Problem 1: Four ghostly galleons – call them E, F, G and H, – sail at night on a ghostly sea so foggy that one side of a ship cannot be seen from the other. Each ship pursues its course steadily, changing neither its speed nor heading. G collides with H amidships; but since they are ghostly galleons they pass through each other with no damage nor change in course. As they part, H's captain hears G's say "Damnation! That's our third collision this night!" A little while later, F runs into H amidships with the same effect (none) and H's captain hears the same outburst from F's. What should H's captain do to reach her original destination, whatever it may be, with no further collision; and why will doing so succeed?

Solution 1: H need only change speed while keeping to the same heading. To see why this works, choose a coordinate system centered upon H and moving with it. As seen on a RADAR screen in H, the other galleons trace straight-line courses at constant speeds. The courses traced by F and G are straight lines through H, with which they have collided. But F and G have also collided with each other; therefore F and G trace the same straight line through H. Having both suffered three collisions, not two, F and G must have collided with E at different times, so E's course must stay in that same straight line. It cannot be aligned along H's course because H suffered collisions amidships, not by the bow or stern. If H changes speed and E does not, their courses will no longer intersect. E's captain has no reason to change speed since he cannot yet know anything about the course nor speed of H.

Problem 2: In an n-dimensional Euclidean space, the vertices of a triangle are at positions joined to the origin by vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$. Explain why the triangle's unoriented area is

$$A := \sqrt{\det(M^T\cdot M)}/2 \quad \text{wherein} \quad M := \begin{bmatrix} 1 & 1 & 1 \\ \mathbf{x-m} & \mathbf{y-m} & \mathbf{z-m} \end{bmatrix} \quad \text{and} \quad \mathbf{m} := (\mathbf{x}+\mathbf{y}+\mathbf{z})/3 .$$

Solution 2: If $n > 3$ choose a new orthonormal basis whose first three vectors span the subspace containing $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$. This reduces n to 1, 2 or 3 ; if $n = 1$ then $A = 0$ trivially. Next, without changing the determinant, subtract the column containing $\mathbf{x}$ from the others to get

$$A = \sqrt{\det\left(\begin{bmatrix} 1 & 0 & 0 \\ \mathbf{x-m} & \mathbf{u} & \mathbf{w} \end{bmatrix}^T \cdot \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{x-m} & \mathbf{u} & \mathbf{w} \end{bmatrix} \right)}/2$$

wherein $\mathbf{u} := \mathbf{y}-\mathbf{x}$ and $w := \mathbf{z}-\mathbf{x}$. Then add a third of each of the second and third columns to the first to get

$$A = \sqrt{\det\left(\begin{bmatrix} 1 & 0 & 0 \\ \mathbf{o} & \mathbf{u} & \mathbf{w} \end{bmatrix}^T \cdot \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{o} & \mathbf{u} & \mathbf{w} \end{bmatrix} \right)}/2 ,$$

which expands to $A = \sqrt{\mathbf{u}^T\mathbf{u}\cdot\mathbf{w}^T\mathbf{w} - (\mathbf{u}^T\mathbf{w})^2}/2 = \|\mathbf{u}\times\mathbf{w}\|/2$ by Lagrange's identity. Since $\|\mathbf{u}\times\mathbf{w}\|/2$ is the area of the triangle in question, the formula for A is vindicated.

An alternative proof moves the origin to $\mathbf{m}$, whereupon $\mathbf{x}+\mathbf{y}+\mathbf{z} = \mathbf{o}$. Next choose a new orthonormal basis whose first two vectors span the subspace containing $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ to reduce n to 2 . Then invoke a well-known formula: $A = |\det(M)|/2 = \sqrt{\det(M^T\cdot M)}/2$ now that $\mathbf{m} = \mathbf{o}$.

Problem 3: A skew-symmetric bilinear operator $\mathbf{W}$ is defined for any linear functional $\mathbf{w}^T \ne \mathbf{o}^T$ thus: $\mathbf{Wxy} := \mathbf{xw}^T\mathbf{y} - \mathbf{yw}^T\mathbf{x} = -\mathbf{Wyx}$. How does the *Range* of $\mathbf{W}$ compare with the *Nullspace* of $\mathbf{w}^T$, and why? (If you think $\mathbf{x}$ and $\mathbf{y}$ are columns, think of $\mathbf{w}^T$ as a row.)

NOTE THAT $\mathbf{xy}$ IN $\mathbf{Wxy}$ IS NOT A VECTOR PRODUCT. IT IS A PAIR OF VECTORS WRITTEN TO SUGGEST AN OBJECT THAT BEHAVES LIKE A PRODUCT LINEAR IN EACH FACTOR. The bilinear form "$\mathbf{Wxy}$" could be written "$\mathbf{W}(\mathbf{z}, ß\mathbf{x} + \mu\mathbf{y}) = ß\mathbf{W}(\mathbf{z}, \mathbf{x}) + \mu\mathbf{W}(\mathbf{z}, \mathbf{y})$" to show its linearity, and "$\mathbf{W}(\mathbf{x}, \mathbf{y}) = -\mathbf{W}(\mathbf{y}, \mathbf{x})$" to show its skew-symmetry but the extra commas and parentheses would clutter the page without clarifying anything.

Solution 3: Range($\mathbf{W}$) = Nullspace($\mathbf{w}^T$) ; here is why: Evidently $\mathbf{w}^T\mathbf{Wxy} = 0$ for all $\mathbf{x}$ and $\mathbf{y}$, so Nullspace($\mathbf{w}^T$) $\supseteq$ Range($\mathbf{W}$). On the other hand, for every $\mathbf{z}$ in Nullspace($\mathbf{w}^T$), so that $\mathbf{w}^T\mathbf{z} = 0$, and for any $\mathbf{m}$ such that $\mathbf{w}^T\mathbf{m} \ne 0$ (such an $\mathbf{m}$ must exist because $\mathbf{w}^T \ne \mathbf{o}^T$), set $\mathbf{u} := \mathbf{m}/\mathbf{w}^T\mathbf{m}$ to infer that $\mathbf{Wzu} = \mathbf{z}$ and hence that Range($\mathbf{W}$) $\supseteq$ Nullspace($\mathbf{w}^T$). Therefore Range($\mathbf{W}$) = Nullspace($\mathbf{w}^T$) .

This conclusion is surprising because the ranges of bilinear operators are not all linear (sub)spaces. For example, Let Δ be a 4-by-4 matrix whose one nonzero element sits in the upper-left corner of Δ and define the bilinear operator $\mathbb{S}XY := X\cdot\Delta\cdot Y^T - Y\cdot\Delta\cdot X^T = -\mathbb{S}YX$ to map 4-by-4 matrices X and Y each linearly into the 16-dimensional space of 4-by-4 matrices. Range($\mathbb{S}$) consists of all skew-symmetric 4-by-4 matrices of rank 2 . Some sums of these have rank 4 so they cannot constitute a linear subspace of 4-by-4 matrices.

The equation $\mathbf{Wxy} = \mathbf{r}$ can be solved for $\mathbf{x}$ and $\mathbf{y}$ given any $\mathbf{r}$ in Range($\mathbf{W}$) as follows: Choose any $\mathbf{m}$ and $\mathbf{u} := \mathbf{m}/\mathbf{w}^T\mathbf{m}$ as in the solution above; then $\mathbf{y} := \mathbf{u}$ and $\mathbf{x} := \mathbf{r}$ solves $\mathbf{Wxy} = \mathbf{Wru} = \mathbf{r}$, though not uniquely.

Problem 4: Suppose an odd number (at least three) of coins have the property that, if any one coin is removed, the rest can be partitioned into two groups each with the same number of coins and also the same total weight. Show that all the coins must have the same weight.

Proof 4: Let the coins' weights be $x_1, x_2, x_3, \ldots, x_{2N+1}$. Then for $i = 1, 2, 3, \ldots, 2N+1$ these weights must satisfy $\sum_j h_{i,j}\cdot x_j = 0$ wherein every $h_{i,j} = \pm 1$ except $h_{ii} = 0$, and for each i there are N coefficients $h_{i,j} = +1$ and N coefficients $h_{i,j} = -1$. Evidently a nontrivial solution for the homogeneous linear equations has $x_1 = x_2 = x_3 = \ldots = x_{2N+1} \ne 0$. Can there be any other kind of nontrivial solution? To see why not, we shall show that the deletion of the last row and column from the (2N+1)-by-(2N+1) matrix $\overline{H} := [h_{i,j}]$ leaves a 2N-by-2N matrix H with $\det(H) \ne 0$, so the choice of a nonzero x_{2N+1} determines uniquely all the other x_j's in a nontrivial solution.

What makes the problem interesting is that we do not know the signs of the nonzero $h_{i,j}$'s . Let's perform our computation of det(H) **mod** 2 . Now every off-diagonal $h_{i,j} \equiv 1$ **mod** 2 , and we find that $H \equiv (uu^T - I)$ **mod** 2 where $u^T = [1, 1, 1, \ldots, 1]$ and I is the *Identity* matrix, all with 2N columns. Then $H^2 \equiv I$ **mod** 2 , so $\det(H) \equiv 1$ **mod** 2 . Consequently $\det(H) \ne 0$, as claimed.

(Adapted from problem B5 of the 1988 Putnam Exam; *cf.* L-S. Hahn (1992) *Math. Magazine* **65** #2 pp. 111-2.)

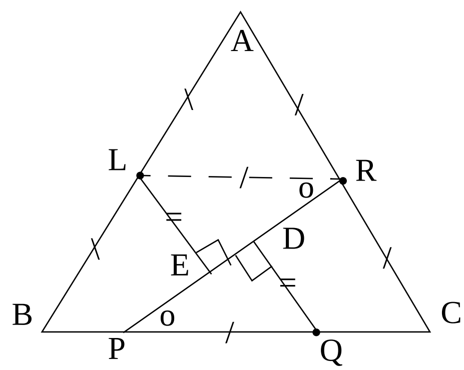
Problem 5: Here is what is known about a linear operator **L** that maps a vector space to itself: No matrix L , that represents **L** in any basis for the space, can have 0 as its first diagonal element, but this element is 3 in at least one such L . What operator is **L** ? Justify your answer.

Solution 5: We shall see why $\mathbf{L} = 3\mathbf{I}$ where **I** is the identity operator. Evidently $\mathbf{L} \neq \mathbf{O}$, so $\mathbf{w}^T\mathbf{L} \neq \mathbf{o}^T$ for some functional $\mathbf{w}^T \neq \mathbf{o}^T$. Suppose now that a vector **v** existed satisfying both $\mathbf{w}^T\mathbf{v} = 1$ and $\mathbf{w}^T\mathbf{L}\mathbf{v} = 0$. Then a basis $\mathbf{B} = [\mathbf{v}, \mathbf{b}_2, \mathbf{b}_3, \dots]$ could be chosen in which $[\mathbf{b}_2, \mathbf{b}_3, \dots]$ is a basis for the subspace annihilated by $\mathbf{w}^T$ (so that every $\mathbf{w}^T\mathbf{b}_j = 0$), and then $\mathbf{w}^T$ would be the first " row " in the inverse basis $\mathbf{B}^{-1}$, whereupon the matrix $L = \mathbf{B}^{-1}\mathbf{L}\mathbf{B}$ that represents **L** in this basis would have $\mathbf{w}^T\mathbf{L}\mathbf{v} = 0$ for its first diagonal element. This can't happen, according to the problem's statement. Therefore no vector **v** can ever satisfy both $\mathbf{w}^T\mathbf{v} = 1$ and $\mathbf{w}^T\mathbf{L}\mathbf{v} = 0$; consequently *every* $\mathbf{w}^T\mathbf{L} = \mu\mathbf{w}^T$ for some scalar $\mu = \mu(\mathbf{w}^T) \neq 0$. This implies $L = \mathbf{B}^{-1}\mathbf{L}\mathbf{B}$ is diagonal for *every* basis **B** . No two diagonal elements can differ without violating the equation $\mathbf{w}^T\mathbf{L} = \mu\mathbf{w}^T$ when $\mathbf{w}^T$ is the difference between their corresponding " rows " in $\mathbf{B}^{-1}$. This makes L a scalar multiple of the identity matrix I , and therefore **L** is a scalar multiple of the identity operator **I** . The scalar 3 is the only scalar consistent with the problem's data.

There is another way to show that L is diagonal: Choose any vector $\mathbf{u} \neq \mathbf{o}$ and set $\mathbf{v} := \mathbf{L}\mathbf{u}$. If **v** were not a scalar multiple of **u** then **u** and **v** would be linearly independent and could be embedded in a basis $\mathbf{B} = [\mathbf{u}, \mathbf{v}, \mathbf{b}_3, \mathbf{b}_4, \dots]$ for the vector space; then the matrix $L = \mathbf{B}^{-1}\mathbf{L}\mathbf{B}$ that represents **L** in this basis would have 0 for its first diagonal element because the first "column" of $\mathbf{B}L = \mathbf{L}\mathbf{B}$ would be $\mathbf{v} = \mathbf{L}\mathbf{u}$, making $[0, 1, 0, 0, 0, \dots, 0]^T$ the first column of L . Since the problem ruled that out, $\mathbf{L}\mathbf{u} = \mu(\mathbf{u})\mathbf{u}$ for every vector u in the space, *etc.* as before.

Problem 6: You are given a large flat layer-cake in the shape of an equilateral triangle, but you wanted a square flat layer-cake. How few straight knife-strokes suffice to cut the triangle into pieces that can be reassembled into a square of the same area? After reassembly the cake's icing can be applied over the cuts and sides of the cake to conceal that it was not square originally. But reassembly must turn no piece over lest mismatched layers reveal that the cake was manipulated.

Solution 6: The minimum number is 2 or 3 . One cut is too few since it creates at most two right-angled corners unless one piece is flipped over like a pancake to produce a rectangle. Here is a way for three cuts to generate four pieces that can be reassembled into a square: The triangle ABC shown below has edges each of length 2 units, and area $\sqrt{3}$ square units. Points L and R are midpoints of AB and AC respectively. First cut through R to P on BC at an angle $\angle LRP = \angle RPC = \arcsin(\sqrt[4]{3}/2) \approx 42.150335^\circ$. Lay point Q on PC at one unit distance from P and put the second cut through Q perpendicular to PR, cutting it at point D . Put the third cut through L perpendicular to PR cutting it at point E . Cuts LE and QD have the same length, $\sqrt[4]{3}/2$ units. With point L fixed, rotate quadrilateral LBPE clockwise through a half-turn (180˚) moving B onto A . With R fixed, rotate triangle RPC counterclockwise through a half-turn moving C onto A . Then, with the moved point Q fixed, turn the moved triangle QDP counterclockwise through a half-turn putting the moved triangle's point P to the moved quadrilateral's point P ; do you see why points P rejoin? The resulting figure is a rectangle two of whose opposite sides have lengths twice LE , the same as twice QD , namely $\sqrt[4]{3}$ units. Since the rectangle's area is the same as the original triangle's, namely $\sqrt{3}$ square units, its four sides must have the same length, namely $\sqrt[4]{3}$ units.



Could two cuts work? We think not, but our "proofs" are too complicated to believe. We hope someone reading this will show us a simple proof that two cuts are too few.

Problem 7: Suppose integers m and n satisfy $n \ge m \ge 3$; and suppose n line segments have nonnegative lengths x_j satisfying $(\sum_1^n x_j)^2 \ge (n - (m-2)^2/m)\cdot\sum_1^n x_j^2$. Prove that any subset of m of these segments in any order can be assembled head-to-tail to form a polygon with m edges.

Proof 7: It so happens that m segments can be the edges of a polygon with m sides if and only if no segment's length exceeds the sum of all the $m-1$ others' lengths. When this condition is satisfied by every subset of m lengths $x_j \ge 0$ chosen from a set of n of them, the vector **x** in *Euclidean* n-space whose coordinates are $\mathbf{x} = [x_1, x_2, x_3, \dots, x_n]$ will be deemed *O.K.* Our task is to prove that all nonnegative vectors **x** satisfying the inequality in question are *O.K.* though perhaps some *O.K.* vectors do not satisfy that inequality. (For example, it is violated when $n \ge m = 3$ and n-vector $\mathbf{x} = [1, 1, 1, \dots, 1, 1/3]$ even though this **x** is *O.K.*)

Let **K** be the set of n-vectors **x** that are *O.K.* and let n-vector $\mathbf{u} := [1, 1, 1, \dots, 1]$. This **u** is *O.K.*, so **u** lies in **K** , which lies in the positive orthant of n-space. We have to characterize **K** algebraically, and prove that it contains a cone-shaped region consisting of all nonnegative n-vectors **x** satisfying the inequality in question, namely " $(\mathbf{u}\bullet\mathbf{x})^2 \ge (n - (m-2)^2/m)\cdot\mathbf{x}\bullet\mathbf{x}$ ". To that end let **W** be the set of all n-vectors **w** whose elements are all zeros except for $m-1$ elements equal to $+1$ and one element equal to -1 . These characterize **K** as the set of nonnegative vectors **x** that satisfy $\mathbf{w}\bullet\mathbf{x} \ge 0$ for every **w** in **W** . There are $m\cdot{}^nC_m := n!/((m-1)!\cdot(n-m)!)$ such vectors **w** , each a *Normal* (perpendicular) to a *Facet* (hyperplane) bounding what turns out to be a *Polyhedral Cone* **K** .

Between vector **u** and each facet of **K** is an angle θ , the complement of the angle between **u** and the normal **w** to that facet, so $\sin(\theta) = \mathbf{w}\bullet\mathbf{u}/(\|\mathbf{w}\|\cdot\|\mathbf{u}\|) = (m-2)/\sqrt{m\cdot n}$. Every nonnegative vector **x** making an angle less than θ with **u** lies inside **K** ; do you see why? Consequently,

if $\mathbf{x} \ge \mathbf{o}$ and $\mathbf{u}\bullet\mathbf{x}/(\|\mathbf{u}\|\cdot\|\mathbf{x}\|) = \cos(\angle(\mathbf{u},\mathbf{x})) \ge \cos(\theta) = \sqrt{1 - (m-2)^2/(m\cdot n)}$ then **x** is *O.K.*

Substituting $\|\mathbf{u}\| = \sqrt{n}$ and squaring the last inequality yields the inequality in question.

For rather more about this subject see "Assembling r-gons Out of n Given Segments" by B.V. Dekster, pp. 44-8 of *Mathematics Magazine* **65** #1 (Feb. 1992). About θ see http://www.cs.berkeley.edu/~wkahan/MathH90/Angles.pdf.

Problem 8: For each $n > 1$ exhibit two n-by-n matrices B and C that allow both equations $B \cdot F = F \cdot B$ and $C \cdot F = F \cdot C$ to be satisfied simultaneously only if F is a scalar multiple of the n-by-n identity I, and prove your matrices B and C have the desired property.

Solution 8: One such pair is $B = J$ and $C = J^T$ where J is the n-by-n matrix whose every element is zero except for the $n-1$ elements just above or to the right of the diagonal, and every one of these is 1, so $J^n = O \neq J^{n-1}$. For example, when $n = 6$,

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

The elements $\phi_{i,j}$ of any F satisfying $J \cdot F = F \cdot J$ must satisfy $\phi_{i+1,j} = \phi_{i,j-1}$ for $1 \le i < n$ and $1 < j \le n$, and $\phi_{n,j-1} = \phi_{i+1,1} = 0$. Consequently this F is upper-triangular with all its nonzero elements $\phi_{i,j}$ determined solely by $j - i \ge 0$; in other words, only if $F = \sum_{0 \le k < n} ß_k \cdot J^k$ is some polynomial in J can $J \cdot F = F \cdot J$. But now $J^T \cdot F - F \cdot J^T = \sum_{0<k<n} ß_k \cdot (J^T \cdot J^k - J^k \cdot J^T)$ can vanish only if every $ß_k = 0$ when $0 < k < n$ because then the two nonzero elements of $J^T \cdot J^k - J^k \cdot J^T$ fall in locations different for different indices k. Therefore only $F = ß_0 \cdot I$ can satisfy both $J \cdot F = F \cdot J$ and $J^T \cdot F = F \cdot J^T$ simultaneously.

This problem was adapted from P.M. Gibson's solved on p. 63 of *The Amer. Math. Monthly* **87** #1 (Jan. 1980). Instead of $B = J$, any n-by-n matrix B satisfying $B^n = O \neq B^{n-1}$ will do as well; then $B = E \cdot J \cdot E^{-1}$ for some E from which $C := E \cdot J^T \cdot E^{-1}$ is constructed.

Problem 9: Prove that if $\mathrm{Rank}(A\cdot B - B\cdot A) \le 1$ then matrices A and B have at least one eigenvector in common. (Hard**!**)

Proof 9: The hypothesis about rank implies that $A\cdot B - B\cdot A = \mathbf{c}\cdot\mathbf{d}^T$ for some vectors $\mathbf{c}$ and $\mathbf{d}$, possibly $\mathbf{o}$. Since $A\cdot B - B\cdot A = (A-\alpha\cdot I)\cdot B - B\cdot(A-\alpha\cdot I)$ for every scalar α, we may replace A by $A - \alpha\cdot I$ for any eigenvalue α of A; in other words, no generality is lost by assuming that 0 is one of the eigenvalues of A. Let $\mathbb{X}$ be the vector space upon which A and B operate; it contains $\mathbb{Z} := \mathrm{Kernel}(A) = \mathrm{Nullspace}(A)$ and $\mathbb{R} := A\mathbb{X} = \mathrm{Range}(A)$, neither of which is just $\mathbf{o}$ lest our problem be trivial. Our proof will establish first that either $\mathbb{Z} \supseteq B\cdot\mathbb{Z}$ or else $\mathbb{R} \supseteq B\cdot\mathbb{R}$.

If $\mathbf{c}\cdot\mathbf{d}^T = O$ then $\mathbb{Z} \supseteq B\cdot\mathbb{Z}$ because $A\cdot\mathbb{Z} = \mathbf{o} = O\cdot\mathbb{Z} = (A\cdot B - B\cdot A)\mathbb{Z} = A\cdot(B\cdot\mathbb{Z})$. In this case $\mathbb{Z}$ is an *Invariant Subspace* of B which must therefore contain at least one of its eigenvectors; this is also an eigenvector of A belonging to its eigenvalue 0. Ditto if $\mathbf{c}\cdot\mathbf{d}^T \ne O$ and $\mathbb{Z} \supseteq B\cdot\mathbb{Z}$.

If $\mathbf{c}\cdot\mathbf{d}^T \ne O$ but not $\mathbb{Z} \supseteq B\cdot\mathbb{Z}$, some $\mathbf{z}$ in $\mathbb{Z}$ must satisfy $\mathbf{c}\cdot\mathbf{d}^T\cdot\mathbf{z} = (A\cdot B - B\cdot A)\mathbf{z} = A\cdot B\mathbf{z} \ne \mathbf{o}$, which places $\mathbf{c}$ in $\mathbb{R}$. This implies that $\mathbb{R} \supseteq (A\cdot B - \mathbf{c}\cdot\mathbf{d}^T)\cdot\mathbb{X} = B\cdot A\cdot\mathbb{X} = B\cdot\mathbb{R}$. In other words, $\mathbb{R}$ is an invariant subspace of A and of B, so each of A and B has at least one eigenvector in $\mathbb{R}$. Now replace A and B by their respective *Restrictions* to $\mathbb{R}$, thus reducing our problem about linear operators acting upon $\mathbb{X}$ to the same problem for linear operators acting upon a smaller space $\mathbb{R}$.

What does "Restriction" mean? To simplify its explanation we suppose that the nullspace $\mathbb{Z}$ of A has dimension 1, and we change coordinates to a new basis of $\mathbb{X}$ that begins with a basis of $\mathbb{R}$ and appends one more vector. (It cannot lie in $\mathbb{R}$; it need not lie in $\mathbb{Z}$ and cannot if $\mathbb{R} \supseteq \mathbb{Z}$.) In the new basis, A, B, $\mathbf{c}$, $A\cdot B - B\cdot A$ and vectors $\mathbf{r}$ in $\mathbb{R}$ are represented by matrices respectively

$$\begin{bmatrix} \bar{A} & \mathbf{a} \\ \mathbf{o}^T & 0 \end{bmatrix}, \quad \begin{bmatrix} \bar{B} & \mathbf{b} \\ \mathbf{o}^T & \beta \end{bmatrix}, \quad \begin{bmatrix} \bar{\mathbf{c}} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \bar{A}\cdot\bar{B} - (\bar{B}\cdot\bar{A}) & \bar{A}\cdot\mathbf{b} + \mathbf{a}\cdot\beta - (\bar{B}\cdot\mathbf{a}) \\ \mathbf{o}^T & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \bar{\mathbf{r}} \\ 0 \end{bmatrix}.$$

Restricting A, B, $\mathbf{c}\cdot\mathbf{d}^T$ and $A\cdot B - B\cdot A$ to act only upon vectors $\mathbf{r}$ in $\mathbb{R}$, thus producing only vectors in $\mathbb{R}$, amounts to deleting the last row and column of each of A, B, $\mathbf{c}\cdot\mathbf{d}^T$ and $A\cdot B - B\cdot A$, thus reducing their dimensions.

This restriction process does not increase the rank of $A\cdot B - B\cdot A$, so the reasoning above can be repeated until either the reduced $\mathbb{Z} \supseteq B\cdot\mathbb{Z}$ or else the reduced $\mathbb{R} \supseteq B\cdot\mathbb{R}$ and has dimension 1.

Problem 10: Assume each diagonal element of a square matrix B exceeds in magnitude the sum of the magnitudes of the off-diagonal elements in the diagonal element's row; prove that B is invertible.

Proof 10: This problem's assertion has three proofs, each illuminating a different approach to matrix theory. One proof is explored in a one-page note on *Diagonal Prominence* posted at `<www.cs.berkeley.edu/~wkahan/MathH110/diagprom.pdf>`. A second proof deduces from *Gershgorin's Circle Theorem* that zero cannot be an eigenvalue of B ; see Problem 11. A third proof comes after the following digression about a particular *Matrix Norm*:

For column vectors $\mathbf{x}$ define $\|\mathbf{x}\|$ to be the biggest of the magnitudes of the components of $\mathbf{x}$. That this norm satisfies the minimum requirements for a norm, namely that $\|\mathbf{x}\| > 0$ except for $\|\mathbf{o}\| = 0$, $\|\lambda\cdot\mathbf{x}\| = |\lambda|\cdot\|\mathbf{x}\|$ for every scalar λ , and $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$, is left to the reader to confirm. Distinguish $\|\mathbf{x}\|$ from $|\mathbf{x}|$, which is the column vector whose every element is the magnitude of the corresponding element of $\mathbf{x}$, whence $\|\mathbf{x}\| = \|\,|\mathbf{x}|\,\|$. Similarly let $|E|$ be the matrix whose every element is the magnitude of the corresponding element of matrix E .

Note that $|E|$ here is a matrix, not the scalar det(E) nor a scalar-valued norm of E .

Now let $\|E\| := \max_{\|\mathbf{x}\|=1} \|E\cdot\mathbf{x}\|$ and confirm that $\|E\| = \|\,|E|\cdot\mathbf{u}\,\|$ in which $\mathbf{u}$ is the column vector whose every element is 1 . Thus, $\|E\|$ is the biggest of the row-sums of $|E|$. Then confirm easily that $\|E\| > 0$ except for $\|O\| = 0$, $\|\lambda\cdot E\| = |\lambda|\cdot\|E\|$ for every scalar λ , and $\|E+F\| \le \|E\| + \|F\|$. Finally, $\|E\cdot F\| \le \|E\|\cdot\|F\|$ because $\|E\cdot F\mathbf{x}\| = \big(\|E\cdot(F\cdot\mathbf{x})\|/\|F\cdot\mathbf{x}\|\big)\cdot\|F\cdot\mathbf{x}\| \le \|E\|\cdot\|F\|\cdot\|\mathbf{x}\|$ if $F\cdot\mathbf{x} \ne \mathbf{o}$.

The norms $\|\ldots\|$ above are often written " $\|\ldots\|_\infty$ " to distinguish them from Euclidean and other norms.

Back to the given problem: Let D := Diag(B) be the diagonal matrix obtained from B by setting all its off-diagonal elements to zeros. Then $E := I - D^{-1}\cdot B$ has only zeros on its diagonal, and $\|E\| < 1$ because of the problem's assumption. Consequently the infinite series $\sum_{n\ge0} E^n$ must converge to $(I - E)^{-1}$ and therefore $B^{-1} = D^{-1}\cdot(I - E)^{-1}$ exists, as the problem asserts.

Problem 11: With any n-by-n matrix B whose elements are $ß_{i,j}$ associate n closed disks Δ_k in the complex plane: Δ_k is centered at $ß_{k,k}$ and has radius $\rho_k := \sum_{j\neq k} |ß_{k,j}|$ for k = 1, 2, … and n . Explain why each eigenvalue of B lies in at least one of the disks Δ_k .

Solution 11: Each eigenvalue λ of B has an eigenvector **x** , a nonzero column of elements ξ_j satisfying $B\cdot\mathbf{x} = \lambda\cdot\mathbf{x}$; *i.e.*, $\sum_j ß_{i,j}\cdot\xi_j = \lambda\cdot\xi_i$ for i = 1, 2, … and n . Let ξ_k be the element of **x** with the biggest magnitude; $0 \neq |\xi_k| \geq |\xi_j|$ for every j . Then

$$|\lambda - ß_{k,k}|\cdot|\xi_k| = |(\lambda - ß_{k,k})\cdot\xi_k| = |\sum_{j\neq k} ß_{k,j}\cdot\xi_j| \leq \sum_{j\neq k} |ß_{k,j}|\cdot|\xi_j| \leq (\sum_{j\neq k} |ß_{k,j}|)\cdot|\xi_k| = \rho_k\cdot|\xi_k| ,$$

which puts λ into Δ_k , as claimed, after $|\xi_k|$ is cancelled out.

Actually, each *Connected Component* of $\cup_k \Delta_k$ contains as many eigenvalues of B as the component has disks Δ_k . This is *Gershgorin's Circle Theorem*, and is proved by exploiting the continuity of B's eigenvalues as functions of its elements. Let $B(\tau)$ have the same diagonal elements $ß_{j,j}$ as B has but off-diagonal elements $\tau\cdot ß_{i,j}$. As τ increases from 0 to 1 the eigenvalues of $B(\tau)$ run continuously from the disks' centers $ß_{k,k}$ to the eigenvalues of B . The radius $\rho_k(\tau)$ of disk $\Delta_k(\tau)$ grows from $\rho_k(0) = 0$ to $\rho_k(1) = \rho_k$. As a set of points, the eigenvalues of $B(\tau)$ remain trapped in $\cup_k \Delta_k(\tau)$, and this proves Gershgorin's theorem. It depends upon the continuity of eigenvalues, which is harder to prove because an eigenvalue may fail to be a differentiable function wherever its multiplicity changes. All that is a story for another day.

Problem 12 : The specification of a function is incomplete without a specification of its domain. This problem shows how an infinitesimal change in its domain can change a function *f* utterly. In particular the *Open Interval*]–1, 1[consisting of all real numbers x strictly between ±1 (and often denoted elsewhere by the overworked notation " **(**–1, 1**)** ") differs infinitesimally from the *Closed Interval* **[**–1, 1**]** , which includes its end-points ±1 too.

… Continued …

For any fixed positive constant $\lambda < 1$, and for any real variable z in $]-1, 1[$, let

$$\sigma(z) := \lambda\cdot(1-|z|) > 0 \quad \text{and interval} \quad J(z) := [z-\sigma(z),\ z+\sigma(z)] \subset\ ]-1, 1[.$$

If $f(x)$ is a continuous function that satisfies $\int_{J(z)} f(\xi)\cdot d\xi = 0$ for every z in $]-1, 1[$, must $f(x) \equiv 0$? Why? Your answers will depend upon the domain of f, so suppose each of …

12.0 : The domain of f is $]-1, 1[$. **12.1 :** The domain of f is $[-1, 1]$.

separately and answer appropriately for each of these two cases.

Solution 12 : If the domain of f is the closed interval $[-1, 1]$ then $f(x) \equiv 0$; otherwise f can take nonzero values in its open domain $]-1, 1[$. This problem's solution requires a *Half-Open* subinterval $]-1, X]$ or $[X, 1[$ of the domain to be covered by an infinite collection of end-to-end abutting intervals $J(z_k)$ for suitably chosen centers z_k depending upon variable X in $]-1, 1[$.

Here is how the centers z_k are chosen when X lies in $]-1, 0]$:

$$-1 < z_{k+1} < z_k := (1-\lambda)^{k-1}\cdot(1+X)/(1+\lambda)^k - 1 < 0 \quad \text{for } k = 1, 2, 3, \dots \text{ in turn.}$$

Consequently only the left end-point of

$$J(z_k) = [\, (1-\lambda)^k\cdot(1+X)/(1+\lambda)^k - 1 ,\quad (1-\lambda)^{k-1}\cdot(1+X)/(1+\lambda)^{k-1} - 1\,]$$

overlaps the right end-point of $J(z_{k+1})$. The interiors of different intervals $J(z_k)$ are disjoint, and $\cup_{k\ge1} J(z_k) =]-1, X]$, so $0 = \sum_{k\ge1} \int_{J(z_k)} f(\xi)\cdot d\xi = \int_{-1}^{X} f(\xi)\cdot d\xi$ provided the last integral exists.

Similarly, when X lies in $[0, 1[$ the choices $z_k := 1 - (1-\lambda)^{k-1}\cdot(1-X)/(1+\lambda)^k$ yield both $\cup_{k\ge1} J(z_k) = [X, 1[$ and $0 = \sum_{k\ge1} \int_{J(z_k)} f(\xi)\cdot d\xi = \int_{X}^{1} f(\xi)\cdot d\xi$ provided the last integral exists.

When f is continuous on the closed interval $[-1, 1]$ both of those last integrals exist, and then the derivatives of the equations $\int_{-1}^{X} f(\xi)\cdot d\xi = 0$ and $\int_{X}^{1} f(\xi)\cdot d\xi = 0$ imply that $f(x) \equiv 0$ for all x in $]-1, 1[$ and thus throughout $[-1, 1]$ too by continuity on this domain.

When f is continuous on the open interval $]-1, 1[$ but not on the closed interval $[-1, 1]$, violently rapid and unbounded oscillations of $f(x)$ as x approaches the intervals' ends may allow neither of those last integrals to exist though $\int_{J(z)} f(\xi)\cdot d\xi = 0$ for every z in $]-1, 1[$. This happens to the following example:

This example's f is the derivative $f(x) := F'(x)$ of an odd function $F(x) \equiv -F(-x)$ constructed to be *Continuously Differentiable* at all x in $]-1, 1[$ and satisfy thereon

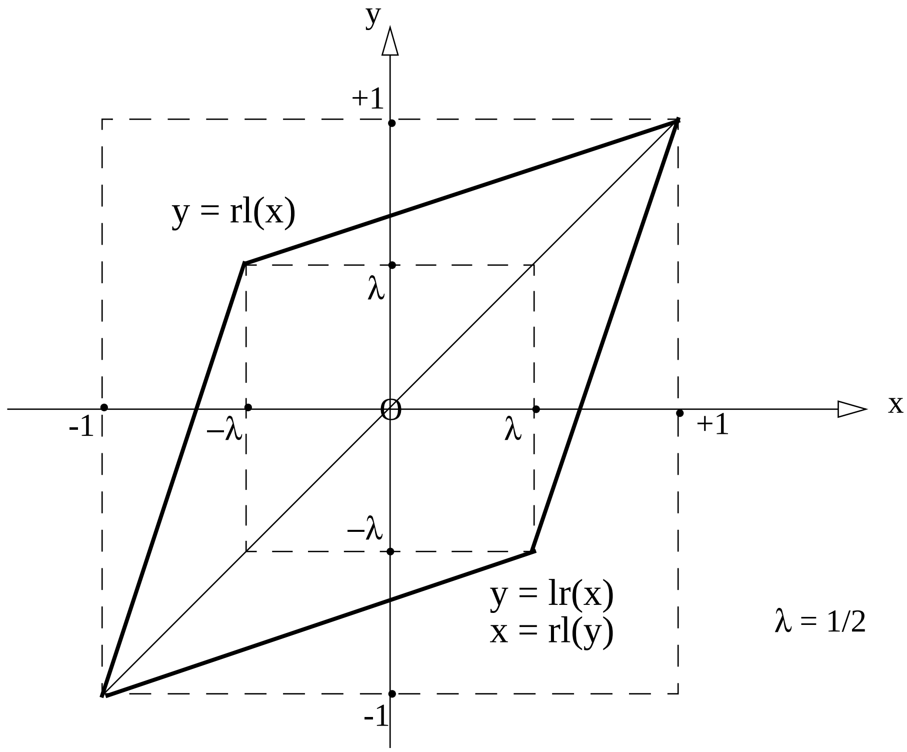
$$\int_{J(x)} f(\xi)\cdot d\xi \;=\; F(x+\lambda\cdot(1-|x|)) - F(x-\lambda\cdot(1-|x|)) \;\equiv 0 \;\text{ but not }\; f(x) \equiv 0 .$$

$F(x)$ will be constructed as a nontrivial solution of the functional equation $F(jr(x)) \equiv F(jl(x))$ in which $jl(x) := x - \lambda\cdot(1-|x|)$ and $jr(x) := x + \lambda\cdot(1-|x|) = -jl(-x)$ are the left- and right-hand end-points of interval $J(x) = [jl(x), jr(x)]$. Functions jl and jr have inverses: If y is in $[-1, 1]$,

$$y = jl(x) \quad \text{just when} \quad x = lj(y) := \big(y + \lambda - \lambda\cdot|y+\lambda|\big)/(1-\lambda^2) ;$$

$$y = jr(x) \quad \text{just when} \quad x = rj(y) := \big(y - \lambda + \lambda\cdot|y-\lambda|\big)/(1-\lambda^2) = -lj(-y) .$$

Then $lr(y) := jl(rj(y)) = \big(y\cdot(\lambda^2+1) - 2\lambda + 2\lambda\cdot|y-\lambda|\big)/(1-\lambda^2)$ and $rl(y) := jr(lj(y)) = -lr(-y)$ turn one functional equation $F(jr(x)) \equiv F(jl(x))$ into two equations $F(y) \equiv F(lr(y))$ and $F(rl(y)) \equiv F(y)$ for F that will amount to the same thing because $F(y) \equiv -F(-y)$. These functions lr and rl are piecewise linear with fixed-points $lr(\pm1) = rl(\pm1) = \pm1$ resp., notches $lr(\lambda) = -\lambda$ and $rl(-\lambda) = \lambda$, and elsewhere derivatives $lr'(y) = (1 + \lambda\cdot signum(y-\lambda))/(1 - \lambda\cdot signum(y-\lambda))$ and $rl'(y) = lr'(-y)$ wherein $signum(z) := z/|z| = \pm1$ except $signum(0) := 0$. And since $lr(rl(x)) \equiv rl(lr(x)) \equiv x$ (after laborious algebraic simplification), functions lr and rl are each the inverse of the other; this fact is obvious from their graphs:



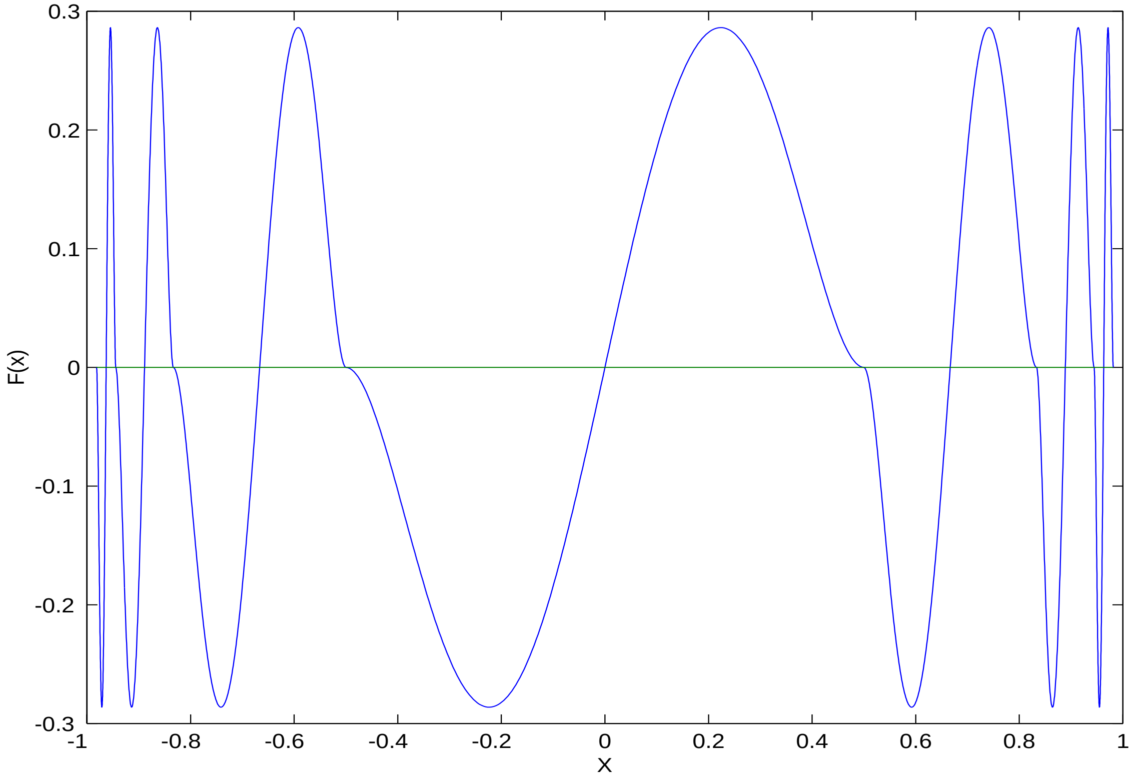
These functions will determine a sequence $\{H_k\}_{k\ge1}$ of abutting subintervals $H_k := [h_{k-1}, h_k]$ that, together with their reflections $-H_k := [-h_k, -h_{k-1}]$, cover all of $]-1, 1[$, and on which F(x) will be defined by a recurrence. The first endpoint $h_0 := -\lambda$ and then they continue with $h_1 = \lambda \le h_k := rl(h_{k-1}) = 1 - (1-\lambda)^k/(1+\lambda)^{k-1} > h_{k-1}$ for $k \ge 1$. Thus every $H_{k+1} = rl(H_k)$.

Define $F(x) := x\cdot(x^2 - \lambda^2)^2/\lambda^5 = -F(-x)$ when x is in $\pm H_1 = [-\lambda, \lambda]$, and $F(x) := F(lr(x))$ when x is in H_{k+1}, so $lr(x)$ is in H_k, for $k = 1, 2, 3, \dots$ in turn. And $F(x) := -F(-x)$ when x is in $-H_{k+1}$. Note that F is continuously differentiable throughout $]-1, 1[$ because $F(\pm h_k) = 0$ and $f(\pm h_k) = F'(\pm h_k) = 0$. Moreover $f(\pm x) = F'(x) = F'(lr(x))\cdot lr'(x) = f(lr(x))\cdot(1+\lambda)/(1-\lambda)$ when x is in H_{k+1}, just as $1 - H_{k+1} = (1 - H_k)\cdot(1-\lambda)/(1+\lambda)$.

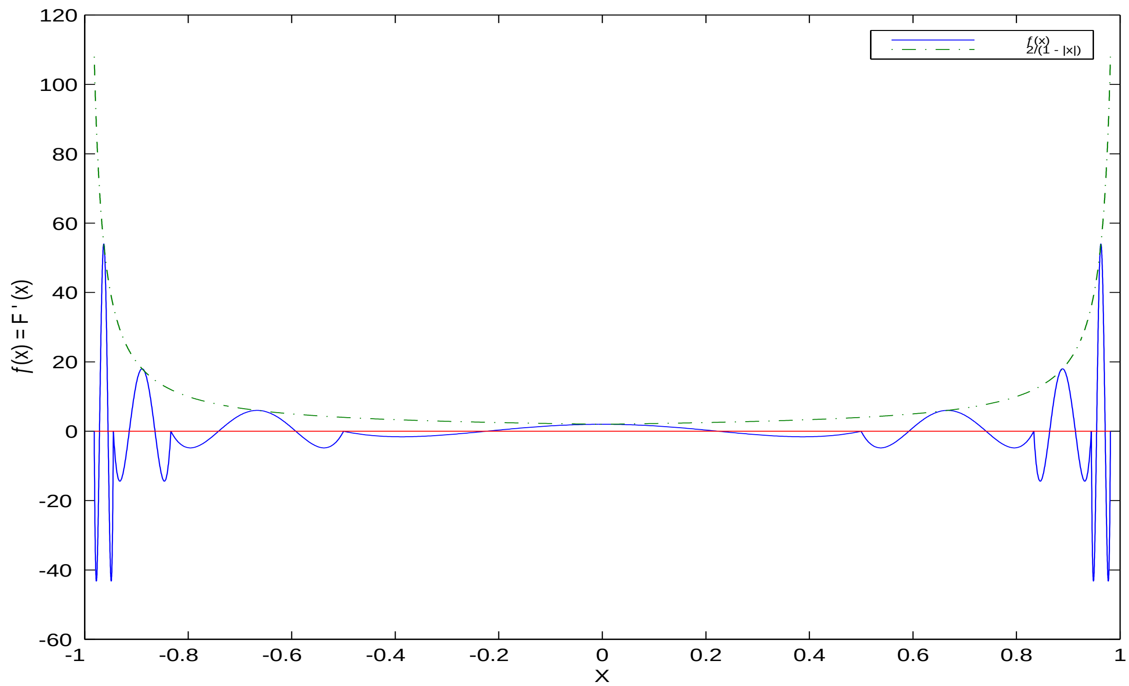
The phrase "Continuously Differentiable" is not quite redundant since derivatives can be only *Darboux Continuous*.

But this F(x) oscillates rapidly between $\pm16/\sqrt{3125}$, and infinitely rapidly as x approaches either end of $]-1, 1[$; then the amplitudes of the oscillations of $f(x) = F'(x)$ tend to infinity as fast as $1/(\lambda\cdot(1-|x|))$. Both F(x) and $f(x)$ are graphed on the next page. Thus neither $\int_{-1}^{X} f(\xi)\cdot d\xi$ nor $\int_{X}^{1} f(\xi)\cdot d\xi$ exists though $\int_{J(z)} f(\xi)\cdot d\xi = 0$ for every z in $]-1, 1[$.

Incomplete Graph of F(x) for $\lambda = 1/2$



Incomplete Graphs of $f(x) = F'(x)$ and $1/(\lambda\cdot(1 - |x|))$ for $\lambda = 1/2$



This problem was adapted from "A Problem of the Pompeiu Type" by K.W. Thompson & T. Schonbek, pp. 32-36 of *The Amer. Math. Monthly* **87** #1 (Jan. 1980). A 2-dimensional generalization appeared on the 1977 Putnam exam.

Problem 13: Assume integers $N > K > 0$, and set column N-vector $\mathbf{n} := [1, 2, 3, \dots, N]^T$. Let P and Q be N-by-N *Permutation Matrices*, so $P\cdot\mathbf{n}$ just rearranges the order of the elements of $\mathbf{n}$. Determine column K-vectors $\mathbf{r}$ and $\mathbf{c}$, and column $(N-K)$-vectors $\bar{\mathbf{r}}$ and $\bar{\mathbf{c}}$, thus: $\begin{bmatrix}\mathbf{r}\\ \bar{\mathbf{r}}\end{bmatrix} := P\cdot\mathbf{n}$ and $\begin{bmatrix}\mathbf{c}\\ \bar{\mathbf{c}}\end{bmatrix} := Q\cdot\mathbf{n}$. Suppose N-by-N matrix B has an inverse $E := B^{-1}$, and define K-by-K submatrix $B_{\mathbf{r},\mathbf{c}}$ by first selecting from B the rows with indices in $\mathbf{r}$ in its order, and then keeping only the columns with indices in $\mathbf{c}$ in its order. For example, take $N := 5$, $K := 2$, $\mathbf{r} := [3, 1]^T$ and $\bar{\mathbf{r}} := [4, 5, 2]^T$, and $\mathbf{c} := [2, 4]^T$ and $\bar{\mathbf{c}} := [5, 1, 3]^T$, so $B_{\mathbf{r},\mathbf{c}} = \begin{bmatrix}\beta_{32} & \beta_{34}\\ \beta_{12} & \beta_{14}\end{bmatrix}$. Do similarly to obtain $B_{\mathbf{r},\bar{\mathbf{c}}}$ and $E_{\bar{\mathbf{c}},\bar{\mathbf{r}}}$ *etc*. Now prove that $\det(B_{\mathbf{r},\mathbf{c}}) = \det(B)\cdot\det(E_{\bar{\mathbf{c}},\bar{\mathbf{r}}})\cdot\det(P)\cdot\det(Q)$.

Proof 13: This formula to be proved was first published in 1834 by Jacobi in a rather more complicated notation. The following short proof partitions matrices thus:

$$\text{The inverse of } P\cdot B\cdot Q^T = \begin{bmatrix}B_{\mathbf{r},\mathbf{c}} & B_{\mathbf{r},\bar{\mathbf{c}}}\\ B_{\bar{\mathbf{r}},\mathbf{c}} & B_{\bar{\mathbf{r}},\bar{\mathbf{c}}}\end{bmatrix} \text{ is } Q\cdot E\cdot P^T = \begin{bmatrix}E_{\mathbf{c},\mathbf{r}} & E_{\mathbf{c},\bar{\mathbf{r}}}\\ E_{\bar{\mathbf{c}},\mathbf{r}} & E_{\bar{\mathbf{c}},\bar{\mathbf{r}}}\end{bmatrix}.$$

Therefore

$\begin{bmatrix}I & O\\ E_{\bar{\mathbf{c}},\mathbf{r}} & E_{\bar{\mathbf{c}},\bar{\mathbf{r}}}\end{bmatrix}\cdot\begin{bmatrix}B_{\mathbf{r},\mathbf{c}} & B_{\mathbf{r},\bar{\mathbf{c}}}\\ B_{\bar{\mathbf{r}},\mathbf{c}} & B_{\bar{\mathbf{r}},\bar{\mathbf{c}}}\end{bmatrix} = \begin{bmatrix}B_{\mathbf{r},\mathbf{c}} & B_{\mathbf{r},\bar{\mathbf{c}}}\\ O & I\end{bmatrix}$, and consequently $\det(E_{\bar{\mathbf{c}},\bar{\mathbf{r}}})\cdot\det(P\cdot B\cdot Q^T) = \det(B_{\mathbf{r},\mathbf{c}})$, whence follows the claimed result. Note that $\det(P) = \pm 1$ according to the parity of P. Likewise for Q.