Putnam 2014

A1 Prove that every nonzero coefficient of the Taylor series of $(1-x+x^2)e^x$ about $x=0$ is a rational number whose numerator (in lowest terms) is either 1 or a prime number.

Solution. We consider the expansion $(1-x+x^2)\sum_{i=0}^{\infty}\frac{1}{i!}x^i$. The coefficient of the constant term is 1 and the coefficient of the x-term is $\frac{1}{2}-1=-\frac{1}{2}$. The conefficient of x^{k+1} for all $k\geq 1$ is given by $\frac{1}{(k+1)!}-\frac{1}{k!}+\frac{1}{(k-1)!}=\frac{1-(k+1)+k(k+1)}{(k+1)!}=\frac{k^2}{(k+1)!}=\frac{k}{(k-1)!(k+1)}$ If k is prime we are done. Now assume that it's not. If k is not a prime power, write $k=ab$ with $1<a,b<k$ and $\gcd(a,b)=1$ (for example, let p to be a prime divisor of k and let r to be the maximum power of p dividing k. Then since k is not a prime power, $p^r<k$ and $k/(p^r)$ is relatively prime to p^r by the maximality of r). Since $a,b<k$, $a|(k-1)!$ and $b|(k-1)!$ and with $\gcd(a,b)=1$, this implies that $k=ab|(k-1)!$. Otherwise, $k=p^r$ for some prime p and $r\geq 2$. Using the formula $v_p((r)!)=\sum_{i=1}^{\infty}\lfloor\frac{r}{p^i}\rfloor$, we have $v_p((k-1)!)=\sum_{i=1}^{\infty}\lfloor\frac{k-1}{p^i}\rfloor=\sum_{i=1}^{\infty}\lfloor\frac{p^r-1}{p^i}\rfloor=p^{r-1}-1+p^{r-2}-1+\cdots+(p-1)\geq 1+1+\cdots+1=r-1$ since $p\geq 2$. Thus $p^{r-1}|(k-1)!$ and so when taking the lowest term the numerator can either be 1 or p.

A2 Let A be the $n\times n$ matrix whose entry in the i-th row and j-th column is

$$\frac{1}{\min(i,j)}$$

for $1\leq i,j\leq n$. Compute $\det(A)$.

Answer. $\frac{1}{n[(n-1)!]^2}$

Solution. We use the well-known matrix identity that row reduction preserves determinant, and we will do row reduce profusely. For brevity, we will denote $f(i)=\frac{1}{i}$ for all $i\geq 1$. Denoting a_{ij} as the i-th row and the j-th column. Then we have $a_{ij}=f(\min(i,j))$.

Now, for each iteration stepped i, denoting row i as r_i and we will do $r_j:=r_j-r_i$ for all $j\geq i$. We show that after the k-th iteration below would be the value for a_{ij}:

- For $i=1$, we have $a_{ij}=f(1)$ as always.
- For $i\leq k$, we have $a_{ij}=0$ for all $j<i$, and $a_{ij}=f(i)-f(i-1)$ for all $j\geq i$.
- For $i>k$, $a_{ij}=0$ for $j\leq k$, and $a_{ij}=f(\min(i,j))-f(k)$ otherwise.

To prove this by inducting on k, the base case is given when all the numbers after the first row are subtracted by the corresponding number in the first row, so for $i>1$, a_{ij} becomes $a_{ij}=f(\min(i,j))-f(1)$, and the condition above is satisfied. Suppose that the conjecture holds after k-th step for some k. At $k+1$-th step, all rows after the $k+1$-th row is subtracted against the corresponding index in $k+1$-th row. The $k+1$-th row is given by the following:

$$\begin{pmatrix}0 & \cdots 0 & f(k+1)-f(k) & \cdots f(k+1)-f(k)\end{pmatrix}$$

where the first k entries are 0. Now after the $k+1$-th iteration, for all $i>k+1$, if $j\leq k$ then a_{ij} becomes $0-0=0$ and if $j>k$ we have a_{ij} becomes $(f(\min(i,j))-f(k))-f(k+1)-f(k)=f(\min(i,j))-f(k+1)$. This entry is 0 if $j=k+1$ since $i>k+1$. For all $i\leq k$ the rows are unaffected by this row reduction, so we still have $a_{ij}=0$ for all $j<i$, and $a_{ij}=f(i)-f(i-1)$ for all $j\geq i$. Thus the claim is proven.

To finish the proof, after $n-1$ iterations, we have, for all $j<i$, $a_{ij}=0$. Thus A is no upper triangular, and the determinant is simply the product of the diagonal entries. We also

have $a_{ii} = 1$ for $i = 1$ and $f(i) - f(i-1)$ for $i \geq 2$. Now $f(i) - f(i-1) = \frac{1}{i} - \frac{1}{i-1} = -\frac{1}{i(i-1)}$. Hence we have

$$\det(A) = \prod_{i=2}^{n} -\frac{1}{i(i-1)} = -(-1)^{n-1} = \frac{1}{n[(n-1)!]^2}$$

A4 Suppose X is a random variable that takes on only nonnegative integer values, with $E[X] = 1$, $E[X^2] = 2$, and $E[X^3] = 5$. (Here $E[Y]$ denotes the expectation of the random variable Y.) Determine the smallest possible value of the probability of the event $X = 0$.

Answer. $\frac{1}{3}$.

Solution. (by Kiran Kedlaya, modified) We let $a_i = P(X = i)$ for each $i \geq 0$ (which means $a_i \geq 0$ for each i). Consider the power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$, and we have $f(1) = 1$. By the problem condition, we also have

- $f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$, so $f'(1) = \sum_{n=0}^{\infty} n a_n = E(X) = 1$
- $f''(x) = \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2}$, so $f''(1) = \sum_{n=0}^{\infty} n(n-1) a_n = \sum_{n=0}^{\infty} n^2 a_n - \sum_{n=0}^{\infty} n a_n = E(X^2) - E(X) = 2 - 1 = 1$
- $f'''(x) = \sum_{n=0}^{\infty} n(n-1)(n-2) a_n x^{n-3}$ so $f'''(1) = \sum_{n=0}^{\infty} (n^3 - 3n^2 + 2n) a_n = E(X^3) - 3E(X^2) + 2E(X) = 5 - 3(2) + 2(1) = 1$

Now we can rearrange $f(x)$ into $f(x) = \sum_{n=0}^{\infty} f^{(n)}(1) \frac{(x-1)^n}{n!}$, i.e. the Taylor's series. We also have, by Taylor's series, $f(x) = f(1) + f'(1)(x-1) + f''(1)\frac{(x-1)^2}{2!} + f'''(1)\frac{(x-1)^3}{3!} + f^{(4)}(c)\frac{(x-1)^4}{4!}$, with c some value in $(1, x)$ or $(x, 1)$ depending whether $x < 1$ or $1 < x$. Thus in particular $a_0 = f(0) = f(1) - f'(1) + \frac{f''(1)}{2} - \frac{f'''(1)}{6} + \frac{f^{(4)}(c)}{24} = 1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{f^{(4)}(c)}{24} = \frac{1}{3} + \frac{f^{(4)}(c)}{24}$ for some $c \in (0, 1)$. We also note that $f^{(4)}(x) = \sum_{n=0}^{\infty} n(n-1)(n-2)(n-3) a_n x^{n-4}$ and for $x \geq 0$ and $n \geq 0$, the quantities $n(n-1)(n-2)(n-3)$, a_n, and x^{n-4} are all nonnegative. Thus $f^{(4)}(x) \geq 0$ for all $x \geq 0$, and in particular $\frac{f^{(4)}(c)}{24} \geq 0$. Thus we have $f(0) = \frac{1}{3} + \frac{f^{(4)}(c)}{24} \geq \frac{1}{3}$, with equality holding when $a_0 = \frac{1}{3}, a_1 = \frac{1}{2}$ and $a_3 = \frac{1}{6}$ and $a_n = 0$ for other n's.

B2 Suppose that f is a function on the interval $[1, 3]$ such that $-1 \leq f(x) \leq 1$ for all x and $\int_1^3 f(x)\,dx = 0$. How large can $\int_1^3 \frac{f(x)}{x}\,dx$ be?

Answer. $\ln \frac{4}{3}$.

Solution. Equality can be attained by taking $f(j) = 1$ for all $1 \leq j < 2$ and $f(j) = -1$ for all $2 \leq j \leq 3$. We show that this is the maximum by the following: if $g(x)$ is defined as $\int_1^x f(y)\,dy$, we have $g(1) = g(3) = 0$. Also since $f(x) \in [-1, 1]$ for all $x \in [1, 3]$, and by Mean value theorem, we have , for every x in the said interval, $g'(c) = f(c) = \frac{g(x) - g(1)}{x - 1}$ for

some constant c in the interval $(1, x)$, so $|\frac{g(x)}{x-1}| \leq 1$. Similarly $|\frac{g(x)}{x-3}| \leq 1$. This means that $g(x) \leq x - 1$ and $g(x) \leq 3 - x$ must hold simultaneously. Using this fact and integrating by parts give:

$$\begin{aligned}
\int_1^3 \frac{f(x)}{x}\, dx &= \frac{g(x)}{x}|_1^3 + \int_1^3 \frac{g(x)}{x^2}\, dx \\
&= (0 - 0) + \int_1^3 \frac{g(x)}{x^2}\, dx \\
&\leq \int_1^2 \frac{x-1}{x^2}\, dx + \int_2^3 \frac{3-x}{x^2}\, dx \\
&= [\ln x + \frac{1}{x}]_1^2 + [-\frac{3}{x} - \ln x]_2^3 \\
&= \ln 2 - \ln 1 + \frac{1}{2} - 1 + \frac{3}{2} - 1 - \ln 3 + \ln 2 \\
&= \ln \frac{4}{3}
\end{aligned}$$

as desired.

B3 Let A be an $m \times n$ matrix with rational entries. Suppose that there are at least $m + n$ distinct prime numbers among the absolute values of the entries of A. Show that the rank of A is at least 2.

Solution. By the theorem of unique prime factorization, if p, q, r, s are prime numbers with $pq = rs$ then $p = r$, $q = s$ or $p = s, q = r$ (so the four numbers cannot be pairwise distinct). The fact that there's at least one prime (and hence nonzero) number in A implies that the rank of A cannot be zero, so we can now assume that the rank of A is 1, which is equivalent to assuming that there exists rational numbers $a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_n$ such that $A_{ij} = a_i b_j$.

Now consider a graph (V, E) with n vertices. and we consider adding coloured edge by the following mechanism: for a row i, if $x_1 < x_2 < \cdots < x_k$ are the all the indices such that A_{ix_j} are among the $m+n$ distinct prime numbers, then we add an edge coloured i between x_j and x_{j+1} for each $1 \leq j \leq k-1$. This means if row i has i_k prime numbers the there will be $i_k - 1$ edges coloured i. Our colouring also ensures that there will be no monochromatic cycle in our graph, and there are at least $\sum_{k=1}^{m}(i_k - 1) = (\sum_{k=1}^{m} i_k) - m = (m + n - m) = n$.

We first see what happens if there are two vertices c_1, c_2 with two edges coloured k_1 and k_2. This means $A_{k_i c_j} = a_{k_i} b_{c_j}$ are all prime numbers for all combinations of $i \in \{1, 2\}$ and $j \in \{1, 2\}$. Notice also that $A_{k_1 c_1} A_{k_2 c_2} = a_{k_1} b_{c_1} a_{k_2} b_{c_2} = a_{k_1} b_{c_2} a_{k_1} b_{c_2} = A_{k_1 c_2} A_{k_2 c_1}$, contradicting that the four prime numbers must be pairwise distinct.

Hence we know that there is at most an edge between two vertices, and since there are exactly n vertices and at least n edges, there exists a cycle comprising at least two different colours (since we have proven that there cannot be a monochromatic cycle above). Let $x_1, x_2, \cdots, x_k$ to be the cycle, with $x_i x_{i+1}$ connected by colour r_i for each $1 \leq i \leq k$. For each i, $A_{r_i x_i}$ and $A_{r_i x_{i+1}}$ are both primes, and let $p_{r_i x_i}, p_{r_i x_{i+1}}$ be the primes. Now $\frac{p_{r_i x_i}}{p_{r_i x_{i+1}}} = \frac{A_{x_i i}}{A_{x_i (i+1)}} = \frac{a_{x_i} b_i}{a_{x_i} b_{i+1}} = \frac{b_i}{b_{i+1}}$ (the fact that both entries are prime, i.e. nonzero, means that we don't have to worry about the validity of division). Thus we have

$$1 = \prod_{i=1}^{k} \frac{b_i}{b_{i+1}} = \prod_{i=1}^{k} \frac{p_{r_i x_i}}{p_{r_i x_{i+1}}}$$

and by the theorem of unique prime factorization, $\prod_{i=1}^{k} p_{ii}$ and $\prod_{i=1}^{k} p_{(i+1)i}$ also implies that $\{p_{r_i x_i} : 1 \leq i \leq k\} = \{p_{r_i x_{i+1}}\} : 1 \leq i \leq k\}$. Since $p_{r_i x_i}$ corresponds to the entry (r_i, x_i)

and $p_{r_i x_{i+1}}$ the entry (r_i, x_{i+1}), and each $x_1, x_2, \cdots, x_k$ assumed to be distinct and each of the $m+n$ primes are distinct, we have $p_{r_i x_{i+1}} = p_{r_{i-1} x_i}$, $r_i = r_{i-1}$, so $r_1 = r_2 = \cdots = r_k$. This also means that the only possibility is all edges of the cycle coloured the same colour r_1, contradiction.

B4 Show that for each positive integer n, all the roots of the polynomial

$$\sum_{k=0}^{n} 2^{k(n-k)} x^k$$

are real numbers.

Solution. Let $f(x)$ be the polynomial, which obviously takes positive values when $x \geq 0$. Consider, now, all x's with $x = -(2^m)$. Then

$$f(x) = \sum_{k=0}^{n} 2^{k(n-k)} x^k = \sum_{k=0}^{n} 2^{k(n-k)} (-2^m)^k = \sum_{k=0}^{n} (-1)^k 2^{k(n+m-k)}$$

We first notice that when k varies, $k(n+m-k)$ takes maximum value when $k = \frac{n+m}{2}$. For this reason, we focus on $m = -n, -n+2, \cdots, n-2, n$, whereby $k(n+m-k) = (\frac{n+m}{2})^2 - (\frac{n+m}{2} - k)^2$, and therefore $f(x) = f(-2^m) = 2^{(\frac{n+m}{2})^2} \sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2}$. Our only interest is the sign of this term, and since the sign of $2^{(\frac{n+m}{2})^2} \sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2}$ is the same as the sign of $\sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2}$, we will focus on the latter.

We isolate the cases $n \leq 2$ first. For $n = 1$ all we have is $x+1$ so $x = -1$ is a solution, obviously. When $n = 2$ we have $x^2 + 2x + 1 = (x+1)^2$, so -1 is a double root. Thus we only deal with $n = 3$ here. We recall that if $a_1, a_2, \cdots, a_k$ are distinct nonnegative numbers then $\sum_{i=1}^{k} 2^{-a_i} < \sum_{i=1}^{\infty} 2^{-i} = 1$. Now we have the following cases to consider:

- Case 1: $m = \pm n$. In the $+n$ case we have

  $$\sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2} = \sum_{k=0}^{n} (-1)^k 2^{-(n-k)^2} = (-1)^n + (-1)^{n-1} 2^{-1} + (-1)^{n-2} 2^{-4} + \cdots + (-1)^0 2^{-n^2}$$

  and by the lemma we had, $|(-1)^{n-1} 2^{-1} + (-1)^{n-2} 2^{-4} + \cdots + (-1)^0 2^{-n^2}| \leq 2^{-1} + 2^{-4} + \cdots + 2^{-n^2} < 1$ so $(-1)^{n-1} 2^{-1} + (-1)^{n-2} 2^{-4} + \cdots + (-1)^0 2^{-n^2} \in (-1, 1)$ which means $\sum_{k=0}^{n} (-1)^k 2^{-(n-k)^2}$ has the same sign as $(-1)^n$. Similarly, when $m = -n$ the expression

  $$\sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2} = \sum_{k=0}^{n} (-1)^k 2^{-k^2}$$

  has the same sign as $(-1)^0 = 1$ (i.e. positive).

- Case 2: now $-n < m < n$ and n has the same parity as n. Then $\sum_{k=0}^{n} (-1)^k 2^{-(\frac{n+m}{2}-k)^2}$ has the following form:

  $$(-1)^0 2^{-(\frac{n+m}{2})^2} + (-1)^1 2^{-(\frac{n+m}{2}-1)^2} + \cdots + (-1)^{(n+m)/2} 2^0 + \cdots + (-1)^n 2^{-(\frac{n+m}{2}-n)^2}$$

W.L.O.G. assume $m \leq 0$; the other case is symmetric to this. We notice that $(\frac{n+m}{2} - k)^2 = (\frac{n+m}{2} - (n+m-k))^2$, and moreover $n+m$ is even so k and $n+m-k$ has the same parity. This means we can group these terms together for $k = 0, 1, \cdots, \frac{n+m}{2} - 1$ to get

$$\sum_{i=0}^{\frac{n+m}{2}-1} ((-1)^i + (-1)^{n+m-i}) 2^{-(\frac{n+m}{2}-i)^2} + (-1)^{\frac{n+m}{2}} + \sum_{i=n+m+1}^{n} (-1)^i 2^{-(\frac{n+m}{2}-i)^2}$$

$$= (-1)^{\frac{n+m}{2}} + 2 \sum_{i=0}^{\frac{n+m}{2}-1} (-1)^i 2^{-(\frac{n+m}{2}-i)^2} + \sum_{i=n+m+1}^{n} (-1)^i 2^{-(\frac{n+m}{2}-i)^2}$$

$$= (-1)^{\frac{n+m}{2}} + 2(-1)^{\frac{n+m}{2}-1} 2^{-1} + 2 \sum_{i=0}^{\frac{n+m}{2}-2} (-1)^i 2^{-(\frac{n+m}{2}-i)^2} + \sum_{i=n+m+1}^{n} (-1)^i 2^{-(\frac{n+m}{2}-i)^2}$$

$$2 \sum_{i=0}^{\frac{n+m}{2}-2} (-1)^i 2^{-(\frac{n+m}{2}-i)^2} + \sum_{i=n+m+1}^{n} (-1)^i 2^{-(\frac{n+m}{2}-i)^2}$$

$$\sum_{i=0}^{\frac{n+m}{2}-2} (-1)^i 2^{-(\frac{n+m}{2}-i)^2+1} + \sum_{i=n+m+1}^{n} (-1)^i 2^{-(\frac{n+m}{2}-i)^2}$$

(basically, the two terms beside $(-1)^{(n+m)/2}$ are $(-1)^{(n+m)/2-1}2^{-1} + (-1)^{(n+m)/2+1}2^{-1}$ and therefore vanishes). We recognize that the exponents $-(\frac{n+m}{2} - i)^2 + 1$ with $i = 0, \cdots, \frac{n+m}{2} - 2$ are different numbers in the range $[-(\frac{n+m}{2})^2 + 1, -3]$ and $-(\frac{n+m}{2} - i)^2$ with $i = n+m+1, n$ are different numbers in the range $[-(\frac{m-n}{2})^2, -(-\frac{m+n}{2} - 1)^2]$ and $-(-\frac{m+n}{2} - 1)^2 < -(\frac{n+m}{2})^2 + 1$ are disjoint, which means together all these exponents represent different negative numbers. Therefore by the lemma above, the sign with follow the dominating one, i.e. $(-1)^{(n+m+2)/2}$, i.e. $(-1)^{(n+m)/2}$. This conclusion will hold for $m > 0$ too.

Summarizing above, we know that when $x = 2^m$ for $m = -n, -n+2, \cdots, n-2, n$, $f(x)$ follows the sign of $(-1)^{(n+m)/2}$. In particular, $(-1)^{(n+m)/2}$ and $(-1)^{(n+m+2)/2}$ have different signs, so there is a root between $(-2^{m+2}, -2^m)$. Considering $m = -n, \cdots, -n+2$ we know that there are roots in the intervals $(-2^n, -2^{n-2}), (-2^{n-2}, \cdots, -2^{n-4}), \cdots, (-2^{-n+2}, -2^{-n})$ which are n disjoint intervals, hence at least n real roots. On the other hand, f is a polynomail with degree n, hence only at n roots in total. Thus all roots are real.