Problem 1: Prove that no rectangular array of real numbers $ß_{ij}$ can satisfy the inequality $\sqrt{\sum_i (\sum_j ß_{ij})^2} \ge \sum_j \sqrt{\sum_i ß_{ij}^2}$ unless $ß_{ij} = \xi_i \cdot \mu_j$ for some two arrays ξ_i and $\mu_j \ge 0$.

Proof 1: Let column vector $\mathbf{b}_j$ have elements $ß_{ij}$; its *Euclidean* norm is $\|\mathbf{b}_j\| = \sqrt{\sum_i ß_{ij}^2}$. The problem's inequality says that $\|\sum_j \mathbf{b}_j\| \ge \sum_j \|\mathbf{b}_j\|$, whereas this norm's *Triangle Inequality* says that $\|\mathbf{b} + \mathbf{y}\| < \|\mathbf{b}\| + \|\mathbf{y}\|$ for all column vectors $\mathbf{b}$ and $\mathbf{y}$ of the same dimension unless one is a nonnegative scalar multiple of the other, in which case $\|\mathbf{b} + \mathbf{y}\| = \|\mathbf{b}\| + \|\mathbf{y}\|$. Now set $\mathbf{b} := \mathbf{b}_1$ and $\mathbf{y} := \sum_{j>1} \mathbf{b}_j$ to infer first, by induction, that the problem's inequality must actually be an equality $\|\sum_j \mathbf{b}_j\| = \sum_j \|\mathbf{b}_j\|$ and, second, that then either $\mathbf{b}_1 = \mathbf{y} \cdot \mu$ for some scalar $\mu \ge 0$ or else $\mathbf{y} = \mathbf{o}$. In the trivial case that $\mathbf{x} := \sum_j \mathbf{b}_j = \mathbf{o}$ the problem's (in)equality forces every $\mathbf{b}_j = \mathbf{o}$, so disregard this case henceforth. Now $\mathbf{x} = \mathbf{b}_1 + \mathbf{y} \ne \mathbf{o}$ whence $\mathbf{b}_1 = \mathbf{x} \cdot \mu_1$ either for $\mu_1 := \mu/(\mu+1)$ if $\mathbf{y} \ne \mathbf{o}$ or else $\mu_1 := 1$ if $\mathbf{y} = \mathbf{o}$. The same reasoning works for any $\mathbf{b}_j$ in place of $\mathbf{b}_1$ to prove that each $\mathbf{b}_j = \mathbf{x} \cdot \mu_j$ for some $\mu_j \ge 0$, and then the elements of $\mathbf{x}$ provide the problem's ξ_i .

An alternative more computational proof starts from first principles without assuming the triangle inequality. Define $\|\ldots\|$ and $\mathbf{x} := \sum_j \mathbf{b}_j$ as before, so the problem's inequality still says that $\|\mathbf{x}\| = \|\sum_j \mathbf{b}_j\| \ge \sum_j \|\mathbf{b}_j\|$; and suppose further that $\mathbf{x} \ne \mathbf{o}$ since the alternative makes every $\mathbf{b}_j = \mathbf{o}$ trivially. Next define $\mu_j := \mathbf{x}^T \cdot \mathbf{b}_j / \mathbf{x}^T \mathbf{x} = \mathbf{x}^T \cdot \mathbf{b}_j / \|\mathbf{x}\|^2$ and $\mathbf{r}_j := \mathbf{b}_j - \mathbf{x} \cdot \mu_j = (\mathbf{I} - \mathbf{x} \cdot \mathbf{x}^T / \|\mathbf{x}\|^2) \cdot \mathbf{b}_j$, whence follows $\mathbf{x}^T \cdot \mathbf{r}_j = 0$ and then $\|\mathbf{b}_j\|^2 = \|\mathbf{r}_j + \mathbf{x} \cdot \mu_j\|^2 = \|\mathbf{r}_j\|^2 + \mu_j^2 \cdot \|\mathbf{x}\|^2 \ge \mu_j^2 \cdot \|\mathbf{x}\|^2$; therefore every $\|\mathbf{b}_j\| \ge \mu_j \cdot \|\mathbf{x}\|$. But now the sum of all terms $\|\mathbf{b}_j\| - \mu_j \cdot \|\mathbf{x}\|$, all nonnegative, satisfies

$$0 \le \sum_j (\|\mathbf{b}_j\| - \mu_j \cdot \|\mathbf{x}\|) = \sum_j \|\mathbf{b}_j\| - \sum_j \mu_j \cdot \|\mathbf{x}\| \le \|\sum_j \mathbf{b}_j\| - \mathbf{x}^T \cdot \sum_j \mathbf{b}_j / \|\mathbf{x}\| = \|\mathbf{x}\| - \|\mathbf{x}\| = 0 ,$$

whence every $\|\mathbf{b}_j\| = \mu_j \cdot \|\mathbf{x}\|$ and then, because $\mu_j^2 \cdot \|\mathbf{x}\|^2 = \|\mathbf{b}_j\|^2 = \|\mathbf{r}_j\|^2 + \mu_j^2 \cdot \|\mathbf{x}\|^2$ above, every $\mathbf{r}_j = \mathbf{o}$ and thus each $\mathbf{b}_j = \mathbf{x} \cdot \mu_j$ for $\mu_j = \|\mathbf{b}_j\| / \|\mathbf{x}\| \ge 0$, as claimed.

Problem 2: For any real x and $p > 0$ the *Remainder* or *Residue* $r := x \textbf{ mod } p$ is the least nonnegative number for which $(x - r)/p$ is an integer, so $0 \le r < p$. A set $\{x_j\}$ of real numbers x_j is called "Dense" in an interval if its every nonempty *Open* subinterval contains infinitely many members of that set. (The endpoints of an *Open* subinterval are excluded from it.) Prove that the set $\{ \sin(n) \}$ generated as n runs through all integers is dense in the interval between -1 and 1 . (You may take for granted that π is irrational. Still, this classical problem is not easy.)

Proof 2: Problem 2 demands a proof that the set $\{ n \textbf{ mod } 2\pi \}$ generated as n runs through all integers is dense in the interval between 0 and 2π , which $\sin(\ldots)$ maps continuously (twice) onto the interval between -1 and 1 . To simplify notation let the irrational number $\mu := 1/(2\pi)$, and let *Fractional Part* function $f(x) := x \textbf{ mod } 1$ for any real x , so $0 \le f(x) < 1$. Now the set $\{ f(n \cdot \mu) \}$ generated as n runs through all integers will be proved dense between 0 and 1 .

Since μ is irrational, $0 < f(n \cdot \mu) = 1 - f(-n \cdot \mu) < 1$ for every positive integer n . No two values of $f(n \cdot \mu)$ can coincide because, were $f(n \cdot \mu) = f(m \cdot \mu)$ for integers m and $n \ne m$, then $(n - m) \cdot \mu$ would be an integer (do you see why?) and μ would have to be rational. In the closed interval $[0, 1]$ containing the infinite set $\{ ƒ(n·µ) \}$ of points generated as n runs through all positive integers, there must be at least one *Condensation Point* æ . This means $0 \le æ \le 1$ and, for every tiny positive $\Delta < 1/2$, the inequality $|ƒ(n·µ) – æ| < \Delta/2$ is satisfied by infinitely many integers $n > 0$. Let M and $N > M$ be any two of them; then $|ƒ(N·µ) – ƒ(M·µ)| < \Delta$. This inequality implies either $0 < \delta := ƒ((N–M)·µ) < \Delta$ or else $0 < \delta := ƒ((N–M)·(–µ)) < \Delta$; do you see why? Let's treat the latter case since the former is easier to treat. For every positive integer $k < 1/\delta$ the value of $k·\delta = ƒ(k·(N–M)·(–µ)) = 1 – ƒ(k·(N–M)µ)$ falls strictly between 0 and 1 ; do you see why? Inside $[0, 1]$ every open subinterval of width Δ (no matter how tiny, but tinier than $1/2$) contains at least one of the values $k·\delta$ and therefore at least one value $ƒ(n·µ)$ too.

Actually, the fractional parts $ƒ(n·µ)$ are *Distributed Uniformly* in the interval between 0 and 1 in this sense: Let $Ç(N, w)$ count the number of values in the set $\{ ƒ(µ), ƒ(2µ), ƒ(3µ), …, ƒ(N·µ) \}$ that fall into some chosen open subinterval of width w (so $0 < w < 1$) inside that interval. Then, no matter which subinterval has been chosen, $Ç(N, w)/N \to w$ as $N \to +\infty$. For proofs see Ch. XXIII of *An Introduction to the Theory of Numbers* (4th ed.) by G.H. Hardy and E.M. Wright (1960, Oxford Univ. Press) ; or see Ch. 3 of I. Niven's *Diophantine Approximations* (1963, Wiley, New York). (But watch out! They write just " (x) " for " ƒ(x) " .).

Problem 3: What is $\mathrm{Lim}_{n\to\infty} (1 + 1/2)·(1+ 1/4)·(1+ 1/16)·(1+ 1/256)·(…)·(1+ 1/2^{2^n})$? Why?

Solution 3: The limit is 2 . Here is why: Observe that at $n = 2$, for example,

$$(1 + 2^{–1})·(1+ 2^{–2})·(1+ 2^{–4}) = 2^{–7} + 2^{–6} + 2^{–5} + 2^{–4} + 2^{–3} + 2^{–2} + 2^{–1} + 2^{–0} = 2 – 2^{–7} .$$

For the sake of a proof by induction, suppose for some $n \ge 1$ that the product of n factors is

$$p_n := (1 + 2^{–1})·(1+ 2^{–2})·(1+ 2^{–4})·(…)·(1+ 2^{-2^{n-1}}) = 2 – 2^{1-2^n} .$$

Then

$$p_{n+1} = p_n·(1 + 2^{-2^n}) = (2 – 2^{1-2^n})·(1 + 2^{-2^n}) = 2 – 2^{1-2^{n+1}} ,$$

which advances the hypothesis about p_n to p_{n+1} . Now let $n \to \infty$ to get $\mathrm{Lim}\ p_n = 2$.

Problem 4: For a chosen constant $ß > 0$ and any initial $x_0 > 0$ define this sequence $\{x_n\}_{n>0}$:

$x_{n+1} := (x_n + ß/x_n)/2$. for $n = 0, 1, 2, 3, …$ in turn. Does $\mathrm{Lim}_{n\to\infty} x_n$ exist? Why?

Solution 4: Yes, $\mathrm{Lim}_{n\to\infty} x_n = \sqrt{ß}$ and convergence is fast. Why? Confirm these assertions:

$$(x_{n+1} – \sqrt{ß})/(x_{n+1} + \sqrt{ß}) = \big((x_n – \sqrt{ß})/(x_n + \sqrt{ß})\big)^2 = \big((x_0 – \sqrt{ß})/(x_0 + \sqrt{ß})\big)^{2^{n+1}} \to 0 \text{ as } n \to \infty .$$

Until several years ago, all electronic computers obtained $\sqrt{ß}$ by computing x_n starting from a cleverly chosen x_0 up to a small integer n dependent upon the arithmetic's precision. This procedure was devised by Heron of Alexandria in the first century AD. Before that, Egyptian priests accepted x_1 as an adequate approximation. Now that computers' memories are so huge, and division is so much slower than multiplication and addition, square roots of modest precision are computed by interpolation in tables.

Problem 5: Suppose $p(z)$ is a polynomial with real or complex coefficients, and suppose a convex polygon P includes all the zeros, real and complex, of $p(z)$. Show that P also includes all the zeros of the derivative $p'(z)$.
(This problem is intended for students who have learned something about complex variables, perhaps from Math. 185, but the solution is accessible also to students who have not taken that course. All you have to know is that, in a vector space, a *Convex Body* contains the closed line segment joining any two points in this body; and the *Convex Hull* of a set of points is the smallest convex body containing all the set's points.)

Solution 5: Let the factorization $p(x) = \mu\cdot(x - z_1)\cdot(x - z_2)\cdot(x - z_3)\cdot(\ldots)\cdot(x - z_n)$ for a suitable constant $\mu \neq 0$ employ all the zeros z_k of $p(z)$; some of them may be repeated. Then take the *Logarithmic Derivative* $(\log(p(x)))' = p'(x)/p(x) = \sum_k 1/(x - z_k) = \sum_k (\bar{x} - \bar{z}_k)/|x - z_k|^2$ wherein the overbar on $\bar{x}$ denotes the complex conjugate of x, so $x\cdot\bar{x} = |x|^2$. Each zero y of $p'(x)$ that is not also a zero of $p(x)$ satisfies an equation $0 = \sum_k (\bar{y} - \bar{z}_k)/|y - z_k|^2$ whose complex conjugate tells us that $y = (\sum_k z_k/|y - z_k|^2)/(\sum_k 1/|y - z_k|^2)$, which exhibits this zero y of $p'(x)$ as a *Positively Weighted Average* of the zeros z_k of $p(x)$. Consequently (do you see why?) y lies in the *Convex Hull* of the zeros of $p(x)$, which is contained in polygon P.

Problem 5's assertion is known as "Lucas' Theorem".

Problem 6: Let $\lfloor x\rfloor :=$ (the largest integer no larger than x) for any real x. For $0 < y < 1$ and any integer $n > 1$ obtain a much simplified expression for $\sum_{k\geq 1} (\lfloor n^k\cdot y\rfloor \textbf{ mod } n)/n^k$.

Solution 6: The expression simplifies to y. Treat n as the *Radix* for arithmetic; for example, if $y = 3/4$ and $n =$ ten then $y =$ "0.75_{ten}" but if $n =$ two then $y =$ "0.11_{two}". In general,

$$y = \text{“}0.y_1\,y_2\,y_3\ldots y_k\ldots{}_n\text{”} = \sum_{k\geq 1} y_k/n^k \text{ for integers (“digits”) } y_k \text{ each in } 0 \leq y_k \leq n-1 .$$

In the ambiguous case when $y_L < n-1 = y_k$ for every $k > L \geq 1$ we replace this nonterminating expansion by the terminated expansion with a (new y_L) := 1 + (old y_L) and (new y_k) := 0 for all $k > L$ without changing the value of y. (Do you see why?)

Now it is evident that $n^K\cdot y = \sum_{1\leq k\leq K-1} y_k\cdot n^{K-k} + y^K + h$ for some nonnegative fraction $h < 1$, whence $\lfloor n^K\cdot y\rfloor \textbf{ mod } n = y_K$ for each $K \geq 1$, so $\sum_{k\geq 1} (\lfloor n^k\cdot y\rfloor \textbf{ mod } n)/n^k = \sum_{k\geq 1} y_k/n^k = y$.