The purpose of this document is to provide students with more examples of good mathematical exposition, taking account of all necessary details, with clarity given priority over brevity. I have chosen some of the problems on the 2009 Putnam Exam for that purpose.

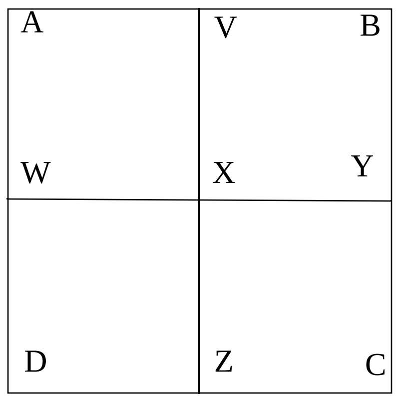
<u>Problem A1</u>

Let f be a real-valued function on the plane such that for every square ABCD in the plane, $f(A) + f(B) + f(C) + f(D) = 0$. Does it follow that $f(P) = 0$ for all points P in the plane ?

Solution A1

As stated this problem has a trivial solution: Allow ABCD to be a degenerate square with $A = B = C = D = P$ to discover that $4{\cdot}f(P) = 0$. The intended problem should say " … such that, for every *nondegenerate* square ABCD in the plane *with distinct vertices*, $f(A) + \dots$ "; and this problem's solution is offered below. In fact, the solution works if this problem allows only nondegenerate squares ABCD restricted to have sides parallel to one given square's.

To see why every $f(P) = 0$, partition any chosen nondegenerate square ABCD into four similar squares with vertices also at the midpoints of the edges of ABCD and at its center thus:



Let's abbreviate $f(A) := a$, $f(B) := b$, … , $f(Y) := y$ and $f(Z) := z$. The problem gives us five equations, one per square, thus:

$$\begin{array}{lr}
a + b + c + d = 0 & [1] \\
a + v + w + x = 0 & [2] \\
b + v + x + y = 0 & [3] \\
c + x + y + z = 0 & [4] \\
d + w + x + z = 0 & [5]
\end{array}$$

From these we compute [2] – [3] + [4] – [5] to get simply

$$a - b + c - d = 0 \qquad [6]$$

and then compute [1] ± [6] to reduce everything to these two equations

$$a + c = 0 \quad \text{and} \quad b + d = 0 .$$

These tell us that, at any two distinct points on a straight line parallel to a diagonal of the given square, the two values of f sum to zero. Therefore f is constant on every line parallel to a diagonal of the given square, and the constant must be zero as the problem claimed.

Yes, $f(P) = 0$ at every point P in the plane, even if the squares allowed are restricted to those with sides parallel to one given square's sides.

Continued …

Dan Wang's solution to Problem A1 works if the allowed squares ABCD have sides parallel either to one given square's sides, or else to its diagonals. He observed that VWZY is such a square, so it provides this equation too:

$$v + w + y + z = 0 \qquad [7]$$

Now compute [2] + [3] + [4] + [5] – [1] – 2·[7] to infer that $x = 0$.

Point X is the center of squares that could have been placed anywhere, so $f(X) = 0$ at every point X in the plane, answering problem A1's question affirmatively.

<u>Problem A2</u>

Functions f, g, h are differentiable on some open interval around 0 and satisfy the equations and initial conditions

$$f' = 2f^2gh + 1/(gh)\ ,\quad g' = fg^2h + 4/(fh)\ ,\quad h' = 3fgh^2 + 1/(fg)\ ,\quad f(0) = g(0) = h(0) = 1\ .$$

Find an explicit formula for $f(x)$, valid in some open interval around 0 .

Solution A2

Hereunder is why $f(x) = \left(\sqrt{2}\cdot\sin(6x + \pi/4)/(1 - \sin(12\cdot x))\right)^{1/6}$ for all $|x|$ small enough.

Set $p(x) := f(x)\cdot g(x)\cdot h(x)$ and take its *logarithmic derivative* to find

$$p'/p = f'/f + g'/g + h'/h = 6(p + 1/p)\ ;\quad \text{and}\quad p(0) = 1\ .$$

The solution of this differential equation for p is $p(x) = \tan(6x + \pi/4)$. Substitution into the given differential equation for f reduces it to

$$\log(f(x))' = f'/f = 2p + 1/p = 2\cdot\tan(6x + \pi/4) + \cot(6x + \pi/4)\ .$$

A symbolic integration turns this into, say,

$$6\cdot\log(f(x)) = \log\left(\sqrt{2}\cdot\sin(6x + \pi/4)/(1 - \sin(12\cdot x))\right)$$

after a tedious trigonometric simplification of expressions obtained from

$$\int\tan(y)\,dy = \log(\sec(y)) \quad\text{and}\quad \int\cot(y)\,dy = \log(\sin(y))\ .$$

However, the tedious simplification is not required to satisfy the problem’s demand for “ an explicit formula for $f(x)$ ”, so any one of infinitely many algebraically equivalent formulas, no matter how complicated, will serve as well. I would hate to have to grade this problem’s submitted solutions without a computerized algebra system like MAPLE or DERIVE competent enough to perform trigonometric simplifications.

Problem A6

Let $f : [0, 1]^2 \to \mathbb{R}$ be a continuous function on the closed unit square such that $\partial f/\partial x$ and $\partial f/\partial y$ exist and are continuous on the interior $(0, 1)^2$. Let

$$a = \int_0^1 f(0, y)\,dy\,, \quad b = \int_0^1 f(1, y)\,dy\,, \quad c = \int_0^1 f(x, 0)\,dx\,, \quad \text{and} \quad d = \int_0^1 f(x, 1)\,dx\,.$$

Prove or disprove: There must be a point (x_o, y_o) in $(0, 1)^2$ such that

$$\frac{\partial}{\partial x} f(x_o,y_o) = b - a \quad \text{and} \quad \frac{\partial}{\partial y} f(x_o,y_o) = d - c\,.$$

Solution A6

No such point (x_o, y_o) need exist though often it does. A disproof of the problem's allegation requires first that a suitable example f be found, and then that the last two equations be shown to have no solution (x_o, y_o) in the open square $(0, 1)^2$. The process may become easier to understand, or at least easier to appreciate, after a change in notation:

Identify column-vector $\mathbf{v} := \begin{bmatrix} x \\ y \end{bmatrix}$ with motions in the (x, y)-plane and reinterpret $f(\mathbf{v}) := f(x, y)$ as a scalar-valued real function of a real vector argument. $f(\mathbf{v})$'s derivative is the row-vector $f'(\mathbf{v}) = [\partial f/\partial x, \partial f/\partial y]$ because it satisfies the infinitesimal equation $df(\mathbf{v}) = f'(\mathbf{v})\cdot d\mathbf{v}$ (which merely abbreviates the *Chain Rule* $df(\mathbf{v})/d\tau = f'(\mathbf{v})\cdot d\mathbf{v}/d\tau$ valid for *every* differentiable vector-valued function $\mathbf{v}(\tau)$ of a real scalar variable τ while $\mathbf{v}(\tau)$ runs in the domain of f).

The problem's definition of $b - a$ can be rewritten

$$b - a = \int_0^1 (f(1, y) - f(0, y))\,dy = \int_0^1\int_0^1 f_x(\xi, y)\,d\xi\,dy \quad \text{wherein } f_x = \partial f/\partial x\,.$$

Similarly for $d - c = \int_0^1\int_0^1 f_y(x, \eta)\,d\eta\,dx$. Therefore row vector $[b - a, d - c]$ can be written

$[b - a,\ d - c]$ = Average of $f'(\mathbf{v})$ over the unit square in the $\mathbf{v}$-plane.

This turns Problem A6 into a special case of a more general question:

> Given a real scalar-valued differentiable function $f(\mathbf{v})$ of a vector argument $\mathbf{v}$, and given a region $\mathbf{S}$ in the domain of f, must a point $\mathbf{v}_o$ exist in $\mathbf{S}$ where
>
> $$f'(\mathbf{v}_o) = \text{Average}(\, f'(\mathbf{v}) \text{ over } \mathbf{v} \text{ in } \mathbf{S}\,)\ ?$$

With rare exceptions, the answer is "NO, NOT IN GENERAL".

One of those exceptions is so important it is taught in every introductory class in Calculus:

> The Mean Value Theorem of the Derivative (a corollary of Rolle's Theorem):
>
> > Given an interval $\mathbf{S}$ in the domain of a differentiable real-valued function f of a real argument, there is a point inside interval $\mathbf{S}$ where the derivative f' takes the same value as this derivative's average over interval $\mathbf{S}$.

More generally, a counter-example that justifies the unexceptional answer "NO" is the length function $f(\mathbf{v}) := \|\mathbf{v}\| := \sqrt{(\mathbf{v}^T\cdot\mathbf{v})}$ on a Euclidean vector-space of dimension 2 or more. The derivative $f'(\mathbf{v}) = \mathbf{v}^T/\|\mathbf{v}\|$, and $\|f'(\mathbf{v})^T\| = 1$, for all $\mathbf{v} \ne \mathbf{o}$; but unless $\mathbf{S}$ is a segment of a ray emanating from $\mathbf{o}$ the average of $f'(\mathbf{v})^T$ has length $\|\text{Average}\| < 1$, so there is no $\mathbf{v}_o$ in $\mathbf{S}$.

Problem B1

Show that every positive rational number can be written as a quotient of products of factorials of (not necessarily distinct) primes. For example, $10/9 = 2!\cdot 5!/(3!)^3$.

Solution B1

(This problem's statement is slightly untidy because $1 = 1/1$ is not here a quotient of products of factorials of primes since 1 is not a prime. Either rewrite $1 = 2!/2!$, or allow "products" to include possibly empty products defined to be 1 by convention. I have chosen the latter.)

Here is a proof: Start with the prime factorization of each rational number into a product of powers (some perhaps negative) of primes. For example, $10/9 = 2^1\cdot 5^1\cdot 3^{-2}$. The problem's assertion is proved by induction on the biggest prime that appears in such a factorization. The assertion is obviously true for every rational number r that is a power of 2 . Now let $p > 2$ be the biggest prime that appears in the prime factorization of any other rational number r , and let $m \neq 0$ be its exponent. Then $s := r\cdot(p!)^{-m}$ is a rational number in whose prime factorization appear only primes (if any) smaller than p . The induction hypothesis supplies an expression for s as a quotient of products (some perhaps empty) of factorials of (not necessarily distinct) primes. And then $r = s\cdot(p!)^m$ must also have the form demanded by the problem. End of proof.

Problem B2

A game involves jumping to the right on the real number line. If a and b are real numbers and $b > a$, the cost of jumping from a to b is $b^3 - a \cdot b^2$. For what real numbers c can one travel from 0 to 1 in a finite number of jumps with total cost exactly c ?

Solution B2

The set of all total costs c constitute an interval $1/3 < c \le 1$. Here is why:

The cost $b^3 - a \cdot b^2 = (b - a) \cdot b^2$ is the area of a rectangle of height b^2 erected over the interval $a \le x \le b$ on the real axis. The total cost c is the *Riemann Sum* of rectangular areas that overestimate the integral $\int_0^1 x^2\,dx = 1/3$, which is the area under the parabolic graph of $y = x^2$ over the interval $0 \le x \le 1$, as we shall see hereunder.

For any integer $n \ge 1$ partition the interval $0 \le x \le 1$ into n subintervals $x_{j-1} \le x \le x_j$ where $1 \le j \le n$ and $0 = x_0 < x_1 < \ldots < x_n = 1$. The total cost c of jumping from $x_0 = 0$ to x_1 to … to x_{n-1} to $x_n = 1$ is $c = Ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\}) := \sum_{1 \le j \le n} (x_j - x_{j-1}) \cdot x_j^2$. Observe that if $n = 1$ then $c = Ç(\{0, 1\}) = 1$ but otherwise, when $n > 1$, then $c < \sum_{1 \le j \le n} (x_j - x_{j-1}) \cdot 1^2 = 1$.

Now set $ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\}) := \sum_{1 \le j \le n} (x_j - x_{j-1}) \cdot x_{j-1}^2 < Ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\})$. This ç is the Riemann Sum that underestimates the area under the parabola because

$(b - a) \cdot a^2 < \int_a^b x^2\,dx = (b^3 - a^3)/3 = (b - a) \cdot (b^2 + b \cdot a + a^2)/3 < (b - a) \cdot b^2$ while $0 \le a < b$,

and therefore $c = Ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\}) > \int_0^1 x^2\,dx = 1/3 > ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\})$.

So far we have established that $1/3 < c \le 1$, among other things.

Next we shall see how to bring c down arbitrarily close to $1/3$. Given any $n > 1$ choose the uniform partition with every $x_j := j/n$ to find that

$$0 < c - 1/3 = Ç - 1/3 < Ç - ç = \sum_{1 \le j \le n} (x_j - x_{j-1}) \cdot (x_j^2 - x_{j-1}^2) = \sum_{1 \le j \le n} (2j - 1)/n^3 = 1/n ,$$

which can be made arbitrarily tiny by choosing n big enough.

Our final task is to establish that c can take *every* real value in the interval $1/3 < c \le 1$. To that end observe that $c = Ç(\{x_0, x_1, \ldots, x_{n-1}, x_n\})$ is a continuous function of its arguments $x_0, x_1, \ldots, x_{n-1}, x_n$ subject to the constraints $0 = x_0 \le x_1 \le \ldots \le x_n = 1$. On this closed domain Ç takes every value between any two that it takes. One of its values is $Ç(\{0, 0, \ldots, 0, 1\}) = 1$, its maximum value. Another value, between $1/3$ and $1/3 + 1/n$, comes arbitratily close to $1/3$ when n is big enough. Therefore c ranges throughout $1/3 < c \le 1$, as claimed.

<u>Problem B5</u>

Let $f : (1, \infty) \to \mathbb{R}$ be a differentiable function such that $f'(x) = (x^2 - f(x)^2)/\big(x^2\cdot(1 + f(x)^2)\big)$ for all $x > 1$. Prove that $f(x) \to \infty$ as $x \to \infty$.

Solution B5

Let me replace "f" by "y" because I shall have another use for "f". The problem becomes …

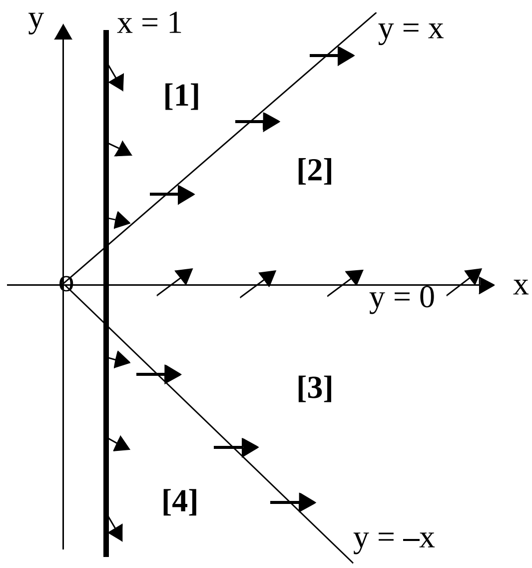
Suppose a real-valued differentiable $y(x)$ satisfies $y'(x) = (x^2 - y(x)^2)/\big(x^2\cdot(1 + y(x)^2)\big)$ for all $x > 1$. Prove that $y(x) \to \infty$ as $x \to \infty$.

The proof will invoke repeatedly a classical differential inequality for solutions $y(x)$ and $Y(x)$ of differential equations $y' = f(x, y)$ and $Y' = F(x, Y)$ respectively that says …

> If finite solutions $y(x)$ and $Y(x)$ both exist throughout an interval $\xi \le x \le \Xi$ whereon both $f(x, u)$ and $F(x, u)$ are continuous functions of both arguments provided u stays between y and Y inclusive, and if thereon EITHER OR BOTH
> $Y(\xi) \ge y(\xi)$ and $F(x, u) > f(x, u)$ OR $Y(\xi) > y(\xi)$ and $F(x, u) \ge f(x, u)$,
> then $Y(x) > y(x)$ throughout $\xi < x \le \Xi$.

(Why not simplify the two alternative hypotheses to one that requires just $Y(\xi) \ge y(\xi)$ and $F(x, u) \ge f(x, u)$? We could do so here because our functions f and F will be differentiable as well as continuous. But, in general, if f and F are merely continuous then they need not determine their respective solutions $y(x)$ and $Y(x)$ uniquely, and then the desired conclusion $Y(x) \ge y(x)$ could be falsified. See any text titled "Differential Inequalities".)

$f(x, y) := (x^2 - y^2)/\big(x^2\cdot(1 + y^2)\big)$ throughout Problem B5's proof, but it will choose $F(x, y)$ differently as needs arise. The proof will trace the passage of trajectories of all solutions $y(x)$ through four regions into which we shall partition the half-plane $x > 1$ of the (x, y)-plane:

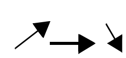


In region **[1]**, $y > x > 1$

In region **[2]**, $0 \le y \le x > 1$.

In region **[3]**, $0 > y > -x < -1$.

In region **[4]**, $y \le -x < -1$.

Little arrows like  show the directions in which trajectories $y(x)$ cross each region's boundaries.

The proof's first task is to infer that every trajectory $y(x)$ ultimately enters region **[2]** as x increases from 1 towards $+\infty$. To this end occasions will arise to notice that

$$\partial f(x, y)/\partial(y^2) = -(1 + x^2)/\big(x^2\cdot(1 + y^2)^2\big) < 0 < \partial f(x, y)/\partial(x^2) = y^2/\big(x^4\cdot(1 + y^2)\big)\ ;$$

this tells us that $f(x, y)$ is a decreasing function of y^2 but an increasing function of x^2 .

While $y(x)$ passes through region **[1]** where $y > x > 1$, we find $y' = f(x, y) < 0$, so $y(x)$ must descend as x increases until the trajectory escapes from region **[1]** into region **[2]** .

While $y(x)$ passes through region **[3]** where $0 > y > -x < -1$, we find $y' = f(x, y) > 0$, so $y(x)$ ascends as x increases. If $y(x)$ did not ultimately escape from region **[3]** across the x-axis into region **[2]** , then $y(x)$ would have to ascend to some limit $\ddot{\mathbf{Y}} < 0$ as $x \to +\infty$; but then $y(x)^2$ would have to descend to $\ddot{\mathbf{Y}}^2 > 0$ forcing $y' = f(x, y) = (x^2 - y^2)/\big(x^2\cdot(1 + y^2)\big)$ to ascend through positive values to its limit $f(+\infty, \ddot{\mathbf{Y}}) = 1/(1 + \ddot{\mathbf{Y}}^2) > 0$ as x increased to $+\infty$, which is impossible to reconcile with the bounded ascent of $y(x)$ to a limit $\ddot{\mathbf{Y}} < 0$. Therefore, if the trajectory of $y(x)$ passes through region **[3]** it ultimately escapes into region **[2]** .

While $y(x)$ passes through region **[4]** where $y \le -x < -1$, we find $-1/x^2 < y' = f(x, y) \le 0$, so $y(x)$ descends as x increases. However, for any $\xi > 1$ and for all $x > \xi$ we find that $y(x) > Y(x) := y(\xi) + 1/x - 1/\xi$; and $Y(x) > -x$ for all sufficiently big x . Consequently the trajectory of $y(x)$ ultimately escapes from region **[4]** into region **[3]** and then into region **[2]** .

Thus we have inferred that every trajectory $y(x)$ ultimately enters region **[2]** as x increases from 1 towards $+\infty$. While $y(x)$ passes through region **[2]** where $0 \le y \le x > 1$, we find $0 \le y' = f(x, y) \le 1$, so $y(x)$ ascends without ever exiting region **[2]** as x increases.

Can the ascent of $y(x)$ be bounded? No; here is why: Suppose for the sake of argument that $y(x)$ ascended to a finite limit $\ddot{\mathbf{y}} > 0$ as $x \to +\infty$. We would find that $y' = f(x, y) > f(x, \ddot{\mathbf{y}})$; and $f(x, \ddot{\mathbf{y}})$ ascends to its limit $f(\infty, \ddot{\mathbf{y}}) = 1/(1 + \ddot{\mathbf{y}}^2) > 0$ as $x \to +\infty$, which is impossible to reconcile with the bounded ascent of $y(x)$ to a finite limit. Therefore $y(x) \to +\infty$ as $x \to +\infty$ just as Problem B5 claims. End of proof.