can pass to the limit when the property can be written as an equality or a non-strict inequality between continuous quantities and is proved on a dense class of objects; Theorem 1 then extends it to all objects.

The checklist (a) to (d) says what actually has to be written down: the parameters of the objects; the continuity of the quantities on the whole class (nonzero denominators, arguments in the domain of the square root or of arccos); a class that is dense in the very class on which the hypothesis holds; the way back from the equality between quantities to the required property. The examples in Section 4 show that each of these checks can fail.

Many of the problems in the text also have a solution without limits, which we have pointed out each time.

Bibliography

[1] V.-Ș. Constantinescu, *Analiză pentru Olimpiadă, clasa a XI-a* (in Romanian), second edition, manuscript, 2026 (forthcoming).

[2] V.-Ș. Constantinescu, *Analiză pentru Olimpiadă, clasa a XII-a* (in Romanian), second edition, manuscript, 2026 (forthcoming).

[3] V.-Ș. Constantinescu, *Algebră pentru Olimpiadă, clasa a XI-a* (in Romanian), second edition, manuscript, 2026 (forthcoming).

[4] V.-Ș. Constantinescu, *Algebră pentru Olimpiadă, clasa a XII-a* (in Romanian), second edition, manuscript, 2026 (forthcoming).

[5] V.-Ș. Constantinescu, *The Romanian Higher Algebra and Analysis Olympiad, Volume 1: Grade 11 Analysis. Complete Theory and Problems*, Second Edition, 2026. https://www.amazon.com/dp/B0HC7VS3B6

[6] V.-Ș. Constantinescu, *The Romanian Higher Algebra and Analysis Olympiad, Volume 2: Grade 12 Analysis. Complete Theory and Problems*, Second Edition, 2026. https://www.amazon.com/dp/B0HC7T685W

[7] V.-Ș. Constantinescu, *The Romanian Higher Algebra and Analysis Olympiad, Volume 3: Linear Algebra. Complete Theory and Problems*, Second Edition, 2026. https://www.amazon.com/dp/B0HC7SKYB8

[8] V.-Ș. Constantinescu, *The Romanian Higher Algebra and Analysis Olympiad, Volume 4: Abstract Algebra. Complete Theory and Problems*, Second Edition, 2026. https://www.amazon.com/dp/B0HHNJ2Q4M

[9] V.-Ș. Constantinescu, *L'Olympiade roumaine d'algèbre et d'analyse supérieures, Volume 1 : Analyse de onzième. Théorie complète et problèmes*, deuxième édition, 2026. https://www.amazon.fr/dp/B0HGZRNJJ4

[10] V.-Ș. Constantinescu, *L'Olympiade roumaine d'algèbre et d'analyse supérieures, Volume 2 : Analyse de douzième. Théorie complète et problèmes*, deuxième édition, 2026. https://www.amazon.fr/dp/B0HJZV48LB

[11] V.-Ș. Constantinescu, *L'Olympiade roumaine d'algèbre et d'analyse supérieures, Volume 3 : Algèbre linéaire. Théorie complète et problèmes*, deuxième édition, 2026. https://www.amazon.fr/dp/B0HHHH6X1Q

[12] V.-Ș. Constantinescu, *L'Olympiade roumaine d'algèbre et d'analyse supérieures, Volume 4 : Algèbre abstraite. Théorie complète et problèmes*, deuxième édition, 2026. https://www.amazon.fr/dp/B0HH111QMB

[13] Romanian National Olympiad (ONM), national round, Piatra-Neamț, 16 April 2022, grade 11, solutions and marking schemes, SSMR. https://ssmr.ro/files/onm2022/finala/cl11_nationala.pdf

[14] Romanian Olympiad, county round (OJM), 2026, grade 11, problems. https://cdn.edupedu.ro/wp-content/uploads/2026/03/Cl_11_a_subiecte_OJM_2026_RO.pdf

[15] Romanian Olympiad, county round (OJM), 2026, grade 11, solutions and marking schemes. https://cdn.edupedu.ro/wp-content/uploads/2026/03/Cl_11_a_solutii_bareme_OJM_2026.pdf

[16] Romanian Olympiad, county round (OJM), 2004, grade 12, problems and solutions (transcribed in the author's ROHS database).

[17] Romanian National Olympiad, county round (OJM), 2025, grade 11, solutions, SSMR. http://ssmr.ro/files/onm2025/faza_municipiu/barem_clasa11.pdf

[18] 11th International Mathematics Competition for University Students (IMC), Skopje, 2004, Day 1, Problems and Solutions. https://www.imc-math.org.uk/archive/problems/2004/IMC_2004_Day_1_Problems_and_Solutions.pdf

[19] Gazeta Matematică, no. 6-8/2025, proposed problem.

[20] Romanian National Mathematical Olympiad, 2009, grade 12, problem 4.

[21] Tournament of Towns, 1990/91, autumn round, basic version, grades 10 to 11, problem 4, on problems.ru, problem 98078. https://problems.ru/view_problem_details_new.php?id=98078

[22] The 64th William Lowell Putnam Mathematical Competition, 2003, problem A2; problems and solutions in K. Kedlaya's archive. https://kskedlaya.org/putnam-archive/2003.tex

[23] 11th International Mathematics Competition for University Students (IMC), Skopje, 2004, Day 2, Problems and Solutions. https://www.imc-math.org.uk/archive/problems/2004/IMC_2004_Day_2_Problems_and_Solutions.pdf

[24] M. Rakovsky, *Progresser en géométrie*, POFM course (Préparation Olympique Française de Mathématiques). https://maths-olympiques.fr/wp-content/uploads/2023/02/cours.pdf

[25] M. Rakovsky, *Progresser en géométrie. Solutions*, POFM. https://maths-olympiques.fr/wp-content/uploads/2023/02/solutions.pdf