*

ISSN 1350-8539

M500 328

0100 0210 0111 1110 1100 2010 2122 0000 0220

2110 0021 1120 0202 2000 2101 0002 1102 2210

1021 2211 1221 2111 2200 0120 0222 2012 2100

2121 2120 0020 2002 1022 1111 0001 1211 2102

2220 2020 1002 0201 2011 1212 1122 2021 2202

0012 1010 0200 1201 2212 0022 1200 2022 2222

1210 0110 1000 2001 1020 0221 0212 0011 1220

1101 2221 0121 0010 2112 1001 1112 2201 0101

0211 1012 1121 1011 1222 0112 0122 1202 0102

The M500 Society and Officers

The M500 Society is a mathematical society for students, staff and friends of the Open University. By publishing M500 and by organizing residential weekends, the Society aims to promote a better understanding of mathematics, its applications and its teaching. Web address: `m500.org.uk`.

The magazine M500 is published by the M500 Society six times a year. It provides a forum for its readers' mathematical interests. Neither the editors nor the Open University necessarily agree with the contents.

The Revision Weekend is a residential Friday to Sunday event providing revision and examination preparation for both undergraduate and postgraduate students. For details, please go to the Society's website.

The Winter Weekend is a residential Friday to Sunday event held each January for mathematical recreation. For details, please go to the Society's website.

Editor – *Tony Forbes*

Advice for authors We welcome contributions to M500 on virtually anything related to mathematics and at any level from trivia to serious research. Please send material for publication to the Editor, above. We prefer an informal style and we usually edit articles for clarity and mathematical presentation. For more information, go to `m500.org.uk/magazine/`.

M500 AGM 2026 The M500 Society AGM will take place on 17th March 2026 at 7:30 pm. The intention is to run the meeting as an online Zoom call and details will be provided 4 weeks prior to the event together with proxy forms for those members unable to attend.

Anyone wishing to nominate themselves as an Officer of the Society should submit their nomination by email to secretary@m500.org.uk no later than 12th February 2026. Anyone wishing to propose a motion for consideration at the AGM should submit the motion no later than 25th February 2026 to secretary@m500.org.uk.

There are opportunities to assist the Committee with communicating the activities of the Society to the student community. If any member is interested in this, they should contact Colin Aldridge at secretary@m500.org.uk.

M500 Revision Weekend 2026 We are planning to run the Revision Weekend in 2026 over 8th – 10th May at Kents Hill Park, Milton Keynes. For further details and an application form please check the website, or e-mail the Revision Weekend Organizer, weekend@m500.org.uk.

Linear programming bounds in the main coding-theory problem

Reinhardt Messerschmidt

A topic that is mentioned but not covered in [1] (the prescribed textbook for the OU course M836 on coding theory) is the use of linear programming bounds in the main coding theory problem. This article will show how a linear programming bound can be used to prove that $A_2(12,3) = 256$. Proofs of $A_2(8,3) = 20$ and $A_2(12,5) = 32$, using similar arguments, are given as examples in [2]. Unless otherwise stated, we will use the notation and terminology in [1].

Krawtchouck polynomials

Suppose n is a positive integer. For every $i, j \in \{0, \dots, n\}$, let

$$K_{ji}^{(n)} \;=\; \sum_{k=0}^{n} (-1)^k \binom{i}{k} \binom{n-i}{j-k},$$

with the following convention: if s is a nonnegative integer and t is an integer such that $t < 0$ or $t > s$, then $\binom{s}{t} = 0$. It is convenient to let the sum range over $k \in \{0, \dots, n\}$, but note that the k-th term is nonzero if and only if $0 \le k \le i$ and $0 \le j-k \le n-i$, that is to say if and only if

$$\max\{0,\ i+j-n\} \;\le\; k \;\le\; \min\{i,\ j\}.$$

The integer $K_{ji}^{(n)}$ is the value at $x = i$ of a j-th degree polynomial $K_j^{(n)}(x)$, the *Krawtchouck polynomial*, which we don't need to define in full for our purposes. We will abbreviate $K_{ji}^{(n)}$ to K_{ji}. We will only be interested in K_{ji} for $j \le n/2$, due to the following symmetry:

Proposition 1. $K_{ji} = (-1)^i K_{n-j,i}$.

Proof. We have

$$\binom{n-i}{(n-j)-k} \;=\; \binom{n-i}{(n-i)-(j+k-i)} \;=\; \binom{n-i}{j+k-i} \;=\; \binom{n-i}{j-(i-k)}$$

and

$$\binom{i}{k} \;=\; \binom{i}{i-k};$$

therefore

$$(-1)^i K_{n-j,i} \;=\; \sum_{k=0}^{n} (-1)^{i+k} \binom{i}{i-k} \binom{n-i}{j-(i-k)}.$$

Substituting $\ell = i - k$,

$$(-1)^i K_{n-j,i} \;=\; \sum_{\ell=i-n}^{i} (-1)^{2i-\ell} \binom{i}{\ell} \binom{n-i}{j-\ell}.$$

Since

$$(-1)^{2i-\ell} \;=\; (-1)^{\ell}, \qquad \ell < 0 \text{ or } \ell > i \implies \binom{i}{\ell} = 0,$$

it follows that

$$(-1)^i K_{n-j,i} \;=\; \sum_{\ell=0}^{n} (-1)^{\ell} \binom{i}{\ell} \binom{n-i}{j-\ell} \;=\; K_{ji}. \qquad \square$$

For every $\boldsymbol{u} \in {F_2}^n$ and $i \in \{0, \ldots, n\}$, let

$$\begin{aligned} W_i(\boldsymbol{u}) \;&=\; \{\boldsymbol{v} \in {F_2}^n : d(\boldsymbol{u}, \boldsymbol{v}) = i\} \;=\; \{\boldsymbol{v} \in {F_2}^n : w(\boldsymbol{u} - \boldsymbol{v}) = i\} \\ W_i \;&=\; W_i(\mathbf{0}) \;=\; \{\boldsymbol{v} \in {F_2}^n : w(\boldsymbol{v}) = i\}. \end{aligned}$$

The next proposition provides a motivation for the definition of K_{ji}.

Proposition 2. *If $\boldsymbol{u} \in W_i$, then*

$$K_{ji} \;=\; \sum_{\boldsymbol{v} \in W_j} (-1)^{\boldsymbol{u} \cdot \boldsymbol{v}}. \tag{1}$$

Proof. For every $k \in \{0, \ldots, n\}$, let

$$W_{jk}(\boldsymbol{u}) \;=\; \{\boldsymbol{v} \in W_j : w(\boldsymbol{u} \cap \boldsymbol{v}) = k\}.$$

If $\boldsymbol{v} \in W_{jk}(\boldsymbol{u})$, then

$$(-1)^{\boldsymbol{u} \cdot \boldsymbol{v}} \;=\; (-1)^{w(\boldsymbol{u} \cap \boldsymbol{v})} \;=\; (-1)^k;$$

therefore

$$\sum_{\boldsymbol{v} \in W_j} (-1)^{\boldsymbol{u} \cdot \boldsymbol{v}} \;=\; \sum_{k=0}^{n} \sum_{\boldsymbol{v} \in W_{jk}(\boldsymbol{u})} (-1)^{\boldsymbol{u} \cdot \boldsymbol{v}} \;=\; \sum_{k=0}^{n} (-1)^k \big|W_{jk}(\boldsymbol{u})\big|.$$

The following procedure generates each element of $W_{jk}(\boldsymbol{u})$ exactly once:

(i) Choose k coordinates from the i coordinates where $\boldsymbol{u}$ is 1. This can be done in $\binom{i}{k}$ ways.

(ii) Choose $j-k$ coordinates from the $n-i$ coordinates where $\boldsymbol{u}$ is 0. This can be done in $\binom{n-i}{j-k}$ ways.

(iii) Form the vector that is 1 in the coordinates chosen in (i)-(ii) and 0 elsewhere.

It follows that

$$\sum_{\boldsymbol{v}\in W_j} (-1)^{\boldsymbol{u}\cdot\boldsymbol{v}} \;=\; \sum_{k=0}^{n} (-1)^k \binom{i}{k}\binom{n-i}{j-k} \;=\; K_{ji}. \qquad \square$$

Distance distribution of a code

For every (n, M, d)-code C and $i \in \{0, \ldots, n\}$, let

$$\sigma_i(C) \;=\; \frac{1}{M}\Big|\big\{(\boldsymbol{u}, \boldsymbol{v}) \in C \times C : d(\boldsymbol{u}, \boldsymbol{v}) = i\big\}\Big|.$$

We will abbreviate $\sigma_i(C)$ to σ_i. The vector $(\sigma_0, \ldots, \sigma_n)$ is the *distance distribution* of C. It has the following basic properties:

$$\sigma_0 \;=\; 1, \qquad \sigma_1 \;=\; \cdots \;=\; \sigma_{d-1} \;=\; 0, \tag{2}$$

$$\sigma_0 + \cdots + \sigma_n \;=\; M. \tag{3}$$

Constraints for linear program

We will only need the following simple result from linear programming:

Proposition 3 (The weak duality theorem)**.** *Suppose $A \in \mathbb{R}^{q\times p}$, $\boldsymbol{b} \in \mathbb{R}^q$, and $\boldsymbol{c} \in \mathbb{R}^p$. If $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^q$ are such that*

$$\boldsymbol{x} \;\geq\; \boldsymbol{0}, \qquad A\boldsymbol{x} \;\geq\; -\boldsymbol{b}, \qquad \boldsymbol{y} \;\geq\; \boldsymbol{0}, \qquad A^T\boldsymbol{y} \;\leq\; -\boldsymbol{c},$$

then

$$\boldsymbol{x}^T\boldsymbol{c} \;\leq\; \boldsymbol{b}^T\boldsymbol{y}.$$

Proof.

$$\boldsymbol{x}^T\boldsymbol{c} \;\leq\; \boldsymbol{x}^T(-A^T\boldsymbol{y}) \;=\; (-A\boldsymbol{x})^T\boldsymbol{y} \;\leq\; \boldsymbol{b}^T\boldsymbol{y}. \qquad \square$$

The next two propositions establish some inequalities that will be combined into a system of the form $A\boldsymbol{x} \geq -\boldsymbol{b}$ in the next section, in preparation for an application of the weak duality theorem.

Proposition 4. *If C is a binary (n, M, d)-code, then*

$$\sum_{i=d}^{n} K_{ji}\sigma_i \ \geq \ -K_{j0}. \tag{4}$$

Proof. We have

$$\sigma_i \ = \ \frac{1}{M}\sum_{\boldsymbol{u}\in C}\ \sum_{\boldsymbol{v}\in C\cap W_i(\boldsymbol{u})} 1;$$

therefore

$$\sum_{i=0}^{n} K_{ji}\sigma_i \ = \ \frac{1}{M}\sum_{i=0}^{n}\sum_{\boldsymbol{u}\in C}\ \sum_{\boldsymbol{v}\in C\cap W_i(\boldsymbol{u})} K_{ji}.$$

Since $\boldsymbol{v} \in W_i(\boldsymbol{u})$ if and only if $\boldsymbol{u}-\boldsymbol{v} \in W_i$, it follows from (1) that

$$\sum_{i=0}^{n} K_{ji}\sigma_i \ = \ \frac{1}{M}\sum_{i=0}^{n}\sum_{\boldsymbol{u}\in C}\ \sum_{\boldsymbol{v}\in C\cap W_i(\boldsymbol{u})}\ \sum_{\boldsymbol{w}\in W_j}(-1)^{(\boldsymbol{u}-\boldsymbol{v})\cdot\boldsymbol{w}}.$$

Note that $(-1)^{(\boldsymbol{u}-\boldsymbol{v})\cdot\boldsymbol{w}} = (-1)^{\boldsymbol{u}\cdot\boldsymbol{w}}(-1)^{\boldsymbol{v}\cdot\boldsymbol{w}}$. Rearranging the order of summation,

$$\sum_{i=0}^{n} K_{ji}\sigma_i \ = \ \frac{1}{M}\sum_{\boldsymbol{w}\in W_j}\sum_{\boldsymbol{u}\in C}(-1)^{\boldsymbol{u}\cdot\boldsymbol{w}}\left(\sum_{i=0}^{n}\ \sum_{\boldsymbol{v}\in C\cap W_i(\boldsymbol{u})}(-1)^{\boldsymbol{v}\cdot\boldsymbol{w}}\right);$$

therefore

$$\sum_{i=0}^{n} K_{ji}\sigma_i \ = \ \frac{1}{M}\sum_{\boldsymbol{w}\in W_j}\left(\sum_{\boldsymbol{u}\in C}(-1)^{\boldsymbol{u}\cdot\boldsymbol{w}}\right)\left(\sum_{\boldsymbol{v}\in C}(-1)^{\boldsymbol{v}\cdot\boldsymbol{w}}\right);$$

therefore

$$\sum_{i=0}^{n} K_{ji}\sigma_i \ = \ \frac{1}{M}\sum_{\boldsymbol{w}\in W_j}\left(\sum_{\boldsymbol{u}\in C}(-1)^{\boldsymbol{u}\cdot\boldsymbol{w}}\right)^2 \ \geq \ 0.$$

The result now follows from (2). □

Proposition 5. *If C is a binary (n, M, d)-code with $d \geq 3$, then*

$$\sigma_{n-1} \;\leq\; 1. \tag{5}$$

Proof. Suppose that the codewords

$$\boldsymbol{u} \;=\; (u_1, \ldots, u_n), \qquad \boldsymbol{v} \;=\; (v_1, \ldots, v_n), \qquad \boldsymbol{w} \;=\; (w_1, \ldots, w_n)$$

are such that $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{u}, \boldsymbol{w}) = n - 1$. Let

$$I \;=\; \{i : u_i = v_i \text{ and } u_i \neq w_i\}, \qquad J \;=\; \{i : u_i \neq v_i \text{ and } u_i = w_i\}.$$

We have

$$|I| \;\leq\; \bigl|\{i : u_i = v_i\}\bigr| \;=\; 1, \qquad |J| \;\leq\; \bigl|\{i : u_i = w_i\}\bigr| \;=\; 1;$$

therefore

$$d(\boldsymbol{v}, \boldsymbol{w}) \;=\; |I| + |J| \;\leq\; 2 \;<\; d;$$

therefore $\boldsymbol{v} = \boldsymbol{w}$. It follows that $|C \cap W_{n-1}(\boldsymbol{u})| \leq 1$ for every $\boldsymbol{u} \in C$; therefore

$$\sigma_{n-1} \;=\; \frac{1}{M} \sum_{\boldsymbol{u} \in C} \bigl|C \cap W_{n-1}(\boldsymbol{u})\bigr| \;\leq\; \frac{1}{M} \sum_{\boldsymbol{u} \in C} 1 \;=\; 1. \qquad \square$$

Main result

We are now ready to prove the main result:

Proposition 6. *$A_2(12, 3) = 256$.*

Proof. Bounded from below by 256: The binary Hamming code with redundancy 4 is a $[15, 11, 3]$-code. It has a 4×15 parity-check matrix

$$H \;=\; \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \end{bmatrix}$$

with the following properties:

(i) its rows are linearly independent,

(ii) any two of its columns are linearly independent,

(iii) the sum of its first two columns is equal to its fifth column, hence these three columns are linearly dependent.

Deleting the last column of H gives a matrix which still satisfies (i)-(iii), and which is therefore the parity-check matrix of a $[14, 10, 3]$-code. Repeating this procedure two more times gives a $[12, 8, 3]$-code, which has $2^8 = 256$ codewords. It follows that $A_2(12, 3) \geq 256$.

Bounded from above by 256: Suppose C' is a binary $(12, M, 3)$-code. Let C be the code obtained from C' by adding an overall parity check; therefore C is a $(13, M, 4)$-code. Since $d(\boldsymbol{u}, \boldsymbol{v})$ is even for every $\boldsymbol{u}, \boldsymbol{v} \in C$, we have

$$\sigma_5 \;=\; \sigma_7 \;=\; \cdots \;=\; \sigma_{13} \;=\; 0. \tag{6}$$

Let

$$A \;=\; \begin{bmatrix} K_{1,4} & K_{1,6} & \cdots & K_{1,12} \\ K_{2,4} & K_{2,6} & \cdots & K_{2,12} \\ \vdots & \vdots & & \vdots \\ K_{6,4} & K_{6,6} & \cdots & K_{6,12} \\ 0 & 0 & \cdots & -1 \end{bmatrix} \;=\; \begin{bmatrix} 5 & 1 & -3 & -7 & -11 \\ 6 & -6 & -2 & 18 & 54 \\ -10 & -6 & 14 & -14 & -154 \\ -29 & 15 & -5 & -25 & 275 \\ -9 & 15 & -25 & 63 & -297 \\ 36 & -20 & 20 & -36 & 132 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

and

$$\boldsymbol{b} \;=\; [K_{1,0},\; K_{2,0},\; \ldots,\; K_{6,0},\; 1]^T \;=\; [13,\; 78,\; 286,\; 715,\; 1287,\; 1716,\; 1]^T$$
$$\boldsymbol{c} \;=\; [1,\; 1,\; 1,\; 1,\; 1]^T, \qquad \boldsymbol{x} \;=\; [\sigma_4,\; \sigma_6,\; \sigma_8,\; \sigma_{10},\; \sigma_{12}]^T.$$

It follows from (2)-(6) that

$$\boldsymbol{x} \;\geq\; 0, \qquad A\boldsymbol{x} \;\geq\; -\boldsymbol{b}, \qquad \boldsymbol{x}^T\boldsymbol{c} \;=\; M - 1.$$

Now observe that the vector

$$\boldsymbol{y} \;=\; (1/6)[3,\; 3,\; 1,\; 1,\; 0,\; 0,\; 256]^T$$

satisfies

$$\boldsymbol{y} \;\geq\; 0, \qquad A^T\boldsymbol{y} \;\leq\; -\boldsymbol{c}, \qquad \boldsymbol{b}^T\boldsymbol{y} \;=\; 255. \tag{7}$$

It follows from the weak duality theorem that $M \leq 256$. Since this holds for every binary $(12, M, 3)$-code, we have $A_2(12, 3) \leq 256$. $\square$

The vector $\boldsymbol{y}$ in the above proof can be found by solving the linear program

$$\text{minimize } \boldsymbol{b}^T\boldsymbol{y}, \text{ subject to } \boldsymbol{y} \geq 0 \text{ and } A^T\boldsymbol{y} \leq -\boldsymbol{c}$$

with the `linprog` function in the Python package `scipy.optimize`. For the purposes of the proof, we only need to verify that $\boldsymbol{y}$ satisfies (7), we don't really need to know where it came from.

We can quickly establish two more values of $A_2(\cdot, 3)$ from the main result:

Proposition 7. *$A_2(13,3) = 512$ and $A_2(14,3) = 1024$.*

Proof. We constructed a binary $[13,9,3]$-code in the proof of the previous proposition; therefore

$$A_2(13,3) \;\geq\; 2^9 \;=\; 512.$$

By exercise 2.2 in [1],

$$A_2(13,3) \;\leq\; 2A_2(12,3) \;=\; 2(256) \;=\; 512;$$

therefore $A_2(13,3) = 512$. In the same way, $A_2(14,3) = 1024$. □

References

[1] R. Hill, *A First Course in Coding Theory*, Oxford University Press, Oxford, 1986.

[2] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Publishing Company, Amsterdam, New York, Oxford, 1977.

Problem 328.1 – Product and sum

David Sixsmith

Suppose n is a positive integer. How many ways are there to choose the product abc where a, b, c are natural numbers that sum to n?

Solitaire SET

Tamsin Forbes and Tony Forbes

The SET game

SET, marketed by Set Enterprises, Inc., is a card game for 1 or more players. We give here a brief description. Alternatively, details are available in [1], or you can acquire the game and read the instructions.

The are 81 cards with various colourful designs on them. They represent points (a, b, c, d), $a, b, c, d \in \{0, 1, 2\}$, in AG$(4, 3)$, the affine geometry of dimension 4 over $\mathbb{Z}_3$. A line, which the game's inventor calls a *SET*, consists of three cards that sum to $(0, 0, 0, 0)$ modulo 3. For notational convenience, we follow the convention in [1]: SET is the game; a *SET* is a line in AG$(4, 3)$.

The general idea is that in a competitive environment players attempt to identify *SET*s. The rules are simple.

Play begins with the dealer dealing 12 cards from the deck face-up on to the table. Thereafter, players compete with each other to find *SET*s. There is no taking of turns. If you think you have identified a *SET*, you call "Set!" and you are then obliged—with a suitable forfeit if you fail—to remove a *SET* from the table, which if necessary is replenished by the dealer to bring the number back up to 12. If there is no *SET* on the table, the dealer deals three more cards if possible. The following array shows a typical state of the game.

(2,1,0,2)	$(1,2,0,1)$	$(0,1,2,1)$	$(2,2,0,2)$	$(2,2,1,2)$	$(2,0,1,1)$
$(0,0,1,2)$	**(1,0,2,1)**	**(0,2,1,0)**	$(0,2,0,0)$	$(1,0,1,0)$	$(1,1,1,2)$

There are eight *SET*s on the table, one which is highlighted. See if you can identify the other seven.

A game ends when there are no more cards to deal and no more *SET*s to take. The winner is the person with the most *SET*s.

Our encoding of the SET cards is the same as in the book, [1, Table 8.1]. A card is represented by the vector (number, colour, filling, shape) with the elements coded as:

number:	$3 \to 0$,	$1 \to 1$,	$2 \to 2$;
colour:	green $\to 0$,	purple $\to 1$,	red $\to 2$;
filling:	empty $\to 0$,	hatched $\to 1$,	solid $\to 2$;
shape:	diamond $\to 0$,	oval $\to 1$,	squiggle $\to 2$.

For instance, $(0, 1, 1, 0)$ is the card depicting 3 purple hatched diamonds.

The solitaire game

We are interested in the solitaire version of SET. There is one player, yourself. Your aim is to achieve a perfect score, 27 *SET*s; that is, the game ends with no cards left on the table—all clear. The front cover shows such a deal, where at each stage while there are undealt cards, if there are two or more *SET*s on the table, you take the first in the ordering defined by (1) on page 10. However, there is no guarantee that a game can finish all clear. For an example of one that cannot, see [1, Section 10.7].

At any stage of a solitaire game the way ahead is clear if the table is *SET*-free or contains just one *SET*. The fundamental problem is how to make good choices on those occasions where there are two or more *SET*s on the table.

According to the book, a run of 100000000 games where *SET*s are chosen at random produced a mere 1.22 per cent all clear, [1, Table 10.4]. We claim that we can increase this hit rate to something better than 10 per cent. Before explaining our algorithm, we must be very precise about the way the game is played. The rules are as in the usual many-player version, but we emphasise the following clarifications.

(i) At each stage of the game, you must either *prove* there is no *SET* on the table,[1] or take exactly one *SET* from the table, which is immediately replenished to 12 cards if necessary and possible. Passing is forbidden if the table contains a *SET*. This rule is vital. Without it the solitaire game becomes pointless. Just deal all 81 cards (pretending not to notice any *SET*s) and then partition them into 27 *SET*s.

(ii) Obviously you can see[2] the cards on the table.

(iii) You can see all the *SET*s taken from the table.

(iv) From (ii) and (iii), you can deduce the undealt cards. However, you are not allowed to know how they are ordered—that would be cheating.

(v) You can see all possible *SET*s formed from the cards on the table.

(vi) You can compute all possible partitions into *SET*s of the undealt cards together with the cards on the table.

(vii) You must make your choice within a finite time. We suggest 3 years for a human, 3 billion clock-ticks for an electronic computer.

---

[1] *The table* refers to the playing area on to which the cards are dealt. To avoid possible confusion, you should place all *SET*s taken and all undealt cards on a grand piano.

[2] Here and elsewhere, *see* means determine the points in AG$(4, 3)$ that are represented by the cards. (This might not be possible if the cards in your view are face down.)

Definitions

To avoid random processes in our proposed algorithm, we need to specify how to order cards and *SET*s. If $x = (x_1, x_2, x_3, x_4)$, $x_1, x_2, x_3, x_4 \in \{0, 1, 2\}$, is a card in its vector form, we define

$$N(x) \;=\; N((x_1, x_2, x_3, x_4)) \;=\; 27x_1 + 9x_2 + 3x_3 + x_4.$$

Thus N maps the SET deck on to $\{0, 1, \ldots, 80\}$ and defines an ordering: if x and y are cards, then

$$x < y \;\text{ if }\; N(x) < N(y).$$

The extension to *SET*s is as expected. If $\{a, b, c\}$, $a < b < c$, and $\{d, e, f\}$, $d < e < f$, are distinct *SET*s, then

$$\{a, b, c\} < \{d, e, f\} \;\text{ if }\; a < d \text{ or } a = d \text{ and } b < e. \tag{1}$$

If x and y are distinct cards, we denote by $x * y$ the unique card that makes a *SET*, $\{x, y, x * y\}$. Also we say that

the pair $\{x, y\}$ is *dead* if $x * y$ is amongst the *SET*s taken;

the pair $\{x, y\}$ is *live* if $x * y$ is in the undealt part of the deck.

Be aware that not live is not the same as dead.

The basic strategy

Proceed as usual if there is no *SET* or exactly one *SET* on the table. However, when there are two or more *SET*s on the table you must make a choice.

Let $\mathcal{X}$ be the collection of all candidate *SET*s that would maximize the number of live pairs on the table if the *SET* were taken.

If $\mathcal{X}$ consists of just one *SET*, choose it. Otherwise select from $\mathcal{X}$ those *SET*s that would minimize the number of dead pairs on the table if the *SET* were taken. If a tie-break is still needed, choose the first *SET* according to the ordering defined by (1). Remove the chosen *SET*.

The idea is that live pairs are good and worth preserving for possible future *SET*s. On the other hand, dead pairs are bad and we should try to get rid of as many as possible when we choose our *SET*.

The general strategy $S(d)$, $d \geq 0$

We find that the basic strategy works well; the proportion of games ending all clear is increased from 1.22% to more than 5.0%. However, it turns out that we can do significantly better.

For the general strategy, proceed with the basic strategy until there are d undealt cards, where d is a not too large multiple of 3. Upon arriving at this stage in the game we may assume that either there are at least 12 cards on the table or there are no undealt cards.

Construct all partitions into *SET*s of the cards on the table together with the undealt cards.

If there are no partitions, the game must end not all clear.

Otherwise we argue as follows.

If $d = 0$, then the table is neatly partitioned into *SET*s, all available for taking. The game ends all clear.

Next, assume $d = 3$. Then there is a partition of at least 15 cards (3 undealt and at least 12 on the table) into at least 5 *SET*s. Of these, at least 2 are on the table since the *SET*s are mutually disjoint and therefore at most 3 can include undealt cards. Removing one of the *SET*s from the table reduces the situation to $d = 0$, above, and the game ends all clear.

Finally, assume $d \geq 6$.

Suppose one of the partitions, $\mathcal{P}$ say, has $d/3 - 1$ of its *SET*s on the table. Now it is possible to choose a *SET* from $\mathcal{P}$ at the next $d/3 - 1$ turns to leave 3 undealt cards and at least 12 cards on the table. Then case $d = 3$ applies, and the game ends all clear.

On the other hand, if there is no such partition $\mathcal{P}$, we do the basic strategy followed immediately by general strategy $S(d - 3)$.

Short cuts and refinements

In each of the following cases we assume there exists a partition into *SET*s of the undealt cards together with the cards on the table.

(i) If there are 0 or 3 undealt cards, the game ends all clear—as already explained.

(ii) If there are 6 undealt cards and at least 15 cards on the table, the game ends all clear. The partition contains at least 7 *SET*s of which one will be on the table since at most 6 can contain undealt cards.

(iii) When there are 6 undealt cards and the table is *SET*-free, the game ends all clear. Since 3 cards must be dealt, the undealt card count is reduced to 3 for (i) to apply.

(iv) When there are 9 undealt cards and the table is *SET*-free, the game ends all clear. Since 3 cards must be dealt, the undealt card count goes to 6 and the table card count goes to at least 15. Now (ii) applies.

Results

The following shows results for general strategy $S(d)$ augmented with (i)–(iv) of Short cuts and refinements.

d	games	all clear	%
0	200000	14082	7.0
3	200000	14948	7.5
6	200000	18394	9.2
9	500000	52793	10.6
12	500000	53196	10.6

The recommended strategy is $S(9)$ for an all-clear rate of about 10.6 per cent. (Recall that choosing *SET*s at random achieves only about 1.22 per cent.) We consider the apparent small improvement offered by $S(12)$ to be insignificant. Moreover, we find that the cost of implementing $S(d)$ increases considerably when d goes from 9 to 12.

Perhaps one of the UK's new AI data centres might like to verify our conclusions for 100000000 games, say, surely a more useful application of computer resources than writing essays for lazy students.

Reference

[1] Liz McMahon, Gary Gordon, Hannah Gordon and Rebecca Gordon, *The Joy of SET*, Princeton University Press, 2017; reviewed in M500 **276**.

Problem 328.2 – Powers of 3

Tony Forbes

There is overwhelming experimental evidence that large powers of 3 do not like being too close to products of 2 consecutive integers. This observation suggests a problem. The equation

$$3^a - b(b-1) \;=\; c, \qquad 0 \le c \le \sqrt{b}$$

has three solutions in non-negative integers:

$$(a,b,c) \;=\; (0,1,1),\; (1,2,1),\; (5,16,3).$$

Either prove that there are no more solutions, or find another. If that looks too difficult, we would nevertheless be interested in restricting c to specific constants, $c = 0,1,2,3$, say, especially $c = 3$, which is relevant to Problem 188.4 – Sixteen tarts.

Solution 261.2 – Trigonometric double integral

Show that

$$\int_0^{\pi/2}\int_0^{\pi/2} \frac{\sin^2\phi\,d\phi\,d\theta}{(2-\sin^2\theta)^{1/2}(2-\sin^2\phi)^{3/2}} \;=\; \frac{\pi}{4}.$$

J. M. Selig

In M500 **327** Peter Fletcher uses Claude, an AI software, to find the Taylor expansion of the integrands. He then integrates these term-by-term, again using AI. Numerically summing a large number of terms then gives a solution very close to $\pi/4$. Using AI in mathematics is an interesting development, one gaining distinguished support, see [1]. However, here I would like to use more traditional methods.

We begin with the same first step, splitting the double integral into two ordinary integrals:

$$I \;=\; \int_0^{\pi/2} \frac{d\theta}{(2-\sin^2\theta)^{1/2}} \cdot \int_0^{\pi/2} \frac{\sin^2\phi\,d\phi}{(2-\sin^2\phi)^{3/2}}.$$

Let's treat these separately, so put

$$I_1 \;=\; \int_0^{\pi/2} \frac{d\theta}{(2-\sin^2\theta)^{1/2}} \quad\text{and}\quad I_2 \;=\; \int_0^{\pi/2} \frac{\sin^2\phi\,d\phi}{(2-\sin^2\phi)^{3/2}}.$$

Dividing I_1 by $2^{1/2}$ and I_2 by $2^{3/2}$ gives

$$I_1 \;=\; \frac{1}{2^{1/2}}\int_0^{\pi/2} \frac{d\theta}{(1-\frac{1}{2}\sin^2\theta)^{1/2}} \quad\text{and}\quad I_2 \;=\; \frac{1}{2^{3/2}}\int_0^{\pi/2} \frac{\sin^2\phi\,d\phi}{(1-\frac{1}{2}\sin^2\phi)^{3/2}}.$$

The first integral here can be recognised as a complete elliptic integral of the first kind, $I_1 = (1/\sqrt{2})K(1/\sqrt{2})$. In turn, this can be expressed in terms of a Gaussian hypergeometric function,

$$I_1 \;=\; \frac{1}{\sqrt{2}}K(1/\sqrt{2}) \;=\; \frac{\pi}{2^{3/2}}\,{}_2F_1\left(\frac{1}{2},\frac{1}{2};1;\frac{1}{2}\right),$$

where ${}_2F_1(a,b;c,z)$ is the Gaussian hypergeometric function. From Wikipedia [2], we have Gauss's second summation theorem,

$${}_2F_1\left(a,b;\frac{1}{2}(1+a+b);\frac{1}{2}\right) \;=\; \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}(1+a+b)\right)}{\Gamma\left(\frac{1}{2}(1+a)\right)\Gamma\left(\frac{1}{2}(1+b)\right)},$$

where $\Gamma(z)$ is the gamma function. This satisfies, $\Gamma(z+1) = z\Gamma(z)$, $\Gamma(1) = 1$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Now setting $a = b = 1/2$ we see that $\frac{1}{2}(1 + a + b) = \frac{1}{2}(1 + \frac{1}{2} + \frac{1}{2}) = 1$, so that Gauss's second summation theorem applies and we can write

$$I_1 \;=\; \frac{\pi}{2^{3/2}}\,{}_2F_1\left(\frac{1}{2}, \frac{1}{2}; 1; \frac{1}{2}\right) \;=\; \left(\frac{\pi}{2}\right)^{3/2} \frac{1}{\Gamma\left(\frac{3}{4}\right)^2}.$$

Turning now to the second integral I_2, we make the substitution, $t = \sin^2\phi$. So that

$$\frac{dt}{d\phi} \;=\; 2\sin\phi\,\cos\phi \;=\; 2t^{1/2}(1-t^2)^{1/2}.$$

The integral becomes

$$I_2 \;=\; \frac{1}{2^{3/2}} \int_0^1 \frac{t^{1/2}(1-t^2)^{-1/2}\,dt}{2(1-\frac{1}{2}t^2)^{3/2}}.$$

Again this is a standard integral, from Wikipedia again [3], we have

$$\int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-zt)^a}\,dt \;=\; \frac{\Gamma(b)\Gamma(c-b)}{\Gamma(c)}\,{}_2F_1\big(a, b; c; z\big).$$

So from the above, we can solve for the parameters; clearly $a = 3/2$ and $b - 1 = 1/2$, so $b = 3/2$. Then $c - b - 1 = -1/2$ so $c = 1 - 1/2 + 3/2 = 2$. This gives us

$$I_2 \;=\; \frac{1}{2^{5/2}} \frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2})}{\Gamma(2)}\,{}_2F_1\left(\frac{3}{2}, \frac{3}{2}; 2; \frac{1}{2}\right) \;=\; \frac{\pi}{2^{7/2}}\,{}_2F_1\left(\frac{3}{2}, \frac{3}{2}; 2; \frac{1}{2}\right).$$

Again, the Gauss summation theorem applies, since $\frac{1}{2}\left(1 + \frac{3}{2} + \frac{3}{2}\right) = 2$; so

$${}_2F_1\left(\frac{3}{2}, \frac{3}{2}; 2; \frac{1}{2}\right) \;=\; \frac{\Gamma(\frac{1}{2})\Gamma(2)}{\Gamma(\frac{5}{4})^2} \;=\; \frac{2^4\pi^{1/2}}{\Gamma(\frac{1}{4})^2}.$$

Putting all of this together we get

$$I \;=\; I_1.I_2 \;=\; \left(\frac{\pi}{2}\right)^{3/2} \frac{1}{\Gamma\left(\frac{3}{4}\right)^2} \times \frac{\pi}{2^{7/2}} \times \frac{2^4\pi^{1/2}}{\Gamma(\frac{1}{4})^2} \;=\; \frac{\pi^3}{2\Big(\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right)\Big)^2}.$$

The final part is to use the Legendre duplication formula

$$\Gamma(z)\Gamma\left(z+\frac{1}{2}\right) \;=\; 2^{1-2z}\sqrt{\pi}\Gamma(2z),$$

see [4]. With $z = 1/4$ we have

$$\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) \;=\; 2^{1/2}\pi.$$

So finally we get the result,

$$I \;=\; \frac{\pi^3}{2\Big(\Gamma\big(\frac{1}{4}\big)\Gamma\big(\frac{3}{4}\big)\Big)^2} \;=\; \frac{\pi^3}{2\Big(2^{1/2}\pi\Big)^2} \;=\; \frac{\pi}{4}.$$

This seems like a lot of fiddling with gamma functions. Perhaps there is a more direct method to get the result; that would take real intelligence though!

References

[1] AIMO Prize, "Terence Tao at IMO 2024: AI and Mathematics", YouTube, `https://www.youtube.com/watch?v=e049IoFBnLA`, Last accessed 7/10/2025.

[2] Wikipedia contributors, "Hypergeometric function, Values at special points, values at $z = 1/2$", Wikipedia, The Free Encyclopedia, `https://en.wikipedia.org/wiki/Hypergeometric_function#Values_at_z_=_1/2`, Last accessed 7/10/2025.

[3] Wikipedia contributors, "Hypergeometric function, Integral formulas, Euler type", Wikipedia, The Free Encyclopedia, `https://en.wikipedia.org/wiki/Hypergeometric_function#Euler_type`, Last accessed 7/10/2025.

[4] Wikipedia contributors, "Gamma function, General", Wikipedia, The Free Encyclopedia, `https://en.wikipedia.org/wiki/Gamma_function#General`, Last accessed 7/10/2025.

Solution 326.5 – Dot Product

Suppose $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real vector space of dimension d. Show that

$$(\mathbf{u}\cdot\mathbf{v})^2 \;=\; \operatorname{trace}\big((\mathbf{u}\mathbf{u}^{\mathrm{T}})(\mathbf{v}\mathbf{v}^{\mathrm{T}})\big).$$

As usual, we consider $\mathbf{u}$ and $\mathbf{v}$ to be either ordered sets of d elements or $d\times 1$ matrices. In the latter case their transposes $\mathbf{u}^{\mathrm{T}}$ and $\mathbf{v}^{\mathrm{T}}$ are $1\times d$ matrices. Therefore when we compute $\mathbf{u}\mathbf{u}^{\mathrm{T}}$ and $\mathbf{v}\mathbf{v}^{\mathrm{T}}$ using matrix multiplication the results are $d\times d$ matrices. On the other hand, $\mathbf{u}\cdot\mathbf{v}$ is a number, the sole entry of the 1×1 matrix $\mathbf{u}^{\mathrm{T}}\mathbf{v}$.

J. M. Selig

A simple way to show this relation uses the identity

$$\operatorname{trace}\big(AB\big) \;=\; \operatorname{trace}\big(BA\big),$$

where A is any $m\times n$ matrix and B is any $n\times m$ matrix.

Now $(\mathbf{u}\cdot\mathbf{v}) = (\mathbf{u}^{\mathrm{T}}\mathbf{v})$; so thinking of this as a 1×1 matrix we can write

$$(\mathbf{u}\cdot\mathbf{v})^2 \;=\; \operatorname{trace}\big((\mathbf{u}^{\mathrm{T}}\mathbf{v})(\mathbf{u}^{\mathrm{T}}\mathbf{v})\big).$$

Then, since the dot product is commutative and matrix multiplication is associative,

$$(\mathbf{u}\cdot\mathbf{v})^2 \;=\; \operatorname{trace}\big((\mathbf{u}^{\mathrm{T}}\mathbf{v})(\mathbf{v}^{\mathrm{T}}\mathbf{u})\big) \;=\; \operatorname{trace}\big(\mathbf{u}^{\mathrm{T}}(\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{u})\big).$$

Using the relation $\operatorname{trace}\big(AB\big) = \operatorname{trace}\big(BA\big)$ and the associativity of matrix multiplication again, this becomes

$$(\mathbf{u}\cdot\mathbf{v})^2 \;=\; \operatorname{trace}\big((\mathbf{u}\mathbf{u}^{\mathrm{T}})(\mathbf{v}\mathbf{v}^{\mathrm{T}})\big).$$

If you are happy with the relation concerning the traces of the matrices we are done. If not, I need to prove the formula.

We start with the simplest case of a pair of vectors thought of as $1\times d$ matrices, the relation is simply,

$$\operatorname{trace}\big(\mathbf{u}\mathbf{v}^{\mathrm{T}}\big) \;=\; \operatorname{trace}\big(\mathbf{v}^{\mathrm{T}}\mathbf{u}\big).$$

This is easy to see, on the left we have the trace of a $d\times d$ matrix,

$$\operatorname{trace}\begin{pmatrix} v_1u_1 & v_1u_2 & v_1u_3 & \cdots & v_1u_d \\ v_2u_1 & v_2u_2 & v_2u_3 & \cdots & v_2u_d \\ v_3u_1 & v_3u_2 & v_3u_3 & \cdots & v_3u_d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_du_1 & v_du_2 & v_du_3 & \cdots & v_du_d \end{pmatrix} \;=\; \sum_{i=1}^{d} v_iu_i.$$

While on the left we have the trace of a 1×1 matrix,

$$\operatorname{trace}\bigl(u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_dv_d\bigr) \;=\; \sum_{i=1}^{d} u_iv_i.$$

which is clearly the same as the right-hand side of the relation.

This can now be generalised to the case of arbitrary matrices. Suppose A has n columns

$$A \;=\; \Bigl(\mathbf{a}_1 \Bigm| \mathbf{a}_2 \Bigm| \mathbf{a}_3 \Bigm| \cdots \Bigm| \mathbf{a}_n\Bigr),$$

so that B has n rows,

$$B \;=\; \begin{pmatrix} \mathbf{b}_1^{\mathrm{T}} \\ \mathbf{b}_2^{\mathrm{T}} \\ \mathbf{b}_3^{\mathrm{T}} \\ \vdots \\ \mathbf{b}_n^{\mathrm{T}} \end{pmatrix}.$$

Then we have

$$AB \;=\; (\mathbf{a}_1\mathbf{b}_1^{\mathrm{T}}) + (\mathbf{a}_2\mathbf{b}_2^{\mathrm{T}}) + (\mathbf{a}_3\mathbf{b}_3^{\mathrm{T}}) + \cdots + (\mathbf{a}_n\mathbf{b}_n^{\mathrm{T}}).$$

Recall that the trace is a linear operator on matrices; $\operatorname{trace}\bigl(X+Y\bigr) = \operatorname{trace}\bigl(X\bigr) + \operatorname{trace}\bigl(Y\bigr)$. So the trace of this matrix is

$$\operatorname{trace}\bigl(AB\bigr) \;=\; \sum_{i=1}^{n} \operatorname{trace}\bigl(\mathbf{a}_i\mathbf{b}_i^{\mathrm{T}}\bigr) \;=\; \sum_{i=1}^{n} \operatorname{trace}\bigl(\mathbf{b}_i^{\mathrm{T}}\mathbf{a}_i\bigr) \;=\; \sum_{i=1}^{n} \mathbf{b}_i^{\mathrm{T}}\mathbf{a}_i,$$

using the relation we found for vectors. On the other hand we have that

$$BA \;=\; \begin{pmatrix} \mathbf{b}_1^{\mathrm{T}}\mathbf{a}_1 & \mathbf{b}_1^{\mathrm{T}}\mathbf{a}_2 & \mathbf{b}_1^{\mathrm{T}}\mathbf{a}_3 & \cdots & \mathbf{b}_1^{\mathrm{T}}\mathbf{a}_n \\ \mathbf{b}_2^{\mathrm{T}}\mathbf{a}_1 & \mathbf{b}_2^{\mathrm{T}}\mathbf{a}_2 & \mathbf{b}_2^{\mathrm{T}}\mathbf{a}_3 & \cdots & \mathbf{b}_2^{\mathrm{T}}\mathbf{a}_n \\ \mathbf{b}_3^{\mathrm{T}}\mathbf{a}_1 & \mathbf{b}_3^{\mathrm{T}}\mathbf{a}_2 & \mathbf{b}_3^{\mathrm{T}}\mathbf{a}_3 & \cdots & \mathbf{b}_3^{\mathrm{T}}\mathbf{a}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{b}_n^{\mathrm{T}}\mathbf{a}_1 & \mathbf{b}_n^{\mathrm{T}}\mathbf{a}_2 & \mathbf{b}_n^{\mathrm{T}}\mathbf{a}_3 & \cdots & \mathbf{b}_n^{\mathrm{T}}\mathbf{a}_n \end{pmatrix}.$$

The diagonal elements of this matrix sum to give

$$\operatorname{trace}\bigl(BA\bigr) \;=\; \mathbf{b}_1^{\mathrm{T}}\mathbf{a}_1 + \mathbf{b}_2^{\mathrm{T}}\mathbf{a}_2 + \mathbf{b}_3^{\mathrm{T}}\mathbf{a}_3 + \cdots + \mathbf{b}_n^{\mathrm{T}}\mathbf{a}_n \;=\; \sum_{i=1}^{n} \mathbf{b}_i^{\mathrm{T}}\mathbf{a}_i,$$

the same as our result for $\operatorname{trace}\bigl(AB\bigr)$.

Solution 324.5 – Tiles

> I have $2n$ tiles in a line. I wish to colour n of them red, and the other n blue. For aesthetic reasons I do not allow there to be more than m consecutive tiles the same colour. How many ways can I effect this colouring?
>
> If it helps, when $m = 1$ the solution is 2. Once I have decided the colour of the first tile, the rest are determined. On the other hand, when $m \geq n$, in which case there is no aesthetic restriction, the solution is $2n$ choose n, in other words $(2n)!/(n!n!)$.

Reinhardt Messerschmidt

I will give a solution for the case $m = 2$. Let x_n be the number of ways of colouring $2n$ tiles such that n are red, n are blue, and no more than 2 consecutive tiles have the same colour.

Let r be the number of runs of the first colour. The number of runs of the second colour is $r - \delta$ for some $\delta \in \{0, 1\}$. We must have

$$2(r - \delta) \;\geq\; n,$$

otherwise the number of tiles of the second colour will fall short of n; therefore $r \geq \lceil n/2 \rceil + \delta$. We must also have $r \leq n$, otherwise the number of tiles of the first colour will exceed n.

Let s and t be the number of runs of length 2 of the first and second colour respectively. We have

$$2s + (r - s) \;=\; n, \qquad 2t + (r - \delta - t) \;=\; n;$$

therefore

$$s \;=\; n - r, \qquad t \;=\; n - r + \delta.$$

The number of ways of choosing s runs from the r runs of the first colour and t runs from the $r - \delta$ runs of the second colour is

$$\binom{r}{s}\binom{r-\delta}{t} \;=\; \binom{r}{n-r}\binom{r-\delta}{n-r+\delta}.$$

There are two choices for the first colour; therefore

$$x_n \;=\; 2\sum_{\delta=0}^{1}\;\sum_{r=\lceil n/2\rceil+\delta}^{n}\binom{r}{n-r}\binom{r-\delta}{n-r+\delta}.$$

Using the identity

$$2\binom{r}{n-r}\binom{r-1}{n-r+1} = \left[\binom{r}{n-r}+\binom{r-1}{n-r+1}\right]^2-\binom{r}{n-r}^2-\binom{r-1}{n-r+1}^2,$$

we find

$$x_n \;=\; 1+\binom{\lceil n/2\rceil}{n-\lceil n/2\rceil}^2+\sum_{r=\lceil n/2\rceil+1}^{n}\left[\binom{r}{n-r}+\binom{r-1}{n-r+1}\right]^2.$$

For small values of n, the formula gives the following.

n	1	2	3	4	5	6	7	8	9	10	11
x_n	2	6	14	34	84	208	518	1296	3254	8196	21700

This is sequence A177790 in [1]. For $n = 4$, the $34/2 = 17$ colourings for which the first colour is red are as follows.

r	δ	s	t	$\binom{r}{s}\binom{r-\delta}{t}$	Colourings		
2	0	2	2	1	`rrbbrrbb`		
3	0	1	1	9	`rrbbrbrb` `rrbrbbrb` `rrbrbrbb`	`rbbrrbrb` `rbrrbbrb` `rbrrbrbb`	`rbbrbrrb` `rbrbbrrb` `rbrbrrbb`
3	1	1	2	3	`rrbbrbbr`	`rbbrrbbr`	`rbbrbbrr`
4	0	0	0	1	`rbrbrbrb`		
4	1	0	1	3	`rbbrbrbr`	`rbrbbrbr`	`rbrbrbbr`

References

[1] *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`, accessed 13 July 2025.

Solution 325.1 – A sieve for squares

> Write down a sequence of infinitely many 1s. Then:
>
> for $k = 2$ to ∞:
>
> for $j = 1$ to ∞:
>
> add 1 mod 2 to the (jk)-th number in the sequence.
>
> Show that the nth number in the sequence is 1 if and only if n is an integer square—or find a counter-example.

David Sixsmith

The problem is equivalent, and slightly easier to imagine, if we start with a sequence of infinitely many zeros, and then allow the first variable k to go from 1 to ∞.

Note then that the Nth symbol in the sequence gets modified every time $N = jk$, for integers j, k. In other words the Nth symbol is modified every time N has a factor (including 1 and itself). So the question is equivalent to asking which integers have an odd number of factors.

This is exactly the square numbers. An easy way to see this is as follows. If N is not square, then factors of N occur in pairs – d and N/d, and these are not equal. However, if N is a square then it has one factor which is its square root, and all other factors occur in pairs in the same way.

Problem 328.3 – Integer expression

Tony Forbes

Let n be an even positive integer. Show that

$$\frac{n!}{(n/2)!\; 5^{\lfloor n/9+1/2 \rfloor}\; 2^{n/2}}$$

is an integer except when $n \in \{14, 24\}$. Or find another counter-example. The exponent of 5 in the denominator rounds $n/9$ to the nearest integer.

Incidentally, I note that

$$\frac{52!}{26!} \;=\; 1.999997098 \times 10^{41},$$

which leads me to ask: Can you find more examples of $(2m)!/m! \approx$ integer times a power of 10?

Solution 324.4 – Sum

Show that

$$\frac{1}{1\cdot 3\cdot 5\cdot 7}+\frac{1}{9\cdot 11\cdot 13\cdot 15}+\frac{1}{17\cdot 19\cdot 21\cdot 23}+\dots \;=\; \frac{(2-\sqrt{2})\pi}{192}.$$

Henry Ricardo

The general term is $a_k = 1/\left((8k+1)(8k+3)(8k+5)(8k+7)\right)$, and we have

$$\begin{aligned}
a_k \;&=\; -\frac{1}{48(8k+7)}+\frac{1}{16(8k+5)}-\frac{1}{16(8k+3)}+\frac{1}{48(8k+1)}\\
&=\; \frac{1}{384}\left(\frac{1}{k+1/8}-\frac{1}{k+7/8}\right)+\frac{1}{128}\left(\frac{1}{k+5/8}-\frac{1}{k+3/8}\right)\\
&=\; \frac{1}{384}\left[\left(\frac{1}{k+1}-\frac{1}{k+7/8}\right)-\left(\frac{1}{k+1}-\frac{1}{k+1/8}\right)\right]\\
&\quad+\frac{1}{128}\left[\left(\frac{1}{k+1}-\frac{1}{k+3/8}\right)-\left(\frac{1}{k+1}-\frac{1}{k+5/8}\right)\right].
\end{aligned}$$

Now we can write

$$\sum_{k=0}^{\infty} a_k \;=\; \frac{1}{384}\left(\psi\left(\frac{7}{8}\right)-\psi\left(\frac{1}{8}\right)\right)+\frac{1}{128}\left(\psi\left(\frac{3}{8}\right)-\psi\left(\frac{5}{8}\right)\right), \qquad (1)$$

where we have used the series definition of Euler's digamma function:

$$\psi(x) \;=\; -\gamma+\sum_{n=0}^{\infty}\left(\frac{1}{n+1}-\frac{1}{n+x}\right), \quad x\neq 0,-1,-2,\dots.$$

Next, the reflection formula $\psi(1-x)-\psi(x)=\pi\cot\pi x$ and a tangent half-angle formula yield

$$\begin{aligned}
\psi(7/8)-\psi(1/8) \;&=\; \pi\cot(\pi/8) \;=\; \pi(\sqrt{2}+1),\\
\psi(3/8)-\psi(5/8) \;&=\; \pi\cot(5\pi/8) \;=\; -\pi/\left(\sqrt{3+2\sqrt{2}}\right),
\end{aligned}$$

so that (1) becomes

$$\frac{(\sqrt{2}+1)\pi}{384} - \frac{\pi}{128\left(\sqrt{3+2\sqrt{2}}\right)} \;=\; \frac{(2-\sqrt{2})\pi}{192}.$$

Solution 324.8 – Random reals

> I choose n real numbers uniformly at random in the interval $[0,1]$. What is the probability the sum is less than one? What is the probability the sum of the squares is less than one? How does the ratio of these probabilities behave as $n \to \infty$?

Reinhardt Messerschmidt

Let $(X_1, X_2, \dots)$ be a sequence of independent random variables that are each uniformly distributed on $[0,1]$, and let

$$Y_n \;=\; X_1 + \dots + X_n, \qquad Z_n \;=\; {X_1}^2 + \dots + {X_n}^2.$$

We will use induction on n to show that $\mathbb{P}[Y_n < y] = y^n/n!$ for every $n \in \{1,2,\dots\}$ and $y \in [0,1]$. The base case is clear. For the inductive case, suppose $n > 1$ and $\mathbb{P}[Y_{n-1} < y] = y^{n-1}/(n-1)!$ for every $y \in [0,1]$. The result follows from

$$\begin{aligned}\mathbb{P}[Y_n < y] \;&=\; \mathbb{P}[Y_{n-1} + X_n < y] \;=\; \int_0^y \mathbb{P}[Y_{n-1} < y - x]\, dx \\ &=\; \int_0^y \frac{(y-x)^{n-1}}{(n-1)!}\, dx \;=\; \left.\frac{-(y-x)^n}{n!}\right|_{x=0}^{x=y} \;=\; y^n/n!\,.\end{aligned}$$

Let B_n be the unit ball in $\mathbb{R}^n$. The event $Z_n < 1$ occurs if and only if $(X_1, \dots, X_n) \in B_n$; therefore

$$\mathbb{P}[Z_n < 1] \;=\; \text{volume}(B_n \cap [0,1]^n) \;=\; \frac{\text{volume}(B_n)}{2^n}.$$

By the formula for the volume of the unit ball (corollary 2.55 in [1]),

$$\mathbb{P}[Z_n < 1] \;=\; \frac{\pi^{n/2}}{2^n \Gamma(n/2+1)}.$$

It follows that

$$\frac{\mathbb{P}[Y_n < 1]}{\mathbb{P}[Z_n < 1]} \;=\; \frac{2^n \Gamma(n/2+1)}{\pi^{n/2} \Gamma(n+1)}.$$

By Legendre's duplication formula for the gamma function (the corollary in section 12.15 of [2]),

$$\Gamma\left(2\left(\frac{n+1}{2}\right)\right) = \pi^{-1/2}2^{2(n+1)/2-1}\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n+1}{2}+\frac{1}{2}\right)$$
$$= \pi^{-1/2}2^n\Gamma\big((n+1)/2\big)\Gamma\big(n/2+1\big);$$

therefore
$$\frac{\mathbb{P}[Y_n<1]}{\mathbb{P}[Z_n<1]} = \frac{1}{\pi^{(n-1)/2}\Gamma\big((n+1)/2\big)}.$$

This converges to 0 as $n \longrightarrow \infty$.

References

[1] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., John Wiley & Sons, New York, 1999.

[2] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 3rd ed., Cambridge University Press, Cambridge, 1920.

David Sixsmith

1 Solution – for the sums

For each natural number n, let t_n denote the probability the sum is less than one, and let s_n denote the probability the sum of the squares is less than one. We will begin by considering the t_n, and see how to translate the problem into a problem of Euclidean geometry.

For the moment, suppose that $n = 2$, so that we are choosing two numbers x_1, x_2 uniformly at random in the interval $[0, 1]$. We can then consider the two numbers as coordinates in the plane, so that the problem is equivalent to choosing a point (x_1, x_2) uniformly at random in the square $[0, 1]^2$, and then asking that $x_1 + x_2 < 1$.

Since the square has area one, all we need is the area of the triangle, nearest the origin, formed by intersecting the square with the line $x_1 + x_2 = 1$. This is easily seen to be one half.

Things get harder to imagine in n-dimensional space (for me at least). We are asked to choose a coordinate $(x_1, x_2, \ldots, x_n)$ uniformly at random in the n-cube $[0, 1]^n$, such that $\sum_1^n x_i < 1$.

Again, since the n-cube has volume one, we obtain the probability by calculating the volume of the n-dimensional object where all the coordinates are positive and the sum of the coordinates is at most one. Call this object C_n. We claim that it has n-dimensional volume $n!$.

We prove by induction. The case $n = 1$ is clear. Suppose we have proved for $n = k$. Consider the intersection, I_ξ, of C_{k+1} with the hyper-plane

$$x_{k+1} \;=\; 1 - \xi, \qquad \text{for } \xi \in [0, 1].$$

In I_ξ all the coordinates are positive, and $\sum_{i=1}^{k} x_i = \xi$. Hence I_ξ is a copy C_k, linearly scaled by a factor of ξ, and so, by the induction hypothesis, has $n-1$-dimensional volume $\xi^k/k!$. Integrating, we obtain that the volume of C_{k+1} is equal to

$$\int_0^1 \frac{\xi^k}{k!}\, d\xi \;=\; \frac{1}{(k+1)!},$$

which completes the induction. Hence we have deduced that

$$t_n \;=\; \frac{1}{n!}.$$

2 Solution – for the sums of the squares

We can now calculate the s_n. By an argument as above, we can see that s_n is the volume of the part of the unit sphere, centred at the origin, where all coordinates are positive. Since the unit sphere has volume $\pi^{n/2}/\Gamma(1 + n/2)$ (see Wikipedia) we can deduce that

$$s_n \;=\; \frac{\pi^{n/2}}{2^n \Gamma\left(1 + \frac{n}{2}\right)}.$$

Use of Stirling's formula (or even just quick calculations) shows that s_n tends to zero much more slowly than t_n. In fact, Stirling's formula gives

$$\frac{s_n}{t_n} \;\sim\; \sqrt{2} \cdot (cn)^{n/2} \qquad \text{as } n \to \infty,$$

where $c = \pi/2e \approx 0.5779$.

Problem 328.4 – Diophantine equation

Tony Forbes

The equation

$$3 \cdot 9^a - 2b^2 \;=\; 1$$

has this solution in positive integers: $a = 2$, $b = 11$. Are there any others?

Solution 314.6 – Square

> There is a square with vertices A, B, C, D, $AB \| CD$, $BC \| AD$, point E bisects AD, and F is the point where a line through C meets BE at 90 degrees. Show that $|AB| = |DF|$.

Tony Forbes

Whilst idly browsing past issues of M500 I found this interesting problem. I don't recall anyone submitting a solution. So I thought I would have a go in my usual brute-force style. The details should be quite easy to follow—especially if you draw a diagram.

Assume $|AB| = 1$. Let

$$\alpha \;=\; \angle ABE \;=\; \arctan\frac{1}{2} \quad \text{and} \quad \beta \;=\; \angle DEF \;=\; \frac{\pi}{2} + \alpha.$$

Then

$$|DE| \;=\; \frac{1}{2}, \quad |BE| \;=\; \frac{\sqrt{5}}{2} \quad \text{and} \quad |BF| \;=\; \sin\alpha \;=\; \frac{1}{\sqrt{5}}.$$

Hence

$$|EF| \;=\; |BE| - |BF| \;=\; \frac{\sqrt{5}}{2} - \frac{1}{\sqrt{5}} \quad \text{and} \quad \cos\beta \;=\; -\frac{1}{\sqrt{5}}.$$

We now have enough material to invoke the cosine rule:

$$\begin{aligned} |DF|^2 \;&=\; |DE|^2 + |EF|^2 - 2|DE||EF|\cos\beta \\ &=\; \frac{1}{4} + \left(\frac{\sqrt{5}}{2} - \frac{1}{\sqrt{5}}\right)^2 - \left(\frac{\sqrt{5}}{2} - \frac{1}{\sqrt{5}}\right)\left(-\frac{1}{\sqrt{5}}\right). \end{aligned}$$

The manner in which the ghastly mess on the right simplifies to 1 suggests there must be a neater way to solve the problem. I leave it for someone else to find.

Problem 328.5 – Sum

Tony Forbes

Show that

$$\frac{1}{1\cdot 2^2} + \frac{1}{5\cdot 6^2} + \frac{1}{9\cdot 10^2} + \ldots \;=\; \frac{4\pi - \pi^2 + \log 256}{32}.$$

Contents

04.00

Problem 328.6 – Integral

Compute $\displaystyle\int_0^x \tanh(\log t)\,dt, \quad x \geq 0.$

Front cover A SET deal that ends all clear. See page 8.