THE UNIVERSITY OF
SYDNEY

Sydney University Mathematical Society Problem Competition 2010

1. For any positive integer n, let $D(n)$ be the number obtained by writing next to each other the usual decimal expressions for $2n$ and for n, in that order. For example, $D(10) = 2010$ and $D(627) = 1254627$. Show that there are infinitely many n for which $D(n)$ is a perfect square.

Solution. By definition, we have $D(n) = (2 \times 10^k + 1)n$, where k is the number of (decimal) digits of n. This will be a perfect square if and only if we have

$$2 \times 10^k + 1 = ma^2, \quad n = mb^2,$$

for some m, a, b, so we try to find such m, a, b. It is easy to see that a has to be at least 7 (a cannot be 1, since $mb^2 = n < 10^k < ma^2$; and if $2 \leq a \leq 6$, then $2 \times 10^k + 1$ can never be divisible by a^2). Taking powers of 10 modulo 49, we find that $10^{19} \equiv 24 \pmod{49}$ and $10^{42} \equiv 1 \pmod{49}$. (The latter is a special case of the Fermat–Euler theorem, because $\phi(49) = 42$.) So if k is any positive integer satisfying $k \equiv 19 \pmod{42}$, then

$$2 \times 10^k + 1 \equiv 2 \times 24 + 1 \equiv 0 \pmod{49},$$

which means that we can define $a = 7$ and $m = \dfrac{2 \times 10^k + 1}{49}$. If we then set $b = 2$ so that $n = 4m$, then

$$10^{k-1} < n = \frac{80}{49} \times 10^{k-1} + \frac{4}{49} < 2 \times 10^{k-1},$$

so n does indeed have k digits (and, incidentally, its first digit is 1). Since there are obviously infinitely many k such that $k \equiv 19 \pmod{42}$, this produces infinitely many n such that $D(n)$ is a perfect square.

2. Start with any nonempty string of (lowercase) letters. Apply the following operation: remove the first letter, and then after every other letter in the string, insert the letter which succeeds that letter in the alphabet, except that you should not insert anything after `z`. For example, the string `fsaazn` becomes `stababzno` after applying this operation, and that becomes `tuabbcabbcznoop` after applying the operation again. Show that, no matter what the initial string is, repeating this operation eventually results in the empty string.

Solution. As one would imagine, this result has nothing to do with the specific 26-letter alphabet that we customarily use. We will in fact prove it for an alphabet of any size, by induction on the size of the alphabet. However, for notational convenience, we continue to let `z` denote the last letter of the alphabet, and write the beginning of the alphabet as `a`, `b`, $\cdots$.

The base case is clear: if the alphabet has only one letter, namely `z`, then we never insert anything, so the operation is just removing the first letter, and if we repeat that we eventually make the string empty. Hence we can assume that our alphabet has at least two letters, and that the result is known for all smaller alphabets.

Now imagine applying the operation repeatedly to a string. Notice that the number of `a`’s in the string never increases, since `a` is never inserted. Consider the portion of the string which comes before the first a. Whatever is happening later in the string, the effect of the operation on this initial portion is just the same as if it were a word in its own right; a word, moreover, which uses the smaller alphabet b, $\cdots$, z. So by the induction hypothesis, this initial portion must eventually become empty. At that point, the first a is the first letter, which is removed in the next application of the operation. So the number of a's in the string eventually does decrease, and must in the long run become zero. At that point, the whole word uses the smaller alphabet b, $\cdots$, z, and by the induction hypothesis it must eventually become empty. This completes the induction step.

This operation (specifically, the number of iterations required to produce the empty word from a string of a's) has been studied by W. Dison and T. Riley in their paper 'Hydra groups'.

3. Define $f(x)$ to be the sum of the series $\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \cdots$, which converges for all real x. Show that for any positive integer m, the following equation holds for all real x:

$$f(x) + f(x + \frac{2\pi}{m}) + f(x + \frac{4\pi}{m}) + f(x + \frac{6\pi}{m}) + \cdots + f(x + \frac{2(m-1)\pi}{m}) = \frac{1}{m} f(mx).$$

Solution. The left-hand side of the desired equation is

$$\begin{aligned}
\sum_{j=0}^{m-1} f(x + \frac{2j\pi}{m}) &= \sum_{j=0}^{m-1} \sum_{n\geq 1} \frac{\sin n(x + \frac{2j\pi}{m})}{n^2} \\
&= \sum_{n\geq 1} \frac{\sum_{j=0}^{m-1} \sin(nx + j\frac{2n\pi}{m})}{n^2} \\
&= \sum_{n\geq 1} \frac{\sin(nx)\sum_{j=0}^{m-1}\cos(j\frac{2n\pi}{m}) + \cos(nx)\sum_{j=0}^{m-1}\sin(j\frac{2n\pi}{m})}{n^2}.
\end{aligned}$$

Now if ζ denotes the complex number $\exp(\mathbf{i}\frac{2n\pi}{m})$, we have $\zeta^m = 1$. Hence

$$\begin{aligned}
\sum_{j=0}^{m-1} \cos(j\frac{2n\pi}{m}) + \mathbf{i}\sin(j\frac{2n\pi}{m}) &= \sum_{j=0}^{m-1} \zeta^j \\
&= \begin{cases} m, & \text{if } \zeta = 1, \\ \frac{\zeta^m - 1}{\zeta - 1}, & \text{if } \zeta \neq 1, \end{cases} \\
&= \begin{cases} m, & \text{if } n = km \text{ for some integer } k, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Thus our expression becomes

$$\sum_{\substack{n\geq 1\\ n=km}} \frac{m\sin(nx)}{n^2} = \sum_{k\geq 1} \frac{m\sin(kmx)}{k^2m^2} = \frac{1}{m} f(mx),$$

as required.

The function $f(x)$, which up to a slight scaling of variables is known as the *Lobachevsky function*, has several interesting properties. It is the imaginary part of $Li_2(e^{ix})$, where Li_2 denotes the dilogarithm function $Li_2(z) = \sum_{n\geq 1} \frac{z^n}{n^2}$. There is an integral formula

$$f(x) = -2\int_0^x \ln|2\sin(\frac{\theta}{2})|\, \mathrm{d}\theta.$$

In particular, $f'(x) = -2\ln|2\sin(\frac{x}{2})|$ whenever x is not an integer multiple of 2π.

4. The *Catalan numbers* are defined by the recursion $c_n = c_0c_{n-1} + c_1c_{n-2} + \cdots + c_{n-1}c_0$, with $c_0 = 1$. Determine the sum of the series $\displaystyle\sum_{n=0}^{\infty} \frac{c_n}{2^{4n}(2n+3)}$.

Solution. It is a well-known fact in analysis that the binomial power series

$$\sum_{n\geq 0} \binom{\frac{1}{2}}{n} x^n = 1 + \sum_{n\geq 1} (-1)^{n-1} \frac{1\cdot 3\cdot 5\cdots(2n-3)}{2^n n!} x^n$$

has radius of convergence 1 and value $\sqrt{1+x}$. (Although we won't need this, Abel's theorem implies that it actually converges uniformly to $\sqrt{1+x}$ on the whole closed interval $[-1,1]$, since $\sum_{n\geq 1} \frac{1\cdot 3\cdot 5\cdots(2n-3)}{2^n n!}$ converges by Raabe's test.) For all $x \in (-1,1)$, we have

$$\begin{aligned} 1 + x &= \left(1 + \sum_{n\geq 0} \binom{\frac{1}{2}}{n+1} x^{n+1}\right)^2 \\ &= 1 + 2\sum_{n\geq 0} \binom{\frac{1}{2}}{n+1} x^{n+1} + \sum_{n\geq 0} \left(\sum_{m=0}^{n-1} \binom{\frac{1}{2}}{m+1}\binom{\frac{1}{2}}{n-m}\right) x^{n+1}, \end{aligned}$$

so $\binom{\frac{1}{2}}{n+1} = -\frac{1}{2}\sum_{m=0}^{n-1} \binom{\frac{1}{2}}{m+1}\binom{\frac{1}{2}}{n-m}$ for all $n \geq 1$. Comparing this with the Catalan recurrence, one can easily prove by induction that

$$\binom{\frac{1}{2}}{n+1} = \frac{(-1)^n}{2^{2n+1}} c_n \text{ for all } n \geq 0,$$

which is equivalent to the well-known formula $c_n = \frac{(2n)!}{(n+1)!n!}$. Hence the binomial power series can be rewritten

$$1 + \sum_{n\geq 0} \frac{(-1)^n c_n}{2^{2n+1}} x^{n+1}.$$

It follows that the power series

$$1 - \sum_{n\geq 0} \frac{c_n}{2^{2n+1}} x^{2n+2}$$

also has radius of convergence 1 and value $\sqrt{1-x^2}$ (again, the convergence is actually uniform on the whole of $[-1,1]$). Applying $\int_0^{1/2}$ term-by-term, we get

$$\begin{aligned} \frac{1}{2} - \sum_{n\geq 0} \frac{c_n}{2^{4n+4}(2n+3)} &= \int_0^{1/2} \sqrt{1-x^2}\,\mathrm{d}x \\ &= \frac{\sin^{-1}\frac{1}{2} + \frac{1}{2}\sqrt{1-(\frac{1}{2})^2}}{2} \\ &= \frac{\pi}{12} + \frac{\sqrt{3}}{8}. \end{aligned}$$

Multiplying through by 16 gives the desired result:

$$\sum_{n=0}^{\infty} \frac{c_n}{2^{4n}(2n+3)} = 8 - 2\sqrt{3} - \frac{4\pi}{3}.$$

5. Fix an integer $n \geq 3$.

a) Construct a subset $S \subseteq \{1, 2, \cdots, n\}$ which is as large as possible such that among any three elements of S, there are two which have no common factor greater than 1.

b) Construct a subset $T \subseteq \{1, 2, \cdots, n\}$ which is as large as possible such that among any three elements of T, there are two which have a common factor greater than 1.

In each part, you must prove that no subset with more elements has the specified property.

Solution. In part (a), the following subset obviously satisfies the required property:

$$S_0 = \{1\} \cup \{p_1, p_1^2\} \cup \{p_2, p_2^2\} \cup \cdots \cup \{p_k, p_k^2\} \cup \{p_{k+1}\} \cup \cdots \cup \{p_\ell\},$$

where $p_1, p_2, p_3, \cdots$ denotes the sequence of prime numbers, $k = \pi(\sqrt{n})$ is maximal such that $p_k^2 \leq n$, and $\ell = \pi(n)$ is maximal such that $p_\ell \leq n$. The size of this subset is $\pi(\sqrt{n}) + \pi(n) + 1$.

We now aim to show that S_0 is as large as possible subject to the constraint in part (a). We in fact show something a bit stronger: any subset S of $\{1, 2, \cdots, n\}$ with the property that no three elements of S have a common prime factor has at most $\pi(\sqrt{n}) + \pi(n) + 1$ elements. For every $s \in S$ with $s \neq 1$, let $f(s)$ denote the smallest prime factor of s. This clearly defines a function $f : S \setminus \{1\} \to \{p_1, \cdots, p_\ell\}$. By the property satisfied by S, $|f^{-1}(p_i)| \leq 2$ for all $i = 1, \cdots, k$. But also $|f^{-1}(p_i)| \leq 1$ for $i = k+1, \cdots, \ell$, since if $p_i^2 > n$ then p_i is the only element of $\{2, 3, \cdots, n\}$ whose smallest prime factor is p_i. Hence $|S| \leq 1 + 2k + (\ell - k) = \pi(\sqrt{n}) + \pi(n) + 1$, as claimed.

In part (b), the following subset obviously satisfies the required property:

$$T_0 = \{1 \leq m \leq n \,|\, m \text{ is divisible by either 2 or 3}\}.$$

The size of this subset is $\lfloor \frac{n}{2} \rfloor + \lfloor \frac{n}{3} \rfloor - \lfloor \frac{n}{6} \rfloor$.

We now aim to show that T_0 is as large as possible subject to the constraint in part (b): in other words, that in any subset $U \subseteq \{1, 2, \cdots, n\}$ which has at least $\lfloor \frac{n}{2} \rfloor + \lfloor \frac{n}{3} \rfloor - \lfloor \frac{n}{6} \rfloor + 1$ elements, there exist three elements which are pairwise coprime. It is easy to check that this is true for $3 \leq n \leq 8$. We can then prove it in general by induction, assuming that the statement is known when n is replaced by $n - 6$. Notice that

$$\lfloor \frac{n-6}{2} \rfloor + \lfloor \frac{n-6}{3} \rfloor - \lfloor \frac{n-6}{6} \rfloor + 1 = \lfloor \frac{n}{2} \rfloor + \lfloor \frac{n}{3} \rfloor - \lfloor \frac{n}{6} \rfloor - 3,$$

so if $U \cap \{1, 2, \cdots, n-6\}$ has at least this many elements, we are done. Otherwise, U must contain at least five elements of $\{n-5, n-4, n-3, n-2, n-1, n\}$. If U contains three consecutive numbers in the pattern odd–even–odd, then we are done, so the only remaining cases to consider are where

$$U \cap \{n-5, n-4, n-3, n-2, n-1, n\} = \begin{cases} \{n-5, n-4, n-3, n-1, n\}, & \text{where } n \text{ is odd, or} \\ \{n-5, n-4, n-2, n-1, n\}, & \text{where } n \text{ is even.} \end{cases}$$

If n is odd, then n and $n - 4$ are coprime (because their difference is 4 which is coprime to both of them). If $n \not\equiv 1$ (modulo 3), then $n - 1$ is coprime to both n and $n - 4$; if $n \equiv 1$ (modulo 3), then $n - 3$ is coprime to both n and $n - 4$. So the case of odd n is finished. If n is even, then $n - 1$ and $n - 5$ are coprime. If $n \not\equiv 2$ (modulo 3), then $n - 2$ is coprime to both $n - 1$ and $n - 5$; if $n \equiv 2$ (modulo 3), then $n - 4$ is coprime to both $n - 1$ and $n - 5$. So the case of even n is finished.

6. The sisters Alice, Bess and Cath need to share a circular pizza which has been divided into $2n$ pieces (each a circular sector having an angle of $\frac{1}{n} \times 180°$ at the centre of the pizza), where n is some integer greater than 1. An allocation of the $2n$ pieces to the three girls is acceptable if:

a) there is some diameter d of the pizza (that is, some line through the centre of the pizza) such that Alice's pieces all lie on the same side of d;

b) there is no diameter e of the pizza such that all the pieces on one side of e go to Bess;

c) every sister gets at least one piece.

Show that there are just as many acceptable allocations in which Cath gets an even number of pieces as there are in which she gets an odd number of pieces.

Solution. Let X denote the set of pieces. Specifying an allocation of the pieces to the three sisters is equivalent to specifying two subsets A, B of X such that $A \subseteq B$: namely, we can let A denote the set of pieces allocated to Alice, and B the set of pieces allocated to Alice and Cath together. In this framework, $X \setminus B$ denotes the set of pieces allocated to Bess, and $B \setminus A$ denotes the set of pieces allocated to Cath. Let $\mathcal{D}$ be the set of subsets of X for which there exists a diameter of the pizza such that all the pieces in the subset lie on the same side of that diameter (we allow the empty set to be a member of $\mathcal{D}$). Then the acceptability conditions can be rewritten:

a) $A \in \mathcal{D}$;

b) $B \notin \mathcal{D}$;

c) $A \neq \emptyset$, $B \neq X$ (and $B \neq A$, which is automatic from (1) and (2)).

Now the fact we are asked to prove is equivalent to the equation

$$\sum_{\substack{A,B \subseteq X \\ A \subseteq B \\ A \in \mathcal{D}, A \neq \emptyset \\ B \notin \mathcal{D}, B \neq X}} (-1)^{|B \setminus A|} = 0,$$

since the left-hand side equals the number of acceptable allocations in which Cath gets an even number of pieces, minus the number of acceptable allocations in which she gets an odd number of pieces. Notice that for a fixed subset $A \subseteq X$, not equal to X, we have

$$\sum_{A \subseteq B \subseteq X} (-1)^{|B \setminus A|} = \sum_{E \subseteq X \setminus A} (-1)^{|E|} = \sum_{k=0}^{|X \setminus A|} \binom{|X \setminus A|}{k} (-1)^k = 0,$$

by the Binomial Theorem, and hence

$$\sum_{A \subseteq B \neq X} (-1)^{|B \setminus A|} = -(-1)^{|X \setminus A|}.$$

Similarly, for a fixed $B \subseteq X$, not equal to $\emptyset$, we have

$$\sum_{\emptyset \neq A \subseteq B} (-1)^{|B \setminus A|} = -(-1)^{|B|}.$$

Another obvious fact is that $X \notin \mathcal{D}$, and that if $A \subseteq B$, it is impossible to have both $A \notin \mathcal{D}$ and $B \in \mathcal{D}$. Hence

$$\begin{aligned}
\sum_{\substack{A,B\subseteq X\\ A\subseteq B\\ A\in\mathcal{D},A\neq\emptyset\\ B\notin\mathcal{D},B\neq X}} (-1)^{|B\setminus A|} &= \sum_{\substack{A,B\subseteq X\\ A\subseteq B\\ A\neq\emptyset\\ B\neq X}} (-1)^{|B\setminus A|} - \sum_{\substack{A,B\subseteq X\\ A\subseteq B\\ A\neq\emptyset\\ B\in\mathcal{D}}} (-1)^{|B\setminus A|} - \sum_{\substack{A,B\subseteq X\\ A\subseteq B\\ A\notin\mathcal{D}\\ B\neq X}} (-1)^{|B\setminus A|}\\
&= -\sum_{\emptyset\neq A\neq X} (-1)^{|X\setminus A|} + \sum_{B\in\mathcal{D},B\neq\emptyset} (-1)^{|B|} + \sum_{A\notin\mathcal{D},A\neq X} (-1)^{|X\setminus A|}\\
&= -\sum_{\emptyset\neq A\neq X} (-1)^{|A|} + \sum_{A\in\mathcal{D},A\neq\emptyset} (-1)^{|A|} + \sum_{A\notin\mathcal{D},A\neq X} (-1)^{|A|}\\
&= 0,
\end{aligned}$$

where in the second-last line we used the fact that $|X|$ is even. Notice that we did not need to use the specific definition of $\mathcal{D}$.

7. Find all real numbers x, y, z, t such that

$$x+y+z+t = x^2+y^2+z^2+t^2 = x^3+y^3+z^3+t^3 = x^4+y^4+z^4+t^4.$$

Solution. There are 16 obvious solutions where each of x, y, z, t is either 0 or 1. These are in fact the only solutions, even if we neglect the constraint on $x+y+z+t$. The reason is that the other two equalities imply

$$\begin{aligned}
x^2(x-1)^2 + y^2(y-1)^2 + z^2(z-1)^2 + t^2(t-1)^2 &= x^4 - 2x^3 + x^2 + \cdots + t^4 - 2t^3 + t^2\\
&= (x^4+y^4+z^4+t^4) - 2(x^3+y^3+z^3+t^3) + (x^2+y^2+z^2+t^2)\\
&= 0,
\end{aligned}$$

from which it is clear that $x, y, z, t \in \{0, 1\}$.

This provides a simpler proof of the $n = 4$ case of Problem 10 in the 2006 SUMS Competition (the same argument would also work for $n > 4$). The connection is as follows. If c denotes the common value of the power sums $x+y+z+t$, $x^2+y^2+z^2+t^2$, etc., then Newton's Formula implies that the elementary symmetric polynomials in x, y, z, t are as follows:

$$\begin{aligned}
x+y+z+t &= c,\\
xy+xz+xt+yz+yt+zt &= \frac{c(c-1)}{2},\\
xyz+xyt+xzt+yzt &= \frac{c(c-1)(c-2)}{6},\\
xyzt &= \frac{c(c-1)(c-2)(c-3)}{24}.
\end{aligned}$$

The conclusion is that the polynomial $(X+x)(X+y)(X+z)(X+t)$ equals

$$X^4 + cX^3 + \frac{c(c-1)}{2}X^2 + \frac{c(c-1)(c-2)}{6}X + \frac{c(c-1)(c-2)(c-3)}{24},$$

and we have shown that the only values of c for which this polynomial factorizes completely into linear factors over the real numbers are $c = 0, 1, 2, 3, 4$.

8. Consider $n \times n$ matrices with entries in the field $\mathbb{Z}_p$ of integers modulo p, where p is some prime number. If M is such a matrix, its *characteristic polynomial* is defined to be $\det(xI - M)$ where I denotes the identity matrix, and we say that M is *unipotent* if its characteristic polynomial equals $(x-1)^n$ (that is, M has a single eigenvalue 1 with multiplicity n). Find, in terms of n and p, the smallest positive integer a such that $M^a = I$ for all unipotent matrices M.

Solution. First, note why the question makes sense: if M is unipotent then it must be invertible, so it belongs to the finite group $GL_n(\mathbb{Z}_p)$ of invertible $n \times n$ matrices with entries in $\mathbb{Z}_p$. Therefore it has a finite order $e(M)$, which is defined to be the smallest exponent such that $M^{e(M)} = I$ (and is bounded above by the size of the group); for general e, $M^e = I$ if and only if e is a multiple of $e(M)$. So the minimal exponent a in the question is well defined, and equals the least common multiple of the orders $e(M)$ of unipotent matrices M.

The key result is the Jordan canonical form theorem, which says that any unipotent matrix M is similar to one which is block-diagonal, where every diagonal block is of the form

$$J_m = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \qquad \text{(an } m \times m \text{ matrix).}$$

If two matrices A and B are similar, say $B = XAX^{-1}$, then $A^k = I$ if and only if $B^k = I$, because $B^k = XA^kX^{-1}$; so $e(A) = e(B)$. Moreover, if B is block-diagonal with diagonal blocks $B_1, B_2, \cdots, B_s$, then B^k is block-diagonal with diagonal blocks $B_1^k, B_2^k, \cdots, B_s^k$, so $B^k = I$ if and only if $B_i^k = I$ for all i, which means that $e(B) = \operatorname{lcm}\{e(B_i)\}$. Therefore the quantity we seek is

$$a = \operatorname{lcm}\{e(J_m) \mid 1 \le m \le n\}.$$

Note that $e(J_1) = 1$, so we can assume henceforth that $m \ge 2$.

It is easy to prove by induction that the kth power J_m^k is an upper-triangular matrix where the entries on each diagonal strip are the same: the entries on the diagonal passing through the $(1, j+1)$ entry are all equal to the binomial coefficient $\binom{k}{j}$, interpreted modulo p. So

$$e(J_m) = \min\left\{ k \ge 1 \;\middle|\; \binom{k}{j} \equiv 0 \text{ (mod } p\text{) for all } 1 \le j \le m-1 \right\}.$$

It is well known that $\binom{k}{j}$ is divisible by p if and only if there is at least one carry when the numbers j and $k-j$ are added in base p. In particular, if k has a nonzero p^s digit in its base p expression, there is obviously no carry when p^s and $k - p^s$ are added, so $\binom{k}{p^s}$ is not divisible by p. So to have the property that $\binom{k}{j}$ is divisible by p for all $1 \le j \le m-1$, we must have that k has zero digits in the positions $p^0, p^1, p^2, \cdots, p^{\lfloor \log_p(m-1) \rfloor}$, so $k \ge p^{\lfloor \log_p(m-1) \rfloor + 1} = p^{\lceil \log_p(m) \rceil}$ (the smallest power of p which is greater than or equal to m). Conversely, if $k = p^{\lceil \log_p(m) \rceil}$, there is a carry when j and $k-j$ are added for all $1 \le j \le m-1$. So

$$e(J_m) = p^{\lceil \log_p(m) \rceil}, \text{ implying that } a = p^{\lceil \log_p(n) \rceil}.$$

Note that these formulas give the correct answer when $m = 1$ and $n = 1$ also.

9. A *ring* is a set R with two binary operations $(r,s) \mapsto r+s$ and $(r,s) \mapsto rs$, an operation $r \mapsto -r$, and an element $0 \in R$, satisfying the following axioms:

$$\begin{gathered} r+(s+t)=(r+s)+t, \quad r+s=s+r, \quad r+0=r, \quad r+(-r)=0, \\ r(st)=(rs)t, \quad r(s+t)=rs+rt, \quad (r+s)t=rt+st, \end{gathered}$$

for all $r,s,t \in R$. Show that if R is a ring with the property that $r^4=r$ for all $r \in R$, then R is *commutative* in the sense that $rs=sr$ for all $r,s \in R$.

Solution. We will take for granted the elementary consequences of the ring axioms which can be found in any algebra textbook. Here we are concerned with the consequences of the extra assumption that $x^4=x$ for all $x \in R$. We will prove a number of properties of R in succession, leading up to the desired commutativity.

First note that for any $x \in R$ we have $x=x^4=(-x)^4=-x$, so R has characteristic 2, meaning that subtraction is the same as addition.

We use the notation $[x,y]$ for the commutator $xy-yx=xy+yx$; notice that $[x,y]=0$ if and only if x and y commute, and that $[y,x]=[x,y]$ always. We have the identities

$$(x+y)^2=x^2+y^2+[x,y] \text{ and } [x,[x,y]]=[x^2,y],$$

valid for all $x,y \in R$.

Let S denote the subset $\{x \in R \,|\, x^2=x\}$ (which obviously contains 0). We can generate elements of S as follows: for any $x \in R$, we have

$$(x+x^2)^2=x^2+x^4+[x,x^2]=x^2+x, \text{ so } x+x^2 \in S.$$

We will not need it, but it is also easy to see that $x^3 \in S$.

We claim that if $x \in R$ and $y \in S$, then $[x,y]=0$. To see this, note that

$$(xy+yxy)^2=xyxy+yxy^2xy+[xy,yxy]=xyxy+yxyxy+[xy,yxy]=0.$$

Hence $xy+yxy=(xy+yxy)^4=0$, so $xy=yxy$. By an almost identical argument one can show that $yx=yxy$, so $xy=yx$ as required.

It now follows that S is a subring of R. For if $x,y \in S$, then

$$(x+y)^2=x+y+[x,y]=x+y, \quad (xy)^2=xyxy=x(xy)y=xy,$$

so $x+y, xy \in S$.

We now prove that any $x,y \in R$ commute. Since $x+x^2, y+y^2, (x+y)+(x+y)^2 \in S$ and S is a subring,

$$[x,y]=x+x^2+y+y^2+(x+y)+(x+y)^2 \in S.$$

This implies that $[x^2,y]=[x,[x,y]]=0$. But also $[x+x^2,y]=0$ since $x+x^2 \in S$, so $[x,y]=[x+x^2,y]+[x^2,y]=0$ as required.

This result is a very special case of a more general theorem due to Jacobson, which states that if R is a ring in which every element x equals one of its powers x^n for $n>1$ (where n is allowed to depend on x), then R is commutative. See T. Y. Lam, *A First Course in Noncommutative Rings*, Theorem 12.10.

10. For any 2×2 integer matrix $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, define the *integral column space* $C(A)$ of A to be

$$\left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{Z}^2 \;\middle|\; \begin{pmatrix} u \\ v \end{pmatrix} = s\begin{pmatrix} a \\ c \end{pmatrix} + t\begin{pmatrix} b \\ d \end{pmatrix} \text{ for some } s,t \in \mathbb{Z} \right\}.$$

a) Suppose that $A_1, A_2, \cdots, A_m$ are 2×2 integer matrices such that $\det(A_k) \neq 0$ for every k, and the union $C(A_1) \cup C(A_2) \cup \cdots \cup C(A_m)$ is all of $\mathbb{Z}^2$. Show that

$$\frac{1}{|\det(A_1)|} + \frac{1}{|\det(A_2)|} + \cdots + \frac{1}{|\det(A_m)|} \geq 1.$$

b) Show that for any $\epsilon > 0$, there is a sequence $A_1, A_2, A_3, \cdots$ of 2×2 integer matrices such that $\det(A_k) \neq 0$ for every k, the union $C(A_1) \cup C(A_2) \cup C(A_3) \cup \cdots$ is all of $\mathbb{Z}^2$, and

$$\frac{1}{|\det(A_1)|} + \frac{1}{|\det(A_2)|} + \frac{1}{|\det(A_3)|} + \cdots \leq \epsilon.$$

Solution. Note that $C(A)$ is exactly $\{A\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) \mid \left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) \in \mathbb{Z}^2\}$, or in other words the image of the $\mathbb{Z}$-linear map $\mathbb{Z}^2 \to \mathbb{Z}^2$ represented by A. This is a subgroup of the additive group $\mathbb{Z}^2$. If $\det(A) \neq 0$ (i.e. the linear map $\mathbb{Z}^2 \to \mathbb{Z}^2$ represented by A is injective), then the index of the subgroup $C(A)$ is $|\det(A)|$, in the sense that the quotient group $\mathbb{Z}^2/C(A)$ has $|\det(A)|$ elements. To prove this, one can use the standard fact that there are some matrices X and Y with determinant ± 1 such that $XAY = \left(\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right)$ is diagonal; then it is easy to see that $\mathbb{Z}^2/C(A)$ is isomorphic to

$$\mathbb{Z}^2/C(XAY) = \mathbb{Z}^2/(\mathbb{Z}a_1 \oplus \mathbb{Z}a_2) \cong \mathbb{Z}/\mathbb{Z}a_1 \oplus \mathbb{Z}/\mathbb{Z}a_2, \text{ which has } |a_1 a_2| = |\det(A)| \text{ elements.}$$

Thus, in a loose sense, if you choose an element of $\mathbb{Z}^2$ at random, the probability that it lies in $C(A)$ is $\frac{1}{|\det(A)|}$. This is why the result of (a) is to be expected, and the result of (b) is somewhat surprising: every element lies in one of the subsets $C(A_k)$, and yet the sum of the probabilities of its lying in the various $C(A_k)$'s is arbitrarily small.

The subtlety comes from the infiniteness (or more accurately, the non-compactness) of $\mathbb{Z}^2$, which means it is impossible to make strict sense of these probabilities as stated. The best one can do is to make statements about associated finite sets. For example, in the intersection of $\mathbb{Z}^2$ with the square $[-N, N] \times [-N, N]$ for some positive N, the probability that an element lies in $C(A)$ is indeed approximately $\frac{1}{|\det(A)|}$ ("approximately" because of boundary effects, which would become more negligible as N gets larger). Alternatively, in a finite quotient group $\mathbb{Z}^2/H$ where H is a subgroup contained in $C(A)$, the probability that an element lies in $C(A)/H$ is exactly $\frac{1}{|\det(A)|}$, because of the isomorphism $(\mathbb{Z}^2/H)/(C(A)/H) \cong \mathbb{Z}^2/C(A)$.

To prove part (a), we could use either of these two ideas; the latter is easier. Let H be the intersection $C(A_1) \cap C(A_2) \cap \cdots \cap C(A_m)$, which is another subgroup of $\mathbb{Z}^2$. By induction on m, one can easily prove that H has finite index, or more precisely that $|\mathbb{Z}^2/H| \leq |\det(A_1)||\det(A_2)| \cdots |\det(A_m)|$. The assumption that $\mathbb{Z}^2 = C(A_1) \cup C(A_2) \cup \cdots \cup C(A_m)$ implies that $\mathbb{Z}^2/H = C(A_1)/H \cup C(A_2)/H \cup \cdots \cup C(A_m)/H$, so

$$|C(A_1)/H| + |C(A_2)/H| + \cdots + |C(A_m)/H| \geq |\mathbb{Z}^2/H|.$$

Dividing both sides by $|\mathbb{Z}^2/H|$ and using the isomorphism $(\mathbb{Z}^2/H)/(C(A_k)/H) \cong \mathbb{Z}^2/C(A_k)$ mentioned above, we obtain the result. Note that the same argument would work for an infinite sequence of matrices $A_1, A_2, \cdots$ if it happened that the intersection $C(A_1) \cap C(A_2) \cdots$ had finite index.

The key to part (b) is the observation that for every $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) \in \mathbb{Z}^2$, there are matrices A such that $\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) \in C(A)$ for which $|\det(A)|$ is arbitrarily large. If $u = v = 0$ this is obvious. Otherwise, let $x, y \in \mathbb{Z}$ be any integers such that $uy - vx \neq 0$, and set $A = \left(\begin{smallmatrix} u & Nx \\ v & Ny \end{smallmatrix}\right)$ for a positive integer

N; one has $|\det(A)| = N|uy - vx|$ which can be made arbitrarily large by choosing N large enough.

Since $\mathbb{Z}^2$ is countable, we can list its elements as $p_1, p_2, p_3, \cdots$, and choose A_k to be a matrix such that $p_k \in C(A_k)$ and $|\det(A_k)| \geq 2^k M$ where M is some overall positive constant. Then it is obvious that $\mathbb{Z}^2 = C(A_1) \cup C(A_2) \cup C(A_3) \cup \cdots$, and yet

$$\frac{1}{|\det(A_1)|} + \frac{1}{|\det(A_2)|} + \frac{1}{|\det(A_3)|} + \cdots \leq \frac{1}{2M} + \frac{1}{4M} + \frac{1}{8M} + \cdots = \frac{1}{M},$$

where $\frac{1}{M}$ can be made arbitrarily small by choosing M large enough.