**Problem 1(a):** *Maple V* used to say that the series $S := 1 - 1 + 1 - 1 + 1 - 1 + \ldots$ converged to $1/2$ . It was wrong because $S$ converges to $P/Q$ where $P$ and $Q$ are your favorite odd primes. For instance, if $P = 3$ and $Q = 7$ then we find for all $|x| < 1$ that
$(1+x+x^2)/(1+x+x^2+x^3+x^4+x^5+x^6) = (1-x^3)/(1-x^7) = 1 - x^3 + x^7 - x^{10} + x^{14} - x^{17} + x^{21} - x^{24} + \ldots$
and letting $x \to 1$ on both sides yields $3/7 = S$ in the limit. Right? Why?

**Solution 1(a):** " $3/7 = S$ " is wrong because $S$ does not converge. The faulty reasoning misapplies an important theorem due to N.H. Abel (1802-1829) about the convergence of any power series $\sum_{k\ge0} c_k z^k$ ; it converges to a differentiable function $f(z)$ of a complex argument $z$ inside but not outside the biggest circle centered at $0$ in the complex plane and containing no singularity of $f$ . If the series converges at a point on that circle it converges to the limit $f(z)$ reaches as $z$ moves there along any straight line segment starting inside the circle. But continuity of $f$ around a point on the circle does not guarantee convergence of the series there.

Abel's theorem is proved on pp. 37-40 of *Classical Complex Analysis* by L-S. Hahn & B. Epstein (1996, Jones & Bartlett, Mass.)

**Problem 1(b):** Has the series $1/6 + 1/12 + 1/20 + 1/30 + \ldots + 1/((k+1)k) + \ldots$ two sums? One is $(1/2 - 1/3) + (1/3 - 1/4) + (1/4 - 1/5) + (1/5 - 1/6) + (1/6 - 1/7) + \ldots = 1/2$ . The other sum is $(1 - 5/6) + (5/6 - 3/4) + (3/4 - 7/10) + (7/10 - 2/3) + (2/3 - 9/14) + \ldots = 1$ . Can you explain it?

**Solution 1(b):** The series $1/2 - 1/3 + 1/3 - 1/4 + 1/4 - 1/5 + \ldots + 1/k - 1/(k+1) + \ldots$ converges to $1/2$ but the series $1 - 5/6 + 5/6 - 3/4 + 3/4 - 7/10 + \ldots + (1/2 + 1/k) - (1/2 + 1/(k+1)) + \ldots$ does not converge at all.

**Problem 2:** The cubic polynomial $f(x) := x^3 + \alpha\cdot x^2 + \beta\cdot x + \gamma$ has integer coefficients $\alpha$, $\beta$, $\gamma$ none of which exceeds $99$ in magnitude, and $|f(2+\sqrt{3})| < 0.0001$ . Must $f(2+\sqrt{3}) = 0$ ? Why?

**Proof 2:** Let $\theta := f(2+\sqrt{3})$ ; we shall prove that $\theta = 0$ . First we obtain $(2+\sqrt{3})^2 = 7 + 4\sqrt{3}$ and $(2+\sqrt{3})^3 = 26 + 15\sqrt{3}$ . From these get $\theta = f(2+\sqrt{3}) = 26 + 7\alpha + 2\beta + \gamma + (15 + 4\alpha + \beta)\sqrt{3}$ . The conjugate of $\theta$ is $\Theta := f(2-\sqrt{3}) = 26 + 7\alpha + 2\beta + \gamma - (15 + 4\alpha + \beta)\sqrt{3}$ , and their product turns out to be $\theta\cdot\Theta = (26 + 7\alpha + 2\beta + \gamma)^2 - 3(15 + 4\alpha + \beta)^2$ , an integer. This integer must be small;

$$\begin{aligned} |\theta\cdot\Theta| &< (0.0001)\cdot(|26 + 7\alpha + 2\beta + \gamma| + |15 + 4\alpha + \beta|\sqrt{3}) \\ &< (0.0001)\cdot(26 + 700 + 200 + 100) + (15 + 400 + 100)\cdot\sqrt{3}) \\ &< (0.0001)\cdot(1026 + 515\cdot 2) = 0.2056 . \end{aligned}$$

The only integer that small is zero. Therefore $\theta\cdot\Theta = f(2+\sqrt{3})\cdot f(2-\sqrt{3}) = 0$ , so at least one of $f(2+\sqrt{3})$ and $f(2-\sqrt{3})$ must vanish. Actually both must vanish since $\sqrt{3}$ is irrational.

Consequently $f(x) = (x^2 - 4x + 1)\cdot(x + \gamma)$ for some integer $\gamma$ in $-24 \le \gamma \le 25$ . This problem was modified from one in the 6th *Chinese Girls' Mathematics Olympiad* of 14 Aug. 2007.

**Problem 3:** Suppose positive variables x, y and z satisfy $x^2+y^2+z^2 = 1$. Prove that

$$\sqrt{x^2 + (y^2-z^2)^2/4} + y+z \le \sqrt{3} .$$

**Solution 3:** Define $Q(y, z) := \sqrt{1-y^2-z^2 + (y^2-z^2)^2/4} + y+z = Q(z, y)$ over the circular disk where $y^2+z^2 = 1-x^2 < 1$. We observe that the inequality $Q(y, z) \le \sqrt{3}$ to be proved is almost false since $Q(1/\sqrt{3}, 1/\sqrt{3}) = \sqrt{3}$. This suggests that we prove first $Q(y, z) \le Q((y+z)/2, (y+z)/2)$, and then $Q(w, w) \le \sqrt{3}$ if $2w^2 \le 1$. Note: Point $[(y+z)/2, (y+z)/2]$ lies in the disk whenever $[y, z]$ does because then $2\cdot((y+z)/2)^2 \le y^2 + z^2 < 1$. We shall apply these inequalities again later.

To prove $Q(y, z) \le Q(w, w)$ at $w := (y+z)/2$ we examine

$$Q((y+z)/2, (y+z)/2) - Q(y, z) = \sqrt{1-2\cdot((y+z)/2)^2} - \sqrt{1-y^2-z^2 + (y^2-z^2)^2/4}$$

which has the same sign as has

$$\left(1-2\cdot((y+z)/2)^2\right) - \left(1-y^2-z^2 + (y^2-z^2)^2/4\right) = (y-z)^2/2 - (y^2-z^2)^2/4 = (y-z)^2\cdot\left(2-(y+z)^2\right)/4 ,$$

which the last inequalities above prove positive as needed to prove what was just claimed.

Now differentiate $Q(w, w) = \sqrt{1-2w^2} + 2w$ to find its maximum value $Q(1/\sqrt{3}, 1/\sqrt{3}) = \sqrt{3}$.

A short alternative solution by Roman Vaisberg substitutes $y := (u+v)/\sqrt{2}$ and $z := (u-v)/\sqrt{2}$ into $Q(y, z)$ to get $u = (y+z)/\sqrt{2} > 0$, $v = (y-z)/\sqrt{2}$, $u^2+v^2 = y^2+z^2 < 1$, and

$$G(u, v) := Q\left((u+v)/\sqrt{2}, (u-v)/\sqrt{2}\right) = \sqrt{1-u^2-v^2 + u^2\cdot v^2} + \sqrt{2}\cdot u = \sqrt{(1-u^2)\cdot(1-v^2)} + \sqrt{2}\cdot u .$$

Evidently $G(u, v) \le G(u, 0) = Q(u/\sqrt{2}, u/\sqrt{2}) \le \sqrt{3}$ as above.

This problem was modified from one in the 6th *Chinese Girls' Mathematics Olympiad* of 14 Aug. 2007.

**Problem 4:** The *Monastic Archipelago* consists of several small circular mostly rocky tropical islands in the ocean far from shipping lanes. Each island has a monastery housing a number of monks, different numbers on different islands. After a month of monsoon rains have washed shoreline sand away, each monk dredges back a pile of sand and spends subsequent days mostly praying and meditating at the shore next to his pile. At dusk every monk performs a virtuous act of sharing: He divides his pile into two halves and donates them to his two nearest neighbors along the shore. They do likewise to him, so his next day's sand pile is the average of his two neighbors' piles for the day just ending. As the days pass, the piles tend to become nearly equal on some islands, but not on others. Explain why.

**Solution 4:** The piles approach equality on every island with an odd number of monks, but do so only rarely on islands with even numbers of monks. Here is why: Let N be the number of monks on some one of the islands. $N \ge 2$ because "a number of monks" means more than one. (For more than a millennium, meticulous scholars construed Euclid's slightly sloppy definition of "a number" as "a quantity of things" *in the plural* to imply that neither 0 nor 1 could be numbers.) Assign a different integer j **mod** N to each of the island's monks in such a way that the shoreline neighbors of monk #j **mod** N are monks #j±1 **mod** N. Let $x_{j,k}$ be the amount of sand in the pile next to monk #j **mod** N on day #k after the monsoon rains end. Our problem says that $x_{j,k+1} = (x_{j+1,k} + x_{j-1,k})/2$ with the understanding that the first subscript be construed **mod** N .

Let us rewrite this equation as $\mathbf{x}_{k+1} = (E\cdot\mathbf{x}_k + E^{-1}\cdot\mathbf{x}_k)/2$ in which $\mathbf{x}_k$ is the column vector whose components are $x_{j,k}$ and E is the N-by-N *Circular Shift* matrix exemplified when N = 6 by

$$E = \begin{bmatrix} 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 1&0&0&0&0&0 \end{bmatrix} \quad \text{and} \quad E^{-1} = \begin{bmatrix} 0&0&0&0&0&1\\ 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0 \end{bmatrix} = E^T .$$

Thus we see that $\mathbf{x}_{k+1} = H\cdot\mathbf{x}_k = \dots = H^k\cdot\mathbf{x}_1$ wherein $H := (E + E^{-1})/2 = H^T$ . To understand what $\mathbf{x}_k$ does as $k \to \infty$ requires an appreciation of the eigenvalues and eigenvectors of H . Here it is:

Let $w := e^{2\pi\imath/N} = \cos(2\pi/N) + \imath\cdot\sin(2\pi/N)$ be the *Principal* $N^{th}$ root of 1 ; this means $w^N = 1$ but $w^n \ne 1$ if $0 < n < N$ . The N column vectors $\mathbf{w}_n := [1,\ w^n,\ w^{2n},\ w^{3n},\ \dots,\ w^{(N-1)n}]^T$ for $n = 1, 2, \dots, N-1, N$ are the eigenvectors of E because $E\cdot\mathbf{w}_n = w^n\cdot\mathbf{w}_n$ . The N eigenvalues $w^n$ are distinct, so the eigenvectors $\mathbf{w}_n$ are linearly independent and constitute a basis for the space of N-dimensional *complex* column vectors. They are also eigenvectors of $H = (E + E^{-1})/2$ because $H\cdot\mathbf{w}_n = \theta_n\cdot\mathbf{w}_n$ for the eigenvalues $\theta_n := (w^n + w^{-n})/2 = \cos(2\pi n/N)$ of H . Note that $\mathbf{w}_N = [1,\ 1,\ 1,\ \dots,\ 1]^T$ and, if N is even, $\mathbf{w}_{N/2} = [1,\ -1,\ 1,\ -1,\ \dots,\ 1,\ -1]^T$ .

The basis matrix $B := [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_N]$ must have an inverse. It turns out to be $\overline{B}^T/N$ where the overstroke $\overline{\dots}$ means *Complex Conjugate*; $\overline{\mathbf{w}}_n := [1,\ w^{-n},\ w^{-2n},\ w^{-3n},\ \dots,\ w^{-(N-1)n}]^T$ . To confirm that $B^{-1} = \overline{B}^T/N$ we shall use the *polynomial* $f(z) := (z^N - 1)/(z - 1) = \sum_{0\le j<N} z^j$ whose $f(w^{\pm n}) = 0$ for $0 < n < N$ but $f(w^0) = f(w^{\pm N}) = f(1) = N$ . Consequently for $1 \le m \le N$ too,

$$\overline{\mathbf{w}}_m{}^T\cdot\mathbf{w}_n = \textstyle\sum_{0\le j<N} w^{j\cdot(n-m)} = f(w^{(n-m)}) = \{\, N \text{ if } n = m \text{ but } 0 \text{ otherwise} \,\} ,$$

which explains why $\overline{B}^T\cdot B = N\cdot I$ and thus $B^{-1} = \overline{B}^T/N$ .

An expansion of $\mathbf{x}_1 = B\cdot(B^{-1}\cdot\mathbf{x}_1) = \sum_n \mathbf{w}_n\cdot\xi_n$ as a linear combination of eigenvectors $\mathbf{w}_n$ exists; it has coefficients $\xi_n := \overline{\mathbf{w}}_n\cdot\mathbf{x}_1/N$ . Then $\mathbf{x}_{k+1} = H^k\cdot\mathbf{x}_1 = \sum_n \cos^k(2\pi n/N)\cdot\mathbf{w}_n\cdot\xi_n$ . What this does as $k \to \infty$ is now evident because $\cos^k(2\pi n/N) \to 0$ for all n except $n = N$ and $n = N/2$ .

If N is odd no $n = N/2$ , so $\mathbf{x}_{k+1} \to \mathbf{w}_N\cdot\xi_N = \mathbf{w}_N\cdot(\mathbf{w}_N{}^T\cdot\mathbf{x}_1)/N$ ; all sand piles approach equality with the nonzero average of the piles' initial values. This confirms the solution's first claim.

If N is even, $\mathbf{x}_{k+1}$ comes ever closer to $\mathbf{w}_N\cdot(\mathbf{w}_N{}^T\cdot\mathbf{x}_1)/N + (-1)^k\cdot\mathbf{w}_{N/2}\cdot(\mathbf{w}_{N/2}{}^T\cdot\mathbf{x}_1)/N$ which oscillates around the average with constant amplitude $\mathbf{w}_{N/2}{}^T\cdot\mathbf{x}_1/N$ except in the unlikely event that this amplitude vanished initially. This confirms the solution's second claim. It seems obvious after Steven Lu's observation that, when N is even, the two sums $E_k := \sum_{1\le j\le N/2} x_{2j,k}$ and $D_k := \sum_{1\le j\le N/2} x_{2j-1,k}$ alternate: $E_{k+1} = D_k$ and $D_{k+1} = E_k$. This reveals that when $E_1 \ne D_1$ the sand piles can never approach equality.

**Problem 5:** A solid cube 20x20x20 is built out of bricks each 2x2x1. All are laid with their faces parallel to the cube's faces, though bricks need not all be laid flat. Prove that at least one straight line perpendicular to a face of the cube pierces its interior but no brick's interior. (Hard**!**)

**Solution 5:** Partition each face of the cube into an array of 20x20 unit *squares*. Through every interior corner of these unit squares pass a *needle*, a straight line perpendicular to all four squares through whose common corner the needle passes. There are $3\cdot 19^2 = 1083$ needles. Each of the 2000 bricks in the cube is pierced internally by just one needle, so some needles must pierce more than one brick. We shall discover that at least 83 needles pierce no brick**!**

Consider any one needle. Through it pass two orthogonal planes parallel to faces the needle does not penetrate. These planes cut the cube into four *quadrants*, each of which has a volume that must be an even number of cubic units. The planes may cut through some of the bricks. A brick cut by one plane is cut into *halves* each of volume 2. A brick cut by two planes is cut into *quarters* each of volume 1. Each quadrant contains an even number of quarters since the quadrant's volume is even, so every needle pierces an even number of bricks, if any. 2000 bricks can be pierced by at most 1000 needles, leaving at least 83 needles to pierce no brick.

(This is Jan Mycielski's problem found on pp. 801-2 of *Amer. Math. Monthly* **78** #7 Aug-Sept 1971. The same reasoning succeeds if "20" is replaced by "2", "4", "6", …, or "18", but not by "22". What happens then?)

**Problem 6:** Suppose $\Delta$ is a nondegenerate triangle (its vertices aren't collinear). Prove that …

**6(a):** Three points can be chosen, one on each edge of $\Delta$ but none at a vertex, through which three chosen points no ellipse inscribed in $\Delta$ can possibly pass.

**6(b):** An ellipse can be inscribed in $\Delta$ and touch the midpoints of all its edges.

**Proof 6(a):** To see why no inscribed ellipse need touch the triangle's edges at three points chosen arbitrarily, perform a linear transformation L that maps $\Delta$ onto the triangle $L(\Delta) := \mathit{\Delta}$ in the Cartesian (x, y)-plane whose edges' equations are $y = 1$, $x = 1$ and $2x + y = 0$. An ellipse **E** inscribed in $\Delta$ maps to an ellipse $\boldsymbol{E} := L(\mathbf{E})$ inscribed in $\mathit{\Delta}$. Points where $\boldsymbol{E}$ touches $\mathit{\Delta}$ are mapped by $L^{-1}$ to points where **E** touches $\Delta$. Can an ellipse $\boldsymbol{E}$ be inscribed in $\mathit{\Delta}$ and touch its edges at any chosen points, say (0, 1), (1, 0) and (0, 0) respectively? No; we shall show that such an ellipse $\boldsymbol{E}$ would generate a contradiction. Thus we shall infer that no ellipse **E** inscribed in $\Delta$ can touch it at the points on it to which our chosen points are mapped by $L^{-1}$.

If such an inscribed $\boldsymbol{E}$ existed, let its equation be $f(x, y) = 0$ where

$$f(x, y) := a\cdot x^2 + 2b\cdot x\cdot y + c\cdot y^2 + 2d\cdot x + 2e\cdot y\ .$$

Evidently $f(0, 0) = 0$ already. Also $f(0, 1) = c + 2e = 0$, and $f(1, 0) = a + 2d = 0$. These simplify $f$ to $f(x, y) = a\cdot x^2 + 2b\cdot x\cdot y + c\cdot y^2 - a\cdot x - c\cdot y$. Since $\boldsymbol{E}$ must be tangent to each edge at its contact point, at $(x, y) = (0, 1)$ we would have $dy/dx = -f_x(x, y)/f_y(x, y) = 0$, and this would imply that $2b - a = 0$; similarly at $(1, 0)$ we'd find $dx/dy = -f_y(x, y)/f_x(x, y) = 0$, so $2b - c = 0$. Now $f$ would simplify to $f(x, y) = c\cdot(x^2 + x\cdot y + y^2 - x - y)$ for some arbitrary constant $c \neq 0$. But this would force $dy/dx = -f_x(x, y)/f_y(x, y) = -1$ at $(x, y) = (0, 0)$ although the edge's slope is $dy/dx = -2$. Impossible! Therefore no ellipse $\boldsymbol{E}$ inscribed in our $\mathit{\Delta}$ touches its edges at the three points we chose arbitrarily. Likewise for $\mathbf{E}$ inscribed in $\Delta$.

There must be something special about the points at which an inscribed ellipse $\mathbf{E}$ can touch $\Delta$. Hye-Jin Jang's neat solution for Problem 6(a) reveals what's special: Let $\Delta$ have vertices P, Q and R, and suppose that $\mathbf{E}$ touches edge PQ at r, QR at p, and RP at q; then she shows why $(\|Pr\|/\|rQ\|)\cdot(\|Qp\|/\|pR\|)\cdot(\|Rq\|/\|qP\|) = 1$ is the special condition. It is derived from any invertible linear function L that maps $\mathbf{E}$ to a circle $\boldsymbol{C} := L(\mathbf{E})$ touching $\mathit{\Delta} := L(\Delta)$ at $p = L(p)$ on $QR = L(QR)$, $q = L(q)$ on $RP = L(RP)$, and $r = L(r)$ on $PQ = L(PQ)$. Although lengths of the edges of $\mathit{\Delta}$ may differ from the lengths of corresponding edges of $\Delta$, the ratios of lengths measured in parallel directions are preserved since L maps parallel lines to parallel lines. Thus

$$\|Pr\|/\|rQ\| = \|\mathrm{Pr}\|/\|\mathrm{rQ}\|\ , \quad \|Qp\|/\|pR\| = \|\mathrm{Qp}\|/\|\mathrm{pR}\|\ , \quad \text{and} \quad \|Rq\|/\|qP\| = \|\mathrm{Rq}\|/\|\mathrm{qP}\|\ .$$

But now the two tangents from each vertex of $\mathit{\Delta}$ to $\boldsymbol{C}$ have equal lengths: $\|Pr\| = \|qP\|$, $\|Qp\| = \|rQ\|$ and $\|Rq\| = \|pR\|$, so that $(\|Pr\|/\|rQ\|)\cdot(\|Qp\|/\|pR\|)\cdot(\|Rq\|/\|qP\|) = 1$. Because L preserves these ratios, $(\|\mathrm{Pr}\|/\|\mathrm{rQ}\|)\cdot(\|\mathrm{Qp}\|/\|\mathrm{pR}\|)\cdot(\|\mathrm{Rq}\|/\|\mathrm{qP}\|) = 1$ too, as claimed.

**Proof 6(b):** Given any nondegenerate triangle $\Delta$ in the Euclidean plane, many linear maps L can be found to produce $\Delta = L(\boldsymbol{Æ})$ for some equilateral triangle $\boldsymbol{Æ}$. For instance, if $\Delta$'s vertices are points $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$, and if $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are the vertices of an equilateral triangle $\boldsymbol{T}$, let L be the linear operator that maps vectors $\mathbf{P}-\mathbf{R}$ to $\mathbf{p}-\mathbf{r}$ and $\mathbf{Q}-\mathbf{R}$ to $\mathbf{q}-\mathbf{r}$; do you see why such an L must exist? Then $\Delta = L\cdot\boldsymbol{Æ}$ for an equilateral triangle $\boldsymbol{Æ} := \boldsymbol{T} - \mathbf{R} + L^{-1}\cdot\mathbf{r}$. Now let $\boldsymbol{C}$ be the circle inscribed in $\boldsymbol{Æ}$ and touching all its edges' midpoints. $\mathbf{E} := L\cdot\boldsymbol{C}$ is the ellipse inscribed in $\Delta$ and touching all its edges' midpoints. That was easy.

**Problem 7:** Suppose the triangle $\Delta$ and ellipse $\mathbf{E}$ in Problem 6(b) are drawn in the complex plane. Prove that the foci of $\mathbf{E}$ are the zeros of the derivative of whatever cubic polynomial's zeros are just the vertices of $\Delta$.

You should know that the foci of an ellipse earned their name because light emanating from one focus and reflecting off the ellipse as if it were a mirror converges, thus focussed by the mirror, at the other focus.

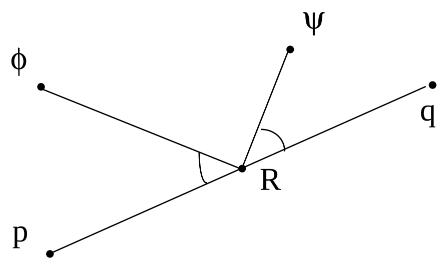
**Proof 7:** Let the triangle $\Delta$ have vertices p, q and r; now these are complex numbers. Let $\phi$ and $\psi$ be the zeros of the derivative of a cubic whose zeros are just p, q and r. The derivative of such a cubic, say $(z-p)(z-q)(z-r)$, is $3(z-\phi)(z-\psi)$; their ratio is the rational function

$$f(z) := d\ln\big((z-p)(z-q)(z-r)\big)/dz$$
$$= 1/(z-p) + 1/(z-q) + 1/(z-r) = 3(z-\phi)(z-\psi)/\big((z-p)(z-q)(z-r)\big)\ .$$

Rather than try to express $\phi$ and $\psi$ explicitly in terms of p, q and r we shall infer what we need to know from the properties of $f$ .

Denote by $\bar{z} := x-\imath y$ the complex conjugate of $z = x+\imath y$ . Since $f(\phi) = 0$ , so does $0 = \bar{f}(\bar{\phi}) = 1/(\bar{\phi}-\bar{p}) + 1/(\bar{\phi}-\bar{q}) + 1/(\bar{\phi}-\bar{r}) = (\phi-p)/|\phi-p|^2 + (\phi-q)/|\phi-q|^2 + (\phi-r)/|\phi-r|^2$ . This equation says that $\phi$ is a positively weighted average of p, q and r ; therefore $\phi$ lies inside $\Delta$ , and the same goes for $\psi$ . (Thus have we verified an instance of *Lucas' Theorem*: The zeros of a polynomial's derivative lie in the convex hull of the polynomial's zeros.)

Now let $P := (q+r)/2$ , $Q := (r+p)/2$ and $R := (p+q)/2$ be the midpoints of $\Delta$'s edges. We know from Problem 6(b) that an ellipse **E** inscribed in $\Delta$ touches its edges at P, Q and R , and we wish to prove that $\phi$ and $\psi$ are **E**'s foci. To do this we use the *Reflection Characterization* of an ellipse: Rays of light from one focus are reflected to the other by the ellipse. Since the edges of $\Delta$ are tangents to **E** at the contact points P, Q and R , we need merely show that two line segments joining $\phi$ and $\psi$ to any contact point make equal angles with the edge there. Consider contact point R and its edge $p \leftrightarrow q$ . Since $R-q = -(R-p) = (p-q)/2$ , $1/(R-q) = -1/(R-p)$ and therefore $f(R) = 1/(R-r)$ . This implies $\big((\phi-R)/(p-R)\big)\cdot\big((\psi-R)/(q-R)\big) = 1/3$ , whence follows $\mathrm{Arg}\big((\phi-R)/(p-R)\big) = -\mathrm{Arg}\big((\psi-R)/(q-R)\big)$ as required the reflection property.



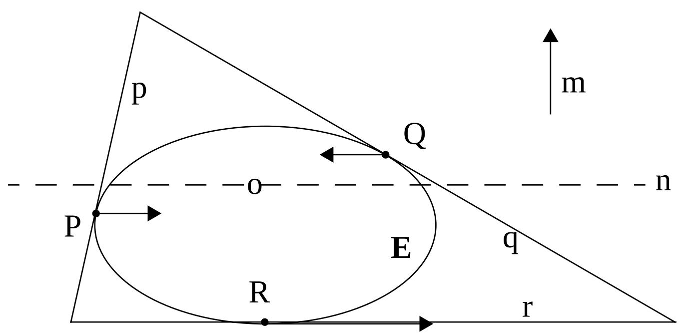

It is not hard to deduce from equations like $f(P) = 1/(P-p)$ that $|P-\phi| + |P-\psi| = |Q-\phi| + |Q-\psi| = |R-\phi| + |R-\psi|$ , which puts P, Q and R on some ellipse ***É*** with foci $\phi$ and $\psi$ . But, until ***É*** is proved not to cross any edge of $\Delta$ , we cannot be sure ***É*** is the desired inscribed ellipse **E** .

How is the Reflection Characterization deduced from some other characterization of an ellipse, say as the locus of points **z** whose distances from the foci $\phi$ and $\psi$ sum to a given constant? Revert now to vector notation for points in the Euclidean plane where length $\|\mathbf{v}\| = \sqrt{(\mathbf{v}^T\mathbf{v})}$ . Then its differential is $d\|\mathbf{v}\| = \mathbf{v}^T d\mathbf{v}/\|\mathbf{v}\|$ provided $\mathbf{v} \ne \mathbf{o}$ . As **z** runs on the ellipse whereon $\|\mathbf{z}-\phi\| + \|\mathbf{z}-\psi\| = \text{constant} > \|\phi-\psi\|$ , the tangent at **z** is in the direction of d**z** satisfying $0 = d(\|\mathbf{z}-\phi\| + \|\mathbf{z}-\psi\|) = \big( (\mathbf{z}-\phi)/\|\mathbf{z}-\phi\| + (\mathbf{z}-\psi)/\|\mathbf{z}-\psi\| \big)^T d\mathbf{z}$ . This says the tangent direction d**z** is perpendicular to the sum of the two unit-vectors $(\mathbf{z}-\phi)/\|\mathbf{z}-\phi\|$ and $(\mathbf{z}-\psi)/\|\mathbf{z}-\psi\|$ directed from the foci to **z** on the ellipse. A nonzero sum of unit vectors always bisects the angle between them; draw the rhombus whose vertices are at **o**, one unit vector, the other, and their sum to see why. This confirms the picture above with $\mathbf{z} = \mathbf{R}$ and d**z** parallel to $\mathbf{q}-\mathbf{p}$ .

Conversely, given two points $\phi$ and $\psi$ , if **z** moves on some curve whose tangent at **z** always bisects the exterior angle between $\mathbf{z}-\phi$ and $\mathbf{z}-\phi$ then the curve must be an ellipse with foci $\phi$ and $\psi$ because $\|\mathbf{z}-\phi\| + \|\mathbf{z}-\psi\|$ stays constant on that curve.

**Problem 8: The Biggest Ellipse in a Triangle**
Prove that, of all ellipses inside any given nondegenerate triangle, *the one* ellipse of largest area touches the triangle's edges at their midpoints. (*Cf.* Problems 6 & 7 above.)

**Proof 8:** Let **E** be *an* ellipse of largest area inscribed in a nondegenerate triangle $\Delta$ . **E** must touch all three edges of $\Delta$ ; otherwise, were some edge untouched, **E** could be translated very slightly towards this edge and away from the other two edges, and then enlarged. Call **E**'s three points of contact P, Q and R in $\Delta$'s edges p, q and r respectively. We shall prove by contradiction that PQ is parallel to r , QR to p , and RP to q , from which will follow via *Similar* triangles that P, Q and R are midpoints of their respective edges.




Suppose PQ were not parallel to r . Then a line n parallel to r through the midpoint o of PQ would separate P and Q ; suppose P were closer than Q to r . A *Shear* S that left n fixed, but moved P slightly away from p and Q slightly away from q , and slid R along r slightly, would change **E** into another ellipse with the same area but out of contact with p and q . But then, according to the previous paragraph, **E**'s area could not be a maximum.

The maximizing ellipse is unique because Problems 6 & 7 determined its foci, and distinct *Confocal* (with the same foci) ellipses never intersect.

What is that *Shear* S ? Put the plane's origin at o on line n , whereupon n, P, Q and R can be treated as vectors. Let m be any nonzero vector perpendicular to n , so that $m^T n = 0$ . Then $S = I + ßnm^T$ is a linear operator with a tiny but nonzero scalar ß . Since Sn = n , this shear leaves the line n fixed, and slides R to $SR = R + ß(m^T R)n$ along the line r parallel to vector n . Now $o = (P+Q)/2$ , so $P = -Q$ ; and $m^T P = -m^T Q \neq 0$ since P and Q lie on opposite sides of line n , not on it. Therefore $SP - P = ß(m^T P)n = -ß(m^T Q)n = -(SQ - Q)$ , which tells us S moves P and Q in opposite directions parallel to n . Choose the sign of ß so that P and Q get moved closer together and therefore into $\Delta$ rather than out of it. If |ß| is tiny enough, S will move **E** slightly to S**E** still inside $\Delta$ and touching edge r but not p nor q . Finally, Area(S**E**) = det(S)·Area(**E**) = Area(**E**)

because det(S) = 1 ; this last equation follows from a matrix identity $\det(I - B\cdot C) = \det\left(\begin{bmatrix} I & B \\ C & I \end{bmatrix}\right) = \det(I - C\cdot B)$

valid whenever the matrix products B·C and C·B are both square. The identity's proof is left to the reader. Or else prove Area(S**E**) = Area(**E**) by cutting **E** into infinitesimally thin slices all parallel to n and unaltered in length and width by S .