GAZETA MATEMATICĂ

SERIA A

ANUL XXXII (CXI) **Nr. 3 – 4/ 2014**

ARTICOLE

Functions for which mixed partial derivatives are distinct

Dumitru Popa[1)]

Abstract. Let $h : \mathbb{R}^3 \to \mathbb{R}$ be a function of class C^1 and let $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x_1, x_2) = \begin{cases} h\left(\dfrac{x_1^2}{x_1^2+x_2^2}, \dfrac{x_2^2}{x_1^2+x_2^2}, x_1 x_2\right) & \text{if} \quad (x_1, x_2) \neq (0,0), \\ 0 & \text{if} \quad (x_1, x_2) = (0,0). \end{cases}$$

We find necessary and sufficient conditions for f to be continuous at $(0,0)$, there exist $\frac{\partial f}{\partial x_1}(0,0)$, be Fréchet differentiable at $(0,0)$ and having partial derivatives $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0)$, $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$. We also show that this result can be extended to a real linear space endowed with a scalar product.

Keywords: Fréchet differentiable, mixed partial derivative, Schwartz and Young theorem, Hilbert spaces, real-valued functions

MSC: Primary 26B05; Secondary 54C30.

Introduction

The Schwarz theorem asserts that if $f : \mathbb{R}^2 \to \mathbb{R}$ is twice Fréchet differentiable at a point $a \in \mathbb{R}^2$, then $f''(a)$ is symmetric, i.e., $f''(a)(x,y) = f''(a)(y,x)$ for all $x, y \in \mathbb{R}^2$. In particular, there exist $\frac{\partial^2 f}{\partial x_1 \partial x_2}(a)$, $\frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$ and $\frac{\partial^2 f}{\partial x_1 \partial x_2}(a) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a)$, see [2, Propoziţia 11, p. 100]. Also the Young theorem asserts that if $f : \mathbb{R}^2 \to \mathbb{R}$ is such that there exist $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$, and there exists and is continuous $\frac{\partial^2 f}{\partial x_1 \partial x_2} : \mathbb{R}^2 \to \mathbb{R}$, then there exists $\frac{\partial^2 f}{\partial x_2 \partial x_1} : \mathbb{R}^2 \to \mathbb{R}$ and $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$; see [2, Propoziţia 10, p. 99]. As it is well known, the Young theorem and Fubini's theorem are equivalent and true; for more details see [1] and the references therein. The standard example of a function $f : \mathbb{R}^2 \to \mathbb{R}$ for which there exist $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0)$, $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$ and are different

[1)]Department of Mathematics, Ovidius University of Constanţa, Romania, dpopa@univ-ovidius.ro

is $f\left(x_1,x_2\right)=\begin{cases}\frac{x_1x_2\left(x_1^2-x_2^2\right)}{x_1^2+x_2^2} \text{if}(x_1,x_2)\neq(0,0)\\ 0 \quad \text{if}(x_1,x_2)=(0,0)\end{cases}$; see [2, p. 113, exercise 1], [5, p. 186, exercise 6], [6, Example 2.7.1, p. 61]. In this paper we extend the above example, and further we show that this can be also extended to the context of real Hilbert spaces.

The notation and definitions used in this paper are standard; see [2, 3, 4, 5, 7].

Throughout this paper we denote by $e_1=(1,0,0)$, $e_2=(0,1,0)$ and $e_3=(0,0,1)$ the standard unit vectors in $\mathbb{R}^3$.

1. The real case

Proposition 1. *Let $h:\mathbb{R}^3\to\mathbb{R}$ be a function of class C^1 and let $f:\mathbb{R}^2\to\mathbb{R}$ defined by*

$$f\left(x_1,x_2\right)=\begin{cases} h\left(\frac{x_1^2}{x_1^2+x_2^2},\frac{x_2^2}{x_1^2+x_2^2},x_1x_2\right) & \text{if} \quad (x_1,x_2)\neq(0,0)\,,\\ 0 & \text{if} \quad (x_1,x_2)=(0,0)\,.\end{cases}$$

Then:
(i) f is continuous at $(0,0)$ if and only if $h\left(y_1,1-y_1,0\right)=0$ for all $y_1\in[0,1]$.
(ii) there exists $\frac{\partial f}{\partial x_1}(0,0)$ if and only if $h(e_1)=0$. In this case $\frac{\partial f}{\partial x_1}(0,0)=0$.
(iii) there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if $h(e_2)=0$. In this case $\frac{\partial f}{\partial x_2}(0,0)=0$.
(iv) f is Fréchet differentiable at $(0,0)$ if and only if $h\left(y_1,1-y_1,0\right)=0$ for all $y_1\in[0,1]$.
(v) there exists $\frac{\partial^2 f}{\partial x_1\partial x_2}(0,0)$ if and only if $h(e_2)=0$. In this case
$\frac{\partial^2 f}{\partial x_1\partial x_2}(0,0)=\frac{\partial h}{\partial y_3}(e_1)$.
(vi) there exists $\frac{\partial^2 f}{\partial x_2\partial x_1}(0,0)$ if and only if $h(e_1)=0$. In this case
$\frac{\partial^2 f}{\partial x_2\partial x_1}(0,0)=\frac{\partial h}{\partial y_3}(e_2)$.

Proof. First we recall a well known result: h is of class C^1 if and only if there exist and are continuous $\frac{\partial h}{\partial y_1},\frac{\partial h}{\partial y_2},\frac{\partial h}{\partial y_3}:\mathbb{R}^3\to\mathbb{R}$.

(i) Let us suppose that f is continuous at $(0,0)$. Then

$$\lim_{(x_1,x_2)\to(0,0)} f\left(x_1,x_2\right)=f(0,0)=0.$$

In particular, $\lim_{x_2\to 0} f\left(0,x_2\right)=0$. Since for all $x_2\neq 0$, $f\left(0,x_2\right)=h(0,1,0)$, it follows that $h(0,1,0)=0$. Let $0<y_1\leq 1$. Then, as it is well known, $\lim_{x_1\to 0} f\left(x_1,x_1\sqrt{\frac{1-y_1}{y_1}}\right)=0$, i.e., $\lim_{x_1\to 0} h\left(y_1,1-y_1,x_1^2\sqrt{\frac{1-y_1}{y_1}}\right)=0$. Since h is continuous,

$$\lim_{x_1\to 0} h\left(y_1,1-y_1,x_1^2\sqrt{\frac{1-y_1}{y_1}}\right)=h\left(y_1,1-y_1,0\right),$$

thus $h(y_1, 1-y_1, 0) = 0$.

Conversely, suppose that $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$.

Let $\psi : [0,1] \times \mathbb{R} \to \mathbb{R}$ be the function

$$\psi(y_1, y_3) = \begin{cases} \dfrac{h(y_1, 1-y_1, y_3) - h(y_1, 1-y_1, 0)}{y_3} & \text{if} \quad y_3 \neq 0 \\ \dfrac{\partial h}{\partial y_3}(y_1, 1-y_1, 0), & \text{if} \quad y_3 = 0 \end{cases}.$$

Let us note that since h and $\frac{\partial h}{\partial y_3}$ are continuous it follows that ψ is continuous. Also $h(y_1, 1-y_1, y_3) - h(y_1, 1-y_1, 0) = y_3 \psi(y_1, y_3)$ for all $(y_1, y_3) \in [0,1] \times \mathbb{R}$ and by hypothesis

$$h(y_1, 1-y_1, y_3) = y_3 \psi(y_1, y_3) \text{ for all } (y_1, y_3) \in [0,1] \times \mathbb{R}.$$

Let $(x_1, x_2) \in [-1,1]^2$ with $(x_1, x_2) \neq (0,0)$. Then

$$f(x_1, x_2) = h\left(\frac{x_1^2}{x_1^2 + x_2^2}, \frac{x_2^2}{x_1^2 + x_2^2}, x_1 x_2\right) = x_1 x_2 \psi\left(\frac{x_1^2}{x_1^2 + x_2^2}, x_1 x_2\right)$$

and

$$\begin{aligned} |f(x_1, x_2)| &= |x_1 x_2| \left|\psi\left(\frac{x_1^2}{x_1^2 + x_2^2}, x_1 x_2\right)\right| \\ &\leq |x_1 x_2| \sup_{(y_1, y_3) \in [0,1] \times [-1,1]} |\psi(y_1, y_3)| = M |x_1 x_2| \end{aligned}$$

(since ψ is continuous and $[0,1] \times [-1,1]$ is a compact set, by the Weierstrass theorem the supremum is finite and attained). From here we get that $\lim_{(x_1,x_2) \to (0,0)} f(x_1, x_2) = 0 = f(0,0)$, i.e., f is continuous at $(0,0)$.

(ii) By definition there exists $\frac{\partial f}{\partial x_1}(0,0)$ if and only if the real function $x_1 \mapsto f(x_1, 0)$ is derivable at 0. Now, $f(x_1, 0) = \begin{cases} h(e_1) & \text{if} \quad x_1 \neq 0 \\ 0 & \text{if} \quad x_1 = 0 \end{cases}$. Thus, as it is easy to prove, there exists $\frac{\partial f}{\partial x_1}(0,0)$ if and only if $h(e_1) = 0$ and moreover, $\frac{\partial f}{\partial x_1}(0,0) = 0$.

(iii) By definition there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if the real function $x_2 \mapsto f(0, x_2)$ is derivable at 0. Since $f(0, x_2) = \begin{cases} h(e_2) & \text{if} \quad x_2 \neq 0 \\ 0 & \text{if} \quad x_2 = 0 \end{cases}$, there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if $h(e_2) = 0$ and moreover, $\frac{\partial f}{\partial x_2}(0,0) = 0$.

(iv) Suppose that f is Fréchet differentiable at $(0,0)$. Then, as it is well known, f is continuous at $(0,0)$ and from (i) we deduce that $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$.

Conversely, assume that $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$. Then, in particular, $h(e_1) = h(1,0,0) = 0$, $h(e_2) = h(0,1,0) = 0$, and from (ii) and (iii) we have $\frac{\partial f}{\partial x_1}(0,0) = 0$, $\frac{\partial f}{\partial x_2}(0,0) = 0$. As it is well known f is Fréchet differentiable at $(0,0)$ if and only if

$$\lim_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2)-f(0,0)-\dfrac{\partial f}{\partial x_1}(0,0)\,x_1-\dfrac{\partial f}{\partial x_2}(0,0)\,x_2}{\sqrt{x_1^2+x_2^2}}=0,$$

i.e., $\lim\limits_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2)}{\sqrt{x_1^2+x_2^2}}=0$. To prove this, we recall that in (i) we have shown that

$$|f(x_1,x_2)|\le M\,|x_1x_2| \text{ for all } (x_1,x_2)\in[-1,1]^2 \text{ with } (x_1,x_2)\ne(0,0)$$

and, since $\frac{|x_1|}{\sqrt{x_1^2+x_2^2}}\le 1$, we deduce

$$\frac{|f(x_1,x_2)|}{\sqrt{x_1^2+x_2^2}}\le M\,|x_2| \text{ for all } (x_1,x_2)\in[-1,1]^2 \text{ with } (x_1,x_2)\ne(0,0).$$

From here we get that $\lim\limits_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2)}{\sqrt{x_1^2+x_2^2}}=0$.

(v) Let us suppose that there exists $\frac{\partial^2 f}{\partial x_1\partial x_2}(0,0)$. In particular, there exists $\frac{\partial f}{\partial x_2}(0,0)$ and by (iii), $h(e_2)=0$ and $\frac{\partial f}{\partial x_2}(0,0)=0$.

Conversely, let us suppose that $h(e_2)=0$. Let $(x_1,x_2)\ne(0,0)$. By the chain rule we have

$$\frac{\partial f}{\partial x_2}=\frac{\partial h}{\partial y_1}\cdot\frac{\partial y_1}{\partial x_2}+\frac{\partial h}{\partial y_2}\cdot\frac{\partial y_2}{\partial x_2}+\frac{\partial h}{\partial y_3}\cdot\frac{\partial y_3}{\partial x_2},$$

i.e.,

$$\frac{\partial f}{\partial x_2}=-\frac{2x_1^2x_2}{\left(x_1^2+x_2^2\right)^2}\cdot\frac{\partial h}{\partial y_1}+\frac{2x_1^2x_2}{\left(x_1^2+x_2^2\right)^2}\cdot\frac{\partial h}{\partial y_2}+x_1\cdot\frac{\partial h}{\partial y_3}.$$

Above we used the usual convention that we wrote $\frac{\partial f}{\partial x_1}$ instead of $\frac{\partial f}{\partial x_1}(x_1,x_2)$ and $\frac{\partial h}{\partial y_1}$ instead of $\frac{\partial h}{\partial y_1}\left(\frac{x_1^2}{x_1^2+x_2^2},\frac{x_2^2}{x_1^2+x_2^2},x_1x_2\right)$, etc. Then

$$\frac{\partial f}{\partial x_2}(x_1,0)=\begin{cases} x_1\cdot\dfrac{\partial h}{\partial y_3}(1,0,0) & \text{for } x_1\ne 0\\ 0 & \text{for } x_1=0\end{cases}=x_1\cdot\frac{\partial h}{\partial y_3}(e_1) \text{ for } x_1\in\mathbb{R}.$$

By definition, we have

$$\frac{\partial^2 f}{\partial x_1\partial x_2}(0,0)=\lim_{x_1\to0}\frac{\dfrac{\partial f}{\partial x_2}(x_1,0)-\dfrac{\partial f}{\partial x_2}(0,0)}{x_1}=\frac{\partial h}{\partial y_3}(e_1).$$

(vi) Let us suppose that $\frac{\partial^2 f}{\partial x_2\partial x_1}(0,0)$ exists. Then, in particular, there exists $\frac{\partial f}{\partial x_1}(0,0)$, which by (ii) gives us $h(e_1)=0$ and $\frac{\partial f}{\partial x_1}(0,0)=0$.

Conversely, let us suppose that $h(e_1) = 0$. Let $(x_1, x_2) \neq (0,0)$. By the chain rule we have

$$\frac{\partial f}{\partial x_1} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial h}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} + \frac{\partial h}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_1},$$

i.e.,

$$\frac{\partial f}{\partial x_1} = \frac{2x_1 x_2^2}{\left(x_1^2 + x_2^2\right)^2} \cdot \frac{\partial h}{\partial y_1} - \frac{2x_1 x_2^2}{\left(x_1^2 + x_2^2\right)^2} \cdot \frac{\partial h}{\partial y_2} + x_2 \cdot \frac{\partial h}{\partial y_3}$$

with the same convention as above. Then

$$\frac{\partial f}{\partial x_1}(0, x_2) = \begin{cases} x_2 \cdot \frac{\partial h}{\partial y_3}(0,1,0) & \text{for} \quad x_2 \neq 0 \\ 0 & \text{for} \quad x_2 = 0 \end{cases} = x_2 \cdot \frac{\partial h}{\partial y_3}(e_2) \text{ for } x_2 \in \mathbb{R}.$$

Thus, by definition we have

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = \lim_{x_2 \to 0} \frac{\frac{\partial f}{\partial x_1}(0, x_2) - \frac{\partial f}{\partial x_1}(0,0)}{x_2} = \frac{\partial h}{\partial y_3}(e_2).$$

□

Taking $h : \mathbb{R}^3 \to \mathbb{R}$, $h(y_1, y_2, y_3) = (y_1 - y_2) y_3$ in Proposition 1, we get the Dieudonné's example.

2. Real Hilbert spaces

One of the main features of the example given in Proposition 1 is that it can be extended to the context of real Hilbert spaces, more precisely to real linear spaces endowed with a scalar product. Even, maybe, for some readers this extension is almost obvious, we give the full details. We recall some definitions and results. Throughout the rest of the paper we denote by H a real linear space endowed with a scalar product $\langle , \rangle$, $\|x\| = \sqrt{\langle x, x \rangle}$ for all $x \in H$ and $B_H = \{x \in H \mid \|x\| \leq 1\}$ is the closed unit ball of H. By $I_H : H \to H$ we denote the identity operator, i.e., $I_H(x) = x$ for $x \in H$ and $L(H) = \{A : H \to H \mid A \text{ is linear and continuous}\}$ endowed with the operator norm.

Let $f : H \to \mathbb{R}$ and $a \in H$. By definition, the function f is *Fréchet differentiable at* a if and only if there exists $f'(a) \in H$ such that

$$\lim_{x \to a} \frac{f(x) - f(a) - \langle f'(a), x - a \rangle}{\|x - a\|} = 0.$$

We need the following results:

R1) Let $\varphi : H \to \mathbb{R}$, $\varphi(x) = \|x\|^2$. Then $\varphi'(a) = 2a$ for all $a \in H$. This follows simply by definition.

R2) If we have the maps $H \xrightarrow{f} \mathbb{R} \xrightarrow{\varphi} \mathbb{R}$ such that f is Fréchet differentiable at $a \in H$, and φ derivable at $f(a)$, then $\varphi \circ f$ is Fréchet differentiable at a and $(\varphi \circ f)'(a) = \varphi'(f(a)) f'(a)$. This is the well known theorem of differentiability of the composition of functions, see [2, 3, 4, 5, 7].

R3) If $\lambda > 0$, then $P : H \to \mathbb{R}$ defined by $P(x) = \frac{1}{\|x\|^2+\lambda}$ is Fréchet differentiable at $x \in H$ and $P'(x) = -\frac{2x}{\left(\|x\|^2+\lambda\right)^2}$. This follows from R1) and R2).

Let $f : H \to H$ and $a \in H$. By definition, the function f is Fréchet differentiable at a if and only if there exists $f'(a) \in L(H)$ such that

$$\lim_{x \to a} \frac{f(x) - f(a) - f'(a)(x-a)}{\|x-a\|} = 0;$$

see [2, 3, 4, 5, 7].

R4) Let $A \in L(H)$ and $f : H \to H$, $f(x) = A(x)$. Then f is Fréchet differentiable at every $a \in H$ and $f'(a) = A$ for all $a \in H$. In particular, if $c \in \mathbb{R}$ and $f : H \to H$ is defined by $f(x) = cx$, then $f'(a) = cI_H$ for all $a \in H$. This follows by definition; see [2, 3, 4, 5, 7].

On the cartesian product $H \times H$ we consider the natural scalar product

$$\langle (x_1, x_2), (y_1, y_2) \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle .$$

Let $f : H \times H \to \mathbb{R}$ and $(a_1, a_2) \in H \times H$. We say that f is *Fréchet differentiable at* (a_1, a_2) *with respect to the first variable* if and only if the function $v_1 : H \to \mathbb{R}$, $v_1(x_1) = f(x_1, a_2)$, is Fréchet differentiable at a_1. Similarly, f is *Fréchet differentiable at* (a_1, a_2) *with respect to the second variable* if and only if the function $v_2 : H \to \mathbb{R}$, $v(x_2) = f(a_1, x_2)$, is Fréchet differentiable at a_2.

Let $f : H \times H \to \mathbb{R}$ and $(a_1, a_2) \in H \times H$. We say that f is *twice Fréchet differentiable at* (a_1, a_2) *with respect to the variables* x_1 *and* x_2 if and only if there exists $\frac{\partial f}{\partial x_2} : H \times H \to H$ and the function $g_1 = \frac{\partial f}{\partial x_2} : H \times H \to H$ is Fréchet differentiable with respect to x_1 at (a_1, a_2). In this case, $\frac{\partial g_1}{\partial x_1}(a) \overset{\text{def}}{=} \frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2) \in L(H)$.

Indeed, let us note that $f : H \times H \to H$ is twice Fréchet differentiable with respect to x_1 at (a_1, a_2) if and only if the function $v_1 : H \to H$ defined by $v_1(x_1) = \frac{\partial f}{\partial x_2}(x_1, a_2)$ is Fréchet differentiable at a_1, so $\frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2) = v_1'(a_1) \in L(H)$.

Similarly, the function $f : H \times H \to \mathbb{R}$ is *twice Fréchet differentiable at* $(a_1, a_2) \in H \times H$ *with respect to the variables* x_2 *and* x_1 if and only if there exists $\frac{\partial f}{\partial x_1} : H \times H \to H$ and the function $g_2 = \frac{\partial f}{\partial x_1} : H \times H \to H$ is Fréchet differentiable with respect to x_2 at (a_1, a_2). In this case, $\frac{\partial g_2}{\partial x_2}(a) \overset{\text{def}}{=} \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2) \in L(H)$.

R5) The chain rule. Let us consider the maps $H \times H \overset{y_1}{\to} \mathbb{R}$, $H \times H \overset{y_2}{\to} \mathbb{R}$, $H \times H \overset{y_3}{\to} \mathbb{R}$, $\mathbb{R}^3 \overset{h}{\to} \mathbb{R}$ and let $f : H \times H \to \mathbb{R}$ be defined by

$$f(x_1, x_2) = h(y_1(x_1, x_2), y_2(x_1, x_2), y_3(x_1, x_2))$$

and $(a_1, a_2) \in H \times H$. If y_1, y_2, y_3 are Fréchet differentiable at (a_1, a_2) and h is Fréchet differentiable at $(y_1(a_1,a_2), y_2(a_1,a_2), y_3(a_1,a_2))$, then f is Fréchet differentiable at (a_1, a_2) and the chain rule holds:

$$\frac{\partial f}{\partial x_1} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial h}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} + \frac{\partial h}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_1},$$
$$\frac{\partial f}{\partial x_2} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} + \frac{\partial h}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_2} + \frac{\partial h}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_2},$$

with the usual convention that by $\frac{\partial f}{\partial x_1}$ (respectively $\frac{\partial h}{\partial y_1}$) we understand $\frac{\partial f}{\partial x_1}(a_1, a_2)$ (respectively $\frac{\partial h}{\partial y_1}(y_1(a_1,a_2), y_2(a_1,a_2), y_3(a_1,a_2))$); see [2, 3, 4, 5, 7]. Let us note that $\frac{\partial h}{\partial y_1} = \frac{\partial h}{\partial y_1}(y_1(a_1,a_2), y_2(a_1,a_2), y_3(a_1,a_2)) \in \mathbb{R}$, $\frac{\partial y_1}{\partial x_1} = \frac{\partial y_1}{\partial x_1}(a_1, a_2) \in H$.

We also need the following well known result.

R6) If $f : H \times H \to \mathbb{R}$ is such that there exists $\lim_{(x,y)\to(0,0)} f(x,y) \in \mathbb{R}$, then $\lim_{x\to 0} f(x, mx) = \lim_{y\to 0} f(ny, y) = \lim_{(x,y)\to(0,0)} f(x,y)$ for all $m, n \in \mathbb{R}$.

With this preparation we are ready to prove the extension of Proposition 1 to Hilbert spaces.

Proposition 2. *Let $(H, \langle , \rangle)$ be a real linear space endowed with a scalar product, $h : \mathbb{R}^3 \to \mathbb{R}$ a function of class C^1 and let $f : H \times H \to \mathbb{R}$,*

$$f(x_1, x_2) = \begin{cases} h\left(\frac{\|x_1\|^2}{\|x_1\|^2+\|x_2\|^2}, \frac{\|x_2\|^2}{\|x_1\|^2+\|x_2\|^2}, \langle x_1, x_2\rangle\right) & \text{if } (x_1,x_2) \neq (0,0) \\ 0 & \text{if } (x_1,x_2) = (0,0) \end{cases}.$$

Then:
(i) f is continuous at $(0,0)$ if and only if $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$.
(ii) there exists $\frac{\partial f}{\partial x_1}(0,0)$ if and only if $h(e_1) = 0$. In this case $\frac{\partial f}{\partial x_1}(0,0) = 0$.
(iii) there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if $h(e_2) = 0$. In this case $\frac{\partial f}{\partial x_2}(0,0) = 0$.
(iv) f is Fréchet differentiable at $(0,0)$ if and only if $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$.
(v) there exists $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0)$ if and only if $h(e_2) = 0$. In this case $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = \frac{\partial h}{\partial y_3}(e_1) I_H$.
(vi) there exists $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$ if and only if $h(e_1) = 0$. In this case $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = \frac{\partial h}{\partial y_3}(e_2) I_H$.

Proof. Again h is of class C^1 if and only if there exist and are continuous $\frac{\partial h}{\partial y_1}, \frac{\partial h}{\partial y_2}, \frac{\partial h}{\partial y_3} : \mathbb{R}^3 \to \mathbb{R}$.

(i) Let us suppose that f is continuous at $(0,0)$. Then

$$\lim_{(x_1,x_2)\to(0,0)} f(x_1, x_2) = f(0,0) = 0.$$

In particular, $\lim_{x_2 \to 0} f(0, x_2) = 0$. Since for all $x_2 \neq 0$, $f(0, x_2) = h(0, 1, 0)$, it follows that $h(0, 1, 0) = 0$. Let $0 < y_1 \leq 1$. Then, from R6 we have $\lim_{x_1 \to 0} f\left(x_1, \sqrt{\frac{1-y_1}{y_1}} x_1\right) = 0$, i.e., $\lim_{x_1 \to 0} h\left(y_1, 1-y_1, \|x_1\|^2 \sqrt{\frac{1-y_1}{y_1}}\right) = 0$. Since h is continuous,

$$\lim_{x_1 \to 0} h\left(y_1, 1-y_1, \|x_1\|^2 \sqrt{\frac{1-y_1}{y_1}}\right) = h(y_1, 1-y_1, 0),$$

thus $h(y_1, 1-y_1, 0) = 0$.

Conversely, suppose that $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0, 1]$.

Let $\psi : [0, 1] \times \mathbb{R} \to \mathbb{R}$ be defined by

$$\psi(y_1, y_3) = \begin{cases} \frac{h(y_1, 1-y_1, y_3) - h(y_1, 1-y_1, 0)}{y_3} & \text{if} \quad y_3 \neq 0, \\ \frac{\partial h}{\partial y_3}(y_1, 1-y_1, 0) & \text{if} \quad y_3 = 0. \end{cases}$$

Let us note that since h and $\frac{\partial h}{\partial y_3}$ are continuous, ψ is continuous. Also $h(y_1, 1-y_1, y_3) - h(y_1, 1-y_1, 0) = y_3 \psi(y_1, y_3)$ for all $(y_1, y_3) \in [0, 1] \times \mathbb{R}$ and by hypothesis

$$h(y_1, 1-y_1, y_3) = y_3 \psi(y_1, y_3) \text{ for all } (y_1, y_3) \in [0, 1] \times \mathbb{R}.$$

Let $(x_1, x_2) \in B_H \times B_H$ with $(x_1, x_2) \neq (0, 0)$. Then

$$\begin{aligned} f(x_1, x_2) &= h\left(\frac{\|x_1\|^2}{\|x_1\|^2 + \|x_2\|^2}, \frac{\|x_2\|^2}{\|x_1\|^2 + \|x_2\|^2}, \langle x_1, x_2 \rangle\right) \\ &= \langle x_1, x_2 \rangle \, \psi\left(\frac{\|x_1\|^2}{\|x_1\|^2 + \|x_2\|^2}, \langle x_1, x_2 \rangle\right). \end{aligned}$$

From Cauchy-Bunyakovsky-Schwarz inequality we have $|\langle x_1, x_2 \rangle| \leq \|x_1\| \|x_2\|$ and since $(x_1, x_2) \in B_H \times B_H$ we deduce $|\langle x_1, x_2 \rangle| \leq 1$, i.e., $\langle x_1, x_2 \rangle \in [-1, 1]$. We have

$$\begin{aligned} |f(x_1, x_2)| &= |\langle x_1, x_2 \rangle| \left|\psi\left(\frac{\|x_1\|^2}{\|x_1\|^2 + \|x_2\|^2}, \langle x_1, x_2 \rangle\right)\right| \\ &\leq |\langle x_1, x_2 \rangle| \sup_{(y_1, y_3) \in [0,1] \times [-1,1]} |\psi(y_1, y_3)| = M \|x_1\| \|x_2\|. \end{aligned}$$

Now we get that $\lim_{(x_1, x_2) \to (0,0)} f(x_1, x_2) = 0 = f(0, 0)$, i.e., f is continuous at $(0, 0)$.

(ii) By definition there exists $\frac{\partial f}{\partial x_1}(0, 0)$ if and only if the function $H \ni x_1 \mapsto f(x_1, 0) \in \mathbb{R}$ is Fréchet differentiable at 0.

Now, $f(x_1, 0) = \begin{cases} h(e_1) & \text{if} \quad x_1 \neq 0 \\ 0 & \text{if} \quad x_1 = 0 \end{cases}$. If this function is Fréchet differentiable at 0, then it is continuous at 0 and thus $h(e_1) = 0$. Conversely, if $h(e_1) = 0$ then $f(x_1, 0) = 0$ for all $x_1 \in H$ and thus $\frac{\partial f}{\partial x_1}(0, 0) = 0$.

(iii) By definition there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if the function $H \ni x_2 \mapsto f(0,x_2) \in \mathbb{R}$ is Fréchet differentiable at 0.

From $f(0,x_2) = \begin{cases} h(e_2) & \text{if } x_2 \neq 0 \\ 0 & \text{if } x_2 = 0 \end{cases}$, as in (ii) we deduce that there exists $\frac{\partial f}{\partial x_2}(0,0)$ if and only if $h(e_2) = 0$ and moreover $\frac{\partial f}{\partial x_2}(0,0) = 0$.

(iv) Let us suppose that f is Fréchet differentiable at $(0,0)$. Then f is continuous at $(0,0)$ and from (i) we deduce $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$.

Conversely, let us suppose that $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$. Then, in particular, $h(e_1) = h(1,0,0) = 0$, $h(e_2) = h(0,1,0) = 0$ and from (ii) and (iii) we get $\frac{\partial f}{\partial x_1}(0,0) = 0$, $\frac{\partial f}{\partial x_2}(0,0) = 0$. As it is well known, f is Fréchet differentiable at $(0,0)$ if and only if

$$\lim_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2) - f(0,0) - \left\langle \frac{\partial f}{\partial x_1}(0,0), x_1 \right\rangle - \left\langle \frac{\partial f}{\partial x_2}(0,0), x_2 \right\rangle}{\sqrt{\|x_1\|^2 + \|x_2\|^2}} = 0$$

(see [2, 3, 4, 5, 7]), which is equivalent to $\lim\limits_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2)}{\sqrt{\|x_1\|^2+\|x_2\|^2}} = 0$. To prove this, recall that since $h(y_1, 1-y_1, 0) = 0$ for all $y_1 \in [0,1]$, in (i) we have shown that

$$|f(x_1,x_2)| \leq M \|x_1\| \|x_2\| \text{ for all } (x_1,x_2) \in B_H \times B_H \text{ with } (x_1,x_2) \neq (0,0).$$

Since $\frac{\|x_1\|}{\sqrt{\|x_1\|^2+\|x_2\|^2}} \leq 1$, we deduce

$$\frac{|f(x_1,x_2)|}{\sqrt{\|x_1\|^2 + \|x_2\|^2}} \leq M \|x_2\| \text{ for all } (x_1,x_2) \in B_H \times B_H \text{ with } (x_1,x_2) \neq (0,0).$$

From here we get $\lim\limits_{(x_1,x_2)\to(0,0)} \frac{f(x_1,x_2)}{\sqrt{\|x_1\|^2+\|x_2\|^2}} = 0$.

(v) Let us suppose that there exists $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0)$. In particular, there exist $\frac{\partial f}{\partial x_2}(0,0)$ and, by (iii), $h(e_2) = 0$ and $\frac{\partial f}{\partial x_2}(0,0) = 0$.

Conversely, let us suppose that $h(e_2) = 0$. Let $(x_1,x_2) \in H \times H$, $(x_1,x_2) \neq (0,0)$. By the chain rule R5, we have

$$\frac{\partial f}{\partial x_2} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} + \frac{\partial h}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_2} + \frac{\partial h}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_2},$$

i.e., by R3

$$\frac{\partial f}{\partial x_2} = -\frac{\partial h}{\partial y_1} \cdot \frac{2\|x_1\|^2 x_2}{\left(\|x_1\|^2 + \|x_2\|^2\right)^2} + \frac{\partial h}{\partial y_2} \cdot \frac{2\|x_1\|^2 x_2}{\left(\|x_1\|^2 + \|x_2\|^2\right)^2} + \frac{\partial h}{\partial y_3} \cdot x_1.$$

Above we used the usual convention that we wrote $\frac{\partial f}{\partial x_1}$ instead of $\frac{\partial f}{\partial x_1}(x_1, x_2)$ and $\frac{\partial h}{\partial y_1}$ instead of $\frac{\partial h}{\partial y_1}\left(\frac{\|x_1\|^2}{\|x_1\|^2+\|x_2\|^2}, \frac{\|x_2\|^2}{\|x_1\|^2+\|x_2\|^2}, \langle x_1, x_2\rangle\right)$, etc. Then

$$\frac{\partial f}{\partial x_2}(x_1, 0) = \begin{cases} \frac{\partial h}{\partial y_3}(1,0,0)\, x_1 & \text{for} \quad x_1 \neq 0 \\ 0 & \text{for} \quad x_1 = 0 \end{cases} = \frac{\partial h}{\partial y_3}(e_1)\, x_1 \text{ for } x_1 \in H.$$

From R4 we deduce $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = \frac{\partial h}{\partial y_3}(e_1)\, I_H$.

(vi) Let us suppose that $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0)$ exists. Then, in particular, there exists $\frac{\partial f}{\partial x_1}(0,0)$ which by (ii) gives us $h(e_1) = 0$ and $\frac{\partial f}{\partial x_1}(0,0) = 0$.

Conversely, let us suppose that $h(e_1) = 0$. Let $(x_1, x_2) \in H \times H$, $(x_1, x_2) \neq (0,0)$. By the chain rule R5 we have

$$\frac{\partial f}{\partial x_1} = \frac{\partial h}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial h}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} + \frac{\partial h}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_1},$$

i.e., by R3

$$\frac{\partial f}{\partial x_1} = \frac{\partial h}{\partial y_1} \cdot \frac{2\|x_2\|^2 x_1}{\left(\|x_1\|^2 + \|x_2\|^2\right)^2} - \frac{\partial h}{\partial y_2} \cdot \frac{2\|x_2\|^2 x_1}{\left(\|x_1\|^2 + \|x_2\|^2\right)^2} + \frac{\partial h}{\partial y_3} \cdot x_2$$

with the same convention as above. Then

$$\frac{\partial f}{\partial x_1}(0, x_2) = \begin{cases} \frac{\partial h}{\partial y_3}(0,1,0)\, x_2 & \text{for} \quad x_2 \neq 0 \\ 0 & \text{for} \quad x_2 = 0 \end{cases} = \frac{\partial h}{\partial y_3}(0,1,0)\, x_2 \text{ for } x_2 \in H.$$

From R4 we deduce $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = \frac{\partial h}{\partial y_3}(e_2)\, I_H$. □

References

[1] A. Aksoy, M. Martelli, Mixed partial derivatives and Fubini's theorem, *Coll. Math. J.*, **33** (2002), 126–130.

[2] N. Boboc, *Analiză matematică*, partea a II-a, Bucureşti, 1993.

[3] H. Cartan, *Calcul differentiel*, Hermann, Paris, 1967.

[4] I. Colojoară, *Analiză matematică*, Editura Didactică şi Pedagogică, Bucureşti, 1983.

[5] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York and London, 1969.

[6] J. J. Duistermaat, J. A. C. Kolk, *Multidimensional real analysis I: Differentiation,* Cambridge University Press, 2004.

[7] D. Popa, *Analiză matematică*, curs, Universitatea Ovidius Constanţa, Constanţa, 1996.

Evaluation of an integral with fractional part function

Ovidiu Furdui[1)]

Abstract. In this article we evaluate the following class of fractional part integrals

$$I_{p,q} = \int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\}^q \mathrm{d}x,$$

where $p, q \geq 1$ are integers and $\{x\}$ denotes the fractional part of x. We prove that $I_{p,q}$ equals a series involving the product of the reciprocal of a special binomial coefficient and an expression involving the Riemann zeta function.

Keywords: Fractional part integrals, binomial coefficients, Riemann zeta function.

MSC: 40A05, 40A10, 11M06

1. Introduction and the main result

In the footnote of Question 892996 in Mathematics Stack Exchange it is mentioned as an open problem, proposed by O. Oloa, the evaluation of the following class of fractional part integrals

$$I_{p,q} = \int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\}^q \mathrm{d}x, \tag{1}$$

where $p, q \geq 1$ are integers and $\{x\}$ denotes the fractional part of x.

In this paper we prove that $I_{p,q}$ equals a series involving the product of the reciprocal of a special binomial coefficient and an expression involving the Riemann zeta function.

The main result of this article is the following theorem.

Theorem 1. *Let $p, q \geq 1$ be integers and let $I_{p,q}$ be the integral in* (1). *Then,*

$$I_{p,q} = \sum_{j=1}^{\infty} \frac{1}{\binom{q+j}{j}} \left(\zeta(j+1) - 1 - \frac{1}{2^{j+1}} - \cdots - \frac{1}{p^{j+1}}\right),$$

where ζ denotes the Riemann zeta function.

Proof. We change variables $x = \frac{1}{y}$ and we get

$$\begin{aligned} I_{p,q} &= \int_p^{\infty} \frac{\{y\}^q}{y^2} \mathrm{d}y = \sum_{k=p}^{\infty} \int_k^{k+1} \frac{(y-k)^q}{y^2} \mathrm{d}y \\ &= \sum_{k=p}^{\infty} \int_0^1 \frac{u^q}{(u+k)^2} \mathrm{d}u = \int_0^1 u^q \left(\sum_{k=p}^{\infty} \frac{1}{(u+k)^2}\right) \mathrm{d}u. \end{aligned}$$

[1)]Department of Mathematics, Technical University of Cluj-Napoca, Romania, Ovidiu.Furdui@math.utcluj.ro, ofurdui@yahoo.com

On the other hand,

$$\frac{1}{(u+k)^2} = \int_0^\infty \mathrm{e}^{-(u+k)t} t\mathrm{d}t,$$

and this implies that

$$\sum_{k=p}^{\infty} \frac{1}{(u+k)^2} = \sum_{k=p}^{\infty} \int_0^\infty \mathrm{e}^{-(u+k)t} t\mathrm{d}t = \int_0^\infty t\mathrm{e}^{-ut} \left(\sum_{k=p}^{\infty} \mathrm{e}^{-kt} \right) \mathrm{d}t$$

$$= \int_0^\infty t \frac{\mathrm{e}^{-(u+p)t}}{1-\mathrm{e}^{-t}} \mathrm{d}t.$$

It follows, based on Tonelli Theorem [4, p. 309], that

$$I_{p,q} = \int_0^1 u^q \left(\int_0^\infty t \frac{\mathrm{e}^{-(u+p)t}}{1-\mathrm{e}^{-t}} \mathrm{d}t \right) \mathrm{d}u = \int_0^\infty t \frac{\mathrm{e}^{-pt}}{1-\mathrm{e}^{-t}} \left(\int_0^1 u^q \mathrm{e}^{-ut} \mathrm{d}u \right) \mathrm{d}t.$$

Let

$$J_q = \int_0^1 u^q \mathrm{e}^{-ut} \mathrm{d}u.$$

Integrating by parts we get the recurrence formula $J_q = -\frac{\mathrm{e}^{-t}}{t} + \frac{q}{t} J_{q-1}$. Let $a_q = J_q \frac{t^q}{q!}$ and we note that $a_q = -\frac{\mathrm{e}^{-t}}{t} \cdot \frac{t^q}{q!} + a_{q-1}$. This implies that

$$\begin{aligned} a_q &= -\frac{\mathrm{e}^{-t}}{t} \left(\frac{t^q}{q!} + \frac{t^{q-1}}{(q-1)!} + \cdots + \frac{t}{1!} \right) + \frac{1-\mathrm{e}^{-t}}{t} \\ &= \frac{\mathrm{e}^{-t}}{t} \left[\mathrm{e}^t - \left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^q}{q!} \right) \right] \\ &= \frac{\mathrm{e}^{-t}}{t} \sum_{j=1}^{\infty} \frac{t^{q+j}}{(q+j)!}. \end{aligned}$$

Thus,

$$J_q = q! \mathrm{e}^{-t} \sum_{j=1}^{\infty} \frac{t^{j-1}}{(q+j)!},$$

and this implies that

$$I_{p,q} = q! \int_0^\infty \frac{\mathrm{e}^{-(p+1)t}}{1-\mathrm{e}^{-t}} \sum_{j=1}^{\infty} \frac{t^j}{(q+j)!} \mathrm{d}t = q! \sum_{j=1}^{\infty} \frac{1}{(q+j)!} \int_0^\infty t^j \frac{\mathrm{e}^{-(p+1)t}}{1-\mathrm{e}^{-t}} \mathrm{d}t. \quad (2)$$

On the other hand,

$$\begin{aligned}
\int_0^\infty t^j \frac{\mathrm{e}^{-(p+1)t}}{1-\mathrm{e}^{-t}}\mathrm{d}t &= \int_0^\infty t^j \mathrm{e}^{-(p+1)t} \sum_{m=0}^\infty \mathrm{e}^{-tm}\mathrm{d}t \\
&= \sum_{m=0}^\infty \int_0^\infty t^j \mathrm{e}^{-(p+1+m)t}\mathrm{d}t \\
&\overset{(p+1+m)t=x}{=} \sum_{m=0}^\infty \frac{1}{(p+1+m)^{j+1}} \int_0^\infty x^j \mathrm{e}^{-x}\mathrm{d}x \qquad (3)\\
&= \Gamma(j+1) \sum_{m=0}^\infty \frac{1}{(p+1+m)^{j+1}} \\
&= j!\left(\zeta(j+1) - 1 - \frac{1}{2^{j+1}} - \cdots - \frac{1}{p^{j+1}}\right).
\end{aligned}$$

Combining (2) and (3) we get that

$$I_{p,q} = \sum_{j=1}^\infty \frac{1}{\binom{q+j}{j}} \left(\zeta(j+1) - 1 - \frac{1}{2^{j+1}} - \cdots - \frac{1}{p^{j+1}}\right),$$

and the theorem is proved. □

The following special cases are worth mentioning.

Corollary 2. ***Special integrals with fractional part.***

(a) *Let $q \geq 1$ be an integer. Then,*

$$\int_0^1 \left\{\frac{1}{x}\right\}^q dx = \sum_{j=1}^\infty \frac{1}{\binom{q+j}{j}} \left(\zeta(j+1) - 1\right).$$

(b) *Let $p \geq 1$ be an integer. Then,*

$$\begin{aligned}
\int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\} dx &= \sum_{j=1}^\infty \frac{1}{j+1} \left(\zeta(j+1) - 1 - \frac{1}{2^{j+1}} - \cdots - \frac{1}{p^{j+1}}\right) \\
&= H_p - \ln p - \gamma,
\end{aligned}$$

where H_p denotes the pth harmonic number and γ is the Euler–Mascheroni constant.

(c) *Let $p \geq 1$ be an integer. Then,*

$$\begin{aligned}
\int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\}^2 dx &= 2\sum_{j=1}^\infty \frac{1}{(j+1)(j+2)} \left(\zeta(j+1) - 1 - \frac{1}{2^{j+1}} - \cdots - \frac{1}{p^{j+1}}\right) \\
&= \ln(2\pi) - \gamma + H_p + 2p\ln p - 2p - 2\ln p!.
\end{aligned}$$

Proof. Part (a) of the corollary follows from the theorem by taking $p = 1$.

In particular, when $q = 1$ we recover an integral of *de la Vallée Poussin* ([1, p. 32], [3, pp. 109–111])

$$\int_0^1 \left\{\frac{1}{x}\right\} \mathrm{d}x = \sum_{j=1}^{\infty} \frac{\zeta(j+1)-1}{j+1} = 1 - \gamma,$$

where the last equality follows by direct calculation or based on [2, Identity (151), p. 174].

(b) The first equality of part (b) follows from the theorem by taking $q = 1$. To prove the second equality we observe that

$$\int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\} \mathrm{d}x = \frac{1}{p}\int_0^1 \left\{\frac{p}{y}\right\} \mathrm{d}y = H_p - \ln p - \gamma,$$

since (see [2, Problem 2.5, p. 100])

$$\int_0^1 \left\{\frac{p}{y}\right\} \mathrm{d}y = p(H_p - \ln p - \gamma). \tag{4}$$

(c) The first equality of part (c) of the corollary follows from the theorem by taking $q = 2$. To prove the second equality we note that

$$\int_0^{\frac{1}{p}} \left\{\frac{1}{x}\right\}^2 \mathrm{d}x = \frac{1}{p}\int_0^1 \left\{\frac{p}{y}\right\}^2 \mathrm{d}y = \ln(2\pi) - \gamma + H_p + 2p\ln p - 2p - 2\ln p!,$$

since (see [2, Problem 2.6, p. 100])

$$\int_0^1 \left\{\frac{p}{y}\right\}^2 \mathrm{d}y = p\left(\ln(2\pi) - \gamma + H_p + 2p\ln p - 2p - 2\ln p!\right), \tag{5}$$

and the corollary is proved. □

Remark 3. We mention that integrals (4) and (5) can be evaluated by direct computation by reducing the integral to a series and then by calculating the nth partial sum of the series [2, pp. 113–114]. Other integrals, single, double or multiple, involving the fractional part function as well as open problems can be found in [2, Chapter 2].

Theorem 4. *Let $m \geq 0$ and $p, q \geq 1$ be integers. Then,*

$$I_{p,m,q} = \int_0^{\frac{1}{p}} x^m \left\{\frac{1}{x}\right\}^q \mathrm{d}x$$

$$= \frac{q!}{(m+1)!}\sum_{j=1}^{\infty} \frac{(m+j)!}{(q+j)!}\left(\zeta(m+j+1) - 1 - \frac{1}{2^{m+j+1}} - \cdots - \frac{1}{p^{m+j+1}}\right),$$

where ζ denotes the Riemann zeta function.

Proof. The proof of this theorem, which is similar to the proof of Theorem 1, is left to the interested reader. □

References

[1] S.R. Finch, *Mathematical Constants*, Encyclopedia of Mathematics and Its Applications 94, Cambridge University Press, New York, 2003.
[2] O. Furdui, *Limits, Series and Fractional Part Integrals*, Problems in Mathematical Analysis, Springer, London, 2013.
[3] J. Havil, *Gamma, Exploring Euler's constant*, Princeton University Press, Princeton, 2003.
[4] H.L. Royden, *Real Analysis*, Third Edition, Prentice Hall, Englewood Cliffs, 1988.
[5] H.M. Srivastava, J. Choi, *Series Associated with the Zeta and Related Functions*, Kluwer Academic Publishers, Dordrecht, 2001.

The limit, continuity and Fréchet differentiability of some functions on $\mathbb{R}^n$

Dumitru Popa[1)]

Abstract. We give necessary and sufficient conditions such that the function $f:\mathbb{R}^n\to\mathbb{R}$ defined by

$$f(x_1,\dots,x_n)=\begin{cases}\frac{x_1^{\alpha_1}\cdots x_n^{\alpha_n}}{\prod\limits_{i=1}^{m}\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}} & \text{if} \quad (x_1,\dots,x_n)\neq(0,\dots,0)\\ 0 & \text{if} \quad (x_1,\dots,x_n)=(0,\dots,0)\end{cases}$$

has a finite limit at $(0,\dots,0)$, is continuous and Fréchet differentiable at $(0,\dots,0)$.

Keywords: Fréchet differentiable, mixed partial derivative, real-valued functions

MSC: Primary 26B05; Secondary 54C30.

In the study of Fréchet differentiability, one of the standard examples is the following: the function $f:\mathbb{R}^2\to\mathbb{R}$ defined by

$$f(x,y)=\begin{cases}\frac{xy}{\sqrt{x^2+y^2}} & \text{if} \quad (x,y)\neq(0,0)\\ 0 & \text{if} \quad (x,y)=(0,0)\end{cases}$$

is continuous at $(0,0)$, there exist $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$, but f is not Fréchet differentiable at $(0,0)$. In this note we study the existence of the limit, continuity and Fréchet differentiability of some functions which extend the above example. The notations are standard, see [1].

Proposition 1. *Let n, m, $\alpha_1,\dots,\alpha_n$ be positive integers, $n\geq 2$, $\beta_1,\dots,\beta_m$, $\gamma_1,\dots,\gamma_m$ be positive real numbers and $f:\mathbb{R}^n\to\mathbb{R}$ defined by*

$$f(x_1,\dots,x_n)=\begin{cases}\frac{x_1^{\alpha_1}\cdots x_n^{\alpha_n}}{\prod\limits_{i=1}^{m}\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}} & \textit{if} \quad (x_1,\dots,x_n)\neq(0,\dots,0),\\ 0 & \textit{if} \quad (x_1,\dots,x_n)=(0,\dots,0).\end{cases}$$

1) Department of Mathematics, Ovidius University of Constanţa, Romania, dpopa@univ-ovidius.ro

Then f has finite limit at $(0,\ldots,0)$ if and only if

$$\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m.$$

Proof. We use the well-known characterization for the existence of the limit of a function defined on $\mathbb{R}^n$: a function $f:\mathbb{R}^n\to\mathbb{R}$ has the limit $l\in\mathbb{R}$ at $(0,\ldots,0)$ if and only if for all sequences $\left(x_k^1,\ldots,x_k^n\right)_{k\in\mathbb{N}}\subset\mathbb{R}^n$, $\left(x_k^1,\ldots,x_k^n\right)\neq(0,\ldots,0)$ for all $k\in\mathbb{N}^*$, and $\lim\limits_{k\to\infty}\left(x_k^1,\ldots,x_k^n\right)=(0,\ldots,0)$, it follows that $\lim\limits_{k\to\infty}f\left(x_k^1,\ldots,x_k^n\right)=l$.

Let us suppose that there exists $\lim\limits_{(x_1,\ldots,x_n)\to(0,\ldots,0)}f\left(x_1,\ldots,x_n\right)=l\in\mathbb{R}$. Since $n\geq 2$, $\lim\limits_{k\to\infty}(\frac{1}{k},\underbrace{0,\ldots,0}_{n-1\text{ times}})=(\underbrace{0,\ldots,0}_{n\text{ times}})$ and $(\frac{1}{k},\underbrace{0,\ldots,0}_{n-1\text{ times}})\neq(0,\ldots,0)$ for all $k\in\mathbb{N}^*$, we deduce $\lim\limits_{k\to\infty}f\left(\frac{1}{k},0,\ldots,0\right)=l$, and using that $f\left(\frac{1}{k},0,\ldots,0\right)=0$ for all $k\in\mathbb{N}^*$, we obtain $l=0$. From $\lim\limits_{k\to\infty}\left(\frac{1}{k},\ldots,\frac{1}{k}\right)=(0,\ldots,0)$ it follows that $\lim\limits_{k\to\infty}f\left(\frac{1}{k},\ldots,\frac{1}{k}\right)=l=0$. Using the equality

$$f\left(\frac{1}{k},\ldots,\frac{1}{k}\right)=\frac{k^{-(\alpha_1+\cdots+\alpha_n)+\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}{n^{\gamma_1+\cdots+\gamma_m}}$$

we obtain $\lim\limits_{k\to\infty}\frac{k^{-(\alpha_1+\cdots+\alpha_n)+\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}{n^{\gamma_1+\cdots+\gamma_m}}=0$ and hence

$$\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m.$$

Thus, if f has finite limit at $(0,\ldots,0)$, then

$$\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m.$$

Conversely, let us suppose that $\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m$.

Let $(x_1,\ldots,x_n)\in\mathbb{R}^n$. We have $0\leq|x_1|^{\beta_i}\leq|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}$ and since $\gamma_i>0$, we deduce

$$0\leq|x_1|^{\beta_i\gamma_i}\leq\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}$$

for all $i=1,\ldots,m$.

Multiplying these inequalities we get

$$|x_1|^{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}\leq\prod_{i=1}^{m}\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}. \qquad (1)$$

Set $g(x_1,\ldots,x_n)=\prod\limits_{i=1}^{m}\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}$. From (1) we obtain $|x_1|\leq\left[g\left(x_1,\ldots,x_n\right)\right]^{\frac{1}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}$ and, since $\alpha_1>0$, we have

$$|x_1|^{\alpha_1}\leq\left[g\left(x_1,\ldots,x_n\right)\right]^{\frac{\alpha_1}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}.$$

Similarly we get

$$|x_i|^{\alpha_i} \le [g(x_1,\dots,x_n)]^{\frac{\alpha_i}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}$$

for all $i=1,\dots,m$.

Multiplying the inequalities we deduce

$$|x_1|^{\alpha_1}\cdots|x_n|^{\alpha_n} \le [g(x_1,\dots,x_n)]^{\frac{\alpha_1+\cdots+\alpha_n}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}}. \qquad (2)$$

Then, for $(x_1,\dots,x_n)\in\mathbb{R}^n-\{(0,\dots,0)\}$ from (2) we deduce

$$|f(x_1,\dots,x_n)| = \frac{|x_1|^{\alpha_1}\cdots|x_n|^{\alpha_n}}{g(x_1,\dots,x_n)} \le [g(x_1,\dots,x_n)]^{\frac{\alpha_1+\cdots+\alpha_n}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}-1}. \qquad (3)$$

Since $\lim\limits_{(x_1,\dots,x_n)\to(0,\dots,0)} g(x_1,\dots,x_n)=0$ and

$$\frac{\alpha_1+\cdots+\alpha_n}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}-1>0,$$

it follows that

$$\lim_{(x_1,\dots,x_n)\to(0,\dots,0)} [g(x_1,\dots,x_n)]^{\frac{\alpha_1+\cdots+\alpha_n}{\beta_1\gamma_1+\cdots+\beta_m\gamma_m}-1}=0. \qquad (4)$$

From (3) and (4) we obtain

$$\lim_{(x_1,\dots,x_n)\to(0,\dots,0)} f(x_1,\dots,x_n)=0.$$

In conclusion, the function f has a finite limit at $(0,\dots,0)$ if and only if $\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m$ and in this case the limit is 0. □

We use in the sequel the result proved in Proposition 1 for the study of the continuity and Fréchet differentiability of f on $\mathbb{R}^n$.

Proposition 2. *Let n, m, α_1, …, α_n be positive integers, $n\ge 2$, β_1, …, β_m, γ_1, …, γ_m be positive real numbers and $f:\mathbb{R}^n\to\mathbb{R}$ defined by*

$$f(x_1,\dots,x_n)=\begin{cases}\dfrac{x_1^{\alpha_1}\cdots x_n^{\alpha_n}}{\prod\limits_{i=1}^m\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}} & \text{if} \quad (x_1,\dots,x_n)\neq(0,\dots,0),\\ 0 & \text{if} \quad (x_1,\dots,x_n)=(0,\dots,0).\end{cases}$$

Then

(i) f is continuous at $(0,\dots,0)$ if and only if

$$\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m.$$

(ii) f is Fréchet differentiable at $(0,\dots,0)$ if and only if

$$\alpha_1+\cdots+\alpha_n>\beta_1\gamma_1+\cdots+\beta_m\gamma_m+1.$$

Proof. (*i*) If f is continuous at $(0,\ldots,0)$, then

$$\lim_{(x_1,\ldots,x_n)\to(0,\ldots,0)} f(x_1,\ldots,x_n) = f(0,\ldots,0).$$

From Proposition 1 it follows that $\alpha_1+\cdots+\alpha_n > \beta_1\gamma_1+\cdots+\beta_m\gamma_m$. Conversely, if $\alpha_1+\cdots+\alpha_n > \beta_1\gamma_1+\cdots+\beta_m\gamma_m$ then we have shown in Proposition 1 that $\lim\limits_{(x_1,\ldots,x_n)\to(0,\ldots,0)} f(x_1,\ldots,x_n) = 0 = f(0,\ldots,0)$ and thus f is continuous at $(0,\ldots,0)$.

(*ii*) As it is well known, see [1], f is Fréchet differentiable at $(0,\ldots,0)$ if and only if there exist $\frac{\partial f}{\partial x_1}(0,\ldots,0)\in\mathbb{R}$, $\ldots$, $\frac{\partial f}{\partial x_n}(0,\ldots,0)\in\mathbb{R}$ and

$$\lim_{(x_1,\ldots,x_n)\to(0,\ldots,0)} \frac{f(x_1,\ldots,x_n) - \sum\limits_{i=1}^{n}\frac{\partial f}{\partial x_i}(0,\ldots,0)\,x_i}{\sqrt{x_1^2+\cdots+x_n^2}} = 0.$$

For $x_1\neq 0$ we have $f(x_1,0,\ldots,0)=0$ and since $f(0,0,\ldots,0)=0$ we obtain $\frac{\partial f}{\partial x_1}(0,\ldots,0)=0$. In a similar way $\frac{\partial f}{\partial x_i}(0,\ldots,0)=0$ for all $i=1,\ldots,n$. Thus f is Fréchet differentiable at $(0,\ldots,0)$ if and only if

$$\lim_{(x_1,\ldots,x_n)\to(0,\ldots,0)} \frac{f(x_1,\ldots,x_n)}{\sqrt{x_1^2+\cdots+x_n^2}} = 0,$$

that is,

$$\lim_{(x_1,\ldots,x_n)\to(0,\ldots,0)} \frac{x_1^{\alpha_1}\cdots x_n^{\alpha_n}}{\left(|x_1|^2+\cdots+|x_n|^2\right)^{\frac{1}{2}}\prod\limits_{i=1}^{m}\left(|x_1|^{\beta_i}+\cdots+|x_n|^{\beta_i}\right)^{\gamma_i}} = 0.$$

By using the Proposition 1 this is equivalent to $\alpha_1+\cdots+\alpha_n > \beta_1\gamma_1+\cdots+\beta_m\gamma_m+1$. □

From Proposition 2 we obtain a different proof of the following result (see [2]).

Corollary 3. *Let $n\geq 2$ be a positive integer and $f:\mathbb{R}^n\to\mathbb{R}$ defined by*

$$f(x_1,\ldots,x_n) = \begin{cases} \frac{x_1\cdots x_n}{x_1^2+\cdots+x_n^2} & \text{if } (x_1,\ldots,x_n)\neq(0,\ldots,0), \\ 0 & \text{if } (x_1,\ldots,x_n)=(0,\ldots,0). \end{cases}$$

Then:

(*i*) *f is continuous at $(0,\ldots,0)$ if and only if $n\geq 3$.*

(*ii*) *f is Fréchet differentiable at $(0,\ldots,0)$ if and only if $n\geq 4$.*

Among many other possible examples we give

Corollary 4. *Let $n\geq 2$ be a positive integer and $f:\mathbb{R}^n\to\mathbb{R}$ defined by*

$$f(x_1,\ldots,x_n) = \begin{cases} \frac{x_1\cdots x_n}{\left(x_1^2+\cdots+x_n^2\right)\left(x_1^4+\cdots+x_n^4\right)} & \text{if } (x_1,\ldots,x_n)\neq(0,\ldots,0), \\ 0 & \text{if } (x_1,\ldots,x_n)=(0,\ldots,0). \end{cases}$$

Then:

(*i*) f *is continuous at* $(0, \ldots, 0)$ *if and only if* $n \geq 7$.

(*ii*) f *is Fréchet differentiable at* $(0, \ldots, 0)$ *if and only if* $n \geq 8$.

References

[1] N. Boboc, *Analiză matematică*, partea a II-a, Bucureşti 1993.
[2] Problem 41, *Gaz. Mat. Ser. A.*, **16** (1998), 52.

Traian Lalescu national mathematical contest for university students, Timişoara 2014

Vasile Pop[1)], Tiberiu Trif[2)]

Abstract. This note presents the solutions to the problems proposed at the 2014 edition of the Traian Lalescu national mathematical contest for university students, hosted by the West University in Timişoara between the 21st and the 24th of May 2014.

Keywords: Function of class C^1, countable set, minimal polynomial, Jordan normal form, rank of a matrix, Sylvester's theorem.

MSC: 11C08, 11C20, 26A27.

Between the 21st and the 24th of May 2014, the national phase of the student contest Traian Lalescu took place in Timişoara.

Over 60 students participated at the contest, representing 12 universities from 6 cities: Bucureşti, Cluj, Constanţa, Craiova, Iaşi and Timişoara.

The contest was divided in 4 sections: A – mathematics faculties, B – technical education, electrical engineering, 1st year, C – technical education, mechanical and construction engineering, 1st year, D – technical education, 2nd year.

The subjects were proposed, discussed and chosen in the morning of the contest, by commissions responsible for each section. There was one member representing each university in each commission.

As far as the organization of the contest is concerned, apart from the contribution of the West University in Timişoara, which provided optimal conditions for the contest, accommodation and meals, the Ministry of Education and Research and the Traian Lalescu Foundation also contributed to the event.

We are next going to present the statements and the solutions to the problems given in sections A and B of the contest. For the official solutions, please refer to the following web page: http://cntl.math.uvt.ro/

[1)] Technical University of Cluj-Napoca, Department of Mathematics, Cluj-Napoca, Romania, `vasile.pop@math.utcluj.ro`

[2)] Babeş-Bolyai University, Faculty of Mathematics and Computer Science, Cluj-Napoca, Romania, `ttrif@math.ubbcluj.ro`

Section A

Problem 1. Let I be a nondegenerate interval of the real axis, let $f : I \to \mathbb{R}$ be a function of class C^1 on I, and let $g : I \to \mathbb{R}$ be the function defined by $g(x) = |f(x)|$. Prove that there exists an at most countable set $E_f \subseteq I$ such that g is differentiable on $I \setminus E_f$. Provide an example of a function $f : [0, 1] \to \mathbb{R}$ such that f is of class C^1 and E_f is infinite.

Tiberiu Trif

Although the problem was considered easy by the members of the jury, no student solved it completely. Two partial solutions were given.

Solution. Set

$$E_f := \{x \in I \mid f(x) = 0 \text{ and } f'(x) \neq 0\}.$$

We claim that g is differentiable on $I \setminus E_f$. Indeed, given any point $x_0 \in I \setminus E_f$, one has either $f(x_0) \neq 0$ or $f(x_0) = f'(x_0) = 0$.

If $f(x_0) \neq 0$, then the continuity of f at x_0 ensures the existence of a positive real number r such that $\operatorname{sgn} f(x) = \operatorname{sgn} f(x_0)$ for all $x \in J$, where $J := (x_0 - r, x_0 + r) \cap I$. Then for all $x \in J$ one has

$$g(x) = \big(\operatorname{sgn} f(x)\big) f(x) = \big(\operatorname{sgn} f(x_0)\big) f(x),$$

whence g is differentiable at x_0.

If $f(x_0) = f'(x_0) = 0$, then $g(x_0) = 0$, whence

$$g'_+(x_0) = \lim_{x \searrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = \lim_{x \searrow x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} \right| = |f'(x_0)| = 0$$

and

$$g'_-(x_0) = \lim_{x \nearrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = -\lim_{x \nearrow x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} \right| = -|f'(x_0)| = 0.$$

Consequently, g is differentiable at x_0 and $g'(x_0) = 0$.

It remains to show that the set E_f is at most countable. Let $x \in E_f$ be arbitrarily chosen. Since $f'(x) \neq 0$ and f' is continuous, one can find an interval J_x centered at x such that $f'(y) \neq 0$ for all $y \in I \cap J_x$. It follows that f is strictly monotone on $I \cap J_x$, whence $f(y) \neq 0$ for all $y \in I \cap J_x \setminus \{x\}$. Set

$$\begin{aligned} a_x &:= \sup\{y \in I \mid y < x \text{ and } f(y) = 0\}, \\ b_x &:= \inf\{y \in I \mid y > x \text{ and } f(y) = 0\}, \end{aligned}$$

with the convention that $a_x = -\infty$ if $f(y) \neq 0$ for all $y < x$ (respectively $b_x = \infty$ if $f(y) \neq 0$ for all $y > x$). After that, set $A_x := \left(\frac{x+a_x}{2}, \frac{x+b_x}{2}\right)$. It is immediately seen that the family of open intervals $(A_x)_{x \in E_f}$ has the following properties:

(i) $x \in A_x$ for all $x \in E_f$;
(ii) $A_x \cap A_y = \emptyset$ for all $x, y \in E_f$ with $x \neq y$.

For every $x \in E_f$ select $q_x \in A_x \cap \mathbb{Q}$. Since the function $x \mapsto q_x$ is injective, we conclude that $\operatorname{card} E_f \leq \operatorname{card} \mathbb{Q} = \aleph_0$.

In the case of the function $f : [0,1] \to \mathbb{R}$, defined by

$$f(x) := \begin{cases} x^3 \sin(\pi/x) & \text{if } \ x \in (0,1], \\ 0 & \text{if } \ x = 0, \end{cases}$$

one has

$$f'(x) = \begin{cases} 3x^2 \sin(\pi/x) - \pi x \cos(\pi/x) & \text{if } \ x \in (0,1], \\ 0 & \text{if } \ x = 0, \end{cases}$$

whence f is continuously differentiable on $[0,1]$. In addition, $\operatorname{card} E_f = \aleph_0$ because $E_f = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\}$. □

Problem 2. How many solutions does the equation $x^{2013} = 1$ have in $\mathbb{Z}_{2014}$?

Alexandru Gica

The members of the jury ranked this problem as a medium one. It was completely solved by only one student (Mădălina Bolboceanu). There were also two partial solutions.

Solution. Note first that $2014 = 2 \times 19 \times 53$ is the prime decomposition of 2014. It is well known that the function $\varphi : \mathbb{Z}_{2014} \to \mathbb{Z}_2 \times \mathbb{Z}_{19} \times \mathbb{Z}_{53}$ defined by $\varphi(x) := (u, v, w)$, where (u, v, w) is the unique triple in $\mathbb{Z}_2 \times \mathbb{Z}_{19} \times \mathbb{Z}_{53}$ such that $x \equiv u \pmod 2$, $x \equiv v \pmod{19}$, and $x \equiv w \pmod{53}$, is a ring isomorphism. We have

$$\begin{aligned} x^{2013} = 1 \quad &\Leftrightarrow \quad \varphi(x^{2013}) = \varphi(1) \quad \Leftrightarrow \quad \varphi(x)^{2013} = (1,1,1) \\ &\Leftrightarrow \quad \begin{cases} u^{2013} = 1 & \text{in } \mathbb{Z}_2, \\ v^{2013} = 1 & \text{in } \mathbb{Z}_{19}, \\ w^{2013} = 1 & \text{in } \mathbb{Z}_{53}. \end{cases} \end{aligned} \tag{1}$$

Therefore, the number n of solutions to $x^{2013} = 1$ in $\mathbb{Z}_{2014}$ equals $n = n_1 n_2 n_3$, where n_i represents the number of solutions to the ith equation in the system (1), $i \in \{1,2,3\}$.

Clearly, $n_1 = 1$ because $u = 1$ is the unique solution to the first equation in (1). On the other hand, if w is a solution to the third equation in (1), then $w \neq 0$. By Fermat's little theorem it follows that $w^{52} = 1$. Since 52 and 2013 are co-prime, there exist two integers a and b such that $52a + 2013b = 1$. Therefore, we have $w = \left(w^{52}\right)^a \left(w^{2013}\right)^b = 1$, hence $w = 1$ is the unique solution to the third equation in (1). Consequently, $n_3 = 1$. Finally, let v be an arbitrary solution to the second equation in (1). Then $v \neq 0$ and $v^{18} = 1$ by Fermat's little theorem. Since $(2013, 18) = 3$, there exist two integers c and d such that $18c + 2013d = 3$. Then we have $v^3 = \left(v^{18}\right)^c \left(v^{2013}\right)^d = 1$.

Conversely, if $v^3 = 1$ in $\mathbb{Z}_{19}$, then $v^{2013} = (v^3)^{671} = 1$. Consequently, the second equation in (1) is equivalent to

$$v^3 = 1 \quad \text{in } \mathbb{Z}_{19}. \tag{2}$$

But (2) is equivalent to $(v-1)(v^2+v+1) = 0$, i.e., to $v = 1$ or

$$v^2 + v + 1 = 0 \quad \text{in } \mathbb{Z}_{19}. \tag{3}$$

Multiplying both sides in (3) by 4, we see that (3) is equivalent to

$$(2v+1)^2 = 16 \text{ in } \mathbb{Z}_{19} \quad \Leftrightarrow \quad (2v-3)(2v+5) = 0 \text{ in } \mathbb{Z}_{19}.$$

The last equation has the solutions $v = 11$ and $v = 7$ in $\mathbb{Z}_{19}$. In conclusion, we have $n_2 = 3$, whence $n = 3$. □

Problem 3. a) Prove that the center of a parallelogram which is inscribed in an ellipse coincides with the center of the ellipse.

b) Prove that if a rectangle is inscribed in an ellipse which is not a circle, then its sides must be parallel to the symmetry axes of the ellipse.

c) Find the smallest area of an ellipse which is circumscribed to a given rectangle.

Gabriel Mincu

The members of the jury ranked this problem as a medium one. It was completely solved by only one student (Eduard Valentin Curcă). There was also one partial solution.

Solution. a) Let E be an ellipse, and let P be a parallelogram inscribed in E. Further, let π be a plane such that π is parallel to the small semi-axis of E, and the projection of E onto π is a circle C. Taking into account that the lines' parallelism is preserved by the projection onto π, it follows that the projection of P onto π is a parallelogram P', which is inscribed in C. But the only parallelograms that can be inscribed in a circle are rectangles. Hence P' must be a rectangle and the center of P' coincides with the center of C. Consequently, the center of P coincides with the center of E.

b) Choose a Cartesian coordinate system whose origin coincides with the rectangle's center and whose axes are parallel to the rectangle's sides. Let $A(x_0, y_0)$, $B(-x_0, y_0)$, $C(-x_0, -y_0)$, and $D(x_0, -y_0)$ be the rectangle's vertices, and let

$$E: \quad ax^2 + bxy + cy^2 + dx + ey + f = 0$$

be the equation of the ellipse. By $A, B, C, D \in E$ it follows that

$$ax_0^2 + bx_0y_0 + cy_0^2 + dx_0 + ey_0 + f = 0, \tag{1}$$
$$ax_0^2 - bx_0y_0 + cy_0^2 - dx_0 + ey_0 + f = 0, \tag{2}$$
$$ax_0^2 + bx_0y_0 + cy_0^2 - dx_0 - ey_0 + f = 0, \tag{3}$$
$$ax_0^2 - bx_0y_0 + cy_0^2 + dx_0 - ey_0 + f = 0. \tag{4}$$

By subtracting side by side the equations (1) and (3), respectively (2) and (4), we get

$$dx_0 + ey_0 = 0 \quad \text{and} \quad - dx_0 + ey_0 = 0,$$

whence $dx_0 = ey_0 = 0$. Therefore, we have $d = e = 0$. The equations (1), (2), (3), and (4) reduce now to

$$\begin{aligned} ax_0^2 + bx_0y_0 + cy_0^2 + f = 0, \\ ax_0^2 - bx_0y_0 + cy_0^2 + f = 0. \end{aligned}$$

By subtracting side by side the last two equations we get $bx_0y_0 = 0$, whence $b = 0$. In conclusion, the equation of E must have the form

$$E: \quad ax^2 + cy^2 + f = 0.$$

This means that the symmetry axes of E coincide with the coordinate axes, hence they are parallel to the rectangle's sides.

c) Choose a Cartesian coordinate system whose origin coincides with the ellipse's center and whose axes coincide with the ellipse's symmetry axes. Then the equation of E is of the form

$$E: \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Let 2ℓ and $2L$ denote the lengths of the sides of the rectangle which is inscribed in E. Since $(L, \ell) \in E$, it follows that

$$\frac{L^2}{a^2} + \frac{\ell^2}{b^2} = 1. \tag{5}$$

So we have to find the smallest value of $\mathcal{A}(E) = \pi ab$, when a and b satisfy (5). Note that

$$\mathcal{A}(E) = \pi L\ell \frac{a}{L} \cdot \frac{b}{\ell} = \frac{\pi L\ell}{\frac{L}{a} \cdot \frac{\ell}{b}} \geq \frac{\pi L\ell}{\frac{1}{2}\left(\frac{L^2}{a^2} + \frac{\ell^2}{b^2}\right)} = 2\pi L\ell.$$

Consequently, the smallest possible area of E equals $2\pi L\ell$ and it is attained when $\frac{L}{a} = \frac{\ell}{b} = \frac{1}{\sqrt{2}}$, i.e., when the semi-axes of the ellipse have the lengths $a = L\sqrt{2}$ and $b = \ell\sqrt{2}$, respectively. □

Problem 4. Let m and n be positive integers, and let $A \in \mathcal{M}_n(\mathbb{C})$ be a matrix such that $A^m = I_n$. Prove that

$$\operatorname{rank}(A - \varepsilon_0 I_n) + \operatorname{rank}(A - \varepsilon_1 I_n) + \cdots + \operatorname{rank}(A - \varepsilon_{m-1} I_n) = n(m-1),$$

where $\{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{m-1}\} = \{z \in \mathbb{C} \mid z^m = 1\}$.

Dan Moldovan and Vasile Pop

Although the members of the jury ranked this problem as a difficult one, it was completely solved by four students (Mihai Florin Barbu, Mădălina Bolboceanu, Eduard Valentin Curcă, and Petre Claudiu Mîndrilă). The solution below was given by Mădălina Bolboceanu.

Solution. Let $m_A \in \mathbb{C}[X]$ denote the minimal polynomial of A, and let $p_A \in \mathbb{C}[X]$ denote the characteristic polynomial of A. Note that the polynomial $f := X^m - 1$ has the simple roots $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{m-1}$. Since $f(A) = 0$, we must have $m_A \mid f$, hence there exist $r \in \mathbb{N}$ as well as $i_1, \ldots, i_r \in \{0, 1, \ldots, m-1\}$, $i_1 < \cdots < i_r$, such that $m_A = (X - \varepsilon_{i_1}) \cdots (X - \varepsilon_{i_r})$. Taking into account the Frobenius theorem (m_A and p_A have the same irreducible factors), it follows that $p_A = (X - \varepsilon_{i_1})^{\alpha_{i_1}} \cdots (X - \varepsilon_{i_r})^{\alpha_{i_r}}$, with $\alpha_{i_1} + \cdots + \alpha_{i_r} = n$. Further, let J be the Jordan normal form of A. Since each factor $X - \varepsilon_{i_j}$ appears in m_A at power one, it follows that all Jordan blocks corresponding to ε_{i_j} have the size of one, their number being α_{i_j}. Consequently, J has the form

$$J = \operatorname{diag}\Big(\underbrace{\varepsilon_{i_1}, \ldots, \varepsilon_{i_1}}_{\alpha_{i_1}}, \underbrace{\varepsilon_{i_2}, \ldots, \varepsilon_{i_2}}_{\alpha_{i_2}}, \ldots, \underbrace{\varepsilon_{i_r}, \ldots, \varepsilon_{i_r}}_{\alpha_{i_r}}\Big),$$

i.e., A is diagonalizable. Let $S \in \mathcal{M}_n(\mathbb{C})$ be an invertible matrix such that $A = S^{-1}JS$. For every $k \in \{0, 1, \ldots, m-1\}$ one has

$$\begin{aligned} \operatorname{rank}(A - \varepsilon_k I_n) &= \operatorname{rank}\big(S^{-1}(J - \varepsilon_k I_n)S\big) = \operatorname{rank}(J - \varepsilon_k I_n) \\ &= \begin{cases} n & \text{if } k \notin \{i_1, \ldots, i_r\}, \\ n - \alpha_{i_j} & \text{if } k = i_j. \end{cases} \end{aligned}$$

Consequently

$$\begin{aligned} \sum_{k=0}^{m-1} \operatorname{rank}(A - \varepsilon_k I_n) &= \sum_{k \notin \{i_1, \ldots, i_r\}} \operatorname{rank}(A - \varepsilon_k I_n) + \sum_{j=1}^{r} \operatorname{rank}(A - \varepsilon_{i_j} I_n) \\ &= n(m - r) + \sum_{j=1}^{r} (n - \alpha_{i_j}) \\ &= nm - nr + nr - \sum_{j=1}^{r} \alpha_{i_j} = nm - n = n(m-1). \end{aligned}$$

Section B

Problem 1. Let $a > 0$, and let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$f(x, y) = e^{-x-y} + a\sqrt{x^2 + y^2} \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

Prove that f has a unique local extremum and that this is a global minimum.

Cristian Ghiu

This is a standard problem concerning the local extremum points of a function of several variables. It was considered easy by the members of the *jury. However, due to the technicalities encountered during the problem solving process, it turned out to be difficult. The scores obtained by the contestants were low.*

Solution. The restriction of f to $\mathbb{R}^2 \setminus \{(0,0)\}$ is a C^1 function, hence every local extremum of f must be either a critical point, or the point $(0,0)$. Every critical point of f is solution to the system

$$\begin{cases} \dfrac{\partial f}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} = 0 \end{cases} \Leftrightarrow \begin{cases} -e^{-x-y} + a\dfrac{x}{\sqrt{x^2+y^2}} = 0 \\ -e^{-x-y} + a\dfrac{y}{\sqrt{x^2+y^2}} = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = y \neq 0 \\ -e^{-2x} + \dfrac{a}{\sqrt{2}} \cdot \dfrac{x}{|x|} = 0 \end{cases} \Leftrightarrow \begin{cases} x = y > 0 \\ e^{-2x} = \dfrac{a}{\sqrt{2}} \end{cases} \tag{1}$$

Since $x > 0$, it follows that $a < \sqrt{2}$, and in this case f has a unique critical point, namely

$$(x,y) = (c,c), \text{ with } c = \frac{1}{2}\ln\frac{\sqrt{2}}{a} > 0.$$

If $a \geq \sqrt{2}$, then f does not have critical points.

Case I. $a \geq \sqrt{2}$. In this case f does not have critical points. Only $(0,0)$ could be a local extremum for f.

From the inequalities

$$\sqrt{2}\sqrt{x^2+y^2} \geq |x+y| \tag{2}$$

$$e^{-x-y} \geq 1 - x - y \tag{3}$$

it follows that

$$\begin{aligned} f(x,y) &= e^{-x-y} + \sqrt{2}\sqrt{x^2+y^2} + (a-\sqrt{2})\sqrt{x^2+y^2} \\ &\geq 1 - x - y + |x+y| + 0 \geq 1 = f(0,0). \end{aligned}$$

Hence $(0,0)$ is a global minimum point for f.

Case II. $a < \sqrt{2}$. In this case (c,c) and $(0,0)$ could be local extremum points for f. Since $a = e^{-2c}\sqrt{2}$, we have

$$f(x,y) = e^{-2c}(e^{2c-x-y} + \sqrt{2}\sqrt{x^2+y^2}). \tag{4}$$

By (2) and

$$e^{2c-x-y} \geq 1 + 2c - x - y \tag{5}$$

we deduce that

$$f(x,y) \geq e^{-2c}(1 + 2c - x - y + |x+y|) \geq e^{-2c}(1+2c) = f(c,c).$$

Therefore, (c,c) is a global minimum point for f.

We claim that $(0,0)$ is not a local extremum for f. Indeed, we have

$$f(x,x)-f(0,0)=e^{-2x}+a\sqrt{2}|x|-1. \tag{6}$$

By (6) it follows that $f(x,x)-f(0,0)>0,\ \forall\, x<0$. Hence $(0,0)$ cannot be a local maximum for f.

On the other hand, if $x>0$, then

$$f(x,x)-f(0,0)=e^{-2x}+a\sqrt{2}x-1=x\left(\frac{e^{-2x}-1}{x}+a\sqrt{2}\right). \tag{7}$$

Since

$$\lim_{x\to 0}\left(\frac{e^{-2x}-1}{x}+a\sqrt{2}\right)=-2+a\sqrt{2}<0,$$

there exists $r>0$ such that

$$\frac{e^{-2x}-1}{x}+a\sqrt{2}<0,\ \forall\, x\in(-r,r)\setminus\{0\}. \tag{8}$$

By (7) and (8) it follows that $f(x,x)-f(0,0)<0,\ \forall\, x\in(0,r)$. Hence $(0,0)$ cannot be a local minimum for f. □

Problem 2. a) Determine $a,b\in\mathbb{R}$ such that

$$\int_0^{\pi}(ax+bx^2)\cos nx\,\mathrm{d}x=\frac{1}{n^2}\quad\text{for all } n\in\mathbb{N}^*.$$

b) Prove that $\lim\limits_{n\to\infty}\left(\frac{1}{1^2}+\frac{1}{2^2}+\cdots+\frac{1}{n^2}\right)=\frac{\pi^2}{6}$.

Cristian Vladimirescu

This problem deals with the computation of the sum of the series $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}$ *by means of Fourier series. It did not raise special difficulties for the students.*

Solution. a) For every $n\in\mathbb{N}^*$ one has

$$\begin{aligned}\frac{1}{n^2}&=\frac{1}{n}\int_0^{\pi}(\sin nx)'(ax+bx^2)\mathrm{d}x=\frac{1}{n^2}\int_0^{\pi}(\cos nx)'(a+2bx)\mathrm{d}x\\&=\frac{1}{n^2}\left[(-1)^n(a+2b\pi)-a\right].\end{aligned}$$

From this equality it follows immediately that $a=-1$ and $b=\dfrac{1}{2\pi}$.

b) By a) we deduce that

$$\sum_{k=1}^{n}\frac{1}{k^2}=\int_0^{\pi}\left(\frac{x^2}{2\pi}-x\right)\sum_{k=1}^{n}\cos kx\mathrm{d}x.$$

Since

$$\sum_{k=1}^{n} \cos kx = \begin{cases} \dfrac{\sin \dfrac{nx}{2} \cos \dfrac{(n+1)x}{2}}{\sin \dfrac{x}{2}}, & x \in (0,\pi], \\ n, & x = 0, \end{cases}$$

$$= \begin{cases} \dfrac{1}{2} \operatorname{ctg} \dfrac{x}{2} \sin nx + \dfrac{1}{2} \cos nx - \dfrac{1}{2}, & x \in (0,\pi], \\ n, & x = 0, \end{cases}$$

it follows that the function $f : [0,\pi] \to \mathbb{R}$, defined by

$$f(x) := \sum_{k=1}^{n} \cos kx \quad \text{for all } x \in [0,\pi],$$

is continuous on $[0,\pi]$; notice that

$$\sum_{k=1}^{n} \frac{1}{k^2} = \int_0^{\pi} \left(\frac{x^2}{2\pi} - x \right) f(x) \mathrm{d}x.$$

The function $g : [0,\pi] \to \mathbb{R}$, defined by

$$g(x) = \begin{cases} \dfrac{1}{2} \left(\dfrac{x^2}{2\pi} - x \right) \operatorname{ctg} \dfrac{x}{2}, & x \in (0,\pi], \\ -1, & x = 0, \end{cases}$$

is continuously differentiable on $[0,\pi]$.

Therefore, we have

$$\sum_{k=1}^{n} \frac{1}{k^2} = \int_0^{\pi} g(x) \sin nx \mathrm{d}x + \int_0^{\pi} \frac{1}{2} \left(\frac{x^2}{2\pi} - x \right) \cos nx \mathrm{d}x -$$
$$- \int_0^{\pi} \frac{1}{2} \left(\frac{x^2}{2\pi} - x \right) \mathrm{d}x$$
$$= \frac{1}{n} \int_0^{\pi} g'(x) \cos nx \mathrm{d}x - \frac{1}{n} + \frac{1}{2n^2} + \frac{\pi^2}{6}$$

and

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \frac{\pi^2}{6},$$

because

$$\lim_{n\to\infty} \frac{1}{n} \left| \int_0^{\pi} g'(x) \cos nx \mathrm{d}x \right| \le \lim_{n\to\infty} \frac{1}{n} \int_0^{\pi} |g'(x)| \mathrm{d}x = 0. \quad \square$$

Problem 3. Consider the vector space $V = C[0,2\pi]$ and the endomorphism $T : V \to V$ defined by

$$T(f)(x) = \int_0^{2\pi} 4\sin^3(x+y)f(y)\mathrm{d}y, \quad f \in C[0,2\pi],\ x \in [0,2\pi].$$

a) Determine 2014 linearly independent functions in $\operatorname{Ker} T$.

b) Determine all nonzero eigenvalues of T, as well as their corresponding eigenvectors.

Vasile Pop

This problem deals with the theory of linear integral operators of the type $T(f)(x) = \int_a^b K(x,y)f(y)\mathrm{d}y$, *in the special case when the kernel* $K(x,y) = \sum_{i=1}^{n} a_i(x)b_i(y)$ *is degenerated. The problem reduces to solving a simple Fredholm integral equation.*

Solution. Since

$$4\sin^3 x = 3\sin x - \sin 3x \quad \text{and} \quad \sin(x+y) = \sin x\cos y + \cos x \sin y,$$

we have

$$\begin{aligned} T(f)(x) &= \left(\int_0^{2\pi} 3\cos y f(y)\mathrm{d}y\right)\sin x + \left(\int_0^{2\pi} 3\sin y f(y)\mathrm{d}y\right)\cos x \\ &\quad + \left(\int_0^{2\pi} -\cos 3y f(y)\mathrm{d}y\right)\sin 3x + \left(\int_0^{2\pi} -\sin 3y f(y)\mathrm{d}y\right)\cos 3x \\ &\overset{(*)}{=} I_1(f)\sin x + I_2(f)\cos x + I_3(f)\sin 3x + I_4(f)\cos 3x,\ x \in [0,2\pi]. \end{aligned}$$

a) $f \in \ker T \ \Leftrightarrow\ I_1(f) = I_2(f) = I_3(f) = I_4(f) = 0$.

Note that the functions $\sin 2x$, $\cos 2x$, $\sin 4x$, $\cos 4x$, $\sin 5x$, $\cos 5x$, $\ldots$, $\sin 1009x$, $\cos 1009x$ are linearly independent and that all the above four integrals vanish because

$$\int_0^{2\pi} \sin kx \sin px\,\mathrm{d}x = \int_0^{2\pi} \cos kx \cos px\,\mathrm{d}x = \int_0^{2\pi} \sin kx \cos px\,\mathrm{d}x = 0$$

for all $k, p \in \mathbb{N}$, $k \neq p$.

b) If λ is a nonzero eigenvalue of T, then the corresponding eigenvectors belong to the image of T: $T(f) = \lambda f \Rightarrow f = T\left(\dfrac{1}{\lambda}f\right)$.

According to $(*)$, the image $g = T(f)$ of every function f is a linear combination of the functions $\sin x$, $\cos x$, $\sin 3x$, $\cos 3x$. Therefore, every eigenvector is of the form

$$f(x) = a\sin x + b\cos x + c\sin 3x + d\cos 3x, \quad x \in [0,2\pi].$$

For such a function we obtain

$$I_1(f) = 3\pi b, \quad I_2(f) = 3\pi a, \quad I_3(f) = -\pi d, \quad I_4(f) = -\pi c.$$

Then the equality $T(f) = \lambda f$, with $\lambda \in \mathbb{R}^*$ eigenvalue of T and f a corresponding eigenvector, is equivalent to the system

$$\lambda a = 3\pi b, \quad \lambda b = 3\pi a, \quad \lambda c = -\pi d, \quad \lambda d = -\pi c.$$

We get

$$(\lambda^2 - 9\pi^2)ab = 0 \quad \text{and} \quad (\lambda^2 - \pi^2)cd = 0.$$

If $\lambda \neq \pm\pi$ and $\lambda \neq \pm 3\pi$ then $a = b = c = d = 0 \Rightarrow f = 0$ (which is not convenient). Thus we have the nonzero eigenvalues $\lambda_1 = \pi$, $\lambda_2 = -\pi$, $\lambda_3 = 3\pi$, $\lambda_4 = -3\pi$ with the corresponding eigenvectors

$$\begin{aligned} f_1(x) &= a(\sin 3x - \cos 3x), \quad & f_2(x) &= a(\sin 3x + \cos 3x), \\ f_3(x) &= a(\sin x + \cos x), & f_4(x) &= a(\sin x - \cos x), \end{aligned}$$

for all $x \in [0, 2\pi]$ and $a \in \mathbb{R}^*$. □

Problem 4. Let a, b, c, and n be positive integers such that

$$0 \le a + b - n \le c \le a \le b \le n.$$

Prove that for every matrix $C \in \mathcal{M}_n(\mathbb{C})$ with $\operatorname{rank} C = c$ there exist two matrices $A, B \in \mathcal{M}_n(\mathbb{C})$ such that $\operatorname{rank} A = a$, $\operatorname{rank} B = b$, and $C = AB$.

Vasile Pop

This problem is a converse of Sylvester's rank inequality: if $A, B \in \mathcal{M}_n(\mathbb{C})$, then $\operatorname{rank} A + \operatorname{rank} B - n \le \operatorname{rank}(AB) \le \min\{\operatorname{rank} A, \operatorname{rank} B\}$. In particular, it asserts that every natural number $c \in [a+b-n, a]$ equals the rank of the product of two matrices whose ranks are a and b, respectively.

Solution. Since $\operatorname{rank} C = c$, there exist two invertible matrices P and Q in $\mathcal{M}_n(\mathbb{C})$ such that

$$C = P \left[\begin{array}{c|c} I_c & 0 \\ \hline 0 & 0 \end{array}\right] Q,$$

where I_c denotes the identity matrix of size c.

We denote by $[a_1, \ldots, a_n]$ the diagonal matrix having on its main diagonal the entries $a_1, \ldots, a_n$. Set $D = [\underbrace{1, \ldots, 1}_{c}, \underbrace{0, \ldots, 0}_{n-c}]$. Then $C = PDQ$.

It suffices to prove that there exist a matrix A_1 of rank a and a matrix B_1 of rank b such that $A_1 B_1 = D$. Then the matrices $A := PA_1$ and $B := B_1 Q$ (with $\operatorname{rank} A = \operatorname{rank} A_1 = a$ and $\operatorname{rank} B = \operatorname{rank} B_1 = b$) satisfy $C = AB$.

Set

$$A_1 := [\underbrace{1, \ldots, 1}_{a}, \underbrace{0, \ldots, 0}_{n-a}] = [\underbrace{1, \ldots, 1}_{c}, \underbrace{1, \ldots, 1}_{a-c}, \underbrace{0, \ldots, 0}_{n-a}]$$

and

$$B_1 := [\underbrace{1, \ldots, 1}_{c}, \underbrace{0, \ldots, 0}_{n-b}, \underbrace{1, \ldots, 1}_{b-c}].$$

Since $a + b - n \le c \iff a - c \le n - b$, it follows that $A_1 B_1 = D$. □

NOTE MATEMATICE

A new proof of Finsler-Hadwiger reverse inequality in non-obtuse triangles

ROBERTO BOSCH[1)]

Abstract. In this note we give a new proof of Finsler-Hadwiger reverse inequality in non-obtuse triangles.

Keywords: reverse Finsler-Hadwiger inequality, non-obtuse triangles, stationary points
MSC: 51M04

Dedicated to professor Henry Ricardo, Medgar Evers College (CUNY), NY, USA.

In general the Finsler-Hadwiger reverse inequality states that in any triangle ABC with sides a, b, c the following inequality is valid

$$a^2+b^2+c^2 \leq 4\sqrt{3}S + k\left[(a-b)^2+(b-c)^2+(c-a)^2\right],$$

where S denotes the area of the triangle ABC and $k=3$ [3]. For non-obtuse triangles the constant k was improved in paper [2] to $k=2$ and later to $k=\frac{6-\sqrt{6}}{2}$, and at the end it is conjectured that $k=\frac{2-\sqrt{3}}{3-2\sqrt{2}}$ is optimal. This conjecture was verified in [1]. Here we present a new proof using calculus and trigonometry.

We shall prove the following result:

Theorem 1. *In any non-obtuse triangle ABC the following inequality holds*

$$a^2+b^2+c^2 \leq 4\sqrt{3}S + \frac{2-\sqrt{3}}{3-2\sqrt{2}}\left[(a-b)^2+(b-c)^2+(c-a)^2\right],$$

where a, b, c are the sides and S is the area. The constant $\frac{2-\sqrt{3}}{3-2\sqrt{2}}$ is optimal and it is attained for a right angled isosceles triangle.

Proof. Using the formulas $a=2R\sin A, b=2R\sin B, c=2R\sin C$ and moving everything to the right side we rewrite the inequality as

$$f(A,B,C) = -\sin^2 A - \sin^2 B - \sin^2 C + 2\sqrt{3}\sin A \sin B \sin C$$

$$+\frac{2-\sqrt{3}}{3-2\sqrt{2}}[(\sin A-\sin B)^2+(\sin B-\sin C)^2+(\sin C-\sin A)^2] \geq 0,$$

with $0 \leq A, B, C \leq \frac{\pi}{2}$ and $A+B+C=\pi$. These restrictions are the intersection of a cube $([0,\frac{\pi}{2}]\times[0,\frac{\pi}{2}]\times[0,\frac{\pi}{2}])$ and a plane, so the resulting

1) Archimedean Academy 12425 SW 72nd St, Miami, FL, USA, 33183, bobbydrg@gmail.com

region in $\mathbb{R}^3$ is a compact set K. Since f is continuous, its minimum on K exists. To find this value we consider the system of equations

$$\frac{\partial f}{\partial A} = \frac{\partial f}{\partial B} = \frac{\partial f}{\partial C} = 0,$$

that is to say

$$\cos A \left[-\sin A + \sqrt{3}\sin B \sin C + \frac{2-\sqrt{3}}{3-2\sqrt{2}}\,(2\sin A - \sin B - \sin C) \right] = 0,$$

$$\cos B \left[-\sin B + \sqrt{3}\sin A \sin C + \frac{2-\sqrt{3}}{3-2\sqrt{2}}\,(2\sin B - \sin A - \sin C) \right] = 0,$$

$$\cos C \left[-\sin C + \sqrt{3}\sin A \sin B + \frac{2-\sqrt{3}}{3-2\sqrt{2}}\,(2\sin C - \sin B - \sin A) \right] = 0.$$

Supposing $A, B, C \neq \frac{\pi}{2}$ and solving the system with the aid of *Maple*, the solutions are found to be $(\pi, 0, 0), (0, \pi, 0), (0, 0, \pi)$ and $(-\pi, \pi, \pi)$, $(\pi, -\pi, \pi)$, $(\pi, \pi, -\pi)$. None of them is on the considered region K, so that $A = \frac{\pi}{2}$ or $B = \frac{\pi}{2}$ or $C = \frac{\pi}{2}$. In any case we are on the boundary of the region. Assuming $C = \frac{\pi}{2}$ (similarly for A and B), we need to prove that

$$-2 + 2\sqrt{3}\sin A \cos A + \frac{2-\sqrt{3}}{3-2\sqrt{2}}\,[4 - 2(\sin A \cos A + \sin A + \cos A)] \geq 0,$$

which is equivalent to

$$\frac{4-2\sqrt{3}}{3-2\sqrt{2}} - 1 + \left(\sqrt{3} - \frac{2-\sqrt{3}}{3-2\sqrt{2}} \right) \sin A \cos A - \frac{2-\sqrt{3}}{3-2\sqrt{2}}(\sin A + \cos A) \geq 0,$$

with $0 \leq A \leq \frac{\pi}{2}$. Define

$$g(A) = \frac{4-2\sqrt{3}}{3-2\sqrt{2}} - 1 + \left(\sqrt{3} - \frac{2-\sqrt{3}}{3-2\sqrt{2}} \right) \sin A \cos A - \frac{2-\sqrt{3}}{3-2\sqrt{2}}(\sin A + \cos A).$$

The function g is increasing on $\left[\frac{\pi}{4}, \frac{\pi}{2}\right]$ because

$$g'(A) = (\cos A - \sin A) \left[\left(\sqrt{3} - \frac{2-\sqrt{3}}{3-2\sqrt{2}} \right) (\cos A + \sin A) - \frac{2-\sqrt{3}}{3-2\sqrt{2}} \right] \geq 0,$$

(note that $\tan A \geq 1$ and $\cos A + \sin A \leq \sqrt{2}$). Observing that one also has $g(A) = g(\frac{\pi}{2} - A)$, it follows

$$g(A) \geq g\left(\frac{\pi}{4}\right) = 0 \quad \text{for} \quad 0 \leq A \leq \frac{\pi}{2}.$$

It just remains to show that $k = \frac{2-\sqrt{3}}{3-2\sqrt{2}}$ is sharp, which is equivalent to find A, B, C such that

$$h(A,B,C) = \frac{\sin^2 A + \sin^2 B + \sin^2 C - 2\sqrt{3}\sin A \sin B \sin C}{(\sin A - \sin B)^2 + (\sin B - \sin C)^2 + (\sin C - \sin A)^2} > k - \varepsilon$$

for any $\varepsilon > 0$. This is clear from $h\left(\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{4}\right) = k$. $\square$

Acknowledgements. The author would like to thank the referee for his valuable suggestions.

References

[1] A. Cipu, Optimal reverse Finsler-Hadwiger inequalities, *Gaz. Mat. Ser. A*, **30(109)** (2012), 62–68.

[2] C. Lupu, C. Mateescu, V. Matei, M. Opincariu, Refinements of the Finsler-Hadwiger reverse inequality, *Gaz. Mat. Ser. A*, **28(107)** (2010), 130–133.

[3] P. G. Popescu, I. V. Maftei, J. L. Diaz-Barrero, M. Dincă, *Inegalităţi Matematice: Modele inovatoare*, Editura Didactică şi Pedagogică, 2007.

A computational proof of the Cayley-Hamilton theorem

CONSTANTIN-NICOLAE BELI[1]

Abstract. We give a computational proof of the Cayley-Hamilton theorem which states that if A is a square matrix over some field K and P_A is its characteristic polynomial $P_A(X) = \det(XI - A)$, then $P_A(A) = 0$.

Keywords: characteristic polynomial, Cayley-Hamilton theorem

MSC: 15A15, 15A24

Let K be a field. If $m, n \geq 1$ then we denote by $\{e_{i,j}^{m,n} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ the canonical basis of $M_{m,n}(K)$, where $e_{i,j}^{m,n}$ has 1 on the (i,j) position and 0 everywhere else. We have $e_{i,j}^{l,m} e_{k,l}^{m,n} = \delta_{j,k} e_{i,l}^{l,n}$.

If $n_0, \ldots, n_s \geq 1$ and for $1 \leq t \leq s$ we have $A_t = (a_{i,j}^t) \in M_{n_{t-1},n_t}(K)$ then $A_t = \sum_{i,j} a_{i,j}^t e_{i,j}^{n_{t-1},n_t}$, where i goes from 1 to n_{t-1} and j from 1 to n_t. We get $A_1 \cdots A_s = \sum a_{i_1,j_1}^1 \cdots a_{i_s,j_s}^s e_{i_1,j_1}^{n_0,n_1} \cdots e_{i_s,j_s}^{n_{s-1},n_s}$, with $1 \leq i_t \leq n_{t-1}$ and $1 \leq j_t \leq n_t$ $\forall t$. But

$$e_{i_1,j_1}^{n_0,n_1} \cdots e_{i_s,j_s}^{n_{s-1},n_s} = \delta_{j_1,i_2} \cdots \delta_{j_{s-1},i_s} e_{i_1,j_s}^{n_0,n_s} = \begin{cases} e_{i_1,j_s}^{n_0,n_s} & \text{if } j_t = i_{t+1} \ \forall 1 \leq t < s, \\ 0 & \text{otherwise}. \end{cases}$$

---

[1] Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania, `constantin.beli@imar.ro`

It follows that $A_1 \cdots A_s = \sum a^1_{k_0,k_1} a^2_{k_1,k_2} \cdots a^s_{k_{s-1},k_s} e^{n_0,n_s}_{k_0,k_s}$, where $1 \leq k_t \leq n_t$ for all t with $0 \leq t \leq s$. In particular, the (i,j) entry of $A_1 \cdots A_s$ is $\sum\limits_{k_0=i,\, k_s=j} a^1_{k_0,k_1} \cdots a^s_{k_{s-1},k_s}$.

When we take $A_1 = \cdots = A_s = A = (a_{i,j}) \in M_n(K)$ we get:

Lemma 1. *For any $s \geq 0$ and $1 \leq i,j \leq n$ the (i,j) entry of A^s is $\sum\limits_{k_0=i,\, k_s=j} \prod\limits_{t=1}^{s} a_{k_{t-1},k_t}$, where $1 \leq k_t \leq n$ for all t with $0 \leq t \leq s$.*

This result also holds when $s = 0$ because, if we make the convention that any sum over an empty set is 0 and any product over an empty set is 1, we have

$$\sum_{k_0=i,\, k_0=j} \prod_{t=1}^{0} a_{k_{t-1},k_t} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Indeed if $i = j$ then $\sum\limits_{k_0=i,\, k_0=j}$ has only one term, corresponding to $k_0 = i$, and this term is $\prod\limits_{t=1}^{0} a_{k_{t-1},k_t} = 1$, and if $i \neq j$ then the sum is taken over the empty set so it is 0.

Given a cyclic permutation $\sigma \in S_n$, we denote by $\{\sigma\}$ the set of all elements of σ, by $\ell(\sigma)$ the length of σ, $\ell(\sigma) = |\{\sigma\}|$, and put $P(\sigma) := \{(h, \sigma(h)) \mid h \in \{\sigma\}\}$. So if $\sigma = (h_1, \ldots, h_q)$ then $\{\sigma\} = \{h_1, \ldots, h_q\}$, $\ell(\sigma) = q$, and $P(\sigma) = \{(h_1, h_2), (h_2, h_3) \ldots, (h_q, h_1)\}$.

Lemma 2. *If $P_A(X) = c_n X^n + \cdots + c_0$ then*

$$c_t = \sum_{\{\sigma_1,\ldots,\sigma_m\}} (-1)^m \prod_{(h,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{h,l},$$

where $\{\sigma_1, \ldots, \sigma_m\}$ covers all sets of mutually disjoint cyclic permutations in S_n such that $\ell(\sigma_1) + \cdots + \ell(\sigma_m) = n - s$.

Proof. We have $XI - A = (a'_{i,j})$, where $a'_{i,i} = X - a_{i,i}$ and $a'_{i,j} = -a_{i,j}$ if $i \neq j$. Hence $P_A(X) = \det(XI - A) = \sum\limits_{\sigma\in S_n} \varepsilon(\sigma) \prod\limits_{h=1}^{n} a'_{h,\sigma(h)} = \sum\limits_{\sigma\in S_n} \varepsilon(\sigma) \sum\limits_{b_1,\ldots,b_n} \prod\limits_{h=1}^{n} b_h$, where $b_h = -a_{h,\sigma(h)}$ if $\sigma(h) \neq h$ and $b_h \in \{-a_{h,h}, X\}$ if $\sigma(h) = h$.

For a given $\sigma \in S_n$ and a choice of $b_1, \ldots, b_n$ the set $\{1, \ldots, n\}$ writes as a disjoint union $J \cup J'$ such that $b_h = -a_{h,\sigma(h)}$ for $h \in J$ and $b_h = X$ for $h \in J'$. Moreover J' is contained in $\mathrm{Fix}(\sigma)$, the set of all fixed points of σ. It follows that

$$P_A(X) = \sum_{\sigma,J,J'} \varepsilon(\sigma) \prod_{h\in J} (-a_{h,\sigma(h)}) X^{|J'|} = \sum_{\sigma,J,J'} (-1)^{|J|} \varepsilon(\sigma) \prod_{h\in J} a_{h,\sigma(h)} X^{n-|J|},$$

where σ covers S_n and (J, J') covers all partitions of $\{1, \ldots, n\}$ into two sets such that $J' \subseteq \mathrm{Fix}(\sigma)$. But for σ, J, J' with these properties σ decomposes uniquely as a product of mutually disjoint cycles $\sigma = \sigma_1 \cdots \sigma_m$ with $\{\sigma_1\} \cup \cdots \cup \{\sigma_m\} = J$. (Every $h \in J \cap \mathrm{Fix}(\sigma)$ is included as a cycle of length one in $\{\sigma_1 \cdots \sigma_m\}$). The mapping $(\sigma, J, J') \mapsto \{\sigma_1 \cdots \sigma_m\}$ is one-to-one as if $\{\sigma_1 \cdots \sigma_m\}$ of mutually disjoint cycles from S_n then σ, J, J' are given by $\sigma = \sigma_1 \cdots \sigma_n$, $J = \{\sigma_1\} \cup \cdots \cup \{\sigma_m\}$ and $J' = \{1, \ldots, n\} \setminus J$. Then we get $|J| = |\{\sigma_1\}| + \cdots + |\{\sigma_m\}| = \ell(\sigma_1) + \cdots + \ell(\sigma_m)$. Therefore

$$(-1)^{|J|}\varepsilon(\sigma) = (-1)^{|J|}(-1)^{(\ell(\sigma_1)-1)+\cdots+(\ell(\sigma_m)-1)} = (-1)^m.$$

Also

$$\begin{aligned}\{(h,\sigma(h)) \mid h \in J\} &= \bigcup_{r=1}^{m} \{(h,\sigma(h)) \mid h \in \{\sigma_r\}\} \\ &= \bigcup_{r=1}^{m} \{(h,\sigma_r(h)) \mid h \in \{\sigma_r\}\} = \bigcup_{r=1}^{m} P(\sigma_r).\end{aligned}$$

In conclusion:

$$P_A(X) = \sum_{\{\sigma_1,\ldots,\sigma_m\}} \left((-1)^m \prod_{(h,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{h,l} \right) X^{n-\ell(\sigma_1)-\cdots-\ell(\sigma_m)}.$$

In order to obtain the monomial $c_s X^s$ of $P_A(X)$ we have to restrict ourselves to terms corresponding to sets $\{\sigma_1, \ldots, \sigma_m\}$ with $n - \ell(\sigma_1) - \cdots - \ell(\sigma_m) = s$, i.e., with $\ell(\sigma_1) + \cdots + \ell(\sigma_m) = n - s$. Hence we get our result. □

From Lemma 1 and 2 we get that the (i,j) entry of the matrix $P_A(A) = c_n A^n + \cdots + c_0 I$, viz., $\sum_{s=0}^{n} c_s A^s(i,j)$, is

$$\begin{aligned}\sum_{s=0}^{n} \left(\sum_{h_0=i,\, h_s=j} \prod_{t=1}^{s} a_{h_{t-1},h_t} \right) \left(\sum_{\ell(\sigma_1)+\cdots+\ell(\sigma_m)=n-s} (-1)^m \prod_{(k,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{k,l} \right) \\ = \sum_{\alpha\in S_{i,j}} F(\alpha),\end{aligned}$$

where $S_{i,j}$ is the set of all $(h_0, \ldots, h_s, \{\sigma_1, \ldots, \sigma_m\})$ with $1 \le h_t \le n$ for all t with $0 \le t \le s$, $h_0 = i$, $h_s = j$, and σ_r are mutually disjoint cycles from S_n such that $s + \ell(\sigma_1) + \cdots + \ell(\sigma_m) = n$ and $F : S_{i,j} \to K$ is given by

$$(h_0, \ldots, h_s, \{\sigma_1, \ldots, \sigma_m\}) \mapsto (-1)^m \prod_{t=1}^{s} a_{h_{t-1},h_t} \prod_{(k,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{k,l}.$$

Remark. If $\alpha = (h_0, \ldots, h_s, \{\sigma_1 \ldots, \sigma_m\}) \in S_{i,j}$ then either there are $0 \le v < u \le s$ with $h_v = h_u$ or $\{h_0, \ldots, h_s\} \cap \bigcup_{r=1}^{m} \{\sigma_r\} \ne \emptyset$. Otherwise we would have $|\{h_0,\dots,h_s\}\cup\bigcup_{r=1}^{m}\{\sigma_r\}| = s+1+\sum_{r=1}^{m}\ell(\sigma_r) = n+1$, which is impossible since $\{h_0,\dots,h_s\}\cup\bigcup_{r=1}^{m}\{\sigma_r\}\subseteq\{1,\dots,n\}$.

Definition. We now define $\phi : S_{i,j}\to S_{i,j}$ as follows.

Let $\alpha\in S_{i,i}$, $\alpha=(h_0,\dots,h_s,\{\sigma_1\dots,\sigma_m\})$. If there are $v<u$ with $h_v=h_u$ we consider such v,u with u minimal. Hence $h_0,\dots,h_{u-1}$ are mutually distinct. We have two cases:

(I) If $\{h_0,\dots,h_{u-1}\}\cap\bigcup_{r=1}^{m}\{\sigma_r\}=\emptyset$ then $h_v,\dots,h_{u-1}$ are mutually distinct and they don't belong to any of $\{\sigma_1\},\dots,\{\sigma_m\}$.

Hence $\sigma_{m+1}:=(h_v,\dots,h_{u-1})$ is a cycle in S_n disjoint form $\sigma_1,\dots,\sigma_m$. Then we define $\phi(\alpha)=(h_0,\dots,h_v=h_u,h_{u+1},\dots,h_s,\{\sigma_1,\dots,\sigma_m,\sigma_{m+1}\})$.

(II) If $\{h_0,\dots,h_{u-1}\}\cap\bigcup_{r=1}^{m}\{\sigma_r\}\neq\emptyset$ (or if v,u are not defined) then let w be minimal such that $h_w\in\bigcup_{r=1}^{m}\{\sigma_r\}$. Let x with $h_w\in\{\sigma_x\}$. Then we write $\sigma_x=(k_1,\dots,k_q)$ with $k_1=h_w$. We define

$$\phi(\alpha)=(h_0,\dots,h_w=k_1,\dots,k_q,k_1=h_w,h_{w+1},\dots,h_s,\{\sigma_1,\dots,\widehat{\sigma}_x.\dots,\sigma_m\})\,.$$

(The existence of w from the case (II) when v,u are not defined is ensured by the Remark above.)

Our definition is good in the sense that $\phi(\alpha)\in S_{i,j}$. Indeed, the condition $h_0=i$, $h_s=j$ from the definition of $S_{i,j}$ is obviously satisfied by $\phi(\alpha)$ in both cases (I) and (II). For the condition $s+\ell(\sigma_1)+\cdots+\ell(\sigma_m)=n$ note that when going from α to $\phi(\alpha)$ in the case (I) by removing $h_v,\dots,h_{u-1}$ from $h_0,\dots,h_s$ we decreased s by $u-v$ but by adding σ_{m+1} to the set $\{\sigma_1,\dots,\sigma_m\}$ we increased $\ell(\sigma_1)+\cdots+\ell(\sigma_m)$ by $u-v=\ell(\sigma_{u+1})$. In the case (II) by adding $k_1,\dots,k_q$ to $h_0,\dots,h_s$ we increased s by q and by removing σ_x from $\{\sigma_1,\dots,\sigma_m\}$ the sum $\ell(\sigma_1)+\cdots+\ell(\sigma_m)$ decreased by $\ell(\sigma_x)=q$. In both cases the sum $s+\ell(\sigma_1)+\cdots+\ell(\sigma_m)=n$ is left unchanged by the transformation $\alpha\mapsto\phi(\alpha)$.

Lemma 3. *(i)* $F\circ\phi=-F$.

(ii) $\phi\circ\phi=\mathrm{id}$.

Proof. Let $\alpha\in S_{i,j}$, $\alpha=(h_0,\dots,h_s,\{\sigma_1,\dots,\sigma_m\})$.

(*i*) If α is in the case (I) of the definition of ϕ then after the transformation $\alpha\mapsto\phi(\alpha)$ the factors $a_{h_v,h_{v+1}},\dots,a_{h_{u-1},h_u}=a_{h_{u-1},h_v}$ were removed from $\prod_{t=1}^{s}a_{h_{t-1},h_t}$ but they were added to $\prod_{(k,l)\in P(\sigma_1)\cup\dots\cup P(\sigma_m)}a_{k,l}$ as $\prod_{(k,l)\in P(\sigma_{m+1})}a_{k,l}=a_{h_v,h_{v+1}}\cdots a_{h_{u-2},h_{u-1}}a_{h_{u-1},h_v}$. If α is in the case (II) then the factors $\prod_{(k,l)\in P(\sigma_x)} a_{k,l} = a_{k_1,k_2}\cdots a_{k_{q-1},k_q}a_{k_q,k_1}$ are removed from $\prod_{(k,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{k,l}$ but they are added to $\prod_{t=1}^{s} a_{h_{t-1},h_t}$.

In both cases the product $\prod_{t=1}^{s} a_{h_{t-1},h_t} \prod_{(k,l)\in P(\sigma_1)\cup\cdots\cup P(\sigma_m)} a_{k,l}$ from the definition of $F(\alpha)$ is preserved. However the factor $(-1)^m$ from $F(\alpha)$ is replaced in $F(\phi(\alpha))$ by $(-1)^{m\pm 1} = -(-1)^m$. Therefore $F(\phi(\alpha)) = -F(\alpha)$.

(ii) We consider the two cases from the definition of ϕ.

If α is in the case (I) then $h_v \in \{\sigma_{m+1}\}$. Since $v \le u-1$ we have that $h_0,\dots,h_v$ are different from each other and $\{h_0,\dots,h_{v-1}\} \cap \bigcup_{r=1}^{m}\{\sigma_r\} = \emptyset$. We also have $\{h_0,\dots,h_{v-1}\}\cap\{\sigma_{m+1}\} = \{h_0,\dots,h_{v-1}\}\cap\{h_v,\dots,h_{u-1}\} = \emptyset$. Therefore $\phi(\alpha)$ is in the case (II) with $w = v$ and $x = m+1$. Then $\phi(\phi(\alpha))$ is obtained from $\phi(\alpha)$ by removing $\sigma_{m+1} = (h_v,\dots,h_{u-1})$ from $\{\sigma_1,\dots,\sigma_m,\sigma_{m+1}\}$ and by inserting the sequence $h_v,\dots,h_{u-1}$ into $h_0,\dots,h_{v-1},h_v = h_u,h_{u+1},\dots,h_s$, between h_{v-1} and h_u.

This means $\phi(\phi(\alpha)) = \alpha$.

If α is in the case (II) then $h_0,\dots,h_w$ are different from each other and $\{h_0,\dots,h_{w-1}\} \cap \bigcup_{r=1}^{m}\{\sigma_r\} = \emptyset$. In particular,

$$\{h_0,\dots,h_{w-1}\}\cap\{k_1,\dots,k_q\} = \{h_0,\dots,h_{w-1}\}\cap\{\sigma_x\} = \emptyset.$$

Thus $h_0,\dots,h_{w-1},h_w = k_1,\dots,k_q$ are different from each other. Also $\{k_1,\dots,h_q\}\cap\bigcup_{r\neq x}\{\sigma_r\} = \{\sigma_x\}\cap\bigcup_{r\neq x}\{\sigma_r\} = \emptyset$.

Hence $\{h_0,\dots,h_{w-1},h_w = k_1,\dots,k_q\}\cap\bigcup_{r\neq x}\{\sigma_r\} = \emptyset$. It follows that $\phi(\alpha)$ is in the case (I) with $v = w$, $u = w+q$. Then $\phi(\phi(\alpha))$ is obtained from $\phi(\alpha)$ by removing the sequence $k_1,\dots,k_q$ from the sequence $h_0,\dots,h_w = k_1,\dots,k_q,k_1 = h_w,h_{w+1},\dots,h_s$ and adding $(k_1,\dots,k_q) = \sigma_x$ to $\{\sigma_1,\dots,\hat{\sigma}_x,\dots,\sigma_m\}$. Hence $\phi(\phi(\alpha)) = \alpha$. □

Now we are in a position to complete our proof for the Cayley-Hamilton theorem.

We define on $S_{i,j}$ the relation $\sim$ given by $\alpha\sim\beta$ if $\beta = \phi^a(\alpha)$ for some a. Then $\sim$ is an equivalence relation and the class of α is $\{\alpha,\phi(\alpha)\}$. Note also that $\alpha\neq\phi(\alpha)$. If $\alpha_1,\dots,\alpha_N$ is a system of representatives for $\sim$ then $S_{i,j} = \{\alpha_1,\phi(\alpha_1),\dots,\alpha_N,\phi(\alpha_N)\}$. It follows that for any i,j the (i,j) entry of $P_A(A)$ is $\sum_{\alpha\in S_{i,j}} F(\alpha) = \sum_{k=1}^{N}(F(\alpha_k)+F(\phi(\alpha_k))) = 0$. Thus $P_A(A) = 0$.

Note added in proof. After the submission of this note, an inductive proof of the Cayley-Hamilton theorem has been published [1].

REFERENCES

[1] N. Anghel, An inductive proof of the Cayley-Hamilton theorem, *Recreaţii Ştiinţifice* **16** (2014), 127–128.

Existence of a Hamiltonian path in a plane configuration

MARIUS CAVACHI[1)]

Abstract. We prove the existence of a hamiltonian path in a plane configuration given by intersecting circles.

Keywords: Circle, hamiltonian path
MSC: 05C10, 05C38

The following result has been previously published in [1] as a problem.

Theorem. *Consider a finite set of plane circles whose interiors have non–empty intersection such that any two circles intersect and any three do not pass through the same point. Then the graph whose vertices are intersection points of circles in the set and whose edges are the resulting circle segments admits a hamiltonian path.*

Proof. Consider for a fixed integer n, two configurations of circles, $\{A_i\}$ and $\{B_i\}$, as in the hypothesis, with $i \in \{1, \ldots, n\}$ and assume that all A_i and B_i contain Q as an interior point. (By hypothesis all A_i contain a point Q in interior and all B_i contain a point Q' in interior. But by translating the configuration $\{B_i\}$ by the vector $\overrightarrow{Q'Q}$ we may assume that $Q' = Q$.) The aim is to take $\{A_i\}$ as a particular configuration that we can show it has a hamiltonian path with certain properties to be described momentarily and then deform it continuously into an arbitrary configuration $\{B_i\}$, while showing that the existence of a hamiltonian path is preserved along the way.

Consider $\mathcal{S}$ a sphere of center O, tangent to the plane $\mathcal{P}$ of the two configurations at Q and let P be the antipodal of Q. Denote by $\pi : \mathcal{S} \setminus \{P\} \to \mathcal{P}$ the stereographic projection from P and denote by $\mathcal{C}_i$ and $\mathcal{D}_i$ the inverse images via π of A_i and B_i respectively. If O_i is the center of $\mathcal{C}_i$, then for $t \in [0, 1]$, construct $\mathcal{C}_i(t)$ to be the circle on $\mathcal{S}$, centered on $O_i(t)$ that is defined on $[OO_i]$ by $|OO_i(t)| = (1 - t)\,|OO_i|$ and such that the plane of $\mathcal{C}_i$ is parallel to that of $\mathcal{C}_i(t)$. Similarly, construct the circles $\mathcal{D}_i(t)$.

By infinitesimally perturbing the two configurations without affecting the existence of a hamiltonian path, we can assume that:

- Any three of the circles $\mathcal{C}_i(1)$ (or $\mathcal{D}_i(1)$) have empty intersection.
- The intersection of any four planes determined by the circles $\mathcal{C}_i$ (or $\mathcal{D}_i$) is also empty.

[1)] Department of Mathematics, Ovidius University of Constanţa, Constanţa, Romania, mcavachi@yahoo.com

If M is in the intersection of two circles $\mathcal{C}_i$, then for $t \in [0;1]$, we have that $M(t) \in [OM]$ defined by $|OM(t)| = (1-t)|OM|$ is contained in the intersection of the two discs bounded by the corresponding circles in the configuration $\mathcal{C}_i(t)$. It follows that for fixed t, any two circles $\mathcal{C}_i(t)$ intersect. Also, any four of the circles $\mathcal{C}_i(t)$ have empty intersection, otherwise if $M(t)$ is a common point, then M constructed by reverting the above procedure lies in the intersection of the planes containing the $\mathcal{C}_i$, contradicting our assumptions.

As t varies, the circles $\mathcal{C}_i$ slide continuously into the circles $\mathcal{C}_i(1)$. We can deform the configuration $\{\mathcal{C}_i(1)\}$ into $\{\mathcal{D}_i(1)\}$ and by reverting the sliding described above, we deform into the configuration $\{\mathcal{D}_i\}$. Applying π, we obtain a continuous deformation of $\{A_i\}$ into $\{B_i\}$ realized by a moving circle configuration $\{E_i(s)\}$ with $s \in [0;1]$ (unrelated to t) having $E_i(0) = A_i$ and $E_i(1) = B_i$. As s varies, the graph determined by the configuration $E_i(s)$ only modifies its isomorphism class when one of the circles crosses an intersection point of two others in the configuration. We remark that for any s, the circles $E_i := E_i(s)$ have pairwise nonempty intersection and that Q is an interior point. (For more details see [3].)

For $k \in \{1, \ldots, n\}$, denote by M_k the set of plane points contained in the interior of at least k of the circles E_i. The sets M_k are easily seen to be open and star shaped with respect to Q. If F_k denotes the boundary of M_k, one notices that $F_1, F_3, F_5, \ldots$ are disjoint cycles and their union contains all the vertices of the graph.

Remark 1. If two circles E_i and E_j cross each other in a point S then there is some k such that of the two segments on E_i and the two segments on E_j based in S one belongs to F_k and the other to F_{k+1}.

We will prove that there exists a hamiltonian path for the graph determined by the configuration E_i displaying the following properties:

(1) For all odd k, the path contains all but one of the edges of the cycle F_k.
(2) The path contains exactly one edge ℓ_k belonging to F_{k+1} that connects F_k to F_{k+2} for $k+2 \leq n$.
(3) The edge ℓ_k does not lie on the same circle in the configuration as any of the two edges it is adjacent to on the path.

Remark 2. The conditions above are very restrictive. Suppose that S', S, T, T' are consecutive points on some cycle F_K with k odd we know that ST is the segment of F_k missing from the hamiltonian path than the point of the path following S', S is uniquely defined. Indeed, by Remark 1, $S'S$ and ST belong to different circles, E_i and E_j, respectively. The other two segments that form at S belong to F_{k-1} or F_{k+1} and they are candidates at being ℓ_k or ℓ_{k-2}, respectively. But one of these segments lies on E_i, same as SS' so it does not qualify by the third rule above. So the remaining segment, say SS'' is the edge ℓ_k or ℓ_{k-2}. Similarly, the point T'' on the path following T', T is uniquely determined. Of course ST can be the segment of F_k missing from the path only if SS'' and TT'' belong one to F_{k-1} and the other to F_{k+1} so that one of them is ℓ_{k-2} and the other ℓ_k.

Also if we now that $\ell_k = ST$ for some segment ST on F_{k+1} then the point S' following T, S on the path is uniquely determined. Indeed, suppose that ST lies on E_i and E_i crosses E_j at S. Then, by Remark 1, of the three other segments that meet at S one lies on E_j and belongs to F_{k+1}. The other two belong to F_k or F_{k+2} an lie one on E_i, the other on E_j. The only one that qualifies to be SS' is the segment on E_j belonging to F_k or F_{k+2}. Similarly the point T' following S, T on the path is uniquely determined. Note that the segment ST belonging to F_{k+1} with k odd qualifies to be ℓ_k only if SS' and TT' with S', T' defined as above belong one to F_k and the other to F_{k+2}.

As a consequence, a hamiltonian path satisfying the three rules is uniquely determined if for some odd k we know ℓ_k or the edge of the cycle F_k missing from the path.

Now we are ready to construct A_i verifying all three rules above. This happens for example when one of the circles in A_i contains no point of intersection of any two other circles in its interior, as in the following pictures which constructs a unique path subject to the conditions 1 - 3:

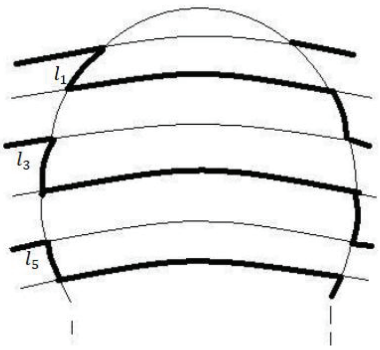

Figure 1

Then it is enough to check that when one of the circles E_i crosses an intersection point of two other circles in the configuration we can still construct a hamiltonian path verifying the properties 1 - 3. Note that at each of these crossings the structure of the edges is not altered with the exception of a small triangle, which collapses to a point at some s and then expands to a different small triangle. At each crossing the status of each large segment of circle will be the same, i.e. it will an edge of the path after the crossing if and only if it was an edge of the path before the crossing. What changes is only the status of the three edges of the small triangles.

Note that the nine segments that occur in each case belong to three consecutive cycles which are F_k, F_{k+1} and F_{k+2} or F_{k-1}, F_k and F_{k+1}, for some odd k. The hamiltonian path may contain all segments belonging to F_k and F_{k+2} that occur or all but one.

The possible cases (or their inverses) are illustrated below in Figures 2-5. (See also Remark 2.) Note that we never try to construct a hamiltonian path for those (finitely many) s for which three circles in the configuration intersect thus leaving the hypothesis of the problem, but we study what happens when we are „near“ such an s, as suggested by the term „cross“. Also, by our assumptions, there is no s such that four of the circles $E_i(s)$ intersect.

As s varies, we move from A_i to B_i and have proved the existence of a hamiltonian path, with the required properties, in the graph associated to the configuration B_i. □

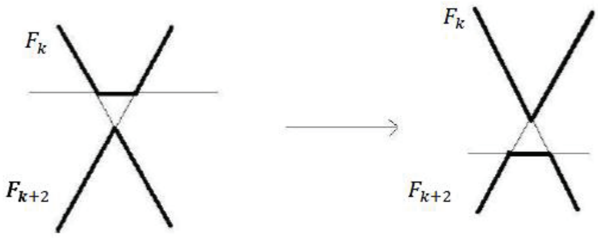

Figure 2

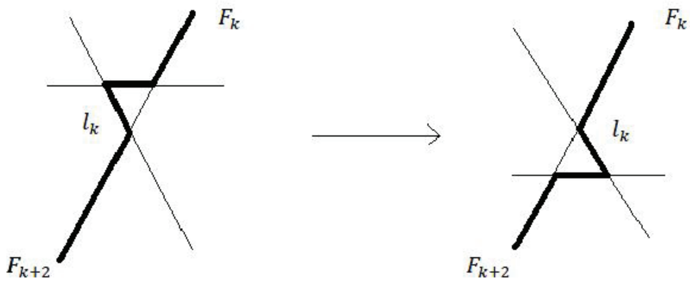

Figure 3

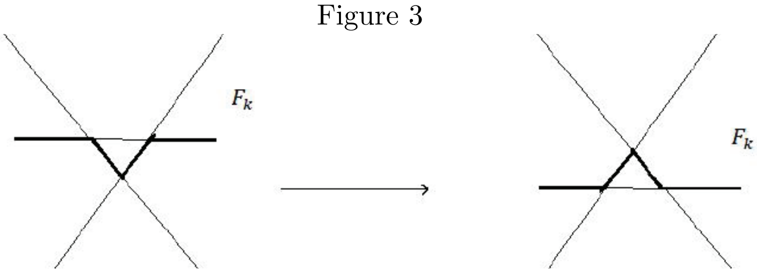

Figure 4

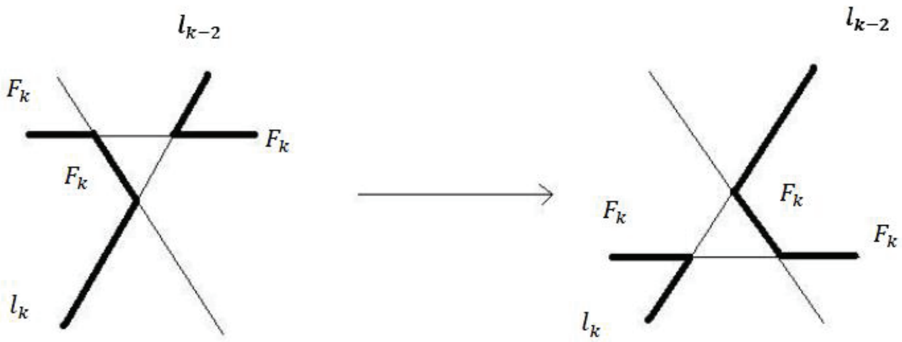

Figure 5

Editor's note. In an editorial comment [2] serious doubts has been raised about the correctness of the above result. We consider the result holds true and decided to publish it as a mathematical note.

References

[1] Problem 11557, *The American Mathematical Monthly*, **118** (2011), 179.
[2] Problem 11557, *The American Mathematical Monthly*, **120** (2013), 755–756.
[3] Problem 383, *GMA* **32(111)** (2014), no. 1–2, 52–54.

PROBLEMS

Authors should submit proposed problems to gmaproblems@rms.unibuc.ro. Files should be in PDF or DVI format. Once a problem is accepted and considered for publication, the author will be asked to submit the TeX file also. The referee process will usually take between several weeks and two months. Solutions may also be submitted to the same e-mail address. For this issue, solutions should arrive before **15th of November 2015**.

PROPOSED PROBLEMS

417. Calculate
$$\int_0^1\int_0^1 \frac{\log(1+x)-\log(1+y)}{x-y}\mathrm{d}x\mathrm{d}y.$$

Proposed by Ovidiu Furdui, Technical University of Cluj-Napoca, Romania and Cornel Vălean, Teremia Mare, Timiş, Romania.

418. (i) Let $R = k[X,Y]/(XY^2)$, k a field. Denote by x and y the residue class of X and Y modulo the ideal (XY^2), respectively. Show that the elements x and $x(1+y)$ are associates, that is, $xR = x(1+y)R$, but there is no invertible element $u \in R$ such that $ux = x(1+y)$.

(ii) Show that we can not find such elements in $\mathbb{Z}/n\mathbb{Z}$.

Proposed by Cornel Băeţica, Faculty of Mathematics and Informatics, University of Bucharest, Bucharest, Romania.

419. Suppose that $n \geq 1$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a function such that the image under f of the interior of any sphere S of codimension 1 is the interior of a sphere of codimension 1 of the same radius. Prove that f is an isometry.

Proposed by Marius Cavachi, Ovidius University of Constanţa, Constanţa, Romania.

420. Let $a, b, c, d \in \mathbb{R}$, $c, d \neq 0$, such that $\frac{a}{c} < \frac{b}{d}$. We consider the Maclaurin expansion $e^{\frac{az+b}{cz+d}} = \sum\limits_{n\geq 0} a_n z^n$.

(i) Find an exact formula as a finite sum for a_n.

(ii) Determine the asymptotic behaviour of a_n as $n \to \infty$.

Try to solve (ii) without using the result from (i).

Proposed by Constantin-Nicolae Beli, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

421. (i) Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a continuous function. Find the value of the limit
$$\lim_{n\to\infty} n^2 \iiint\limits_{x^2+y^2\leq 1,\ 0\leq z\leq 1} \left(\frac{\sqrt{x^2+y^2}+z}{2}\right)^n f(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

(ii) Let $f:\mathbb{R}^4\to\mathbb{R}$ be a continuous function. Find the value of the limit

$$\lim_{n\to\infty} n^2 \iiiint\limits_{x^2+y^2\le 1,\ z^2+t^2\le 1} \left(\frac{\sqrt{x^2+y^2}+\sqrt{z^2+t^2}}{2}\right)^n f(x,y,z,t)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z\mathrm{d}t.$$

Proposed by Dumitru Popa, Ovidius University of Constanţa, Constanţa, Romania.

422. Find a sequence $(x_n)_{n\ge 1}$ of real numbers with the following properties: $x_n \searrow 0$, $\sqrt{n}/(x_1+\cdots+x_n)\to 0$ and $x_{[\sqrt{n}]}/x_n\to 1$, where $[\sqrt{n}]$ denotes the integer part of $\sqrt{n}$.

Can you find a sequence with the above properties, in which $\sqrt{n}$ is replaced by $\ln n$?

Proposed by George Stoica, University of New Brunswick, Saint John, Canada.

423. Determine all differentiable functions $f:[0,\infty)\to\mathbb{R}$ with $f(0)=0$ such that

a) f' is strictly positive and increasing,

b) $\int_0^x (f'(t))^2\mathrm{d}t \ge f(x+f(x))-f(x)\ \forall x\in[0,\infty)$.

Florin Stănescu, Şerban Cioculescu School, Găeşti, Dâmboviţa, Romania.

424. Let $f,g\in\mathbb{C}[X]$ be monic polynomials of the same degree with the property that $|f(z)|=|g(z)|=1$ for an infinity of values of $z\in\mathbb{C}$. Prove that $f=g$.

Proposed by Marius Cavachi, Ovidius University of Constanţa, Constanţa, Romania.

425. Let $n\ge 2$ and $a_1,\dots,a_n\ge 0$ be integers and let $b_1,\dots,b_n$ and λ be positive real numbers. Find the necessary and sufficient condition for the function $f:\mathbb{R}^n\backslash\{(0,\dots,0)\}\to\mathbb{R}$ defined by $f(x_1,...,x_n)=\frac{x_1^{a_1}\cdots x_n^{a_n}}{\left(|x_1|^{b_1}+\cdots+|x_n|^{b_n}\right)^\lambda}$ to have a finite limit in $(0,\dots,0)$.

Proposed by Dumitru Popa, Ovidius University of Constanţa, Constanţa, Romania.

426. Find all functions $f:\mathbb{R}\to\mathbb{R}$, continuous in at least one point, and that satisfy the following inequalities: $f(x-1)\le f(x)-1$, $f(x+\sqrt{2})\le f(x)+\sqrt{2}$ for $x\in\mathbb{R}$.

Proposed by George Stoica, Department of Mathematical Sciences, University of New Brunswick, Canada

427. Let F be a field of characteristic $\neq 2$, and let E/F be a finite multiquadratic extension so that $G:=\mathrm{Gal}(E/F)\cong\mathbb{Z}_2^n$. Let $a\in E^\times$ with the property that $a^{s-1} \in (E^\times)^2$ $\forall s \in G$. Prove that there are $b_s \in E^\times$ with $s \in G$ such that $a^{s-1} = b_s^2$ $\forall s \in G$ and $b_{stu}b_{st}^{-1}b_{su}^{-1}b_{tu}^{-1}b_sb_tb_u = b_u^{(s-1)(t-1)}$ $\forall s,t,u \in G$.

How many $(b_s)_{s\in G} \in (E^\times)^G$ with the properties above exist?
Here we use the exponential notation: if $c \in E^\times$ and $x = \sum_{s\in G} n_s s \in \mathbb{Z}G$ then $c^x := \prod_{s\in G} s(c)^{n_s}$.

Proposed by Constantin-Nicolae Beli, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

428. Show that $f(x) = x^{4n} - 3x^{3n} + 4x^{2n} - 2x^n + 1$ is irreducible in $\mathbb{Z}[x]$ for all integers $n \geq 1$.

Proposed by Cornel Băeţica, Faculty of Mathematics and Informatics, University of Bucharest, Bucharest, Romania.

SOLUTIONS

393. Let A, B, C, D be four distinct points in a plane Π, which are not the vertices of a parallelogram. Let H be one of the half-spaces bounded by Π.

(i) In H we consider the semicircles of diameters AB and CD that are orthogonal on Π. Prove that in H there is exactly one semicircle with the diameter situated on Π that is orthogonal on the two semicircles and on Π.

We denote by $C(AB, CD)$ the semicircle from (i). Define similarly $C(AC, BD)$ and $C(AD, BC)$.

(ii) Prove that $C(AB, CD)$, $C(AC, BD)$, and $C(AD, BC)$ pass through the same point.

(iii) Prove that $C(AB, CD)$, $C(AC, BD)$, and $C(AD, BC)$ are orthogonal on each other.

Proposed by Sergiu Moroianu, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

Solution by the author. We use the Poincaré half-space model of the hyperbolic (or non-euclidean) space. In this model the space is the open euclidean half-space H. The hyperbolic planes are euclidean half-planes inside H orthogonal on Π and bounded by a line lying in Π, or half-sphere inside H bounded by a disc lying in Π. The hyperbolic lines are euclidean half-lines inside H orthogonal on Π, bounded by a point on Π, or semicircles in H orthogonal on Π, bounded by a diameter lying in Π. This representation of the hyperbolic space inside the euclidean space does not preserve the distances but it is conformal, i.e., it preserves the angles. In particular two perpendicular hyperbolic lines will be represented by two euclidean semicircles or half-lines orthogonal on each other in euclidean sense.

The points on Π do not belong to the hyperbolic space, but can be regarded as directions of hyperbolic lines. Two distinct hyperbolic half-lines are said to have the same direction if they become closer and closer (with respect to the hyperbolic distance) when we move towards infinity. This happens precisely when one of the two half-lines is a *last parallel* to the other one in the following sense: if we have a line ℓ and a point $P \notin \ell$ then in the plane containing ℓ and P there are two half-lines ℓ', ℓ'' passing through P with the property that all the lines through P which separate ℓ' from ℓ'' intersect ℓ, while the lines through P which do not separate ℓ' from ℓ'' are parallel to ℓ. Moreover, when we move to infinity in each of the two directions on such a parallel line, the distance to ℓ tends to ∞. The lines ℓ', ℓ'' are called the last parallels to ℓ passing through P. When we move to infinty on ℓ', ℓ'', the points become closer and closer to ℓ (the distance from these points and ℓ tends to 0). A line is uniquely determined by the two directions. In the Poincaré model a line represented by a semicircle of diameter $AB \subset \Pi$ has A and B as directions. A line represented by a vertical half-line bounded by a point $A \in \Pi$ has A as one of the directions and ∞ (the point at infinty) as the other direction. So by $A\infty$ we mean the vertical half-line perpendicular on Π at A.

Two distinct lines ℓ, ℓ' in hyperbolic space can be in precisely one of the following three positions:

(I) $\ell \cap \ell' = \{P\}$ for some point P. In this case the only line orthogonal on both ℓ and ℓ' is the line orthogonal on the plane supporting ℓ and ℓ' passing through P.

(II) $\ell \cap \ell' = \emptyset$ and ℓ, ℓ' are not last parallel to each other (see above), i.e., they don't meet at infinity. In this case there are exactly two points $M \in \ell$, $N \in \ell'$ such that the distance $|MN|$ is minimum (it follows then that MN is orthogonal on both ℓ, ℓ'). Conversely, if $M \in \ell$, $N \in \ell'$ such that $MN \perp \ell, \ell'$ then $|MN|$ is the smallest distance between a point of ℓ and a point of ℓ'. Hence again there is a unique line orthogonal on both ℓ and ℓ'.

(III) ℓ and ℓ' meet at infinity. Then there is no line orthogonal on both ℓ and ℓ'. In our case the semicircles of diameters AB and CD orthogonal on Π are the lines AB and CD from the hyperbolic space and, since $\{A, B\} \cap \{C, D\} = \emptyset$, AB and CD are not last parallel to each other, so they are in one of the cases (I) and (II) above. Therefore there is a unique hyperbolic line $C(AB, CD)$ orthogonal on both AB and CD. If $C(AB, CD)$ is a vertical half-line $O\infty$ then the fact that $O\infty$ is orthogonal on the semicircles of diameters AB and CD means that O is the midpoint of both segments AB and CD, so A, C, B, D are the vertices of a parallelogram, which contradicts the hypothesis. So $C(AB, CD)$ is a semicircle with the diameter on Π orthogonal on Π, as required. Similarly we have the semicircles $C(AC, BD)$ and $C(AD, BC)$.

We denote by M, N, P, Q, R, S the points where the lines $C(AB, CD)$, $C(AC, BD)$ and $C(AD, BC)$ intersect orthogonally AB, CD, AC, BD, AD and BC, respectively. We denote by O, O', O'' the midpoints of MN, PQ

and RS, respectively. Since $M, N \in C(AB, CD)$, we have $O \in C(AB, CD)$ and, similarly, $O' \in C(AC, BD)$ and $O'' \in C(AD, BC)$.

We consider the hyperbolic symmetry σ with respect to the line $C(AB, CD)$. Since the lines AB and CD touch $C(AB, CD)$ orthogonally in M and N the symmetry σ acts on AB and CD as the symmetry with respect to M and N, respectively. Hence σ preserves AB and CD but reverses the directions. Hence σ permutes A, B and C, D. It follows that $\sigma(AC) = BD$ and $\sigma(BD) = AC$. Since $C(AC, BD)$ is the only line orthogonal on AC and BD and it touches AC and BD in P and Q, σ will preserve $C(AC, BD)$, but permutes P and Q. Since $\sigma(P) = Q$ we have that the midpoint O' of PQ belongs to $C(AB, CD)$, so $O' \in C(AB, CD) \cap C(AC, BD)$. Also if $P \neq Q$ then the line $PQ = C(AC, BD)$ is orthogonal on $C(AB, CD)$.

By a similar reasoning, $O \in C(AB, CD) \cap C(AC, BD)$ and if $M \neq N$ then $C(AB, CD) \perp C(AC, BD)$. Hence $O, O' \in C(AB, CD) \cap C(AC, BD)$ and if $M \neq N$ or $P \neq Q$ then $C(AB, CD) \perp C(AC, BD)$. So if $M \neq N$ or $P \neq Q$ then $C(AB, CD) \perp C(AC, BD)$ and $C(AB, CD) \cap C(AC, BD) = \{O\} = \{O'\}$. The remaining case $M = N = O$ and $P = Q = O'$ does not occur. Indeed, if $O \neq O'$ then from $O, O' \in C(AB, CD) \cap C(AC, BD)$ we get $C(AB, CD) = C(AC, BD) =: \ell$. But this implies that ℓ is orthogonal on both AB and AC, which is impossible since AB and AC are in the case (III) above (they meet at infinity at A), so there is no common orthogonal line. If $O = O'$ then AB and AC have a common point $M = P$ and a common direction A, so $AB = AC$ and, so $B = C$. Contradiction.

By the same argument, $O'' = O = O'$, $C(AD, BC) \cap C(AB, CD) = C(AD, BC) \cap C(AC, BD) = \{O''\}$ and also $C(AD, BC) \perp C(AB, CD)$, $C(AC, BD)$. This means that $C(AB, CD)$, $C(AC, BD)$ and $C(AD, BC)$ are orthogonal on each other and they pass through the same point $O = O' = O''$.

Notes. (1) We do not need the condition that A, B, C, D are not the vertices of a parallelogram if we allow $C(AB, CD)$, $C(AC, BD)$ and $C(AD, BC)$ to be vertical half-lines, not only semicircles.

(2) We may obtain a similar result if we choose $D = \infty$. Then the semicircles AD, BD, CD become the vertical lines $A\infty$, $B\infty$ and $C\infty$.

(3) The point O is clearly the hyperbolic midpoint of the segments MN, PQ and RS, but this is not easy to state in terms of euclidean geometry.

394. Find all polynomials $P \in \mathbb{Z}[X]$ such that $a^2+b^2+c^2 \mid f(a)+f(b)+f(c)$ for any $a, b, c \in \mathbb{Z}$

Proposed by Vlad Matei, student, University of Wisconsin, Madison, USA.

Solution by V. Makanin, Sankt Petersburg, Russia. The answer is $f = mX^2$, for $m \in \mathbb{Z}$.

First we have (for $a = b = c = 0$) that 0 divides $3f(0)$, thus $f(0) = 0$. Then $a^2 + b^2 + c^2 \mid f(a) + f(b) + f(c)$ and $a^2 + b^2 + c^2 \mid f(-a) + f(b) + f(c)$ imply $a^2 + b^2 + c^2 \mid f(a) - f(-a)$ for all integer a, b, c, and from here we easily infer that $f(a) = f(-a)$ for all a. As f is an even polynomial, there must exist g (with integer coefficients and with $g(0) = 0$, of course) such that $f(X) = g(X^2)$. The hypothesis becomes $a^2 + b^2 + c^2 \mid g(a^2) + g(b^2) + g(c^2)$ for all $a, b, c \in \mathbb{Z}$.

Now we have that $(a^2 + b^2)^2 + c^2 = (a^2 - b^2)^2 + (2ab)^2 + c^2$ divides $g\big((a^2-b^2)^2\big)+g\big((2ab)^2\big)+g(c^2)$ and $(a^2+b^2)^2+c^2 = (a^2+b^2)^2+c^2+0^2$ divides $g\big((a^2 + b^2)^2\big) + g(c^2) + g(0) = g\big((a^2 + b^2)^2\big) + g(c^2)$, therefore $(a^2 + b^2)^2 + c^2$ divides $g\big((a^2+b^2)^2\big) - g\big((a^2-b^2)^2\big) - g\big((2ab)^2\big)$ for all $a, b, c \in \mathbb{Z}$. Obviously, this implies that for all $a, b \in \mathbb{Z}$ we have

$$g\big((a^2 - b^2)^2 + (2ab)^2\big) = g\big((a^2 + b^2)^2\big) = g\big((a^2 - b^2)^2\big) + g\big((2ab)^2\big).$$

Further let us consider some N of the form $N = 4p_1^2 \cdots p_k^2$, with $p_1, \dots, p_k$ distinct primes. There are 2^k distinct pairs (a, b) of positive integers such that $N = (2ab)^2$. If (a, b) and (c, d) are two such pairs, we have $ab = cd$. If we also have $(a^2 - b^2)^2 = (c^2 - d^2)^2$, we see immediately that either $a = c$ and $b = d$, or $a = d$ and $b = c$. Consequently there still remain 2^{k-1} distinct values of $(a^2 - b^2)^2$ when (a, b) runs over all solutions of $(2ab)^2 = N$. Chose k such that 2^{k-1} is greater than the degree of g; then the equality $g(x + N) = g(x) + g(N)$ is assured for at least $\deg(g) + 1$ values of x (x is of the form $(a^2 - b^2)^2$, with $(2ab)^2 = N$), meaning that it is true for all x. But in that case, by differentiation we get $g'(x + N) = g'(x)$ for all x, yielding that g' (the derivative of g) is constant, hence g is of the form mX, for some integer m (actually $g = mX + n$, but we already know $g(0) = 0$). This gives for f the form $f = mX^2$, and it is easy to check that these polynomials are indeed solutions of the problem. □

395. Let $z_1, z_2, \dots, z_n \geq 1$. Prove the following inequality:

$$\sum_{i=1}^{n} \frac{1}{1 + z_i} + \frac{n(n-2)}{1 + \prod_{i=1}^{n} z_i^{1/n}} \geq (n-1) \sum_{i=1}^{n} \frac{1}{1 + \prod_{j \neq i} z_j^{1/(n-1)}}.$$

Proposed by Cezar Lupu, University of Pittsburgh, Pittsburgh, PA and Ştefan Spătaru, International Computer High School of Bucharest, Bucharest, Romania.

Solution by V. Makanin, Sankt Petersburg, Rusia. Let f be a convex function on the interval I; then for every $x_1, x_2, \dots, x_n \in I$ the inequality

$$\sum_{i=1}^{n} f(x_i) + n(n-2) f\left(\frac{1}{n} \sum_{i=1}^{n} x_i\right) \geq (n-1) \sum_{i=1}^{n} f\left(\frac{1}{n-1} \sum_{1 \leq j \leq n,\, j \neq i} x_j\right)$$

holds. This generalization of Popoviciu's inequality was proved by Vasile Cîrtoaje and can be found, with proof (based on Karamata's inequality), in his book *Algebraic Inequalities*, GIL Publishing House, 2006, pp. 193–195. Now consider the function $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = 1/(1 + e^x)$ and note that f is convex on $[0, \infty)$, having the second derivative

$$f''(x) = \frac{e^x(e^x - 1)}{(1 + e^x)^3} \geq 0 \quad \text{for all } x \geq 0.$$

Then apply the above generalization of Popoviciu's inequality to f and to the numbers $x_i = \log z_i \geq 0$ in order to obtain the desired inequality. □

396. Let F be a field and let V be an F-vector space. We denote, as usual, by $T(V)$, $S(V)$ and $\Lambda(V)$ the tensor, symmetric and exterior algebras over V.

Let $I_{S'}$ be the subgroup of $T(V)$ generated by $x_1 \otimes \cdots \otimes x_n - x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}$ with $x_1, \ldots, x_n \in V$ and $\sigma \in A_n$. Then $I_{S'}$ is a homogeneous ideal in $T(V)$ and we denote by $S'(V) = T(V)/I_{S'}$. Then $S'(V)$ is a graded algebra, $S'(V) = \bigoplus_{n \geq 0} S'^n(V)$. We denote by $\odot$ the product on $S'(V)$. Hence if $x_1, \ldots, x_n \in V$ then the image of $x_1 \otimes \cdots \otimes x_n \in T(V)$ in $S'(V) = T(V)/I_{S'}$ is $x_1 \odot \cdots \odot x_n$.

(i) For $n \geq 1$ let $\rho_{S'^n, S^n} : S'^n(V) \to S^n(V)$ be the linear map given by $x_1 \odot \cdots \odot x_n \mapsto x_1 \cdots x_n$. For any integer n greater than or equal to 2 find a linear map $\rho_{\Lambda^n, S'^n} : \Lambda^n(V) \to S'^n(V)$ such that the short sequence

$$0 \to \Lambda^n(V) \xrightarrow{\rho_{\Lambda^n, S'^n}} S'^n(V) \xrightarrow{\rho_{S'^n, S^n}} S^n(V) \to 0$$

is exact.

(ii) If $F = \mathbb{F}_2$ prove that for any positive integer n there is a linear map $\rho_{S^n, \Lambda^n} : S^n(V) \to \Lambda^n(V)$ with $x_1 \cdots x_n \mapsto x_1 \wedge \cdots \wedge x_n$. If $n = 2, 3$ find a linear map $\rho_{T^{n-1}, S^n} : T^{n-1}(V) \to S^n(V)$ such that the short sequence

$$0 \to T^{n-1}(V) \xrightarrow{\rho_{T^{n-1}, S^n}} S^n(V) \xrightarrow{\rho_{S^n, \Lambda^n}} \Lambda^n(V) \to 0$$

is exact.

Proposed by Constantin-Nicolae Beli, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

Solution by the author. By definition, we have $S(V) = T(V)/I_S$ and $\Lambda(V) = T(V)/I_\Lambda$, where I_S and I_Λ are the bilateral ideals of $T(V)$ generated by $\{x \otimes y - y \otimes x \mid x, y \in V\}$ and $\{x \otimes x \mid x \in V\}$, respectively. It turns out that $I_{S'}$ is also a bilateral ideal. To prove this we must show that $\beta \otimes \alpha \otimes \gamma \in I_{S'}$ whenever α is a generator of $I_{S'}$ and β, γ are generators of $T(V)$. We may take $\alpha = x_1 \otimes \cdots \otimes x_n - x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}$, $\beta = y_1 \otimes \cdots \otimes y_m$ and $\gamma = z_1 \otimes \cdots \otimes z_k$, where $x_i, y_i, z_i \in V$ and $\sigma \in A_n$. Since $\sigma \in A_n$ the mapping $y_i \mapsto y_i$, $x_i \mapsto x_{\sigma(i)}$, $z_i \mapsto z_i$ is an even permutation of $y_1, \ldots, y_m, x_1, \ldots, x_n, z_1, \ldots, z_k$.

Therefore $\beta\otimes\alpha\otimes\gamma = y_1\otimes\cdots\otimes y_m\otimes x_1\otimes\cdots\otimes x_n\otimes z_1\otimes\cdots\otimes z_k - y_1\otimes\cdots\otimes y_m\otimes x_{\sigma(1)}\otimes\cdots\otimes x_{\sigma(n)}\otimes z_1\otimes\cdots\otimes z_k$.

We denote by $I_S^n, I_\Lambda^n, I_{S'}^n$ the homogeneous components of degree n of $I_S, I_\Lambda, I_{S'}$.

Note that if $x_i = x_j$ for some $i \neq j$ then $x_1\otimes\cdots\otimes x_n - x_{\sigma(1)}\otimes\cdots\otimes x_{\sigma(n)}$ belongs to $I_{S'}^n$ for any $\sigma\in S_n$, not merely for $\sigma\in A_n$. Indeed, if $\sigma\in S_n\setminus A_n$ then let τ be the transposition (i,j). Then for any k we have $x_{\tau(k)} = x_k$ ($x_{\tau(i)} = x_j = x_i$, $x_{\tau(j)} = x_i = x_j$ and $\tau(k) = k$ for $k\neq i,j$). It follows that $x_{\tau\sigma(k)} = x_{\sigma(k)}$ $\forall k$. But $\tau,\sigma\in S_n\setminus A_n$, so $\tau\sigma\in A_n$. It follows that $x_1\otimes\cdots\otimes x_n - x_{\sigma(1)}\otimes\cdots\otimes x_{\sigma(n)} = x_1\otimes\cdots\otimes x_n - x_{\tau\sigma(1)}\otimes\cdots\otimes x_{\tau\sigma(n)}\in I_{S'}^n$.

Let $(e_i)_{i\in I}$ be a basis for V. On I we define a total order relation $\leq$. Let $A^n = \{(i_1,\dots,i_n)\in I^n : i_1\leq\ldots\leq i_n\}$, $B^n = \{(i_1,\dots,i_n)\in I^n : i_1<\ldots<i_n\}$ and $C^n = A^n\setminus B^n$. Let $f: B^n\to\Lambda^n(V)$ be given by $f(i_1,\dots,i_n) = e_{i_1}\wedge\cdots\wedge e_{i_n}$ and $h: A^n\to S^n(V)$ be given by $h(i_1,\dots,i_n) = e_{i_1}\cdots e_{i_n}$. Then $(f(b))_{b\in B^n}$ is a basis for $\Lambda^n(V)$ and $(h(a))_{a\in A^n}$ is a basis for $S^n(V)$.

We now assume that $n\geq 2$ and we obtain a basis for $S'^n(V)$. For any $\alpha\in T^n(V)$ we denote by $[\alpha]$ its class in $S'^n(V)$, $[\alpha] = \alpha + I_{S'}^n$. Since $I_{S'}^n$ is generated, as a group, by $x_1\otimes\cdots\otimes x_n - x_{\sigma(1)}\otimes\cdots\otimes x_{\sigma(n)}$, with $x_i\in V$ and $\sigma\in A_n$, it will be generated, as a vector space, by $e_{i_1}\otimes\cdots\otimes e_{i_n} - e_{i_{\sigma(1)}}\otimes\cdots\otimes e_{i_{\sigma(n)}}$, with $(i_1,\dots,i_n)\in I^n$ and $\sigma\in A_n$. On the basis $X = \{e_{i_1}\otimes\cdots\otimes e_{i_n} : (i_1,\dots,i_n)\in I^n\}$ of $T^n(V)$ we define the equivalence relation $\sim$, $e_{i_1}\otimes\cdots\otimes e_{i_n}\sim e_{j_1}\otimes\cdots\otimes e_{j_n}$ if $(j_1,\dots,j_n) = (\sigma(i_1),\dots,\sigma(i_n))$ for some $\sigma\in A_n$. Then $I_{S'}^n$ is generated by elements of the form $\alpha-\beta$, where $\alpha,\beta\in X$, with $\alpha\sim\beta$. It follows that a basis of $S'^n(V) = T^n(V)/I_{S'}^n$ is $\{[\alpha]\mid\alpha\in Y\}$, where Y is a system of representatives for $X/\sim$.

Let $\alpha\in X$, $\alpha = e_{j_1}\otimes\cdots\otimes e_{j_n}$, with $(j_1,\dots,j_n)\in I^n$. Then by arranging $j_1,\dots,j_n$ in increasing order we obtain a sequence $i_1\leq\ldots\leq i_n$, i.e., $(i_1,\dots,i_n)\in A^n$, and there is $\sigma\in S_n$ with $i_s = j_{\sigma(s)}$ $\forall s$. We consider two cases.

If $j_1,\dots,j_n$ are mutually distinct then $i_1<\ldots<i_n$, i.e., $(i_1,\dots,i_n)\in B^n$, and the permutation σ satisfying $(i_1,\dots,i_n) = (j_{\sigma(1)},\dots,(j_{\sigma(n)}))$ is unique. If $\sigma\in A_n$ then $e_{j_1}\otimes\cdots\otimes e_{j_n}\sim e_{i_1}\otimes\cdots\otimes e_{i_n}$. If $\sigma\in S_n\setminus A_n$ then let $\tau = (1,2)$ and we have $\tau\sigma\in A_n$. Since $(i_2,i_1,i_3,\dots,i_n) = (i_{\tau(1)},\dots,i_{\tau(n)}) = (j_{\sigma\tau(1)},\dots,j_{\sigma\tau(n)})$, we have $e_{j_1}\otimes\cdots\otimes e_{j_n}\sim e_{i_2}\otimes e_{i_1}\otimes e_{i_3}\otimes\cdots\otimes e_{i_n}$.

If $j_1,\dots,j_n$ are not mutually distinct then there is some s with $i_s = i_{s+1}$, so $(i_1,\dots,i_n)\in A^n\setminus B^n = C^n$. Let $\tau = (s,s+1)$. Since $i_s = i_{s+1}$, we have $(i_1,\dots,i_n) = (i_{\tau(1)},\dots,i_{\tau(n)}) = (j_{\sigma\tau(1)},\dots,j_{\sigma\tau(n)})$. Therefore in both cases $\sigma\in A_n$ and $\sigma\tau\in A_n$ we have $e_{j_1}\otimes\cdots\otimes e_{j_n}\sim e_{i_1}\otimes\cdots\otimes e_{i_n}$.

In conclusion, a set of representatives for $X_{/\sim}$ is $Y = \{e_{i_1}\otimes\cdots\otimes e_{i_n}, e_{i_2}\otimes e_{i_1}\otimes e_{i_3}\otimes\cdots\otimes e_{i_n} : (i_1,\dots,i_n)\in B^n\}\cup\{e_{i_1}\otimes\cdots\otimes e_{i_n} : (i_1,\dots,i_n)\in C^n\}$. We define $g_1,g_2 : B^n\to S'^n(V)$ and $g: C^n\to S'^n(V)$ by $g_1(i_1,\dots,i_n) = e_{i_1}\odot\cdots\odot e_{i_n}$, $g_2(i_1,\dots,i_n) = e_{i_2}\odot e_{i_1}\odot e_{i_3}\odot\cdots\odot e_{i_n}$, and $g(i_1,\dots,i_n) = e_{i_1}\odot\cdots\odot e_{i_n}$. Since the projection $T^n(V)\to T^n(V)/I_{S'}^n = S'^n(V)$ is given by $[x_1 \otimes \cdots \otimes x_n] = x_1 \odot \cdots \odot x_n$, the basis $\{[\alpha] : \alpha \in Y\}$ of $S'^n(V)$ can be written as $\{g_1(b), g_2(b) : b \in B^n\} \cup \{g(c) : c \in C^n\}$.

Note that one has $I^n_{S'} \subseteq I^n_S$ because both are generated by expressions of the type $x_1 \otimes \cdots \otimes x_n - x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}$, with $\sigma \in A_n$ in the case of $I^n_{S'}$ and with $\sigma \in S_n$ in the case of I^n_s. Thus the map $\rho_{S'^n,S^n} : S'^n(V) = T^n(V)/I^n_{S'} \to S^n(V) = T^n(V)/I^n_S$ given by $x_1 \odot \cdots \odot x_n \mapsto x_1 \cdots x_n$ is well defined and surjective. We write $\rho_{S'^n,S^n}$ in the bases $\{g_1(b), g_2(b) : b \in B^n\} \cup \{g(c) : c \in C^n\}$ and $\{h(a) : a \in A^n\}$ for $S'^n(V)$ and $S^n(V)$, respectively. If $b = (i_1, \dots, i_n) \in B^n$ then $\rho_{S'^n,S^n}(g_1(b)) = \rho_{S'^n,S^n}(g_2(b)) = e_{i_1} \cdots e_{i_n} = h(b)$ and if $c = (i_1, \dots, i_n) \in C^n$ then $\rho_{S'^n,S^n}(g(c)) = e_{i_1} \cdots e_{i_n} = h(c)$. A typical element $\alpha \in S'^n(V)$ has the form $\alpha = \sum_{b\in B^n} (\alpha_b g_1(b) + \beta_b g_2(b)) + \sum_{c\in C} \gamma_c g(c)$, where $\alpha_b, \beta_b, \gamma_c \in F$ are almost all zero. We have $\rho_{S'^n,S^n}(\alpha) = \sum_{b\in B^n} (\alpha_b + \beta_b) h(b) + \sum_{c\in C^n} \gamma_c h(c)$. It follows that $\alpha \in \ker \rho_{S'^n,S^n}$ iff $\alpha_b + \beta_b = 0\ \forall b \in B^n$ and $\gamma_c = 0\ \forall c \in C^n$. Hence

$$\ker \rho_{S'^n,S^n} = \{ \sum_{b\in B^n} (\alpha_b g_1(b) - \alpha_b g_2(b)) : \alpha_b \in F \text{ almost all zero} \}.$$

We now define the map ρ_{Λ^n,S'^n}. Note that the map $\phi : V^n \to S'^n(V)$ given by $(x_1, \dots, x_n) \mapsto x_1 \odot \cdots \odot x_n - x_2 \odot x_1 \odot x_3 \odot \cdots \odot x_n$ is multilinear and alternate. (If $x_i = x_j$ then $x_2 \odot x_1 \odot x_3 \odot \cdots \odot x_n$ writes as $x_{\sigma(1)} \odot \cdots \odot x_{\sigma(n)}$, where $\sigma = (1, 2)$, and it is equal to $x_1 \odot \cdots \odot x_n$, even though $\sigma \in S_n \setminus A_n$ by a remark we made above.) Hence it induces a map $\rho_{\Lambda^n,S'^n} : \Lambda^n(V) \to S'^n(V)$ given by $x_1 \wedge \ldots \wedge x_n \mapsto x_1 \odot \cdots \odot x_n - x_2 \odot x_1 \odot x_3 \odot \cdots \odot x_n$.

We write ρ_{Λ^n,S'^n} in the bases $\{f(b) : b \in B^n\}$ and $\{g_1(b), g_2(b) : b \in B^n\} \cup \{g(c) : c \in C^n\}$ for $\Lambda^n(V)$ and $S'^n(V)$, respectively. If $b = (i_1, \dots, i_n) \in B^n$ we have $\rho_{\Lambda^n,S'^n}(f(b)) = e_{1_1} \odot \cdots \odot e_{i_n} - e_{i_2} \odot e_{1_1} \odot e_{i_3} \odot \cdots \odot e_{i_n} = g_1(b) - g_2(b)$. A typical element of $\Lambda^n(V)$ has the form $\alpha = \sum_{b\in B^n} \alpha_b f(b)$, where $\alpha_b \in F$ are almosts all zero, and we have $\rho_{\Lambda^n,S'^n}(\alpha) = \sum_{b\in B^n} \alpha_b (g_1(b) - g_2(b)) = \sum_{b\in B^n} (\alpha_b g_1(b) - \alpha_b g_2(b))$. Then $\alpha \in \ker \rho_{\Lambda^n,S'^n}$ iff $\alpha_b = 0\ \forall b \in B^n$, i.e., iff $\alpha = 0$. Thus ρ_{Λ^n,S'^n} is injective. We also have $\operatorname{Im} \rho_{\Lambda^n,S'^n} = \{ \sum_{b\in B^n} (\alpha_b g_1(b) - \alpha_b g_2(b)) : \alpha_b \in F \text{ almost all zero}\} = \ker \rho_{S'^n,S^n}$. Therefore the sequence from (i) is exact.

For (ii) we note that the bilateral ideal I_Λ of $T(V)$ contains all expressions of the form $x \otimes y + y \otimes x$ with $x, y \in V$. But since we are in characteristic 2 we have $x \otimes y + y \otimes x = x \otimes y - y \otimes x$, so I_Λ contains the bilateral ideal generated by these expressions, which is I_S. Hence we have a surjective linear map $\rho_{S,\Lambda} : S(V) = T(V)/I_S \to \Lambda(V) = T(V)/I_\Lambda$ given by $x_1 \cdots x_n \mapsto x_1 \wedge \cdots \wedge x_n$. The maps ρ_{S^n,Λ^n} are just the homogeneous components of $\rho_{S,\Lambda}$.

We now assume that $n \geq 2$ and write ρ_{S^n,Λ^n} in terms of the bases $\{h(a) : a \in A^n\}$ and $\{f(b) : b \in B^n\}$. If $a = (i_1, \ldots, i_n) \in A^n = B^n \cup C^n$ then $\rho_{S^n,\Lambda^n}(h(a)) = e_{i_1} \wedge \cdots \wedge e_{i_n}$. If $a \in B^n$ then $\rho_{S^n,\Lambda^n}(h(a)) = f(a)$. If $a \in C^n$ then $i_s = i_{s+1}$ for some s, so $\rho_{S^n,\Lambda^n}(h(a)) = 0$. A typical $\alpha \in S^n(V)$ has the form $\alpha = \sum_{a\in A} \alpha_a h(a)$, where $\alpha_a \in F$ are almost all zero, and we have $\rho_{S^n,\Lambda^n}(\alpha) = \sum_{a\in B^n} \alpha_a f(a)$. Thus $\alpha \in \ker \rho_{S^n,\Lambda^n}$ iff $\alpha_a = 0\ \forall a \in B^n$. It follows that $\ker \rho_{S^n,\Lambda^n} = \{\sum_{c\in C^n} \alpha_c h(c) : \alpha_c \in F$ almost all zero$\}$. Thus $h(c)$, $c \in C^n$, are a basis for $\ker \rho_{S^n,\Lambda^n}$.

Note that all the reasoning above apply for F arbitrary of characteristic 2, not merely $F = \mathbb{F}_2$. From now on we assume that $F = \mathbb{F}_2$.

Take first $n = 2$. Then $C^2 = \{(i,i) : i \in I\}$, so a basis for $\ker \rho_{S^2,\Lambda^2}$ is made of $h(i,i) = e_i^2$ with $i \in I$. We define $\rho_{T^1,S^2} : T^1(V) = V \to S^2(V)$ by $x \mapsto x^2$. We have $\rho_{T^1,S^2}(x+y) = (x+y)^2 = x^2 + 2xy + y^2 = x^2 + y^2 = \rho_{T^1,S^2}(x) + \rho_{T^1,S^2}(y)$ and if $\lambda \in F$ then $\rho_{T^1,S^2}(\lambda x) = (\lambda x)^2 = \lambda x^2 = \lambda \rho_{T^1,S^2}(x)$. (In $F = \mathbb{F}_2$ we have $\lambda^2 = \lambda$.) Thus ρ_{T^1,S^2} is linear. We have $\rho_{T^1,S^2}(e_i) = e_i^2$, so the basis $\{e_i : i \in I\}$ of V is sent bijectively by ρ_{T^1,S^2} to the basis $\{e_i^2 : i \in I\}$ of $\ker \rho_{S^2,\Lambda^2}$. Thus ρ_{T^1,S^2} is a bijection between V and $\ker \rho_{S^2,\Lambda^2}$. Since also ρ_{S^2,Λ^2} is surjective, the sequence

$$0 \to V \xrightarrow{\rho_{T^1,S^2}} S^2(V) \xrightarrow{\rho_{S^2,\Lambda^2}} \Lambda^2(V) \to 0$$

is exact.

If $n = 3$ then $\rho_{T^2,S^3} : T^2(V) = V \otimes V \to S^3(V)$ will be defined as the composition $V \otimes V \xrightarrow{\rho_{T^1,S^2} \otimes 1_V} S^2(V) \otimes V \xrightarrow{m} S^3(V)$, where $m : S^2(V) \otimes V = S^2(V) \otimes S^1(V) \to S^3(V)$ is just the multiplication from $S(V)$. More precisely, ρ_{T^2,S^3} is given by $x \otimes y \mapsto x^2 y$. (We have $m(\rho_{T^1,S^2} \otimes 1_V(x \otimes y)) = m(x^2 \otimes y) = x^2 y$.) An element from the canonical basis $e_{j_1} \otimes e_{j_2}$, with $j_1, j_2 \in I$ is sent by ρ_{T^2,S^3} to $e_{j_1}^2 e_{j_2}$, which can be written as $h(c)$ for some $c \in C^n$. Namely, $c = (j_1, j_1, j_2)$ if $j_1 \leq j_2$ and $c = (j_2, j_1, j_1)$ if $j_1 > j_2$. Conversely, if $c = (i_1, i_2, i_3) \in C^n$ then $i_1 \leq i_2 \leq i_3$ and $i_1 = i_2$ or $i_2 = i_3$. Then $h(c) = e_{i_1} e_{i_2} e_{i_3}$, which can be written as $e_{j_1}^2 e_{j_2} = \rho_{T^2,S^3}(e_{j_1} \otimes e_{j_2})$ for some unique $j_1, j_2 \in I$. Namely $(j_1, j_2) = (i_1, i_3)$ if $i_1 = i_2$ and $(j_1, j_2) = (i_2, i_1)$ if $i_1 < i_2 = i_3$. Hence the basis $\{e_{j_1} \otimes e_{j_2} : j_1, j_2 \in I\}$ of $V \otimes V$ is sent bijectively by ρ_{T^2,S^3} to the basis $\{h(c) : c \in C^3\}$ of $\ker \rho_{S^3,\Lambda^3}$. Thus ρ_{T^2,S^3} is a bijection between $V \otimes V$ and $\ker \rho_{S^3,\Lambda^3}$. Since ρ_{S^3,Λ^3} is also surjective, the sequence $0 \to V \otimes V \xrightarrow{\rho_{T^2,S^3}} S^3(V) \xrightarrow{\rho_{S^3,\Lambda^3}} \Lambda^3(V) \to 0$ is exact. $\square$

397. Let $n \geq 1$ be an integer and let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a function with the property that the image under f of any sphere S of codimension 1 is a shere of codimension 1 of the same radius. Prove that f is an isometry.

Proposed by Marius Cavachi, Ovidius University of Constanţa, Constanţa, Romania.

Solution by the author. If $a, b \in \mathbb{R}^n$, $a \neq b$, then a, b are antipodal points on a sphere S of radius $|b-a|/2$. Then $f(a), f(b)$ belong to $f(S)$, which is again a sphere of radius $|b-a|/2$. Hence $|f(b)-f(a)|$ is less than or equal to the diameter $|b-a|$ of $f(S)$. So the inequality $|f(b)-f(a)| \leq |b-a|$ holds for all $a, b \in \mathbb{R}^n$.

If $d \in f(S)$ is the antipodal point of $f(a)$ then $|d-f(a)| = |b-a|$. Let $c \in S$ such that $f(c) = d$. We have $|b-a| = |d-f(a)| = |f(c)-f(a)| \leq |c-a|$. But a, c belong to the sphere S of radius $|b-a|/2$, so $|c-a| \geq |b-a|$ implies that $|c-a| = |b-a|$ and c is the antipodal point of a on S, i.e., $c = b$. It follows that $d = f(c) = f(b)$, so $|f(b)-f(a)| = |d-f(a)| = |b-a|$. Hence f is an isometry. □

398. Let $A \in \mathcal{M}_n(\mathbb{Q})$ be an invertible matrix.

a) Prove that if for every $k \in \mathbb{N}^*$ there exists $B_k \in \mathcal{M}_n(\mathbb{Q})$ such that $B_k^k = A$, then all the eigenvalues of A are equal to 1.

b) Is the converse of a) true?

Proposed by Victor Alexandru, Cornel Băeţica, Gabriel Mincu, University of Bucharest, Bucharest, Romania

Solution by the authors. a) Let $P_A = \det(XI_n - A)$ be the characteristic polynomial of A. We will denote by $p_1, p_2, \ldots, p_r \in \mathbb{N}$ the primes that divide at least one of the denominators of the coefficients of P_A and put

$$R = \left\{ \frac{a}{s} \in \mathbb{Q} : a \in \mathbb{Z},\ s = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r},\ \alpha_1, \ldots, \alpha_r \in \mathbb{N} \right\}.$$

It is easy to see that R is a subring of $\mathbb{Q}$ and $P_A \in R[X]$.

Let $k \geq 1$ be fixed. We denote by $\alpha_1, \ldots, \alpha_n$ the roots of P_A and by $\beta_1, \ldots, \beta_n$ the roots of P_{B_k}. Then if we order well these roots we have $\alpha_i = \beta_i^k$. (We have $B_k^k = A$.) We denote $P_A = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ and $P_{B_k} = X^n + b_{n-1}X^{n-1} + \cdots + b_0$. Note that $P_{B_k}(X) = (X-\beta_1)\cdots(X-\beta_n)$ divides $P_A(X^k) = (X^n - \alpha_1)\cdots(X^k - \alpha_n) = (X^k - \beta_1^k)\cdots(X^k - \beta_n^k)$. It follows that for any $N \geq 1$ we have $N^n P_{B_k}(X/N) \mid N^{kn} P_A(X^k/N^k)$, i.e., $X^n + Nb_{n-1}X^{n-1} + \cdots + N^n b_0$ divides $X^{kn} + N^k a_{n-1} X^{k(n-1)} + \cdots + N^{kn} a_0$. Since $a_i \in R$, if we take $N = (p_1 \cdots p_r)^s$ for some large enough s we have $N^{ik} a_{n-i} \in \mathbb{Z}$ for $1 \leq i \leq n$, so $N^{kn} P_A(X^k/N^k) \in \mathbb{Z}[X]$. By Gauss's Lemma this implies $N^n P_{B_k}(X) \in \mathbb{Z}[X]$, whence $N^i b_{n-i} \in \mathbb{Z}$ for $1 \leq i \leq n$. But this implies that $b_{n-i} \in R$ for $1 \leq i \leq n$, so $P_{B_k} \in R[X]$.

We now consider a prime $q \in \mathbb{N} \setminus \{p_1, \ldots, p_r\}$. The function $\phi : \mathbb{Z} \to R/qR$ given by $a \mapsto a + qR$ is a ring homomorphism. We have $a \in \ker\phi$ iff $a \in qR$, i.e., iff $a = q\frac{b}{N}$, where N has the form $N = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, so $(q, N) = 1$. Then $q \mid qb = aN$ implies $q \mid a$, so $a \in q\mathbb{Z}$. Conversely, if $a \in q\mathbb{Z}$ then $a \in qR$

(we have $\mathbb{Z} \subseteq R$), so $\ker \phi = q\mathbb{Z}$. Thus ϕ induces an injective morphism $\overline{\phi} : \mathbb{Z}/q\mathbb{Z} \to R/qR$ given by $a + q\mathbb{Z} \mapsto a + qR$. We claim that this morphism is also surjective, so it is an isomorphism. An element in R has the form $\frac{b}{N}$, where N has the form $N = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, so $(q, N) = 1$. Then there is some $a \in \mathbb{Z}$ with $Na \equiv b \pmod{q}$, so $Na = b + qc$ for some $c \in \mathbb{Z}$. It follows that $a = \frac{b}{N} + q\frac{c}{N} \in \frac{b}{N} + qR$, so $\overline{\phi}(a + q\mathbb{Z}) = a + qR = \frac{b}{N} + qR$. So $\overline{\phi}$ is an isomorphism. For any $a \in R$ we denote by $\overline{a}$ its class in $R/qR \cong \mathbb{Z}/q\mathbb{Z} = \mathbb{F}_q$, i.e., $\overline{a} = \overline{\phi}^{-1}(a + qR)$. Similarly, if $P \in R[X]$ we denote by $\overline{P}$ the polynomial in $R/qR[X] \cong \mathbb{F}_q[X]$ obtained by applying $\overline{\phi}^{-1}$ to the coefficients of P.

If $(X - X_1) \cdots (X - X_n) = X^n + S_{n-1}X^{n-1} + \cdots + S_0$ and $(X - X_1^k) \cdots (X - X_n^k) = X^n + T_{n-1}X^{n-1} + \cdots + T_0$ with $S_i, T_i \in \mathbb{Z}[X_1, \ldots, X_n]$ then by Viète's formulas and the fundamental theorem of the symmetric polynomials we have $T_i = Q_i(S_0, \ldots, S_{n-1})$, where $Q_0, \ldots, Q_{n-1} \in \mathbb{Z}[X_0, \ldots, X_{n-1}]$. Since $P_{B_k} = X^n + b_{n-1}X^{n-1} + \cdots + b_0 = (X - \beta_1) \cdots (X - \beta_n)$ and $P_A = X^n + a_{n-1}X^{n-1} + \cdots + a_0 = (X - \beta_1^k) \cdots (X - \beta_n^k)$, we have $b_i = S_i(\beta_1, \ldots, \beta_n)$ and $a_i = T_i(\beta_1, \ldots, \beta_n)$, so $a_i = Q_i(b_0, \ldots, b_{n-1})$ for $0 \leq i \leq n-1$. Let now $\nu_1, \ldots, \nu_n$ be the roots of $\overline{P}_{B_k}$. We have $(X - \nu_1) \cdots (X - \nu_n) = X^n + \overline{b}_{n-1}X^{n-1} + \cdots + \overline{b}_0$ and if we denote $(X - \nu_1^k) \cdots (X - \nu_n^k) = X^n + \overline{c}_{n-1}X^{n-1} + \cdots + \overline{c}_0$ then, by the same reasoning as above, $\overline{c}_i = Q_i(\overline{b}_0, \ldots, \overline{b}_{n-1}) = \overline{Q_i(b_0, \ldots, b_{n-1})} = \overline{a}_i$. Thus $(X - \nu_1^k) \cdots (X - \nu_n^k) = X^n + \overline{a}_{n-1}X^{n-1} + \cdots + \overline{a}_0 = \overline{P}_A$, i.e., the roots of $\overline{P}_A$ are the kth powers of the roots of $\overline{P}_{B_k}$, same as for P_A and P_{B_k}.

We now assume that, besides $q \notin \{p_1, \ldots, p_r\}$, q is not a divisor of the numerator of a_0, that is, $a_0 \notin qR$, so $\overline{a}_0 \neq \overline{0}$. (By hypothesis $a_0 = (-1)^n \det A \neq 0$.) It follows that the roots $\nu_1^k, \ldots, \nu_n^k$ of $\overline{P}_A$ are not zero, so $\nu_i \neq 0\ \forall i$. We take $k = q^m - 1$, where $m = \operatorname{lcm}(1, 2, \ldots, n)$. Now, it is well known that $\mathbb{F}_{q^t}$ is the splitting field of every irreducible polynomial of degree t over $\mathbb{F}_q$. Consequently, every root ν_i of the nth degree polynomial $P_{B_k} \in \mathbb{F}_q[X]$ belongs to $\mathbb{F}_{q^t}$ for some $1 \leq t \leq n$. Then $t \mid m$, so $\nu_i \in \mathbb{F}_{q^t} \subseteq \mathbb{F}_{q^m}$. Since $\nu_i \neq \overline{0}$ we have $\nu_i \in \mathbb{F}_{q^m}^*$. But $k = q^m - 1 = |\mathbb{F}_{q^m}^*|$, so $\nu_i^k = \overline{1}$. It follows that $\overline{P}_A = (X - \nu_i^k) \cdots (X - \nu_i^k) = (X - \overline{1})^n$, which implies that $\overline{a}_i = \overline{(-1)^{n-i}\binom{n}{i}}$, so $a_i - (-1)^{n-i}\binom{n}{i} \in qR$. This means that $a_i - (-1)^{n-i}\binom{n}{i}$ writes as a fraction with the numerator divisible by q. Since this happens for an infinity of primes q (all q not dividing $p_1 \cdots p_r$ or the numerator of a_0) we must have $a_i - (-1)^{n-i}\binom{n}{i} = 0$. It follows that $P_A = X^n + a_{n-1}X^{-1} + \cdots + a_0 = (X - 1)^n$. Hence the conclusion.

b) Yes, the converse of a) is also true. We have $(A - I)^n = P_A(A) = 0$, i.e., $B^n = 0$, where $B = A - I$.

We use the binomial formula $(1+X)^{1/k} = \sum_{i\geq 0} \binom{\frac{1}{k}}{i} X^i$, which implies $1+X = \left(\sum_{i\geq 0} \binom{\frac{1}{k}}{i} X^i\right)^k$. This formula holds not only for $X \in \mathbb{R}$ (or $\mathbb{C}$) with $|X|<1$ but also as an equality of formal series. Since $(1+X)^{1/k} \equiv \sum_{i=0}^{n-1} \binom{\frac{1}{k}}{i} X^i$ (mod X^n), we have $\left(\sum_{i=0}^{n-1} \binom{\frac{1}{k}}{i} X^i\right)^k \equiv 1+X \pmod{X^n}$, so $\left(\sum_{i=0}^{n-1} \binom{\frac{1}{k}}{i} X^i\right)^k = 1+X+X^nQ(X)$ for some $Q \in \mathbb{Z}[X]$.

It follows that $\left(\sum_{i=0}^{n-1} \binom{\frac{1}{k}}{i} B^i\right)^k = I+B+B^nQ(B) = I+B = A$. So we found $B_k \in \mathcal{M}_n(\mathbb{Q})$ with $B_k^k = A$, namely $B_k = \sum_{i=0}^{n-1} \binom{\frac{1}{k}}{i} B^i$. (Note that $B^i = 0$ for $i \geq n$, so we may also write $B_k = \sum_{i\geq 0} \binom{\frac{1}{k}}{i} B^i$.) $\square$

Note. We can remove the condition that A is invertible, but then the necessary and sufficient condition is that μ_A, the minimal polynomial of A, has the form $(X-1)^l$ or $X(X-1)^l$ for some l.

Let m $(0 \leq m \leq n)$ be the multiplicity of the root 0 in P_A. Since the roots of P_A are the kth powers of the roots B_k the multiplicity of 0 in B_k is again m. Moreover, if $P_A = X^m P'_A$ and $P_{B_K} = X^m P'_{B_K}$ then P'_A and P'_{B_K} are monic, of degree $n-m$ and with rational coefficients and the roots of P'_A are kth powers of the roots of P'_{B_K}. Then by the same reasoning from the solution of Problem 398 a) we have $P'_A = (X-1)^{n-m}$. Thus $P_A = X^m(X-1)^{n-m}$.

We may assume that $m \geq 1$ since otherwise $P_A = (X-1)^n$, i.e., $\mu_A = (X-1)^l$ for some l, a case already considered. Now by considering the Jordan canonical form we can write $B_k \sim B'_k \oplus B''_k$, where B'_k is the sum of all Jordan blocks corresponding to the eigenvalue 0 and B''_k is the sum of the Jordan blocks corresponding to nonzero eigenvalues. Since the multiplicity of 0 in P_{B_k} is m, the matrix B'_k is $m \times m$ and $P_{B'_k} = X^m$, so ${B'}_k^m = 0$.

We take $k = m$ and we have $A = B_m^m \sim (B'_m \oplus B''_m)^m = {B'}_m^m \oplus {B''}_m^m = 0_m \oplus {B''}_m^m$. (By 0_m we denote the $m \times m$ zero matrix.) Then $\mu_{0_m} = X$ and $P_{0_m} = X^m$, which, together with $P_A = X^m(X-1)^{n-m}$, implies $P_{{B''}_m^m} = (X-1)^{n-m}$, so $\mu_{{B''}_m^m} = (X-1)^l$ for some $l \leq n-m$. Since $\mu_{0_m} = X$ and $\mu_{{B''}_m^m} = (X-1)^l$, we have $\mu_A = X(X-1)^l$.

We now prove the sufficiency. We keep the notation $B = A - I$ from the solution of part b) of Problem 398.

The case $\mu_A = (X-1)^l$ is just Problem 398 b). Namely, the matrix $B_k \in \mathcal{M}_n(\mathbb{Q})$ satisfying $B_k^k = A$ is $B_k = \sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B^i = \sum_{i\geq 0} \binom{\frac{1}{k}}{i} B^i$.

If $\mu_A = X$ then $A = 0$, so we just take $B_k = 0$.

Suppose now that $\mu_A = X(X-1)^l$ with $l \geq 1$, so $P_A = X^m(X-1)^{n-m}$ with $1 \leq m < n$. We use the Jordan canonical form and therefore write $A = S(A' \oplus A'')S^{-1}$, where $S \in \mathrm{GL}_n(\mathbb{Q})$ and A' and A'' are the sum of the Jordan blocks corresponding to the eigenvalue 0 or 1, respectively. Then $\mu_{A'} = X$, so $A' = 0_m$, and $\mu_{A''} = (X-1)^l$.

Then for any $k \geq 1$ we have $A'' = {B''}_k^k$, where $B''_k = \sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B''^i$, with $B'' = A'' - I_{n-m}$. Then we take $B_k = S(0_m \oplus B''_k)S^{-1}$, for which we have $B_k^k = S(0_m \oplus B''_k)^k S^{-1} = S(0_m^k \oplus {B''}_k^k)S^{-1} = S(0_m \oplus A'')S^{-1} = A$.

In fact B_k can be expressed as a polynomial in A, as in the case $\mu_A = (X-1)^l$. To do this we regard A as a linear function $A : \mathbb{Q}^n \to \mathbb{Q}^n$. Since $\mu_A = X(X-1)^l$ we can write $\mathbb{Q}^n = V_0 \oplus V_1$, where $V_0 = \ker A = \mathrm{Im}\,(A-I)^l$ and $V_1 = \ker(A-I)^l = \mathrm{Im}\, A$ are invariant subspaces of $\mathbb{Q}^n$, i.e., $AV_\lambda \subseteq V_\lambda$ for $\lambda = 0, 1$. We will look for a B_k such that V_λ are invariant relative to B_k. Then $B_k^k = A$ is equivalent to $B_{k|V_\lambda}^k = A_{|V_\lambda}$ for $\lambda = 0, 1$.

On $V_0 = \ker A$ we have $A_{|V_0} = 0$, so when we take $B_{k|V_0} = 0$ we have $B_{k|V_0}^k = A_{|V_0}$. On $V_1 = \ker(A-I)^l$ we have $(A_{|V_1} - I_{|V_1})^k = 0$. Then by the case $\mu_A = (X-1)^l$ if $B_{k|V_i} = \sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B_{|V_1}^i$ then $B_{k|V_1}^k = A_{|V_1}$.

In conclusion, in order that $B_k^k = A$ it is enough that $B_k(x) = 0$ if $x \in V_0$ and $B_k x = \left(\sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B^i\right) x$ if $x \in V_1$. Let $x \in \mathbb{Q}^n$ be arbitrary. Then $x = x_0 + x_1$, where $x_0 = (I-A)^l x$ and $x_1 = (I-(I-A)^l)x = (I-(-B)^l)x$. Note that $x_0 \in \mathrm{Im}\,(A-I)^k = V_0$ and $x_1 = A\left(\sum_{i=1}^{l} \binom{l}{i}(-A)^{i-1}\right) x \in \mathrm{Im}\, A = V_1$. It follows that $B_k x = B_k x_0 + B_k x_1 = 0 + \left(\sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B^i\right) x_1 = \left(\sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B^i\right)(I-(-B)^l)x$. In conclusion, $B_k = \left(\sum_{i=0}^{l-1} \binom{\frac{1}{k}}{i} B^i\right)(I-(-B)^l)$.

There is a direct proof of the fact that B_k defined above works. We have $0 = A(A-I)^l = (B+I)B^l$. Then for any $s \geq l$ we have $(B+1)B^s = 0$, i.e., $B^s = -B^{s+1}$. It follows that $(-B)^l = (-B)^{l+1} = (-B)^{l+2} = \ldots$. Hence for any $s, t \geq l$ we have $(-B)^s = (-B)^t$, so $B^s = (-1)^{t-s}B^t$. In particular, $(-B)^r = (-B)^{2r}$, so $(I-(-B)^l)^2 = I - 2(-B)^l + (-B)^{2r} = I - (-B)^l$, i.e., $I - (-B)^l$ is idempotent. As seen in the solution of Problem 398 b) we have

$\left(\sum_{i=0}^{l-1}\binom{\frac{1}{k}}{i}X^i\right)^k = 1 + X + X^lQ(X)$ for some $Q \in \mathbb{Z}[X]$. It follows that

$$B_k^k = \left(\sum_{i=0}^{l-1}\binom{\frac{1}{k}}{i}B^i\right)^k (I-(-B)^l)^k = (I+B+B^lQ(B))\big(I-(-B)^l\big)$$
$$= I + B - (I+B)(-B)^l + B^l\big(I-(-B)^l\big)Q(B).$$

From $(I+B)(-B)^l = (-1)^l(B+I)B^l = 0$ and

$$B^l\big(I-(-B)^l\big) = (-1)^l(-B)^l\big(I-(-B)^l\big) = (-1)^l\big((-B)^l - (-B)^{2l}\big) = 0$$

it then results $B_k^k = I + B = A$.

Note that B_k also writes as

$$\sum_{i=0}^{l-1}\binom{\frac{1}{k}}{i}B^i\big(I-(-B)^l\big) = \sum_{i=0}^{l-1}\binom{\frac{1}{k}}{i}\big(B^i - (-1)^{l-i}B^l\big).$$

(We have $B^l = (-1)^iB^{l+i}$, so $B^i(-B)^l = (-1)^lB^{l+i} = (-1)^{l-i}B^l$.) Also if $i \geq l$ then $B^i - (-1)^{l-i}B^l = 0$, so we have $B_k = \sum_{i\geq 0}\binom{\frac{1}{k}}{i}\big(B^i - (-1)^{l-i}B^l\big)$.

Also, in both cases when $\mu_A = (X-1)^l$ or $X(X-1)^l$ we have μ_A divides $X(X-1)^n$, so $A(A-I)^n = 0$. It follows that we can use the formula above for B_k with $l = n$: $B_k = \sum_{i=0}^{n-1}\binom{\frac{1}{k}}{i}\big(B^i - (-1)^{n-i}B^n\big) = \sum_{i\geq 0}\binom{\frac{1}{k}}{i}\big(B^i - (-1)^{n-i}B^n\big)$, which works in all cases.

399. Let $n \geq 3$ and let $P = a_nX^n + \cdots + a_0 \in \mathbb{R}[X]$ with $a_i > 0$ $\forall i$ such that all the roots of P' are real. If $0 \leq a < b$ prove that

$$\frac{\int_a^b \frac{1}{P'(x)}\mathrm{d}x}{\int_a^b \frac{1}{P''(x)}\mathrm{d}x} \geq \frac{P'(b)-P'(a)}{P(b)-P(a)}.$$

Proposed by Florin Stănescu, Şerban Cioculescu School, Găeşti, Dâmboviţa, Romania.

Solution by the author. Let $x_1, \ldots, x_{n-1}$ be the roots of P'. Since the coefficients of P' are $ia_i > 0$ for $1 \leq i \leq n$, we have $x_i < 0$ $\forall i$. We have for all positive x $P''(x)/P'(x) = \sum_{j=1}^{n-1}(x-x_1)^{-1}$, whence $(P''(x)/P'(x))' = -\sum_{j=1}^{n-1}(x-x_1)^{-2} < 0$. Hence the function $\phi : [0,\infty) \to (0,\infty)$, $\phi(x) = \frac{P''(x)}{P'(x)}$, is strictly decreasing. Now P'' is strictly increasing and positive on $[0,\infty)$. (It has positive coefficients.) Hence $\frac{1}{P''}$ is positive and strictly decreasing on $[0,\infty)$, same as ϕ. By Chebyshev inequality we have

$$\int_a^b \frac{1}{P'(x)}\mathrm{d}x = \int_a^b \phi(x)\cdot\frac{1}{P''(x)}\mathrm{d}x \geq \frac{1}{b-a}\int_a^b \phi(x)\mathrm{d}x\int_a^b \frac{1}{P''(x)}\mathrm{d}x. \quad (1)$$

On the other hand P' is positive and strictly increasing on $[0,\infty)$. Since ϕ and P' have opposite monotony, by Chebyshev's inequality we also have

$$\int_a^b P''(x)\mathrm{d}x = \int_a^b \phi(x)P'(x)\mathrm{d}x \le \frac{1}{b-a}\int_a^b \phi(x)\mathrm{d}x \int_a^b P'(x)\mathrm{d}x. \qquad (2)$$

From (1) and (2) we conclude that

$$\frac{\int_a^b \frac{1}{P'(x)}\mathrm{d}x}{\int_a^b \frac{1}{P''(x)}\mathrm{d}x} \ge \frac{1}{b-a}\int_a^b \phi(x)\mathrm{d}x \ge \frac{\int_a^b P''(x)\mathrm{d}x}{\int_a^b P'(x)\mathrm{d}x} = \frac{P'(b)-P'(a)}{P(b)-P(a)}.$$

Note that $\int_a^b \phi(x)\mathrm{d}x = \log P'(x)\big|_a^b = \log\left(\frac{P'(b)}{P'(a)}\right)$. □

Solution by V. Makanin, Sankt Petersburg, Russia. We need the following result.

Lemma. Let f and g be continuous real functions defined on an interval I and assuming positive values. Suppose that fg and f/g are both increasing on I. Then for all $a<b$ from I it holds

$$\frac{\int_a^b (f(x))^{-1}\mathrm{d}x}{\int_a^b (g(x))^{-1}\mathrm{d}x} \ge \frac{\int_a^b g(x)\mathrm{d}x}{\int_a^b f(x)\mathrm{d}x}.$$

Proof. Fix, for the moment, $a\in I$ and define

$$H(x) = \int_a^x f(t)\mathrm{d}t \int_a^x \frac{1}{f(t)}\mathrm{d}t - \int_a^x g(t)\mathrm{d}t \int_a^x \frac{1}{g(t)}\mathrm{d}t$$

for all $x \ge a$ in I. Simple calculations yield

$$H'(x) = \int_a^x \left(\frac{f(x)}{g(x)} - \frac{f(t)}{g(t)}\right)\frac{f(x)g(x)-f(t)g(t)}{f(x)f(t)}\mathrm{d}t,$$

and, by the monotonicity and sign hypotheses, one sees that $H'(x)\ge 0$ for all $x\ge a$ in I. Thus H increases for $x \ge a$ and, since $H(a)=0$, we get $H(x)\ge 0$ for all $x\ge a$. But for $x=b$ this is clearly equivalent to the desired inequality (where the denominators are positive). □

Now, for the solution, observe that $f=P'$ and $g=P''$ are increasing and positive on $I=[0,\infty)$ (as polynomial functions with all coefficients positive). On the other hand, P' has all zeros real, thus negative (otherwise P' wouldn't be positive for $x\ge 0$, but this is the case due to its positive coefficients), let them be $-z_1,\dots,-z_{n-1}$, with positive z_i. Then

$$\frac{f(x)}{g(x)} = \frac{P'(x)}{P''(x)} = \left(\frac{1}{x+z_1}+\cdots+\frac{1}{x+z_{n-1}}\right)^{-1}$$

is definite and obviously increasing on $[0,\infty)$. The lemma therefore applies to P' and P'' in place of f and g leading to the desired inequality. □

Note. Here is a detailed proof for the formula for $H'(x)$ from V. Makanin's solution. If $F_1(x)=\int_0^x f(t)\mathrm{d}t$ and $F_2(x)=\int_0^x \frac{1}{f(t)}\mathrm{d}t$ then $F_1'(x)=f(x)$ and $F_2'(x)=\frac{1}{f(x)}$, so the derivative of $F_1(x)F_2(x)=\int_0^x \frac{1}{f(t)}\mathrm{d}t\int_0^x \frac{1}{f(t)}\mathrm{d}t$ is

$$\begin{aligned} f(x)F_2(x)+\frac{1}{f(x)}F_1(x) &= f(x)\int_0^x \frac{1}{f(t)}\mathrm{d}t+\frac{1}{f(x)}\int_0^x f(t)\mathrm{d}t \\ &= \int_0^x \left(\frac{f(x)}{f(t)}+\frac{f(t)}{f(x)}\right)\mathrm{d}x. \end{aligned}$$

A similar formula gives the derivative of $\int_0^x \frac{1}{g(t)}\mathrm{d}t\int_0^x \frac{1}{g(t)}\mathrm{d}t$, so

$$H'(x)=\int_0^x \left(\frac{f(x)}{f(t)}+\frac{f(t)}{f(x)}-\frac{g(x)}{g(t)}-\frac{g(t)}{g(x)}\right)\mathrm{d}x.$$

But one calculates

$$\begin{aligned} \frac{f(x)}{f(t)}+\frac{f(t)}{f(x)}-\frac{g(x)}{g(t)}-\frac{g(t)}{g(x)} &= \frac{(f(x)g(t)-f(t)g(x))(f(x)g(x)-f(t)g(t)}{f(x)f(t)g(x)g(t)} \\ &= \left(\frac{f(x)}{g(x)}-\frac{f(t)}{g(t)}\right)\frac{f(x)g(x)-f(t)g(t)}{f(x)f(t)}. \end{aligned}$$

This gives the formula for $H'(x)$.

400. For nonnegative integer n put $S(n):=\sum_{k=0}^{n}(-2)^k\binom{n}{k}\binom{2n-k}{n-k}$. Prove that $4(n+1)S(n)+(n+2)S(n+2)=0$ and conclude that

$$S_n=\begin{cases}(-1)^{n/2}\binom{n}{n/2} & \text{if } n \text{ is even}\\ 0 & \text{if } n \text{ is odd}\end{cases}.$$

Proposed by Mihai Prunescu, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

Solution by C.N. Beli. For any $\alpha\in\mathbb{Z}$ and any $k\in\mathbb{N}$ we have

$$\begin{aligned} \binom{\alpha}{k} &= \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}=(-1)^k\frac{(-\alpha+k-1)\cdots(-\alpha+1)(-\alpha)}{k!} \\ &= (-1)^k\binom{-\alpha+k-1}{k}. \end{aligned}$$

It follows that $S(n)=(-1)^n\sum_{k=0}^{n}2^k\binom{n}{k}\binom{-n-1}{n-k}=(-1)^n a_n$, where we have put $f:=(1+2x)^n(1+x)^{-n-1}=a_0+a_1x+\cdots$.

Similarly, $S(n+2)=(-1)^{n+2}b_{n+2}=(-1)^n b_{n+2}$, where we have put $g:=(1+2x)^{n+2}(1+x)^{-n-3}=b_0+b_1x+\cdots$. Hence the relation we want to prove is equivalent to $4(n+1)a_n+(n+2)b_{n+2}=0$.

By Liouville's theorem $a_n = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z^{n+1}}\mathrm{d}z$, where γ is a circle of radius $< \frac{1}{2}$ with the center at 0. Similarly for b_{n+2}. Hence,

$$4(n+1)a_n + (n+2)b_{n+2} = \frac{1}{2\pi i}\int_\gamma \big(4(n+1)\frac{f(z)}{z^{n+1}} + (n+2)\frac{g(z)}{z^{n+3}}\big)\mathrm{d}z.$$

It is readily seen that $4(n+1)\frac{f(z)}{z^{n+1}} + (n+2)\frac{g(z)}{z^{n+3}}$ can be written as

$$4(n+1)(1+2z)^n z^{-n-1}(1+z)^{-n-1} + (n+2)(1+2z)^{n+2}z^{-n-3}(1+z)^{-n-3}.$$

Let $\phi = z(1+z)$. We have $\phi' = 1+2z$, so that $(\phi')^2 = 1+4\phi$. Therefore,

$$\begin{aligned}4(n+1)\frac{f(z)}{z^{n+1}} + (n+2)\frac{g(z)}{z^{n+3}} &= 4(n+1)(\phi')^n\phi^{-n-1} + (n+2)(\phi')^{n+2}\phi^{-n-3}\\ &= 4(n+1)(\phi')^n\phi^{-n-1} + (n+2)(\phi')^n(1+4\phi)\phi^{-n-3}\\ &= (\phi')^n((n+2)\phi^{-n-3} + 4(n+2)\phi^{-n-2} + 4(n+1)\phi^{-n-1}).\end{aligned}$$

Let $h = (1+2z)^{n+1}z^{-n-2}(1+z)^{-n-2}(-1-2z-2z^2)$. From

$$h = (\phi')^{n+1}\phi^{-n-2}(-2\phi-1) = (\phi')^{n+1}(-\phi^{-n-2} - 2\phi^{-n-1})$$

we get

$$\begin{aligned}h' &= (n+1)\phi''(\phi')^n\left(-\phi^{-n-2} - 2\phi^{-n-1}\right)\\ &\quad + (\phi')^{n+1}\phi'\left((n+2)\phi^{-n-3} + 2(n+1)\phi^{-n-2}\right)\\ &= 2(n+1)(\phi')^n(-\phi^{-n-2} - 2\phi^{-n-1})\\ &\quad + (\phi')^n(1+4\phi)((n+2)\phi^{-n-3} + 2(n+1)\phi^{-n-2})\\ &= (\phi')^n((n+2)\phi^{-n-3} + 4(n+2)\phi^{-n-2} + 4(n+1)\phi^{-n-1})\\ &= 4(n+1)\frac{f(z)}{z^{n+1}} + (n+2)\frac{g(z)}{z^{n+3}}.\end{aligned}$$

It thus follows that $4(n+1)a_n + (n+2)b_{n+2} = \frac{1}{2\pi i}\int_\gamma h'(z)\mathrm{d}z = 0$. □

401. Prove the following identities:

$$(i)\quad \sum_{p\geq 0} p\binom{2a}{a-p}\binom{2b}{b-p} = \frac{ab}{2(a+b)}\binom{2a}{a}\binom{2b}{b},$$

$$(ii)\quad \sum_{p\geq 0}(2p+1)\binom{2a+1}{a-p}\binom{2b+1}{b-p} = \frac{(2a+1)(2b+1)}{a+b+1}\binom{2a}{a}\binom{2b}{b},$$

with the convention that $\binom{m}{n} = 0$ if $n < 0$ or $n > m$.

Proposed by Ionel Popescu, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

Solution by the author. Note that the general term in the first sum vanishes for p outside the interval $[1, \min\{a,b\}]$ and in the second sum it vanishes for p outside $[0, \min\{a,b\}]$

These identities can be checked with the help of the *zb* package written for Mathematica. For details on this method we refer the reader to [1]. For completeness we give here a more traditional proof.

The first identity is trivial when $a=0$ or $b=0$, so we may assume that $a,b\geq 1$. In this case it is equivalent to $h(a,b):=\sum_{p\geq 0} f(a,b,p)=1$, where

$$f(a,b,p)=\frac{2p(a+b)\binom{2a}{a-p}\binom{2b}{b-p}}{ab\binom{2a}{a}\binom{2b}{b}}.$$

The idea of the *zb* method in our case is to write $f(a+1,b,p)-f(a,b,p)$ in the form $g(a,b,p+1)-g(a,b,p)$ for some g satisfying $g(a,b,p)=0$ for $p=0$ and for $p\gg 0$. Then by summing over p going from 0 to ∞ one gets $h(a,b)-h(a+1,b)=0$, so $h(a,b)$ is independent of a. Then our statement follows from the obvious relation $h(1,b)=1$. (In the sum giving $h(1,b)$ the only nonzero term is $f(1,b,1)$, which calculates easily, $f(1,b,1)=1$.)

So the whole point is to determine the function $g(a,b,p)$ satisfying the conditions above. We refer to [1] for details. Here we just give the results obtained with Mathematica:

$$g(a,b,p)=-\frac{2p(p-1)\binom{2a+1}{a+p}\binom{2b-1}{b-p}}{a(2a+1)\binom{2a}{a}\binom{2b}{b}}.$$

We have $g(a,b,p)=0$ for $p=0$ and for $p\geq\min\{a+2,b+1\}$ and, after dividing by $f(a,b,p)$, the relation $f(a+1,b,p)-f(a,b,p)=g(a,b,p+1)-g(a,b,p)$ is equivalent to $\frac{f(a+1,b,p)}{f(a,b,p)}-1=\frac{g(a,b,p+1)}{f(a,b,p)}-\frac{g(a,b,p)}{f(a,b,p)}$, i.e., to

$$\frac{a(a+1)(a+b+1)}{(a+1-p)(a+1-p)(a+b)}-1=$$

$$=-\frac{(p+1)(b+p)}{2(a+1+p)(a+b)}+\frac{(p-1)(b+p)}{2(a+1-p)(a+b)},$$

which can be easily checked.

The proof of the second statement is done along the same lines. This time one has

$$f(a,b,p)=\frac{(2p+1)(a+b+1)\binom{2a+1}{a-p}\binom{2b+1}{b-p}}{(2a+1)(2b+1)\binom{2a}{a}\binom{2b}{b}}$$

and the function $g(a,b,p)$ satisfying $g(a,b,p)=0$ for $p=0$ and for $p\gg 0$, and $f(a+1,b,p)-f(a,b,p)=g(a,b,p+1)-g(a,b,p)$ is

$$g(a,b,p)=-\frac{p^2(b+1)^2\binom{2a+2}{a+1-p}\binom{2b+1}{b-p}}{\binom{2a+2}{a+1}\binom{2b}{b}}.$$

We leave the calculations to the reader. ◻

REFERENCES

[1] M. Petkovšek, H.S. Wilf, D. Zeilberger, $A = B$, A.K. Peters Ltd., Wellesley, MA (1996).

402. Let a, b, $\lambda \in \mathbb{R}$ and $u : [a, b] \to \mathbb{R}$ be a twice differentiable function with $u'(a) = u'(b) = 0$.

(1) Prove that $u''(c) = \lambda u(c)u'(c)$ for some $c \in (a, b)$.

(2) If moreover $u''(a) = 0$ prove that there exists $d \in (a, b)$ such that $(d - a)u''(d) = u'(d)\big(1 + \lambda(d - a)u(d)\big)$.

Proposed by Cezar Lupu, University of Pittsburgh, USA.

Solution by the author. Let us consider the function $\phi : [a, b] \to \mathbb{R}$ defined by

$$\phi(t) = u'(t) \cdot e^{-\lambda \int_a^t u(x)\mathrm{d}x}, \quad t \in [a, b].$$

A simple calculation of the derivative shows that

$$\phi'(t) = e^{-\lambda \int_a^t u(x)\mathrm{d}x}(u''(t) - \lambda u(t)u'(t)).$$

The condition $u'(a) = u'(b) = 0$ implies that $\phi(a) = \phi(b) = 0$, so, by Rolle's theorem, there exists $c \in (a, b)$ such that $\phi'(c) = 0$, which is equivalent to $u''(c) = \lambda u(c)u'(c)$.

For the second part of the problem, let us notice that $\phi'(a) = \phi'(c) = 0$ and by applying Flett's mean value theorem (see *Math. Gazette* **42** (1958), 38–39), there exists $d \in (a, b)$ such that $\phi'(d) = \frac{\phi(d)-\phi(a)}{d-a}$, which is equivalent to

$$(d - a)e^{-\lambda \int_a^d u(x)\mathrm{d}x}(u''(d) - \lambda u(d)u'(d)) = u'(d)e^{-\lambda \int_a^d u(x)\mathrm{d}x},$$

and thus problem (2) is solved. □

403. A parabola $\mathcal{P}$ has the focus F at distance d from the directrix Δ. Find the maximum length of an arc of $\mathcal{P}$ corresponding to a chord of length L.

Proposed by Gabriel Mincu, University of Bucharest, Romania.

Solution by the author. Let $FE \perp \Delta$, $E \in \Delta$, and let O be the midpoint of EF. We consider a cartesian coordinate system with origin O, the x-axis parallel to Δ, and such that $y_F = \frac{d}{2}$. Then $\Delta = \{(x, -\frac{d}{2}) \mid x \in \mathbb{R}\}$.

The distance from a point of coordinates (x, y) and F is $\sqrt{x^2 + (y - \frac{d}{2})^2}$ and the distance to Δ is $|y + \frac{d}{2}|$. Hence the parabola is given by the equation $\sqrt{x^2 + (y - \frac{d}{2})^2} = |y + \frac{d}{2}|$, i.e., by $y = ax^2$, with $a = \frac{1}{2d}$.

For $x, y \in \mathbb{R}$, $x < y$, we will denote by $\lambda(x, y)$ the length of the arc cut on $\mathcal{P}$ by the points of abscissae x and y. The required maximum will then be the maximum of $\lambda(x, y)$ with the constraint $(y - x)^2 + (ay^2 - ax^2)^2 = L^2$. We will prove that this maximum is reached for $(x, y) = \left(-\dfrac{L}{2}, \dfrac{L}{2}\right)$. To see this, let $x, y \in \mathbb{R}$, $x < y$, and let $M(x, ax^2)$ and $N(y, ay^2)$ be the corresponding points of $\mathcal{P}$. We also denote by P and Q the points of $\mathcal{P}$ which have abscissae $-\frac{L}{2}$ and $\frac{L}{2}$, respectively. We have to analyse several cases:

Case I. If $x \le -\frac{L}{2}$ and $y > \frac{L}{2}$ (or $x < -\frac{L}{2}$ and $y \ge \frac{L}{2}$), the length of the chord MN exceeds L, so the pairs (x, y) of this type have no contribution to the required maximum.

Case II. If $-\frac{L}{2} \le x < y \le \frac{L}{2}$, then $\lambda(x,y) = \int_x^y \sqrt{1+4a^2t^2}\mathrm{d}t \le \int_{-L/2}^{L/2} \sqrt{1+4a^2t^2}\mathrm{d}t = \lambda\left(-\frac{L}{2}, \frac{L}{2}\right)$.

Case III. If $x > -\frac{L}{2}$ and $y > \frac{L}{2}$, let us notice (bearing in mind that $x < y$) that y is uniquely determined by x (since, if N_1 and N_2 were points of $\mathcal{P}$ with abscissae $\dfrac{L}{2} < y_1 < y_2$ and such that $MN_1 = MN_2$, then the isosceles triangle MN_1N_2 would have the obtuse angle $\widehat{MN_1N_2}$, which is contradictory).

Let us notice that in this case some x's may not have a corresponding y, so that the chord MN has length L. Therefore, we must discuss two subcases:

Subcase III.1. If the circle with centre Q and radius L intersects $\mathcal{P}$ only in P and in a point R of abscissa $r > \dfrac{L}{2}$, then the distance from any point $M(x, ax^2)$ of $\mathcal{P}$ with $-\dfrac{L}{2} < x < \dfrac{L}{2}$ to Q is less than L, so there exists $y > \max\left\{\dfrac{L}{2}, x\right\}$ such that $MN = L$. Since the uniqueness of y in the conditions of Case III has been established, we obtain a functional dependence $y = \varphi(x)$, $x \in \left(-\dfrac{L}{2}, \infty\right)$.

Let $\mathcal{S} = \{(z,w) \in \mathbb{R} : z > -\dfrac{L}{2}, w > \dfrac{L}{2}, z < w\}$ and $g : \mathcal{S} \to \mathbb{R}$, $g(z,w) = (w-z)^2 + (aw^2 - az^2)^2 - L^2$. We notice that g is continuously differentiable on $\mathcal{S}$ and $\dfrac{\partial g}{\partial w}(z,w) = 2(w-z)[1 + 2a^2w(z+w)] > 0$. This partial derivative is nowhere zero on $\mathcal{S}$, so, according to the implicit function theorem, in the vicinity of each point (z_0, w_0) such that $g(z_0, w_0) = 0$ we may find continuously derivable functions ψ such that $(z, \psi(z)) \in \mathcal{S}$ and $g(z, \psi(z)) = 0$ for all z in the domain of ψ. The last two conditions imply, given the uniqueness discussed above, that for every such ψ and every x in its domain we have $\psi(x) = \varphi(x)$. The consequence of these considerations is that φ is a continuously differentiable function on $\left(-\dfrac{L}{2}, \infty\right)$.

According to the definition of φ, $(\varphi(x) - x)^2 + a^2(\varphi^2(x) - x^2)^2 = L^2$, whose derivative is $2(\varphi(x) - x)[\varphi'(x) - 1 + 2a^2(\varphi(x) + x)(\varphi(x)\varphi'(x) - x)] = 0$.

Since $\varphi(x) > x$ and $1 + 2a^2\varphi(x)(\varphi(x) + x) > 0$, we obtain

$$\varphi'(x) = \frac{1 + 2a^2x(\varphi(x) + x)}{1 + 2a^2\varphi(x)(\varphi(x) + x)}.$$

Now, $\lambda(x, y) = \lambda(x, \varphi(x)) = \displaystyle\int_x^{\varphi(x)} \sqrt{1 + 4a^2t^2}\mathrm{d}t$, so we consider the function $F : \left(-\dfrac{L}{2}, \infty\right) \to \mathbb{R}$ given by formula $F(x) = \displaystyle\int_x^{\varphi(x)} \sqrt{1 + 4a^2t^2}\mathrm{d}t$. We notice that F is differentiable, with $F'(x) = \sqrt{1 + 4a^2\varphi^2(x)}\varphi'(x) - \sqrt{1 + 4a^2x^2}$, and F can be extended continuously to $\left[-\dfrac{L}{2}, \infty\right)$ by putting $F\left(-\dfrac{L}{2}\right) = \int_{-L/2}^{L/2} \sqrt{1 + 4a^2t^2}\mathrm{d}t$.

From the above we see that $F'(x) < 0$ if and only if $\varphi'(x) < \sqrt{\frac{1+4a^2x^2}{1+4a^2\varphi^2(x)}}$. If $\varphi'(x) < 0$, this relation is obviously verified. If $\varphi'(x) \geq 0$, the inequality is equivalent to $\left(\dfrac{1 + 2a^2x(\varphi(x) + x)}{1 + 2a^2\varphi(x)(\varphi(x) + x)}\right)^2 < \dfrac{1 + 4a^2x^2}{1 + 4a^2\varphi^2(x)}$. After calculations, we see that this relation is equivalent to the obvious inequality $(\varphi(x) + x)(\varphi(x) - x)^3 > 0$.

Consequently, F is a strictly decreasing function on $\left[\dfrac{-L}{2}, \infty\right)$, so that $\lambda(x, y) = F(x) < F\left(-\frac{L}{2}\right) = \lambda\left(-\frac{L}{2}, \frac{L}{2}\right)$ for all (x, y) in the condition of Subcase III.1.

Subcase III.2. If the circle with centre Q and radius L intersects $\mathcal{P}$ in four points: P, R of Subcase III.1 and two other points U, V of abscissae u and v, respectively, and such that $-\dfrac{L}{2} < u \leq v < \dfrac{L}{2}$, the reasoning of Subcase III.1 may still be applied to reach the conclusion that the function given by $x \mapsto \int_x^{\varphi(x)} \sqrt{1 + 4a^2t^2}\mathrm{d}t$ is strictly decreasing on $\left[-\dfrac{L}{2}, u\right)$ and on $[v, \infty)$. Since, according to Case II, the value of this function at v does not exceed $\lambda\left(-\dfrac{L}{2}, \dfrac{L}{2}\right)$, we draw in this case also the conclusion that one has $\lambda(x, y) \leq \lambda\left(-\dfrac{L}{2}, \dfrac{L}{2}\right)$. For a point with the abscissa $x \in [u, v]$ the distance to Q is $\geq L$, so the distance to a point with abscissa $y > \frac{L}{2}$ is $> L$. Hence these poins don't count here.

Case IV. $x < -\dfrac{L}{2}$, $y < \dfrac{L}{2}$ reduces to Case III in view of the symmetry of $\mathcal{P}$ with respect to the y-axis.

The required maximum is therefore $\lambda\left(-\frac{L}{2}, \frac{L}{2}\right) = \int_{-L/2}^{L/2} \sqrt{1 + 4a^2t^2}\mathrm{d}t = 2\int_0^{L/2} \sqrt{1 + 4a^2t^2}\mathrm{d}t$.

The last integral can be computed with the change of variables $\tau = 2at$:

$$\lambda\left(-\frac{L}{2}, \frac{L}{2}\right) = \frac{1}{a}\int_0^{aL} \sqrt{1+\tau^2}\mathrm{d}\tau = \frac{L}{2}\sqrt{1+a^2L^2} + \frac{1}{2a}\ln(aL + \sqrt{1+a^2L^2})$$
$$= \frac{L\sqrt{L^2+4d^2}}{4d} + d\ln\frac{L+\sqrt{L^2+4d^2}}{2d}.$$

404. Let $F : \mathbb{Z}\times\mathbb{Z} \to \mathbb{Z}$ be a function satisfying the following conditions:

1) $\mid F(x, y) \mid \geq \mid x \mid + \mid y \mid \forall x, y \in \mathbb{Z}$.

2) There are $m, n \geq 1$ and the matrices $A = (a_{ij})$, $B = (b_{ij}) \in M_{m,n}(\mathbb{Z})$ such that

$$F(x, y) = \max_{1\leq i\leq m}\min_{1\leq j\leq n}(a_{ij}x + b_{ij}y)\ \forall x, y \in \mathbb{Z}.$$

Prove that either $F(x, y) \geq 0$ for all x, $y \in \mathbb{Z}$ or $F(x, y) \leq 0$ for all x, $y \in \mathbb{Z}$. Give an example of a function F for each of these two cases.

Proposed by Şerban Basarab, Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

Solution by the author. We may extend the function F to the whole $\mathbb{R}\times\mathbb{R}$ (with values in $\mathbb{R}$) by the formula $F(x, y) = \max_{1\leq i\leq m}\min_{1\leq j\leq n}(a_{ij}x + b_{ij}y)$ $\forall x, y \in \mathbb{R}$. This function is obviously continuous. We also have $F(xz, yz) = zF(x, y)$ for any $x, y, z \in \mathbb{R}$, $z > 0$, and $F(0, 0) = 0$.

Now for any $a, b, c \in \mathbb{Z}$, $c > 0$, we have $|a| + |b| \leq F(a, b) = c|F(\frac{a}{c}, \frac{b}{c})|$, so $|\frac{a}{c}| + |\frac{b}{c}| \leq |F(\frac{a}{c}, \frac{b}{c})|$. Thus the inequality $|F(x, y)| \geq |x| + |y|$ holds for any $x, y \in \mathbb{Q}$. By continuity it holds for any $x, y \in \mathbb{R}$. In particular, $F(x, y) \neq 0$ when $(x, y) \neq (0, 0)$. Hence $0 \notin F(\mathbb{R}^2 \setminus \{(0, 0)\})$. Now F is continuous and $\mathbb{R}^2 \setminus \{(0, 0)\}$ is connected. It follows that $F(\mathbb{R}^2 \setminus \{(0, 0)\}) \subseteq \mathbb{R}$ is a connected set, thus an interval. As $0 \notin F(\mathbb{R}^2 \setminus \{(0, 0)\})$, $F(\mathbb{R}^2 \setminus \{(0, 0)\})$ is contained in either $(-\infty, 0)$ or $(0, \infty)$, i.e., $F(x, y) < 0$ when $(x, y) \neq 0$ or $F(x, y) > 0$ when $(x, y) \neq 0$. Since also $F(0, 0) = 0$, we get the conclusion.

Examples of F with $F(x, y) \geq 0$ $\forall x, y \in \mathbb{Z}$ or $F(x, y) \leq 0$ $\forall x, y \in \mathbb{Z}$ are $F(x, y) = |x| + |y|$ and $F(x, y) = -|x| - |y|$, respectively. They clearly satisfy condition 1). For condition 2) if we take $m = 4$, $n = 1$ and $A = (1, 1, -1, -1)^T$, $B = (1, -1, 1, -1)^T$ we get $F(x, y) = \max\{x + y, x - y, -x + y, -x - y\} = |x| + |y|$; if we take $m = 1$, $n = 4$ and $A = (1, 1, -1, -1)$, $B = (1, -1, 1, -1)$ we get $F(x, y) = \min\{x + y, x - y, -x + y, -x - y\} = -|x| - |y|$. □

Erratum.

Due to file mishandling, the print version of *GMA* **32(111)** (2014), no. 1–2, contains two articles with the same title *Again on passing to the limit under integral sign.* The article authored by Mircea Merca is actually titled *An infinite family of inequalities involving cosecant sums.* The Editors apologize to both authors and readers for any inconvenience.