Manifold Mathematics

A Complete Worked Commentary

on the

Joint CSIR–UGC NET Mathematical Sciences

Paper of 17 July 2026

All one hundred and twenty items, treated from first principles.
Every concept developed in full; no step assumed,
no shortcut substituted for an argument.

Volume in the Manifold Mathematics Examination Commentary series

17 August 2026

How this commentary is written

This is not an answer key. An answer key tells you *which* letter is correct; it teaches you nothing you did not already know. What follows is a commentary, and its organising principle is that a solved problem is worth very little unless you also carry away the machinery that solved it.

Accordingly, almost every item below is developed in three movements.

The three movements

First, the concept. Before any computation, the underlying object is defined and its basic properties are established — what a rheonomic constraint is, what completeness of a statistic means, why the Jordan decomposition of a function of bounded variation exists. This part is written so that it stands alone, independently of the question that motivated it.

Second, the computation. The specific problem is then solved, with every algebraic step shown. Where a step uses a standard identity, the identity is stated rather than invoked by name alone.

Third, the discrimination. For a multiple-choice item, knowing why the correct option is correct is only half the work. The remaining options are examined and their specific errors named, because those errors are the ones the setter expects you to make.

On items whose data did not reproduce

The source paper carries a number of items whose essential data — a transition matrix, a pie chart, a design matrix, a pair of statements labelled S_1 and S_2 — did not survive into the text available for this commentary. These items are not silently skipped and they are not silently guessed at.

For each such item the treatment is as follows. The complete theory needed to solve it is developed exactly as if the data were present, because that theory is the reason the item exists. Where the option set alone constrains the answer — and it often does, through internal consistency, dimensional impossibility, or an inequality that eliminates three of the four choices — that constraint is exploited and the reasoning made explicit. Where it does not, the commentary says so plainly and identifies precisely what is missing. A missing datum is recorded, not disguised.

On defective items

Several items in this paper are defective: the arithmetic answer does not appear among the four options, or two options describe the same object, or the stem contains an evident typographical corruption. These are identified and the correct mathematics is given regardless of what the option set permits. In examination conditions such items are usually withdrawn; in preparation they are still worth solving, because the mathematics is sound even when the printing is not.

A consolidated register of defective and incomplete items appears in the final chapter.

Notation

Standard conventions are used throughout: $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{Z}, \mathbb{N}$ for the usual number systems, $\mathbb{F}_q$ for the field with q elements, $M_n(\mathbb{R})$ for $n \times n$ real matrices, $P_n(\mathbb{R})$ for real polynomials of degree at most n (so $\dim P_n(\mathbb{R}) = n+1$), $\mathbb{E}$ and Var for expectation and variance, $\mathbb{P}$ for probability. Vectors are columns; A^{T} is the transpose and A^* the conjugate transpose.

The commentary is written for the candidate who intends to understand the subject, not merely to pass through it. Where a shorter route exists it is mentioned, but the long route is given first.

Contents

CHAPTER 1

Part A — General Aptitude

Part A carries twenty items of which fifteen are to be attempted. The material is elementary, but the marking is unforgiving and the time allowance is short, so the objective in preparation is to make each of these types *automatic* rather than merely solvable.

Question 1. In a competitive examination the success rate of men is 40% and that of women is 50%. There are 60% men and 40% women among the candidates. If a candidate selected at random is successful, the probability that the candidate is a man is **(a)** 4/11 **(b)** 5/11 **(c)** 6/11 **(d)** 7/11.

The concept: inverting a conditional probability

The datum supplied is $\mathbb{P}(\text{success} \mid \text{man})$. The datum demanded is $\mathbb{P}(\text{man} \mid \text{success})$. These are different numbers, and confusing them is the single most common error in elementary probability. The device that converts one into the other is Bayes' theorem.

Bayes' theorem

Let $\{H_1, \ldots, H_k\}$ be a partition of the sample space with $\mathbb{P}(H_i) > 0$, and let E be an event with $\mathbb{P}(E) > 0$. Then

$$\mathbb{P}(H_j \mid E) = \frac{\mathbb{P}(E \mid H_j)\,\mathbb{P}(H_j)}{\displaystyle\sum_{i=1}^{k} \mathbb{P}(E \mid H_i)\,\mathbb{P}(H_i)}.$$

The denominator is the *law of total probability*: it reassembles $\mathbb{P}(E)$ from the branch probabilities. The numerator is one branch of that sum. So the theorem says: *the posterior probability of a hypothesis is the share of the total that its own branch contributes.*

The computation

Write M for "man", W for "woman", S for "successful". The prior distribution over the partition $\{M, W\}$ is

$$\mathbb{P}(M) = 0.60, \qquad \mathbb{P}(W) = 0.40,$$

and the likelihoods are

$$\mathbb{P}(S \mid M) = 0.40, \qquad \mathbb{P}(S \mid W) = 0.50.$$

The two branch contributions are

$$\mathbb{P}(S \cap M) = \mathbb{P}(S \mid M)\mathbb{P}(M) = 0.40 \times 0.60 = 0.24,$$

$$\mathbb{P}(S \cap W) = \mathbb{P}(S \mid W)\mathbb{P}(W) = 0.50 \times 0.40 = 0.20.$$

Hence the total success probability is

$$\mathbb{P}(S) = 0.24 + 0.20 = 0.44,$$

and therefore

$$\mathbb{P}(M \mid S) = \frac{0.24}{0.44} = \frac{24}{44} = \frac{6}{11}.$$

Discrimination

The complementary answer $\mathbb{P}(W \mid S) = 0.20/0.44 = 5/11$ is option **(b)**; it is placed there precisely to catch a candidate who computes the correct total but then reads off the wrong branch. Option **(a)**, $4/11$, is what you obtain if you forget to weight by the priors and simply write $0.4/(0.4+0.5)\cdot$ something; it corresponds to treating the two groups as equally represented.

Sanity check without algebra. Men are the majority (60%) but the weaker group (40% success); women are the minority but stronger. The two effects partially cancel, so the answer should be near $1/2$ but tilted towards men, since $0.6 \times 0.4 = 0.24$ exceeds $0.4 \times 0.5 = 0.20$. And indeed $6/11 \approx 0.545$. Any answer below $1/2$ is immediately wrong.

Answer: **(c)** $6/11$.

Question 2. A sum of money doubles itself in 5 years at a certain rate of simple interest. In how many years will it become 16 times itself at the same rate? **(a)** 40 **(b)** 50 **(c)** 75 **(d)** 80.

The concept: simple interest is linear, not exponential

Simple versus compound growth

Under *simple* interest the interest is computed always on the original principal P, never on accumulated interest. After t years at annual rate r (as a decimal) the amount is

$$A(t) = P(1 + rt),$$

an *affine* function of t. Under *compound* interest the amount is $A(t) = P(1+r)^t$, an exponential function.

The consequence for doubling problems is decisive. Under compounding, doubling times *add*: sixteen-fold growth is four doublings, hence four doubling times. Under simple interest doubling times do *not* add, because the increments of the amount are constant, not proportional.

The computation

Doubling in 5 years means the accumulated interest over 5 years equals the principal:

$$P(1+5r) = 2P \quad\Longrightarrow\quad 5r = 1 \quad\Longrightarrow\quad r = \tfrac{1}{5} = 20\% \text{ per annum.}$$

For the amount to become $16P$ the accumulated interest must be $15P$:

$$P(1+rt) = 16P \quad\Longrightarrow\quad rt = 15 \quad\Longrightarrow\quad t = \frac{15}{r} = 15 \times 5 = 75.$$

Discrimination

Option **(d)**, 80 years, is the answer to "sixteen times the *interest*" or, equivalently, arises from writing 16 instead of 15 for the required multiple of accumulated interest. It is the standard trap: the amount becomes 16 times, so the *interest earned* is 15 times the principal, not 16.

Option **(a)**, 40 years, is the compound-interest answer misapplied: four doublings at 5 years each would be 20 years under compounding; 40 is neither. Option **(b)**, 50, has no natural derivation and is a distractor.

The general rule. Under simple interest, if a sum becomes m times itself in t_m years and n times itself in t_n years at the same rate, then

$$\frac{t_m}{t_n} = \frac{m-1}{n-1}.$$

Here $t_{16}/t_2 = (16-1)/(2-1) = 15$, so $t_{16} = 15 \times 5 = 75$. Memorise the $m-1$; it is the whole content of the problem.

Answer: (c) 75 years.

Question 3. What is the next term in the series A, Z, B, Y, C, X, ...? **(a)** D **(b)** W **(c)** V **(d)** U.

The concept: interleaved sequences

A great many "series" items are not one sequence but two, alternately presented. The diagnostic is to split by parity of position.

The computation

Odd positions: $A, B, C, \ldots$ — the alphabet read forwards. Even positions: $Z, Y, X, \ldots$ — the alphabet read backwards.

The given terms occupy positions 1 through 6; the next term is at position 7, which is odd, and so belongs to the forward strand. That strand has so far produced A (position 1), B (position 3), C (position 5), so its next member is D.

A useful cross-check. The two strands are related by the reflection $k \mapsto 27-k$ on letter positions: $A \leftrightarrow Z$ $(1 \leftrightarrow 26)$, $B \leftrightarrow Y$ $(2 \leftrightarrow 25)$, $C \leftrightarrow X$ $(3 \leftrightarrow 24)$. So the terms pair up as $(A, Z), (B, Y), (C, X), (D, W), \ldots$ The pair beginning at position 7 is (D, W), confirming D. Note that option **(b)**, W, is the *eighth* term — a one-step-ahead error.

Answer: (a) D.

Question 4. A triangle is inscribed in a semicircle of radius r with the diameter as one of its sides. The area of the triangle is maximum when **(a)** the triangle is equilateral **(b)** the triangle is isosceles right-angled **(c)** the triangle is isosceles with base angles 45° **(d)** the height of the triangle is $r/\sqrt{2}$.

The concept: Thales' theorem and a fixed base

Thales' theorem

If AB is a diameter of a circle and C is any other point on the circle, then $\angle ACB = 90°$. Equivalently, the locus of points from which a fixed segment subtends a right angle is the circle on that segment as diameter.

Proof sketch. Let O be the centre. Then $OA = OB = OC = r$, so triangles OAC and OBC are isosceles. Writing $\angle OAC = \angle OCA = \alpha$ and $\angle OBC = \angle OCB = \beta$, the angle sum of $\triangle ABC$ gives $2\alpha + 2\beta = 180°$, whence $\angle ACB = \alpha + \beta = 90°$. □

So every triangle of the described kind is automatically right-angled at the third vertex. The right angle is not something to be arranged; it is forced.

The computation

Place the semicircle as $x^2 + y^2 = r^2$, $y \geq 0$, with the diameter along the x-axis from $(-r, 0)$ to $(r, 0)$. Let the apex be $C = (x, y)$ with $y \geq 0$ on the arc. Then

$$\text{Area} = \tfrac{1}{2} \times (\text{base}) \times (\text{height}) = \tfrac{1}{2} \times 2r \times y = ry.$$

The base is fixed at $2r$, so the area is a strictly increasing function of the height y alone. On the semicircle $y = \sqrt{r^2 - x^2} \leq r$, with equality precisely at $x = 0$. Hence the maximum area is $r \cdot r = r^2$, attained when the apex is at the top of the arc, $C = (0, r)$.

For that triangle the two slant sides have length $\sqrt{r^2 + r^2} = r\sqrt{2}$ each, so it is isosceles; and by Thales the apex angle is 90°. Its base angles are therefore $(180° - 90°)/2 = 45°$ each.

This item is defective. Options **(b)** and **(c)** describe the *same* triangle. A triangle that is isosceles and right-angled has apex angle 90° and base angles 45°; a triangle that is isosceles with base angles 45° has apex angle 90° and is right-angled. There is no triangle satisfying one and not the other. Both are correct descriptions of the maximiser.

Option **(a)** is wrong: an equilateral triangle has all angles 60° and cannot contain the forced right angle. Option **(d)** is wrong: the maximising height is r, not $r/\sqrt{2}$. The value $r/\sqrt{2}$ is the height of the point at 45° around the arc, which is a distractor built from the correct *angle* attached to the wrong *coordinate*.

Answer: **(b)** (equivalently **(c)**; the two options are logically identical and the item as printed admits two correct choices). The maximising triangle is the isosceles right triangle of height r and area r^2.

Question 5. The product of two numbers is 375 and their HCF is 5. The number of such pairs is **(a)** 1 **(b)** 2 **(c)** 3 **(d)** 4.

The concept: factoring out the gcd

The gcd normal form

If $\gcd(a, b) = d$ then we may write $a = dm$, $b = dn$ with $\gcd(m, n) = 1$. Conversely any such m, n produce a pair with gcd exactly d. The coprimality of m and n is not an extra assumption — it is forced, since a common factor of m and n would multiply d and contradict maximality.
This substitution is the standard first move in every "product and HCF" or "LCM and HCF" problem, because it converts a constrained problem into an unconstrained factorisation of a smaller number.

The computation

Write $a = 5m$, $b = 5n$ with $\gcd(m, n) = 1$. Then

$$ab = 25mn = 375 \quad \Longrightarrow \quad mn = 15.$$

Factor $15 = 3 \times 5$. The ordered factorisations of 15 into two positive factors are

$$(1, 15),\ (3, 5),\ (5, 3),\ (15, 1),$$

and every one of them is coprime, because 15 is squarefree with two distinct prime factors, so no prime can appear in both factors.

Since the question asks for *pairs* of numbers, the ordering is immaterial, and the four ordered factorisations collapse to two unordered ones:

$$(m, n) = (1, 15) \ \Rightarrow\ (a, b) = (5, 75), \qquad (m, n) = (3, 5) \ \Rightarrow\ (a, b) = (15, 25).$$

Verification: $5 \times 75 = 375$ with $\gcd = 5$; $15 \times 25 = 375$ with $\gcd = 5$. Both are valid, and there are exactly two.

The general count. If $N = d^2 k$ and we seek unordered pairs with product N and gcd d, the count is half the number of ways of writing k as an ordered product of two coprime factors. That number is $2^{\omega(k)}$ where $\omega(k)$ is the number of distinct prime divisors of k. Here $k = 15$, $\omega = 2$, so $2^2 = 4$ ordered and $4/2 = 2$ unordered. Option **(d)**, 4, is the *ordered* count and is the trap.

Answer: (b) 2.

Question 6. If in a certain code COMPUTE is written as RFUVQPC, how will PRINTER be written? **(a)** SFUOJSQ **(b)** QSJOUFS **(c)** RFUOJSQ **(d)** SFUOJQS.

The concept: composite letter codes

Coding items are built from a small vocabulary of operations: a positional shift (Caesar), a reversal of the word, a reflection of the alphabet ($k \mapsto 27 - k$), or a composition of two of these. The method is to test the operations in order of frequency and to *verify against the option set*, which is often the more reliable source when the stem is misprinted.

The computation

Apply the composite "reverse, then shift each letter forward by one" to PRINTER:

Original	P	R	I	N	T	E	R
Reversed	R	E	T	N	I	R	P
Shift +1	S	F	U	O	J	S	Q

The result is SFUOJSQ, which is exactly option **(a)**. No other option is producible by a natural single or composite operation:

- **(b)** QSJOUFS is +1 with *no* reversal.
- **(c)** RFUOJSQ differs from **(a)** only in the first letter (R instead of S); it is **(a)** with one letter left unshifted.
- **(d)** SFUOJQS is **(a)** with the last two letters transposed.

Options **(c)** and **(d)** are perturbations of the correct answer, which is strong internal evidence that **(a)** is the intended key.

The stem is corrupt. Applying "reverse then +1" to COMPUTE gives:

$$\text{COMPUTE} \xrightarrow{\text{reverse}} \text{ETUPMOC} \xrightarrow{+1} \text{FUVQNPD},$$

whereas the paper prints RFUVQPC. The two agree on the substring FUVQ but not elsewhere. Since the option set is generated cleanly by the reverse-and-shift rule and by nothing else, the rule is the intended one and the printed code word contains a typographical error. Solve from the options, not from the corrupted stem.

Answer: **(a)** SFUOJSQ.

Question 7. There are 5 coins of equal weight and one of them is heavier. Using a balance scale, the minimum number of weighings required to identify the heavier coin is **(a)** 1 **(b)** 2 **(c)** 3 **(d)** 4.

The concept: the information-theoretic lower bound

Counting outcomes

A pan balance has exactly three possible outcomes per weighing: left pan heavier, right pan heavier, or balance. A strategy consisting of w weighings therefore distinguishes at most 3^w situations, because the full record of results is a string of length w over a three-letter alphabet.
If there are N candidate answers, a valid strategy requires

$$3^w \;\geq\; N \qquad \Longrightarrow \qquad w \;\geq\; \lceil \log_3 N \rceil.$$

This is a *lower* bound; it must then be shown to be achievable by exhibiting a strategy.

Note the important premise: we are told one coin is heavier (not merely different), so there are $N = 5$ candidates, not 10.

The lower bound

With $N = 5$ we need $3^w \geq 5$. Since $3^1 = 3 < 5$ and $3^2 = 9 \geq 5$, we obtain $w \geq 2$. So one weighing cannot suffice, and option **(a)** is eliminated on counting grounds alone.

The matching strategy

Label the coins $c_1, \ldots, c_5$.

Weighing 1: c_1 versus c_2.

- If the pans do not balance, the heavier pan holds the heavy coin, and it is identified in one weighing. (We still need the second weighing only in the other branch, so the worst case governs.)
- If they balance, both are genuine and the heavy coin lies in $\{c_3, c_4, c_5\}$.

Weighing 2 (balanced branch): c_3 versus c_4. If one pan drops, that coin is heavy; if they balance, c_5 is heavy.

Thus two weighings always suffice, and by the lower bound two are necessary in the worst case.

A slicker strategy that also achieves the bound. Weigh $\{c_1, c_2\}$ against $\{c_3, c_4\}$. If a pan drops, the heavy coin is one of that pair and one further weighing settles it; if they balance, c_5 is heavy in one weighing. This uses the three-outcome structure more fully and shows why $\lceil \log_3 5 \rceil = 2$ is tight.

Answer: **(b)** 2.

Question 8. Three alarms ring at intervals of 30, 40 and 50 minutes respectively. If they all ring together at 8:00 AM, the next time they will ring together is **(a)** 10:00 AM **(b)** 11:00 AM **(c)** 12:00 noon **(d)** 1:00 PM.

The concept: simultaneity is an LCM problem

Why the LCM

Suppose event i recurs at times $t_0 + ka_i$ for $k = 0, 1, 2, \ldots$. All three coincide at time $t_0 + T$ precisely when T is a common multiple of a_1, a_2, a_3. The *next* coincidence corresponds to the least positive such T, that is, to $\operatorname{lcm}(a_1, a_2, a_3)$.
The computational rule: factor each a_i into primes and take, for each prime, the *highest* power occurring in any of them.

The computation

$$30 = 2 \cdot 3 \cdot 5, \qquad 40 = 2^3 \cdot 5, \qquad 50 = 2 \cdot 5^2.$$

Highest powers: 2^3 (from 40), 3^1 (from 30), 5^2 (from 50). Hence

$$\mathrm{lcm}(30, 40, 50) = 2^3 \cdot 3 \cdot 5^2 = 8 \cdot 3 \cdot 25 = 600 \text{ minutes} = 10 \text{ hours}.$$

Starting from 8:00 AM, the next simultaneous ring is at

$$8{:}00 \text{ AM} + 10 \text{ h} = 6{:}00 \text{ PM}.$$

This item is defective. The correct answer, 6:00 PM, does not appear among the four options, all of which lie within five hours of the start. The mathematics is not in doubt: 600 minutes is unambiguously the least common multiple, and one may verify directly that $600/30 = 20$, $600/40 = 15$, $600/50 = 12$ are all integers while no smaller positive multiple works (any common multiple must be divisible by 8, by 3 and by 25, hence by 600).
For reference, the listed options correspond to other interval triples: 10:00 AM would follow from intervals with LCM 120 (e.g. $30, 40, 60$), and 12:00 noon from LCM 240. If in the examination hall you meet an item of this shape whose answer is absent, compute the LCM, satisfy yourself it is right, and move on.

Answer: None of the printed options. The correct time is **6:00 PM** (lcm = 600 minutes = 10 hours).

Question 9. Pointing to a photograph, a man said, "I have no brother or sister but that man's father is my father's son." Whose photograph is it? **(a)** His own **(b)** His son's **(c)** His father's **(d)** His uncle's.

The concept: resolve the innermost description first

Kinship puzzles are compositions of functions on a family tree, and like all compositions they must be evaluated from the inside out. Write S for the speaker.

The computation

Step 1: evaluate "my father's son". The sons of S's father are exactly S and S's brothers. The stem tells us S has no brother. Therefore

$$\text{my father's son} = S.$$

This clause exists solely to license that identification; without it the phrase would be ambiguous.

Step 2: substitute. The sentence now reads

$$\text{that man's father} = S.$$

Step 3: invert. If S is the father of the man in the photograph, then the man in the photograph is the son of S.

Where candidates go wrong. The reflex is to read "my father's son" as "my father", giving the photograph to the father, option **(c)**. Another frequent slip is to stop after Step 2 and

answer "his own", option **(a)**, forgetting that S was identified as the *father of* the subject, not the subject.
A reliable discipline: rewrite the sentence with the resolved value substituted in, literally, before drawing any conclusion. "That man's father is me" is unambiguous in a way that the original is not.

Answer: **(b)** His son's.

Question 10. Which of the following figures completes the pattern? **(a)** Figure A **(b)** Figure B **(c)** Figure C **(d)** Figure D.

The figures did not reproduce

The stimulus array and the four candidate figures are not present in the available text, so no determinate answer can be given. What can usefully be given is the method, because figure-completion items are drawn from a very short list of transformation families.

The transformation families

Scan the sequence for one of the following, in roughly this order of frequency.
1. Rotation. A constant angular increment, most often 45° or 90°, applied to the whole figure or to one distinguished element. Track a single asymmetric feature (an arrow-head, a shaded cell) rather than the whole picture.
2. Reflection. Alternation about a vertical or horizontal axis. Diagnostic: elements swap left/right but do not change orientation about their own centres.
3. Accretion or deletion. A fixed number of marks is added or removed at each step. Count, do not eyeball.
4. Positional cycling. A shaded cell advances by a fixed number of positions around a closed circuit, so its position is periodic modulo the circuit length.
5. Superposition rules. In a 3×3 matrix, the third entry of each row is obtained from the first two by a logical rule — union, intersection, or symmetric difference of the marks. Test the rule on a *complete* row before applying it to the incomplete one.
6. Compound rules. Two independent features each following their own rule (e.g. the outer shape rotates while the inner shading cycles). Decompose the figure into features and treat each separately.

Answer: Indeterminate: the figures are absent from the source text.

Question 11. The pie chart shows the distribution of a budget. The largest share is allocated to **(a)** Education **(b)** Health **(c)** Defence **(d)** Infrastructure.

The chart did not reproduce

Without the sector angles or labels no answer is determinable.

Reading a pie chart under time pressure

The single fact that governs every pie-chart item is that the whole disc is $360°$ and represents 100%, so

$$\text{percentage} = \frac{\text{sector angle}}{360°} \times 100, \qquad \text{sector angle} = \frac{\text{percentage}}{100} \times 360°.$$

Three landmarks are worth committing to memory because they let you rank sectors by eye without measuring: a quarter of the disc is $90° = 25\%$; a third is $120° \approx 33.3\%$; and 10% is $36°$, roughly the angle of a thin wedge one-tenth of the way round.
For "largest share" items no arithmetic is needed at all — only a comparison of angles. The arithmetic becomes necessary only when the question asks for a *ratio* of two sectors, or for an absolute amount given the total, in which case

$$\text{amount of sector} = \text{total} \times \frac{\text{sector angle}}{360°}.$$

Answer: Indeterminate: the chart is absent from the source text.

Question 12. If the average of 5 consecutive even numbers is 28, then the largest of these numbers is **(a)** 30 **(b)** 32 **(c)** 34 **(d)** 36.

The concept: the mean of a symmetric arithmetic progression

Mean equals middle term

Let $a_1, \ldots, a_n$ be an arithmetic progression with common difference d. If n is odd, say $n = 2k+1$, the terms are symmetric about the middle term a_{k+1}: indeed $a_{k+1-j} + a_{k+1+j} = 2a_{k+1}$ for each j. Summing over $j = 1, \ldots, k$ and adding a_{k+1} gives

$$\sum_{i=1}^{n} a_i = (2k+1)\, a_{k+1} = n\, a_{k+1},$$

so the arithmetic mean is exactly the middle term. Consecutive even numbers form an AP with $d = 2$, so the result applies.

The computation

There are 5 terms, so the middle (third) term equals the mean, namely 28. The five terms are therefore

$$24,\ 26,\ \mathbf{28},\ 30,\ 32,$$

and the largest is 32.

Verification. $24 + 26 + 28 + 30 + 32 = 140$ and $140/5 = 28$. ✓

Algebraic route, for the sceptical. Let the numbers be $x, x+2, x+4, x+6, x+8$. Their sum is $5x + 20$ and their mean is $x + 4 = 28$, so $x = 24$ and the largest is $x + 8 = 32$.

The trap. Option **(a)**, 30, is the fourth term — the answer of a candidate who takes "largest" one place too early, or who mistakenly uses four terms. Option **(d)**, 36, arises from taking 28 as the *smallest* term rather than the mean.

Answer: **(b)** 32.

Question 13. From the given bar diagram, the year with the highest production is **(a)** 2018 **(b)** 2019 **(c)** 2020 **(d)** 2021.

The diagram did not reproduce

Bar-diagram discipline

Three habits prevent almost all bar-chart errors.
Check the axis origin. Charts whose vertical axis starts above zero exaggerate differences enormously. A bar twice as tall may represent a 5% difference. For "which is largest" this does not matter; for "by what factor" it matters entirely.
Distinguish level from change. "Highest production" asks for the tallest bar. "Greatest increase" asks for the largest *gap* between consecutive bars, which is often at a different year, and "greatest percentage increase" asks for the largest ratio, which is usually at a third year again — typically early, when the base is small.
Read the legend on grouped charts. Where each year carries two or more bars, confirm whether the question concerns one series, or their sum, or their difference.

Answer: Indeterminate: the diagram is absent from the source text.

Question 14. In a class of 50 students, 30 play cricket, 25 play football and 10 play both. The number of students who play neither is **(a)** 5 **(b)** 10 **(c)** 15 **(d)** 20.

The concept: inclusion–exclusion for two sets

Inclusion–exclusion, $n = 2$

For finite sets A, B,

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

The subtraction is necessary because elements of $A \cap B$ are counted once in $|A|$ and again in $|B|$; removing $|A \cap B|$ restores each to a single count.
If A, B lie inside a universe U, then by De Morgan $A^c \cap B^c = (A \cup B)^c$, so

$$|\text{neither}| = |U| - |A \cup B|.$$

The computation

Let C be the set of cricket players and F the set of football players, inside a universe of $|U| = 50$ students. We are given $|C| = 30$, $|F| = 25$, $|C \cap F| = 10$. Then

$$|C \cup F| = 30 + 25 - 10 = 45,$$

and hence

$$|\text{neither}| = 50 - 45 = 5.$$

The four-cell breakdown, worth writing out because most follow-up questions ask for one of the other cells:

cricket only	$30 - 10 = 20$
football only	$25 - 10 = 15$
both	10
neither	5
total	50

The trap. Option **(c)**, 15, is "football only"; option **(d)**, 20, is "cricket only". Both are correct answers to questions that were not asked. Option **(b)**, 10, is the intersection, i.e. the datum itself.

Answer: **(a)** 5.

Question 15. A process that increases the entropy of the surroundings is **(a)** Endothermic **(b)** Exothermic **(c)** Isothermal **(d)** Adiabatic.

The concept: entropy transfer by heat

The Clausius relation for the surroundings

For a reversible transfer of heat δq into a body at absolute temperature T, the body's entropy changes by $\mathrm{d}S = \delta q_{\text{rev}}/T$.
Treat the surroundings as a reservoir large enough that its temperature does not change. Then for a process in which the *system* absorbs heat q_{sys} at constant pressure — so that $q_{\text{sys}} = \Delta H$, the enthalpy change — the surroundings lose exactly that heat, and

$$\Delta S_{\text{surr}} = -\frac{\Delta H_{\text{sys}}}{T}.$$

The minus sign is the whole point: heat leaving the system is heat entering the surroundings.

The reasoning

We require $\Delta S_{\text{surr}} > 0$. Since $T > 0$ always, the formula gives

$$\Delta S_{\text{surr}} > 0 \iff -\Delta H_{\text{sys}} > 0 \iff \Delta H_{\text{sys}} < 0.$$

A process with $\Delta H_{\text{sys}} < 0$ releases heat to its surroundings; by definition this is an *exothermic* process.

Discrimination

(a) *Endothermic* has $\Delta H_{\text{sys}} > 0$, so it *absorbs* heat from the surroundings and *decreases* their entropy. This is the exact opposite and is the intended trap.

(c) *Isothermal* specifies only that T is constant; it says nothing about the sign or even the existence of heat flow. An isothermal expansion increases system entropy and decreases that of the surroundings; an isothermal compression does the reverse. The classification is orthogonal to the question.

(d) *Adiabatic* means $\delta q = 0$ by definition, so $\Delta S_{\text{surr}} = 0$ exactly. This is the cleanest elimination in the set: no heat crosses the boundary, so the surroundings' entropy cannot change at all.

Answer: **(b)** Exothermic.

Question 16. A person walks 5 km north, then 3 km east, then 2 km south. The distance from the starting point is **(a)** $\sqrt{10}$ km **(b)** $\sqrt{13}$ km **(c)** $\sqrt{34}$ km **(d)** 6 km.

The concept: displacement is a vector sum

Distance travelled along the path and distance from the start are different quantities. The first is the sum of the leg lengths ($5 + 3 + 2 = 10$ km here); the second is the magnitude of the vector sum of the legs. Only the latter is asked for.

Set up coordinates with east as $+x$ and north as $+y$, starting at the origin. The three legs are the vectors

$$\mathbf{v}_1 = (0, 5), \qquad \mathbf{v}_2 = (3, 0), \qquad \mathbf{v}_3 = (0, -2).$$

The computation

$$\mathbf{d} = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = (0 + 3 + 0,\ 5 + 0 - 2) = (3,\ 3).$$

The north–south components cancel partially: 5 north minus 2 south leaves 3 north. The east–west direction has a single contribution of 3 east. Therefore

$$|\mathbf{d}| = \sqrt{3^2 + 3^2} = \sqrt{18} = 3\sqrt{2} \approx 4.243 \text{ km}.$$

This item is defective. The value $3\sqrt{2} = \sqrt{18}$ is not among the options. The listed values correspond to different final legs, and it is instructive to see exactly which:

no southward leg	$\mathbf{d} = (3, 5)$	$\lvert\mathbf{d}\rvert = \sqrt{34}$	(option **(c)**)
3 km south	$\mathbf{d} = (3, 2)$	$\lvert\mathbf{d}\rvert = \sqrt{13}$	(option **(b)**)
4 km south	$\mathbf{d} = (3, 1)$	$\lvert\mathbf{d}\rvert = \sqrt{10}$	(option **(a)**)
2 km south (as printed)	$\mathbf{d} = (3, 3)$	$\lvert\mathbf{d}\rvert = \sqrt{18}$	(absent)

Option **(d)**, 6 km, corresponds to no natural variant; it is likely a distractor formed from the arithmetic $5 + 3 - 2$.

Answer: None of the printed options. The correct displacement is $\mathbf{3\sqrt{2}}$ **km** ≈ 4.24 km. (Option **(b)** $\sqrt{13}$ would be correct if the final leg were 3 km south.)

Question 17. In a population of 1000, 40% are males. Of the males 20% and of the females 30% are graduates. The percentage of graduates in the population is **(a)** 24 **(b)** 26 **(c)** 28 **(d)** 30.

The concept: the weighted average, and why the naive average fails

Law of total proportion

If a population is partitioned into groups $G_1, \dots, G_k$ with population shares $w_1, \dots, w_k$ (summing to 1), and the proportion possessing an attribute within group G_i is p_i, then the overall proportion is

$$p = \sum_{i=1}^{k} w_i p_i.$$

This is the law of total probability written for proportions. The weights are the *group shares*, not the group rates, and the commonest error in the entire topic is to average the p_i without weights.

The computation, by headcount

Working in absolute numbers is the safest route because it makes each quantity concrete.

$$\text{Males} = 0.40 \times 1000 = 400, \qquad \text{Females} = 1000 - 400 = 600.$$

$$\text{Male graduates} = 0.20 \times 400 = 80, \qquad \text{Female graduates} = 0.30 \times 600 = 180.$$

$$\text{Total graduates} = 80 + 180 = 260 \quad \Longrightarrow \quad \frac{260}{1000} = 26\%.$$

The same computation, by weights

$$p = w_M p_M + w_F p_F = (0.40)(0.20) + (0.60)(0.30) = 0.08 + 0.18 = 0.26.$$

Note that the population size 1000 never entered; the answer is a proportion and is independent of the total. The number 1000 was supplied only to make the headcount route available.

The trap, and a bound worth internalising. The unweighted average $(20 + 30)/2 = 25\%$ is not offered, but a candidate producing it would be drawn to option **(a)**, 24, or **(b)**, 26, at random. The correct value must lie strictly between the two group rates, 20% and 30%, and must lie *closer to the rate of the larger group*. Females are the majority (60%) and have the higher rate (30%), so the answer must exceed 25%. That single observation eliminates **(a)** instantly and, together with the strict upper bound of 30%, leaves only **(b)** and **(c)**.

Answer: **(b)** 26.

Question 18. The mode of the data $2, 3, 4, 5, 5, 6, 7, 8$ is **(a)** 4 **(b)** 5 **(c)** 6 **(d)** 7.

The concept: the three centres, and how they differ

Mean, median, mode

Mode: the value of greatest frequency. It need not be unique (data may be bimodal or multimodal) and need not exist in a useful sense (if all frequencies are equal, every value is a mode, or equivalently there is none). It is the only one of the three that applies to purely nominal data.
Median: the value at the middle of the ordered data. For n odd it is the $\left(\frac{n+1}{2}\right)$th observation; for n even it is the average of the $\frac{n}{2}$th and $\left(\frac{n}{2}+1\right)$th.
Mean: the arithmetic average $\frac{1}{n}\sum x_i$. It is the unique minimiser of $\sum(x_i - c)^2$, whereas the median minimises $\sum |x_i - c|$; this is why the mean is sensitive to outliers and the median is not.

The computation

Tabulate frequencies:

Value	2	3	4	5	6	7	8
Frequency	1	1	1	**2**	1	1	1

Exactly one value, namely 5, occurs more than once. Hence the mode is 5, and the data set is unimodal.

Cross-check with the other centres. There are $n = 8$ observations, so the median is the average of the 4th and 5th ordered values, both of which are 5; the median is 5. The mean is $(2+3+4+5+5+6+7+8)/8 = 40/8 = 5$. So here all three measures of centre coincide at 5 — a coincidence produced by the near-symmetry of the data, and a reassuring confirmation.
The distractors. Option **(a)**, 4, and option **(c)**, 6, are the values immediately flanking the repeated one; option **(d)**, 7, is the value of *highest* magnitude that is not the maximum, and catches a candidate confusing "mode" with a positional statistic.

Answer: **(b)** 5.

Question 19. In an ecosystem, the interaction between a lion and a deer is an example of
(a) Mutualism **(b)** Commensalism **(c)** Predation **(d)** Parasitism.

The concept: classifying interactions by their sign pair

The sign convention

Ecological interactions are classified by the pair of signs recording the effect on the fitness of each participant.

Interaction	Signs	Description
Mutualism	$(+,+)$	both partners benefit
Commensalism	$(+,0)$	one benefits, the other is unaffected
Predation	$(+,-)$	predator benefits, prey is killed
Parasitism	$(+,-)$	parasite benefits, host is harmed but usually not killed
Competition	$(-,-)$	both are harmed
Amensalism	$(0,-)$	one is harmed, the other unaffected

Distinguishing predation from parasitism

Both carry the sign pair $(+,-)$, so the sign convention alone cannot separate them. Three further criteria do:

1. **Fate of the victim.** A predator kills its prey, generally immediately and as part of the act of consumption. A parasite depends on the host remaining alive, since a dead host is a lost habitat.
2. **Relative size.** A predator is typically comparable to or larger than its prey. A parasite is typically much smaller than its host.
3. **Number of victims per lifetime.** A predator consumes many prey individuals over its life. A parasite typically associates with one or a few hosts.

A lion kills the deer, is larger than the deer, and kills many deer over its lifetime. All three criteria point the same way.

Answer: **(c)** Predation.

Question 20. Which of the following is a rheonomic, non-holonomic and conservative system? **(a)** Simple pendulum with fixed support **(b)** Simple pendulum with moving support **(c)** A particle constrained to move on a sphere **(d)** A free particle under gravity.

The concept: the classification of constraints

This item asks for the simultaneous satisfaction of three independent classifications, so all three must be understood separately before the options can be tested.

Holonomic versus non-holonomic

A constraint is *holonomic* if it can be written as an equation among the coordinates and possibly the time,

$$f(\mathbf{r}_1, \ldots, \mathbf{r}_N, t) = 0.$$

Its significance is structural: each independent holonomic constraint reduces the number of degrees of freedom by one and permits the elimination of one coordinate, so that Lagrange's equations may be written in an unconstrained set of generalised coordinates.
A constraint is *non-holonomic* if it cannot be so written. Two families occur:

- *Inequality constraints*, such as $r \leq \ell$ for a bead attached by an inextensible string, or $r \geq a$ for a particle sliding on the outside of a sphere. These bound the configuration space rather than confining it to a submanifold.
- *Non-integrable differential constraints*, such as the rolling condition of a disc on a plane, which relate velocities in a way not obtainable by differentiating any coordinate relation.

Non-holonomic constraints do *not* reduce the number of generalised coordinates, which is precisely why they are harder.

Scleronomic versus rheonomic

A constraint is *scleronomic* if it does not involve the time explicitly: $f(\mathbf{r}) = 0$. It is *rheonomic* if the time appears explicitly: $f(\mathbf{r}, t) = 0$.
The physical content: a scleronomic constraint is a fixed wall or wire; a rheonomic constraint is one that is being moved by an external agency according to a prescribed schedule. The distinction matters because for scleronomic constraints (with velocity-independent potential) the Jacobi integral coincides with the total energy and is conserved, whereas for rheonomic constraints the constraint forces do work and energy is in general not conserved.

Conservative

A system is *conservative* if the applied forces derive from a potential depending only on position, $\mathbf{F} = -\nabla V(\mathbf{r})$. Uniform gravity, with $V = mgz$, is the standard example. Note that "conservative" refers to the *applied* forces; ideal constraint forces do no virtual work and are excluded from the classification.

Testing the options

(a) *Pendulum with fixed support.* The constraint is $x^2 + y^2 = \ell^2$ with ℓ constant: an equation in the coordinates alone. Holonomic and scleronomic. Gravity is conservative. Fails on two counts of three.

(b) *Pendulum with moving support.* If the support is driven along a prescribed path $\mathbf{R}(t)$, the constraint reads $|\mathbf{r} - \mathbf{R}(t)| = \ell$, in which t appears explicitly: *rheonomic.* ✓ If the bob hangs from an inextensible *string* rather than a rigid rod, the physically correct constraint is the inequality

$$|\mathbf{r} - \mathbf{R}(t)| \leq \ell,$$

since a string can go slack but cannot push. An inequality constraint is *non-holonomic.* ✓ The applied force is gravity, hence *conservative.* ✓ All three conditions are met.

(c) *Particle constrained to move on a sphere.* Read as confinement to the surface, $|\mathbf{r}| = a$ is holonomic and scleronomic — fails both. Read as the outside of a sphere, $|\mathbf{r}| \geq a$ is non-holonomic but still scleronomic — fails rheonomic. Under no reading does it satisfy all three.

(d) *Free particle under gravity.* There are no constraints at all, so neither classification applies. Conservative, but vacuously unconstrained.

A point of genuine subtlety. If the pendulum in **(b)** is taken as a rigid rod, the constraint is the *equality* $|\mathbf{r} - \mathbf{R}(t)| = \ell$, which is holonomic (though rheonomic), and then no option satisfies all three conditions. The intended reading is therefore the string, whose slackness makes the constraint an inequality. This is worth knowing because CSIR has used both conventions in different years; when in doubt, choose the reading under which some option is correct.

Answer: **(b)** Simple pendulum with moving support.

CHAPTER 2

Part B — Core Mathematical Sciences

Part B carries forty items of which twenty-five are to be attempted, each with a single correct option. The range is the whole syllabus: analysis, algebra, topology, differential equations, numerical analysis, statistics and operations research. Several items in this section reference data — a matrix, a pair of statements, a design — that did not reproduce in the source text; those are treated according to the policy set out in the front matter.

Question 21. Match List-I with List-II and select the correct answer. **(a)** A-1, B-2, C-3, D-4 **(b)** A-2, B-1, C-4, D-3 **(c)** A-3, B-4, C-1, D-2 **(d)** A-4, B-3, C-2, D-1.

The lists did not reproduce

Neither List-I nor List-II is present, so no matching can be performed.

Strategy for matching items

Matching items are almost never solved by four independent identifications. The efficient procedure exploits the option structure.
Step 1: find your strongest single identification. Locate the one pairing you are certain of — say $B \to 3$.
Step 2: filter. Discard every option inconsistent with it. In a four-option set a single confident pairing typically removes two or three options.
Step 3: break the remaining tie with your second-strongest identification.
Observe that in the option set above, each of the four permutations is a distinct bijection $\{A, B, C, D\} \to \{1, 2, 3, 4\}$, and no two options agree on *any* single letter. Hence one confident pairing determines the answer outright. This is the most favourable possible structure and is worth checking for before doing any work: scan the options first, identify which letter is most discriminating, and solve only that one.

Answer: Indeterminate: the lists are absent from the source text.

Question 22. The solution of the PDE $x(2z+y)z_x - y(2z+x)z_y = (x-y)(2z+x+y)$ with $z(1, \sqrt{8}) = 1$ is **(a)** $z = \frac{x+y}{2}$ **(b)** $z = \frac{xy}{x+y}$ **(c)** $z = \frac{x^2+y^2}{x+y}$ **(d)** $z = \frac{x-y}{2}$.

The concept: Lagrange's method for a quasilinear first-order PDE

Lagrange's equations

Consider $P(x,y,z)z_x + Q(x,y,z)z_y = R(x,y,z)$. The graph $z = z(x,y)$ is a surface whose normal is $(z_x, z_y, -1)$, and the equation says exactly that $(P,Q,R)\cdot(z_x,z_y,-1) = 0$; that is, the vector field (P,Q,R) is *tangent* to the solution surface at every point.
Consequently every solution surface is a union of integral curves of the field, the *characteristics*, obtained from

$$\frac{\mathrm{d}x}{P} = \frac{\mathrm{d}y}{Q} = \frac{\mathrm{d}z}{R}.$$

If $u(x,y,z) = c_1$ and $v(x,y,z) = c_2$ are two functionally independent first integrals of this system, the general solution is $\Phi(u,v) = 0$ for arbitrary Φ, or equivalently $v = \varphi(u)$.

The method of multipliers

First integrals are found by choosing functions ℓ, m, n with

$$\ell P + mQ + nR = 0,$$

for then $\ell\,\mathrm{d}x + m\,\mathrm{d}y + n\,\mathrm{d}z = 0$ along characteristics, and if this form is exact (or made exact) it integrates to a first integral. The art lies entirely in guessing the multipliers, and the productive guesses are $(1,1,1)$, $(1,-1,0)$, $(1/x,1/y,1/z)$, (x,y,z), and (yz,zx,xy).

Applying the method

Here $P = x(2z+y)$, $Q = -y(2z+x)$, $R = (x-y)(2z+x+y)$.

Multiplier set $(1,1,0)$*-adjusted.* Compute

$$P + Q = x(2z+y) - y(2z+x) = 2xz + xy - 2yz - xy = 2z(x-y).$$

Choosing $n = -2z/(2z+x+y)$ makes $P + Q + nR = 2z(x-y) - 2z(x-y) = 0$, giving

$$\mathrm{d}x + \mathrm{d}y - \frac{2z\,\mathrm{d}z}{2z+x+y} = 0. \tag{22.1}$$

Multiplier set $(1/x,\ 1/y,\ \cdot)$. Compute

$$\frac{P}{x} + \frac{Q}{y} = (2z+y) - (2z+x) = y - x,$$

so with $n = 1/(2z+x+y)$ we get $\frac{P}{x} + \frac{Q}{y} + nR = (y-x) + (x-y) = 0$, giving

$$\frac{\mathrm{d}x}{x} + \frac{\mathrm{d}y}{y} + \frac{\mathrm{d}z}{2z+x+y} = 0. \tag{22.2}$$

Introduce $s = x+y$ and $u = 2z+s$. Relation (22.1) reads $\mathrm{d}s = 2z\,\mathrm{d}z/u$, i.e.

$$\frac{\mathrm{d}s}{\mathrm{d}z} = \frac{2z}{2z+s},$$

a homogeneous first-order ODE. Setting $s = vz$,

$$v + z\frac{\mathrm{d}v}{\mathrm{d}z} = \frac{2}{2+v} \quad\Longrightarrow\quad z\frac{\mathrm{d}v}{\mathrm{d}z} = \frac{2-2v-v^2}{2+v},$$

which separates and integrates (after the partial-fraction split $2+v = -\frac{1}{2}(-2-2v)+1$) to

$$-\tfrac{1}{2}\ln\left|2-2v-v^2\right| + \frac{1}{2\sqrt{3}}\ln\left|\frac{\sqrt{3}+1+v}{\sqrt{3}-1-v}\right| = \ln z + C.$$

This item is corrupt. The first integral just obtained involves $\sqrt{3}$ logarithms and bears no relation to the four rational candidates offered. Direct substitution confirms that *none* of the four options satisfies the printed equation. Take option **(b)**, the most plausible-looking: with $z = xy/(x+y)$ one computes $z_x = y^2/(x+y)^2$, $z_y = x^2/(x+y)^2$, whence

$$x(2z+y)z_x - y(2z+x)z_y = \frac{xy\,(y-x)(2z+x+y)}{(x+y)^2},$$

which equals $(x-y)(2z+x+y)$ only if $-xy/(x+y)^2 = 1$, impossible for real $x, y > 0$. A numerical check at $(x,y) = (2,1)$ disposes of option **(c)**: the left side is $166/27$ while the right side is $171/27$.
Nor does the initial condition help: at $(x,y) = (1, \sqrt{8}) = (1, 2\sqrt{2})$ the four candidates take the values 1.914, 0.739, 2.351 and -0.914, none of which is 1. Both the equation and the data are inconsistent with the option set.

Answer: Indeterminate; the item as printed is inconsistent. No listed option satisfies the stated PDE, and none satisfies the stated initial condition. Retain the *method*: form the auxiliary equations, hunt for multipliers, and verify any candidate by direct substitution — substitution is always available and always decisive.

Question 23. For the Markov chain with the given transition matrix, the transient states are **(a)** only 1 **(b)** only 2 **(c)** 1 and 2 **(d)** none.

The transition matrix did not reproduce

Recurrence and transience

Let $\{X_n\}$ be a Markov chain on a countable state space with transition matrix P. For a state i define the first return time $T_i = \inf\{n \geq 1 : X_n = i\}$ and $f_{ii} = \mathbb{P}(T_i < \infty \mid X_0 = i)$.
The state i is *recurrent* if $f_{ii} = 1$ and *transient* if $f_{ii} < 1$. Equivalently, in terms of the n-step probabilities $p_{ii}^{(n)}$:

$$i \text{ recurrent} \iff \sum_{n \geq 1} p_{ii}^{(n)} = \infty, \qquad i \text{ transient} \iff \sum_{n \geq 1} p_{ii}^{(n)} < \infty.$$

The proof of the equivalence is the renewal identity: the expected number of returns is $f_{ii}/(1 - f_{ii})$, finite exactly when $f_{ii} < 1$.

The practical test, via communicating classes

Recurrence is a *class property*: if $i \leftrightarrow j$ then i and j are both recurrent or both transient. So the algorithm is:
1. Draw the directed graph on the states, with an edge $i \to j$ whenever $p_{ij} > 0$.
2. Find the strongly connected components; these are the communicating classes.
3. A class is *closed* if no edge leaves it. In a *finite* chain: every closed class is recurrent, and every non-closed class is transient.

The finiteness matters. On an infinite state space a closed class may still be transient — the simple symmetric random walk on $\mathbb{Z}^3$ is the standard example — and the graph test is no longer sufficient.

What the option set tells us. Options **(a)**, **(b)** and **(c)** between them name only states 1 and 2 as candidates for transience, so the chain has at least three states and states $3, 4, \ldots$ are recurrent. Option **(d)** asserts that the chain is irreducible and recurrent throughout. Since a finite chain must possess at least one recurrent class, the option "all states transient" is correctly absent from the list — a small consistency check on the setter.

Answer: Indeterminate: the transition matrix is absent from the source text.

Question 24. The set $B = \{A \in M_2(\mathbb{R}) : A^2 = A\}$ is **(a)** closed but not open **(b)** open but not closed **(c)** both open and closed **(d)** neither open nor closed.

The concept: preimages of closed sets under continuous maps

Identify $M_2(\mathbb{R})$ with $\mathbb{R}^4$ carrying its usual topology; matrix entries are then coordinates, and any map defined by polynomials in the entries is continuous.

The closedness criterion

If $F : X \to Y$ is continuous and $C \subseteq Y$ is closed, then $F^{-1}(C)$ is closed in X. In particular the solution set of any system of polynomial equations in $\mathbb{R}^n$ — an *algebraic set* — is closed, being the preimage of the closed set $\{0\}$ under a continuous polynomial map.

B is closed

Define $F : M_2(\mathbb{R}) \to M_2(\mathbb{R})$ by $F(A) = A^2 - A$. Each entry of $F(A)$ is a quadratic polynomial in the four entries of A; explicitly, for $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$A^2 - A = \begin{pmatrix} a^2 + bc - a & b(a + d - 1) \\ c(a + d - 1) & d^2 + bc - d \end{pmatrix}.$$

Hence F is continuous, $\{0\}$ is closed in $M_2(\mathbb{R}) \cong \mathbb{R}^4$, and $B = F^{-1}(\{0\})$ is closed.

B is not open: B has empty interior

We show that B contains no open ball, by exhibiting for each $A \in B$ a sequence of non-idempotent matrices converging to A.

Recall that idempotents satisfy $A^2 = A$, so their minimal polynomials divide $x^2 - x = x(x - 1)$ and all eigenvalues lie in $\{0, 1\}$. In particular $\det A \in \{0, 1\}$ and $\operatorname{tr} A \in \{0, 1, 2\}$ — already a severe restriction on a four-dimensional space.

Now let $A \in B$ and consider $A_t = (1 + t)A$ for small $t \neq 0$. Then

$$A_t^2 - A_t = (1 + t)^2 A^2 - (1 + t)A = (1 + t)A\big[(1 + t) - 1\big] = t(1 + t)A.$$

If $A \neq 0$ this is nonzero for all small $t \neq 0$, so $A_t \notin B$ while $A_t \to A$. If $A = 0$, take instead $A_t = tI$, for which $A_t^2 - A_t = (t^2 - t)I \neq 0$ for $0 < t < 1$. Either way, every neighbourhood of every point of B contains points outside B, so B has empty interior and is certainly not open.

What B actually looks like. Classify 2×2 idempotents by rank. Rank 0 gives $A = 0$; rank 2 forces A invertible, and $A^2 = A$ then gives $A = I$. Rank 1 idempotents are the projections onto a line along a complementary line, parametrised by $\operatorname{tr} A = 1$ and $\det A = 0$:

$$A = \begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}, \qquad a(1-a) - bc = 0.$$

This is a quadric surface, a two-dimensional submanifold of $\mathbb{R}^4$. So

$$B = \{0\} \sqcup \{I\} \sqcup \{\text{a 2-dimensional quadric}\},$$

a closed set of measure zero. It is not connected — $\{0\}$ and $\{I\}$ are isolated points — and it is unbounded, since b may be taken arbitrarily large with $a = 1$, $c = 0$.

Eliminating (c). $\mathbb{R}^4$ is connected, so its only clopen subsets are $\emptyset$ and $\mathbb{R}^4$ itself. Since B is neither, it cannot be both open and closed. This single observation kills option **(c)** without any computation.

Answer: **(a)** closed but not open.

Question 25. The Gauss–Seidel iteration matrix for the system has trace equal to **(a)** 0 **(b)** 1/2 **(c)** 1/3 **(d)** 1/4.

The system did not reproduce

The splitting framework

Write $A = D - L - U$ where D is the diagonal of A, $-L$ its strictly lower triangular part and $-U$ its strictly upper triangular part (so L and U are strictly triangular with the signs absorbed). To solve $Ax = b$ iteratively, split $A = M - N$ and iterate

$$Mx^{(k+1)} = Nx^{(k)} + b \quad \Longrightarrow \quad x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b.$$

The *iteration matrix* is $T = M^{-1}N$, and the error satisfies $e^{(k+1)} = Te^{(k)}$, so $e^{(k)} = T^k e^{(0)}$. Convergence for every starting vector holds if and only if the spectral radius satisfies $\rho(T) < 1$.

Jacobi: $M = D$, $N = L + U$, so $T_J = D^{-1}(L + U)$.

Gauss–Seidel: $M = D - L$, $N = U$, so $T_{GS} = (D - L)^{-1}U$.

What can be said about the trace without the matrix

Proposition 2.1. For Gauss–Seidel, the $(1,1)$ entry of T_{GS} is always 0. Consequently T_{GS} always has 0 as an eigenvalue, and for a 2×2 system with $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\operatorname{tr} T_{GS} = \frac{bc}{ad}.$$

Proof. $(D-L)^{-1}$ is lower triangular and U is strictly upper triangular, so

$$(T_{GS})_{ii} = \sum_k \left[(D-L)^{-1}\right]_{ik} U_{ki}.$$

Since U is strictly upper triangular, $U_{ki}=0$ unless $k<i$; and since $(D-L)^{-1}$ is lower triangular, its (i,k) entry vanishes unless $k \le i$. For $i=1$ there is no $k<1$, so $(T_{GS})_{11}=0$.

For the 2×2 case, $D-L = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$ and $U = \begin{pmatrix} 0 & -b \\ 0 & 0 \end{pmatrix}$, whence

$$(D-L)^{-1} = \frac{1}{ad}\begin{pmatrix} d & 0 \\ -c & a \end{pmatrix}, \qquad T_{GS} = \frac{1}{ad}\begin{pmatrix} 0 & -bd \\ 0 & bc \end{pmatrix}.$$

The trace is $bc/(ad)$. □

Reading the options. For a 2×2 system the trace equals $bc/(ad)$, which also equals the nonzero eigenvalue, so $\rho(T_{GS}) = |bc/(ad)|$ and the method converges precisely when $|bc| < |ad|$ — exactly the condition of strict diagonal dominance in this small case. Note also the classical relation $\rho(T_{GS}) = \rho(T_J)^2$ for 2×2 systems, since $\rho(T_J)^2 = |bc/(ad)|$.
Options **(b)**, **(c)**, **(d)** are the values $1/2$, $1/3$, $1/4$, all of which correspond to convergent iterations; option **(a)**, 0, would require $bc=0$, i.e. a triangular system, for which Gauss–Seidel converges in a single step.

Answer: Indeterminate: the system is absent from the source text. For a 2×2 system the trace is $bc/(ad)$.

Question 26. If T is a nilpotent operator on $P_{30}(\mathbb{R})$ with $T^2=0$, the maximum possible rank of T is **(a)** 10 **(b)** 15 **(c)** 20 **(d)** 25.

The concept: what $T^2=0$ says about the subspace lattice

Rank–nullity and the square-zero condition

For a linear operator T on a finite-dimensional space V,

$$\operatorname{rank} T + \dim\ker T = \dim V. \qquad \text{(Rank–nullity)}$$

The condition $T^2=0$ has an exact geometric translation:

$$T^2 = 0 \iff T(Tv)=0 \ \ \forall v \iff \operatorname{im} T \subseteq \ker T.$$

That is, *the image is contained in the kernel.* Everything the operator produces, it subsequently annihilates.

The computation

First fix the dimension. $P_{30}(\mathbb{R})$ denotes real polynomials of degree at most 30, with basis $1, x, x^2, \ldots, x^{30}$; hence

$$\dim P_{30}(\mathbb{R}) = 31.$$

This off-by-one is the most common error in the item and it changes the answer.

Write $r = \operatorname{rank} T$. From $\operatorname{im} T \subseteq \ker T$,

$$r = \dim \operatorname{im} T \ \leq \ \dim \ker T = 31 - r,$$

so $2r \leq 31$, giving $r \leq 15.5$. Since r is an integer,

$$r \leq 15.$$

Attainability

The bound is achieved. Order the basis as $e_1, \ldots, e_{15}, f_1, \ldots, f_{15}, g$ (any relabelling of the 31 monomials) and define

$$T(e_i) = f_i \ (1 \leq i \leq 15), \qquad T(f_i) = 0, \qquad T(g) = 0.$$

Then $\operatorname{im} T = \operatorname{span}\{f_1, \ldots, f_{15}\}$ has dimension 15, and $\ker T = \operatorname{span}\{f_1, \ldots, f_{15}, g\}$ has dimension 16, so indeed $\operatorname{im} T \subseteq \ker T$ and $T^2 = 0$. The rank is 15.

The Jordan picture. An operator with $T^2 = 0$ has Jordan form consisting solely of blocks of size 1 and 2 with eigenvalue 0. If there are k blocks of size 2 and m of size 1, then $2k + m = 31$ and $\operatorname{rank} T = k$. Maximising k means minimising $m \geq 0$; since 31 is odd, $m \geq 1$ and $k \leq 15$. The same bound, seen structurally.
The general statement. On an n-dimensional space, $T^2 = 0$ forces $\operatorname{rank} T \leq \lfloor n/2 \rfloor$. Here $\lfloor 31/2 \rfloor = 15$. A candidate who takes $\dim P_{30} = 30$ obtains $\lfloor 30/2 \rfloor = 15$ as well — the same answer by luck — but the same slip on P_{29} or P_{31} would be fatal.

Answer: **(b)** 15.

Question 27. The particular integral of the PDE $z_{xx} - z_{yy} = \mathrm{e}^{x+y}$ is **(a)** $\frac{\mathrm{e}^{x+y}}{2}$ **(b)** $\frac{\mathrm{e}^{x+y}}{4}$ **(c)** $\frac{x\mathrm{e}^{x+y}}{2}$ **(d)** $\frac{y\mathrm{e}^{x+y}}{2}$.

The concept: operator methods and the resonant case

Write $D = \partial/\partial x$ and $D' = \partial/\partial y$. The equation is

$$(D^2 - D'^2)z = \mathrm{e}^{x+y}, \qquad \text{i.e.} \qquad (D - D')(D + D')z = \mathrm{e}^{x+y}.$$

The exponential shift rule and its failure

For a polynomial operator $F(D, D')$ with constant coefficients,

$$F(D, D')\, \mathrm{e}^{ax+by} = F(a, b)\, \mathrm{e}^{ax+by},$$

because each differentiation simply multiplies by a or b. Hence the particular integral is

$$\frac{1}{F(D, D')} \mathrm{e}^{ax+by} = \frac{\mathrm{e}^{ax+by}}{F(a, b)} \qquad \text{provided } F(a, b) \neq 0.$$

When $F(a,b)=0$ the formula fails: the forcing term is a solution of the homogeneous equation, and the response is *resonant*. One then factors out the offending linear factor and multiplies by x (or by y), exactly as in the ODE case where e^{ax} forcing a system with root a produces xe^{ax}.

Here $F(D,D')=D^2-D'^2$, so $F(1,1)=1-1=0$: we are in the resonant case, and the naive answers $e^{x+y}/2$ and $e^{x+y}/4$ (options **(a)** and **(b)**) cannot be right. They are precisely what a candidate obtains by ignoring the resonance and dividing by something plausible.

The computation, done twice

Route 1: peel off the factors. Since $(D-D')e^{x+y}=e^{x+y}-e^{x+y}=0$, the resonance sits in the factor $D-D'$; the factor $D+D'$ is harmless, with $F(1,1)=1+1=2$. So

$$z_p=\frac{1}{(D-D')}\cdot\frac{1}{(D+D')}e^{x+y}=\frac{1}{(D-D')}\cdot\frac{e^{x+y}}{2},$$

and the resonant inversion of $D-D'$ contributes a factor x:

$$z_p=\frac{x\,e^{x+y}}{2}.$$

Route 2: verify by substitution. This is the only step that requires no faith. Put $z=cxe^{x+y}$.

$$z_x=ce^{x+y}+cxe^{x+y}=c(1+x)e^{x+y},$$
$$z_{xx}=ce^{x+y}+c(1+x)e^{x+y}=c(2+x)e^{x+y},$$
$$z_y=cxe^{x+y},\qquad z_{yy}=cxe^{x+y}.$$

Hence

$$z_{xx}-z_{yy}=c(2+x)e^{x+y}-cxe^{x+y}=2c\,e^{x+y}.$$

Setting $2c=1$ gives $c=\frac{1}{2}$ and $z_p=\frac{1}{2}xe^{x+y}$.

Why option (d) is also a legitimate particular integral — and why it is still not the answer. Try $z=cye^{x+y}$: then $z_{xx}=cye^{x+y}$ and $z_{yy}=c(2+y)e^{x+y}$, so $z_{xx}-z_{yy}=-2c\,e^{x+y}$, requiring $c=-\frac{1}{2}$. Thus $-\frac{1}{2}ye^{x+y}$ is also a particular integral, but with the *opposite* sign to option **(d)**. Option **(d)** as printed, $+\frac{1}{2}ye^{x+y}$, satisfies $z_{xx}-z_{yy}=-e^{x+y}$ and is therefore wrong.

The two genuine particular integrals differ by $\frac{1}{2}(x+y)e^{x+y}$, which one checks is a solution of the homogeneous equation — as it must be, since particular integrals are unique only modulo the complementary function.

Answer: (c) $\dfrac{xe^{x+y}}{2}$.

Question 28. The singular solution of $p=z+q$ is **(a)** $z=ce^{x}$ **(b)** $z=0$ **(c)** $z=e^{x+c}$ **(d)** $z=c$.

The concept: complete, general and singular integrals

The three kinds of integral

For a first-order PDE $F(x, y, z, p, q) = 0$:
Complete integral: a two-parameter family $f(x, y, z, a, b) = 0$ of solutions.
General integral: obtained by setting $b = \varphi(a)$ for an arbitrary function φ and eliminating a between $f = 0$ and $\partial f/\partial a = 0$; this is the envelope of a one-parameter subfamily.
Singular integral: the envelope of the *entire* two-parameter family, obtained by eliminating a and b from

$$f = 0, \qquad \frac{\partial f}{\partial a} = 0, \qquad \frac{\partial f}{\partial b} = 0.$$

Geometrically it is a surface tangent to every member of the family, and it is a genuine solution not belonging to the family. If the elimination is impossible, there is no singular integral.

Finding a complete integral

The equation $p = z + q$ involves p, q, z but not x, y explicitly, which is the standard type $F(z, p, q) = 0$. For such equations one seeks a solution depending on x and y only through $u = x + ay$:

$$z = z(u), \qquad p = z'(u), \qquad q = a\, z'(u).$$

Substituting:

$$z' = z + az' \quad \Longrightarrow \quad (1-a)z' = z \quad \Longrightarrow \quad \frac{\mathrm{d}z}{z} = \frac{\mathrm{d}u}{1-a} \quad (a \neq 1).$$

Integrating,

$$\ln z = \frac{u}{1-a} + \text{const}, \qquad \text{i.e.} \qquad \boxed{z = b\,\exp\left(\frac{x+ay}{1-a}\right)}$$

with two arbitrary constants a and b. This is a complete integral.

Finding the singular integral

Write $f(x, y, z, a, b) = z - b\exp\left(\frac{x+ay}{1-a}\right) = 0$. Differentiate with respect to the parameters:

$$\frac{\partial f}{\partial b} = -\exp\left(\frac{x+ay}{1-a}\right) = 0.$$

An exponential never vanishes, so this equation has no solution — *unless* we allow the degenerate branch obtained by letting $b \to 0$, which collapses the whole family to

$$z \equiv 0.$$

And $z \equiv 0$ is indeed a solution: $p = q = 0$ and $z = 0$, so $p = z + q$ reads $0 = 0 + 0$. ✓

Thus the singular solution is the identically zero surface, the limiting envelope of the family as the amplitude parameter degenerates.

Discrimination

(a) $z = ce^x$: this is a member of the complete integral, obtained by setting $a = 0$, $b = c$. It is a particular solution, not the singular one.

(c) $z = \mathrm{e}^{x+c} = \mathrm{e}^c\mathrm{e}^x$: the same family as **(a)** with the constant written multiplicatively. Again a member, not the envelope.

(d) $z = c$ constant: then $p = q = 0$, so the equation demands $0 = c + 0$, forcing $c = 0$. So **(d)** is correct *only* in the case $c = 0$, which is option **(b)** stated precisely. Option **(b)** is the exact statement.

Answer: **(b)** $z = 0$.

Question 29. The value of $z(2, 0)$ for the given Cauchy problem is **(a)** 1 **(b)** 2 **(c)** e **(d)** e^2.

The Cauchy problem did not reproduce

Solving a Cauchy problem by characteristics

For $a(x, y, z)z_x + b(x, y, z)z_y = c(x, y, z)$ with data $z = z_0(s)$ prescribed on an initial curve $\Gamma :\ (x_0(s), y_0(s))$, the procedure is:
1. Solve the characteristic ODE system

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = b, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = c,$$

with initial conditions $x(0) = x_0(s)$, $y(0) = y_0(s)$, $z(0) = z_0(s)$. This gives $x = X(s, t)$, $y = Y(s, t)$, $z = Z(s, t)$.
2. Invert the map $(s, t) \mapsto (x, y)$ and substitute into Z.
3. Evaluate at the required point.

The transversality condition

Step 2 is possible near $t = 0$ precisely when the Jacobian

$$J = \det \begin{pmatrix} x_0'(s) & a \\ y_0'(s) & b \end{pmatrix} = x_0'(s)\, b - y_0'(s)\, a$$

is nonzero on Γ; that is, when the initial curve is nowhere tangent to the characteristic direction (a, b). If Γ is characteristic at a point, the solution either fails to exist there or fails to be unique. This is the single most frequently examined point in the theory of first-order PDEs.

Reading the options. The presence of e and e^2 among the choices signals that the characteristic equation for z is of the form $\mathrm{d}z/\mathrm{d}t = \lambda z$, whose solution is exponential, rather than $\mathrm{d}z/\mathrm{d}t =$ constant, which would yield the linear values 1 and 2. The target point $(2, 0)$ lying on the x-axis further suggests initial data given on $x = 0$ or on $y = 0$.

Answer: Indeterminate: the Cauchy problem is absent from the source text.

Question 30. Which of the following is a holomorphic function on $\mathbb{C}$? **(a)** $|z|^2$ **(b)** $\bar{z}$ **(c)** e^z **(d)** $z^2 + \bar{z}$.

The concept: the Cauchy–Riemann equations

Complex differentiability

Write $f = u + \mathrm{i}v$ with $z = x + \mathrm{i}y$. Then f is complex-differentiable at a point if and only if u, v are real-differentiable there and satisfy the Cauchy–Riemann equations

$$u_x = v_y, \qquad u_y = -v_x.$$

f is *holomorphic* on an open set if it is complex-differentiable at every point of that set; *entire* means holomorphic on all of $\mathbb{C}$.
An equivalent and often faster formulation uses the Wirtinger derivative

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}\right),$$

for which the Cauchy–Riemann equations are exactly $\partial f/\partial \bar{z} = 0$. Slogan: *a holomorphic function is one that does not depend on* $\bar{z}$. Any explicit occurrence of $\bar{z}$, of $|z|$, of $\operatorname{Re} z$ or of $\operatorname{Im} z$ that does not cancel is fatal.

Testing each option

(a) $f(z) = |z|^2 = z\bar{z} = x^2 + y^2$. Here $u = x^2 + y^2$, $v = 0$. Then $u_x = 2x$ and $v_y = 0$, so the first CR equation forces $x = 0$; $u_y = 2y$ and $-v_x = 0$ force $y = 0$. Thus f is complex-differentiable *only at the origin* and nowhere holomorphic (holomorphy requires differentiability on an open set, and $\{0\}$ is not open). Wirtinger: $\partial(z\bar{z})/\partial\bar{z} = z$, vanishing only at $z = 0$.

(b) $f(z) = \bar{z} = x - \mathrm{i}y$. Here $u = x$, $v = -y$, so $u_x = 1$ while $v_y = -1$: the first CR equation fails everywhere. Nowhere differentiable. Wirtinger: $\partial\bar{z}/\partial\bar{z} = 1 \neq 0$.

(c) $f(z) = \mathrm{e}^z = \mathrm{e}^x \cos y + \mathrm{i}\,\mathrm{e}^x \sin y$. Then

$$u_x = \mathrm{e}^x \cos y, \quad v_y = \mathrm{e}^x \cos y, \quad u_y = -\mathrm{e}^x \sin y, \quad v_x = \mathrm{e}^x \sin y,$$

so $u_x = v_y$ and $u_y = -v_x$ at every point of $\mathbb{C}$. The partials are continuous, so f is entire, with $f'(z) = u_x + \mathrm{i}v_x = \mathrm{e}^x \cos y + \mathrm{i}\mathrm{e}^x \sin y = \mathrm{e}^z$.

(d) $f(z) = z^2 + \bar{z}$. Wirtinger: $\partial f/\partial\bar{z} = 0 + 1 = 1 \neq 0$, so nowhere differentiable. (Directly: the z^2 part is entire and the $\bar{z}$ part is nowhere differentiable, and a sum of a differentiable and a non-differentiable function is non-differentiable.)

A note on the printed option (d). If the source is read as $z^2 + z$ rather than $z^2 + \bar{z}$, then **(d)** would be a polynomial and hence entire, giving two correct options. The intended reading must be $z^2 + \bar{z}$, since otherwise the item is ill-posed. The exponential e^z remains the unambiguous answer under either reading only if **(d)** is read with the conjugate; note this ambiguity when reviewing.

Answer: **(c)** e^z.

Question 31. The radius of convergence of the power series $\sum n!\, z^n$ is **(a)** 0 **(b)** 1 **(c)** ∞ **(d)** $1/\mathrm{e}$.

The concept: the two standard formulae

Cauchy–Hadamard and the ratio test

For $\sum a_n z^n$ the radius of convergence is given universally by the **Cauchy–Hadamard formula**

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n},$$

with the conventions $1/0 = \infty$ and $1/\infty = 0$. The series converges absolutely for $|z| < R$, diverges for $|z| > R$, and may do either on $|z| = R$.
When the limit exists, the **ratio formula**

$$\frac{1}{R} = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$$

is usually easier, and for factorials it is enormously easier.

The computation

Here $a_n = n!$. The ratio test gives

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)!}{n!} = n+1 \longrightarrow \infty,$$

so $1/R = \infty$ and hence $R = 0$.

Root test, for confirmation. By Stirling, $n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$, so

$$(n!)^{1/n} \sim \frac{n}{\mathrm{e}}\,(2\pi n)^{1/(2n)} \sim \frac{n}{\mathrm{e}} \longrightarrow \infty,$$

again giving $R = 0$.

What $R = 0$ means

The series converges only at $z = 0$, where it is the single term $a_0 = 0! = 1$. For any $z \neq 0$ the terms $n!\,z^n$ satisfy $|n!z^n| \to \infty$, so the necessary condition for convergence (terms tending to zero) fails outright.

Why the other options are what they are. $R = 1$ arises from $\sum z^n$; $R = \infty$ from $\sum z^n/n!$, the exponential, which converges everywhere; and $R = 1/\mathrm{e}$ from $\sum n^n z^n/n!$, whose coefficients satisfy $(n^n/n!)^{1/n} \to \mathrm{e}$. These four are the standard family, and it is worth memorising the whole quartet:

$$\sum n! z^n : \ R = 0, \qquad \sum z^n : \ R = 1, \qquad \sum \frac{z^n}{n!} : \ R = \infty, \qquad \sum \frac{n!}{n^n} z^n : \ R = \mathrm{e}.$$

The last of these is the subject of Question 37 below.

Answer: **(a)** 0.

Question 32. In logistic regression the odds ratio corresponding to a unit increase in the predictor is **(a)** e^{β} **(b)** β **(c)** $1+\beta$ **(d)** $\beta/(1-\beta)$.

The concept: the logit link and why it is chosen

Odds, log-odds, and the logistic model

For a probability $p \in (0,1)$ the *odds* are $p/(1-p)$, a number in $(0,\infty)$, and the *log-odds* or *logit* is

$$\operatorname{logit}(p) = \ln \frac{p}{1-p} \in (-\infty, \infty).$$

The logit is chosen as the link function precisely because it maps the bounded interval $(0,1)$ onto the whole real line, so a linear predictor $\beta_0 + \beta_1 x$ can be fitted without any risk of producing a fitted probability outside $[0,1]$.
The logistic regression model is

$$\ln \frac{p(x)}{1-p(x)} = \beta_0 + \beta_1 x \qquad \Longleftrightarrow \qquad p(x) = \frac{\mathrm{e}^{\beta_0+\beta_1 x}}{1+\mathrm{e}^{\beta_0+\beta_1 x}}.$$

The computation

Let $\Omega(x) = p(x)/(1-p(x))$ denote the odds at predictor value x. From the model,

$$\Omega(x) = \exp(\beta_0 + \beta_1 x).$$

The odds ratio comparing $x+1$ to x is

$$\mathrm{OR} = \frac{\Omega(x+1)}{\Omega(x)} = \frac{\exp\big(\beta_0 + \beta_1(x+1)\big)}{\exp(\beta_0 + \beta_1 x)} = \exp(\beta_1).$$

Two features deserve emphasis. First, the answer is e^{β_1} and not β_1: the coefficient lives on the log-odds scale and must be exponentiated to be interpreted on the odds scale. Second, the ratio is *independent of* x — the odds are multiplied by the same factor e^{β_1} regardless of the starting value. This constancy is the defining interpretive property of the logit link and the reason logistic coefficients are reported as odds ratios.

Reading the sign. $\beta_1 > 0$ gives $\mathrm{e}^{\beta_1} > 1$: the odds increase. $\beta_1 = 0$ gives $\mathrm{OR} = 1$: no association. $\beta_1 < 0$ gives $\mathrm{e}^{\beta_1} < 1$: the odds decrease. Hence testing $H_0 : \beta_1 = 0$ is the same as testing $H_0 : \mathrm{OR} = 1$, which is why confidence intervals for odds ratios are computed on the log scale and then exponentiated (producing asymmetric intervals about the point estimate).
The distractors. Option **(d)**, $\beta/(1-\beta)$, mimics the algebraic *form* of odds but applies it to the coefficient rather than to a probability — a category error, since β need not lie in $(0,1)$. Option **(c)**, $1+\beta$, is the first-order approximation $\mathrm{e}^{\beta} \approx 1 + \beta$, valid only for small β and not the exact answer.

Answer: **(a)** e^{β}.

Question 33. For the ANOVA model, the residual sum of squares has degrees of freedom
(a) $n-1$ **(b)** $n-k$ **(c)** $k-1$ **(d)** $n-k+1$.

The concept: degrees of freedom as a dimension count

The orthogonal decomposition

In one-way ANOVA with k groups and n observations in total, write y_{ij} for observation j in group i. The fundamental identity is

$$\underbrace{\sum_{i,j}(y_{ij}-\bar{y})^2}_{SS_{\text{total}}} = \underbrace{\sum_{i} n_i(\bar{y}_i-\bar{y})^2}_{SS_{\text{between}}} + \underbrace{\sum_{i,j}(y_{ij}-\bar{y}_i)^2}_{SS_{\text{within}}=SS_{\text{residual}}} .$$

The cross-term vanishes because the deviations within each group sum to zero — this is Pythagoras in $\mathbb{R}^n$, with the three sums of squares being squared lengths of mutually orthogonal projections.

Degrees of freedom are the dimensions of the corresponding subspaces.

- Total: the data lie in $\mathbb{R}^n$; subtracting the grand mean projects onto the orthogonal complement of the all-ones vector, dimension $n-1$.
- Between: the group-mean vector lies in the k-dimensional space spanned by the group indicators; removing the grand mean leaves $k-1$.
- Within: the residual space is the orthogonal complement of the group-indicator space in $\mathbb{R}^n$, of dimension $n-k$.

And indeed $(n-1)=(k-1)+(n-k)$, matching the sum-of-squares identity term by term.

The answer

The residual (within-groups) sum of squares has $n-k$ degrees of freedom.

Why this matters beyond bookkeeping. Under normality with common variance σ^2,

$$\frac{SS_{\text{within}}}{\sigma^2} \sim \chi^2_{n-k}, \qquad \frac{SS_{\text{between}}}{\sigma^2} \sim \chi^2_{k-1} \text{ under } H_0,$$

and the two are independent by Cochran's theorem (orthogonal projections of a spherical Gaussian). The test statistic is therefore

$$F = \frac{SS_{\text{between}}/(k-1)}{SS_{\text{within}}/(n-k)} \sim F_{k-1,\,n-k} \quad \text{under } H_0.$$

Getting the degrees of freedom wrong invalidates the entire test, which is why this apparently clerical question is asked so often.

The unbiased estimator of σ^2 is $\hat{\sigma}^2 = SS_{\text{within}}/(n-k)$, the mean square error. Dividing by n instead of $n-k$ gives the maximum-likelihood estimator, which is biased downwards.

Answer: **(b)** $n-k$.

Question 34. A sufficient statistic for the parameter of the exponential family is **(a)** the sample mean **(b)** the sample variance **(c)** the order statistics **(d)** the likelihood ratio.

The concept: sufficiency and the factorisation theorem

Sufficiency

A statistic $T = T(X_1, \ldots, X_n)$ is *sufficient* for θ if the conditional distribution of the sample given $T = t$ does not depend on θ. Informally: once T is known, the remaining randomness in the data carries no further information about θ.

Fisher–Neyman factorisation theorem. T is sufficient for θ if and only if the joint density factors as

$$f(\mathbf{x}; \theta) = g\big(T(\mathbf{x}), \theta\big)\, h(\mathbf{x}),$$

where h does not involve θ and g depends on the data only through T.

The exponential family

A family of densities is a k-parameter exponential family if

$$f(x; \boldsymbol{\theta}) = h(x)\, c(\boldsymbol{\theta})\, \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta})\, t_j(x)\right).$$

For an i.i.d. sample the joint density is

$$f(\mathbf{x}; \boldsymbol{\theta}) = \Big(\prod_i h(x_i)\Big) c(\boldsymbol{\theta})^n \exp\left(\sum_{j=1}^{k} w_j(\boldsymbol{\theta}) \sum_{i=1}^{n} t_j(x_i)\right),$$

so by factorisation the vector $T = \big(\sum_i t_1(x_i), \ldots, \sum_i t_k(x_i)\big)$ is sufficient. For a one-parameter family with $t(x) = x$ — which covers the exponential, Poisson, Bernoulli, and normal-with-known-variance cases — the sufficient statistic is $\sum_i X_i$, equivalently the sample mean $\bar{X}$, since the two are in one-to-one correspondence for fixed n.

Discrimination

(a) *Sample mean.* Correct for the standard one-parameter exponential family with natural statistic $t(x) = x$. It is moreover *minimal* sufficient and, for a full-rank family, complete.

(b) *Sample variance.* Sufficient for σ^2 in a normal model with *known* mean, but not for the parameter of the one-parameter exponential family in question. In the $\text{Exp}(\lambda)$ model, S^2 is not sufficient.

(c) *Order statistics.* Always sufficient — for any i.i.d. sample, the order statistics retain all information since the joint density is symmetric in the arguments. But they are a vector of length n and are almost never *minimal*. The item asks for the sufficient statistic in the useful sense of data reduction, and the order statistics achieve essentially none.

(d) *Likelihood ratio.* Not a statistic for a single parameter at all — it is a function of two parameter values as well as the data, and is the object of the Neyman–Pearson lemma, not of sufficiency theory.

Answer: **(a)** the sample mean.

Question 35. The MLE and MOM estimators for the given distribution coincide when **(a)** the distribution is exponential **(b)** the distribution is normal **(c)** both of the above **(d)** neither.

The concept: two estimation principles

Method of moments

Equate the first k population moments (which are functions of the parameters) to the corresponding sample moments and solve. Formally, solve

$$\mu_r(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} X_i^r, \qquad r = 1, \dots, k.$$

It is computationally trivial and consistent, but generally inefficient and can produce estimates outside the parameter space.

Maximum likelihood

Maximise $L(\boldsymbol{\theta}) = \prod_i f(X_i; \boldsymbol{\theta})$, or equivalently the log-likelihood $\ell = \sum_i \ln f(X_i; \boldsymbol{\theta})$. Under regularity conditions the MLE is consistent, asymptotically normal, and asymptotically efficient (attaining the Cramér–Rao bound), and it is invariant under reparametrisation.

The exponential case

Let $X_1, \dots, X_n \sim \mathrm{Exp}(\lambda)$ with density $f(x) = \lambda \mathrm{e}^{-\lambda x}$ for $x > 0$.

MOM. The population mean is $\mathbb{E}X = 1/\lambda$. Equating to $\bar{X}$:

$$\frac{1}{\hat{\lambda}} = \bar{X} \quad \Longrightarrow \quad \hat{\lambda}_{\text{MOM}} = \frac{1}{\bar{X}}.$$

MLE. The log-likelihood is

$$\ell(\lambda) = n \ln \lambda - \lambda \sum_i X_i, \qquad \ell'(\lambda) = \frac{n}{\lambda} - \sum_i X_i = 0 \quad \Longrightarrow \quad \hat{\lambda}_{\text{MLE}} = \frac{n}{\sum X_i} = \frac{1}{\bar{X}}.$$

(Second derivative $-n/\lambda^2 < 0$ confirms a maximum.) The two coincide.

The normal case

Let $X_i \sim N(\mu, \sigma^2)$ with both parameters unknown.

MOM. $\mathbb{E}X = \mu$ and $\mathbb{E}X^2 = \mu^2 + \sigma^2$. Solving,

$$\hat{\mu}_{\text{MOM}} = \bar{X}, \qquad \hat{\sigma}^2_{\text{MOM}} = \frac{1}{n}\sum X_i^2 - \bar{X}^2 = \frac{1}{n}\sum (X_i - \bar{X})^2.$$

MLE. The log-likelihood is

$$\ell = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum (X_i - \mu)^2.$$

Setting $\partial\ell/\partial\mu = \frac{1}{\sigma^2}\sum(X_i - \mu) = 0$ gives $\hat{\mu} = \bar{X}$. Then $\partial\ell/\partial\sigma^2 = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(X_i - \bar{X})^2 = 0$ gives

$$\hat{\sigma}^2_{\text{MLE}} = \frac{1}{n}\sum (X_i - \bar{X})^2.$$

Again the two coincide, with the *same* divisor n in both.

The reason behind the coincidence. Both distributions belong to exponential families whose natural sufficient statistics are precisely the sample moments being matched — $\sum X_i$ for the exponential, and $(\sum X_i, \sum X_i^2)$ for the normal. When the natural sufficient statistic *is* the moment vector, the likelihood equations reduce exactly to the moment equations, and the two methods must agree.
Where they diverge. For Uniform$(0, \theta)$, MOM gives $\hat{\theta} = 2\bar{X}$ while MLE gives $\hat{\theta} = \max X_i$ — radically different, with the MLE far superior (variance $O(n^{-2})$ versus $O(n^{-1})$). Note also that neither normal estimator is the unbiased S^2 with divisor $n-1$; both use n.

Answer: **(c)** both of the above.

Question 36. Which of the following functions is entire? **(a)** $\sin z$ **(b)** z^2 **(c)** e^z **(d)** $|z|^2$.

The concept

Entire functions

f is *entire* if it is holomorphic on the whole of $\mathbb{C}$. Equivalently, f has a power series expansion $\sum a_n z^n$ with infinite radius of convergence, valid on all of $\mathbb{C}$.
The class of entire functions is closed under addition, multiplication and composition, and contains all polynomials, e^z, $\sin z$, $\cos z$, $\sinh z$, $\cosh z$. It does *not* contain $1/z$ (pole at 0), $\log z$ (branch cut), $\sqrt{z}$ (branch point), $\tan z$ (poles at $\pi/2 + k\pi$), or anything involving $\bar{z}$, $|z|$, $\operatorname{Re} z$, $\operatorname{Im} z$.

Testing

(a) $\sin z = \sum_{n\geq 0}(-1)^n z^{2n+1}/(2n+1)!$, radius of convergence ∞. **Entire.**

(b) z^2 is a polynomial. **Entire.**

(c) $e^z = \sum_{n\geq 0} z^n/n!$, radius ∞. **Entire.**

(d) $|z|^2 = z\bar{z}$. As established in Question 30, this satisfies the Cauchy–Riemann equations only at the origin, so it is nowhere holomorphic and certainly not entire. **Not entire.**

This item is defective. Three of the four options — **(a)**, **(b)** and **(c)** — are entire, and only **(d)** is not. A single-answer item cannot have three correct options.
The item was almost certainly intended to read "Which of the following is **not** entire?", in which case the answer is **(d)**. Under the printed wording, any of **(a)**, **(b)**, **(c)** is a correct response.
Compare Question 30, where the option set is properly constructed: there, $|z|^2$, $\bar{z}$ and $z^2 + \bar{z}$ all fail and only e^z succeeds.

Answer: Defective as printed: **(a)**, **(b)** and **(c)** are all entire. The only non-entire function is **(d)** $|z|^2$.

Question 37. The radius of convergence of $\sum \frac{n!}{n^n} z^n$ is **(a)** 0 **(b)** 1 **(c)** e **(d)** ∞.

The computation by the ratio test

With $a_n = n!/n^n$,

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{n!} = \frac{(n+1)\, n^n}{(n+1)^{n+1}} = \frac{n^n}{(n+1)^n} = \left(\frac{n}{n+1}\right)^n = \frac{1}{\left(1+\frac{1}{n}\right)^n}.$$

Since $(1+1/n)^n \to \mathrm{e}$,

$$\frac{1}{R} = \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{\mathrm{e}} \qquad \Longrightarrow \qquad R = \mathrm{e}.$$

The computation by the root test

By Stirling's formula $n! \sim \sqrt{2\pi n}\,(n/\mathrm{e})^n$,

$$a_n = \frac{n!}{n^n} \sim \frac{\sqrt{2\pi n}\, n^n \mathrm{e}^{-n}}{n^n} = \sqrt{2\pi n}\ \mathrm{e}^{-n},$$

so

$$a_n^{1/n} \sim (2\pi n)^{1/(2n)}\, \mathrm{e}^{-1} \longrightarrow \mathrm{e}^{-1},$$

because $(2\pi n)^{1/(2n)} = \exp\!\left(\frac{\ln(2\pi n)}{2n}\right) \to \mathrm{e}^0 = 1$. Hence $1/R = 1/\mathrm{e}$ and $R = \mathrm{e}$, in agreement.

The limit $\lim (n!)^{1/n}/n = 1/\mathrm{e}$ is worth knowing independently of Stirling. It follows from the Cesàro-type result that if $b_n > 0$ and $b_{n+1}/b_n \to L$ then $b_n^{1/n} \to L$: apply it with $b_n = n!/n^n$, whose ratio we computed above to tend to $1/\mathrm{e}$.
Behaviour on the circle $|z| = \mathrm{e}$**.** Take $z = \mathrm{e}$; the terms are $n!\mathrm{e}^n/n^n \sim \sqrt{2\pi n} \to \infty$, so the series diverges. By the same estimate the terms have modulus $\sim \sqrt{2\pi n}$ everywhere on $|z| = \mathrm{e}$, so the series diverges at *every* boundary point.

Answer: **(c)** e.

Question 38. In the dual of the given LPP, the number of variables is **(a)** 2 **(b)** 3 **(c)** 4 **(d)** 5.

The primal problem did not reproduce

The duality correspondence

For the primal in standard form

$$\max\ c^{\mathsf T} x \quad \text{s.t.} \quad Ax \le b,\ x \ge 0, \qquad A \in \mathbb{R}^{m\times n},$$

the dual is

$$\min\ b^\mathsf{T}y \quad \text{s.t.} \quad A^\mathsf{T}y \geq c,\ y \geq 0.$$

The correspondence is a transposition, and the counting rule follows immediately:

Primal	$\longleftrightarrow$	Dual
n variables		n constraints
m constraints		m variables
maximise		minimise
ith constraint is $\leq$		ith variable ≥ 0
ith constraint is $=$		ith variable unrestricted
jth variable ≥ 0		jth constraint is $\geq$
jth variable unrestricted		jth constraint is $=$

The rule needed here: *the number of dual variables equals the number of primal constraints* (not counting the sign restrictions $x \geq 0$, which are structural rather than constraints proper).

Why duality holds — weak and strong

Weak duality. If x is primal-feasible and y dual-feasible then

$$c^\mathsf{T}x \leq (A^\mathsf{T}y)^\mathsf{T}x = y^\mathsf{T}Ax \leq y^\mathsf{T}b,$$

using $c \leq A^\mathsf{T}y$ with $x \geq 0$ for the first inequality and $Ax \leq b$ with $y \geq 0$ for the second. So every dual feasible value bounds every primal feasible value.
Strong duality. If either problem has a finite optimum, so does the other and the optimal values are equal. The proof rests on Farkas' lemma or on the termination of the simplex method.
Complementary slackness. At optimality, $y_i > 0 \Rightarrow (Ax)_i = b_i$ and $x_j > 0 \Rightarrow (A^\mathsf{T}y)_j = c_j$. This is the practical tool for reading off one solution from the other.

A common trap. A constraint written as an equality contributes one dual variable (unrestricted in sign), not two. A candidate who converts each equality into a pair of inequalities before dualising will double-count and typically select an option one or two too large.

Answer: Indeterminate: the primal problem is absent. The number of dual variables equals the number of primal constraints.

Question 39. The congruence $x^2 \equiv a \pmod p$ has solutions when **(a)** $\left(\frac{a}{p}\right) = 1$ **(b)** $\left(\frac{a}{p}\right) = -1$ **(c)** $a \equiv 0 \pmod p$ **(d)** both (a) and (c).

The concept: the Legendre symbol

Definition and the trichotomy

Let p be an odd prime. The Legendre symbol is defined by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a, \\ 1 & \text{if } p \nmid a \text{ and } a \text{ is a quadratic residue mod } p, \\ -1 & \text{if } p \nmid a \text{ and } a \text{ is a quadratic non-residue mod } p. \end{cases}$$

Crucially the value 0 is a *separate* case, not folded into either of the others. This is the whole point of the question.

Case analysis

Case 1: $\left(\frac{a}{p}\right) = 1$. By definition a is a quadratic residue: there exists x with $x^2 \equiv a$. Moreover there are exactly two such x modulo p, namely $\pm x_0$, and they are distinct since p is odd. *Solutions exist.*

Case 2: $\left(\frac{a}{p}\right) = -1$. By definition no solution exists. *No solutions.* Option **(b)** is therefore wrong.

Case 3: $a \equiv 0 \pmod p$. The congruence is $x^2 \equiv 0 \pmod p$. Since p is prime, $p \mid x^2$ implies $p \mid x$, so $x \equiv 0$ is the unique solution. *Solutions exist* — exactly one, rather than two. Option **(c)** is therefore correct as far as it goes, but incomplete.

Since both Case 1 and Case 3 yield solutions, the complete answer is the disjunction, option **(d)**.

Euler's criterion makes the trichotomy computable:

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod p.$$

For $p \nmid a$, Fermat gives $a^{p-1} \equiv 1$, so $a^{(p-1)/2}$ is a square root of 1 modulo p and hence is ± 1; it is $+1$ exactly when a is a residue. For $p \mid a$ the expression is 0.

The residue count. Exactly $(p-1)/2$ of the nonzero residues are quadratic residues and $(p-1)/2$ are not, because the squaring map on $(\mathbb{Z}/p\mathbb{Z})^\times$ is two-to-one. Together with the class $\{0\}$ this accounts for all p residue classes.

Solution counts, summarised: 2 solutions when the symbol is 1; 1 solution when $a \equiv 0$; 0 solutions when the symbol is -1. Total over all a: $2 \cdot \frac{p-1}{2} + 1 = p$, as it must be, since squaring maps $\mathbb{Z}/p\mathbb{Z}$ onto its image with these multiplicities.

Answer: **(d)** both (a) and (c).

Question 40. A complete sufficient statistic for the exponential distribution is **(a)** $\sum X_i$ **(b)** $\prod X_i$ **(c)** $\max X_i$ **(d)** $\min X_i$.

The concept: completeness

Complete statistics

A statistic T with family of distributions $\{P_\theta\}$ is *complete* if for every measurable g,

$$\mathbb{E}_\theta\big[g(T)\big] = 0 \quad \text{for all } \theta \qquad \Longrightarrow \qquad g(T) = 0 \text{ almost surely for all } \theta.$$

Interpretation: the family of distributions of T is rich enough that no nontrivial function of T can have identically zero mean. Completeness is what makes sufficiency *useful* for constructing optimal estimators.
Lehmann–Scheffé. If T is complete and sufficient and $\delta(T)$ is an unbiased estimator of $\tau(\theta)$, then $\delta(T)$ is the *unique* uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$.
Basu's theorem. A complete sufficient statistic is independent of every ancillary statistic.

The computation

Let $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ i.i.d. The joint density is

$$f(\mathbf{x}; \lambda) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right) \prod_i \mathbf{1}_{\{x_i > 0\}}.$$

Sufficiency follows immediately by Fisher–Neyman with $T = \sum X_i$: take $g(T, \lambda) = \lambda^n \mathrm{e}^{-\lambda T}$ and $h(\mathbf{x}) = \prod \mathbf{1}_{\{x_i>0\}}$.
Completeness follows from the general theory of full-rank exponential families: writing the density in canonical form with natural parameter $\eta = -\lambda$ and natural statistic $t(x) = x$, the natural parameter space $\{\eta < 0\}$ contains an open interval, so the family is of full rank and $T = \sum X_i$ is complete.
Directly: $T \sim \text{Gamma}(n, \lambda)$, and $\mathbb{E}_\lambda[g(T)] = 0$ for all $\lambda > 0$ reads

$$\int_0^\infty g(t)\, \frac{\lambda^n t^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda t}\, \mathrm{d}t = 0 \quad \Longrightarrow \quad \int_0^\infty \big[g(t) t^{n-1}\big] \mathrm{e}^{-\lambda t}\, \mathrm{d}t = 0 \ \ \forall \lambda > 0.$$

The left side is the Laplace transform of $g(t)t^{n-1}$; by uniqueness of the Laplace transform this function is zero a.e., hence $g \equiv 0$ a.e. on $(0, \infty)$.

Discrimination

(b) $\prod X_i$ is a function of the data but is not sufficient here — the likelihood depends on the data through the *sum*, not the product. (The product is the sufficient statistic for a *Pareto* or *lognormal* family.)

(c) $\max X_i$ is sufficient and complete for $\text{Uniform}(0, \theta)$, where the support depends on the parameter. For the exponential the support is $(0, \infty)$ regardless of λ, and the maximum is not sufficient.

(d) $\min X_i$ is the relevant statistic for a *shifted* exponential $f(x) = \lambda \mathrm{e}^{-\lambda(x-\mu)}$, $x > \mu$, where it is sufficient for the threshold μ. For the unshifted family it is not sufficient; indeed $n \min X_i \sim \text{Exp}(\lambda)$, which carries information about λ but loses most of the sample's.

Using Lehmann–Scheffé here. Since $T = \sum X_i \sim \text{Gamma}(n, \lambda)$ has $\mathbb{E}[1/T] = \lambda/(n-1)$, the statistic $(n-1)/T$ is unbiased for λ and, being a function of a complete sufficient statistic, is the UMVUE. Note that the MLE n/T is *biased* — a useful reminder that maximum likelihood and unbiasedness are different criteria.

Answer: **(a)** $\sum X_i$.

Question 41. Let G be the group of invertible 2×2 matrices over $\mathbb{Z}/11\mathbb{Z}$. The order of a 3-Sylow subgroup of G is **(a)** 3 **(b)** 9 **(c)** 27 **(d)** 81.

The concept: Sylow subgroups and the order of $GL_n(\mathbb{F}_q)$

Sylow's first theorem

Let $|G| = p^a m$ with $p \nmid m$. Then G possesses a subgroup of order p^a, called a Sylow p-subgroup. Its order is therefore *the full power of p dividing* $|G|$ — determined entirely by the factorisation of $|G|$, with no further information about G required.
Consequently the entire problem reduces to computing $|G|$ and extracting the 3-part.

The order of the general linear group

$$|GL_n(\mathbb{F}_q)| = \prod_{k=0}^{n-1} (q^n - q^k).$$

Reason. An invertible matrix is determined by its columns $v_1, \ldots, v_n$, which must be linearly independent. Choose v_1 to be any nonzero vector: $q^n - 1$ ways. Choose v_2 outside the span of v_1 (a one-dimensional subspace with q elements): $q^n - q$ ways. In general, having chosen $v_1, \ldots, v_k$ spanning a k-dimensional subspace of size q^k, there are $q^n - q^k$ choices for v_{k+1}. Multiplying gives the formula.

The computation

Here $n = 2$, $q = 11$, and $\mathbb{Z}/11\mathbb{Z} = \mathbb{F}_{11}$ since 11 is prime. So

$$|G| = |GL_2(\mathbb{F}_{11})| = (11^2 - 1)(11^2 - 11) = (121 - 1)(121 - 11) = 120 \times 110.$$

Factor each piece:

$$120 = 2^3 \cdot 3 \cdot 5, \qquad 110 = 2 \cdot 5 \cdot 11.$$

Hence

$$|G| = 2^3 \cdot 3 \cdot 5 \cdot 2 \cdot 5 \cdot 11 = 2^4 \cdot 3 \cdot 5^2 \cdot 11 = 13200.$$

The exact power of 3 dividing $|G|$ is $3^1 = 3$. Therefore every Sylow 3-subgroup of G has order 3.

Consistency check on the number of Sylow subgroups. By Sylow's third theorem, $n_3 \equiv 1 \pmod 3$ and $n_3 \mid |G|/3 = 4400$. The divisors of 4400 congruent to 1 modulo 3 include $1, 4, 16, 22, 25, 55, 100, \ldots$. In fact $n_3 = n_3(GL_2(\mathbb{F}_{11}))$ is large; the point of the check is only that $n_3 = 1$ is not forced, so the Sylow 3-subgroup is not normal.
Where the trap lies. A candidate who writes $|G| = (11^2-1)(11^2-11)$ but then miscomputes $110 = 11 \cdot 10 = 11 \cdot 2 \cdot 5$ as containing a factor 3 will inflate the answer. Another common error is to use $|M_2(\mathbb{F}_{11})| = 11^4$ — the count of *all* matrices, invertible or not — which contains no factor of 3 at all and would suggest that G has no 3-torsion, contradicting Cauchy's theorem applied to the visible element of order 3 in $\mathbb{F}_{11}^\times$ (since $3 \mid 10$? no — and this contradiction is itself instructive: the order 3 elements of G do not come from scalars).

Answer: **(a)** 3.

Question 42. The sum of all incongruent positive solutions ≤ 102 of $90x \equiv 12 \pmod{102}$ is **(a)** 334 **(b)** 351 **(c)** 367 **(d)** 391.

The concept: linear congruences

Solvability and solution count

The congruence $ax \equiv b \pmod{m}$ has a solution if and only if $d = \gcd(a, m)$ divides b. When it does, there are exactly d incongruent solutions modulo m, and they are obtained as follows. Divide through by d to obtain $\frac{a}{d}x \equiv \frac{b}{d} \pmod{\frac{m}{d}}$, in which $\gcd(a/d, m/d) = 1$, so a/d is invertible modulo m/d and there is a unique solution $x \equiv x_0 \pmod{m/d}$. The d solutions modulo m are then

$$x_0, \quad x_0 + \frac{m}{d}, \quad x_0 + \frac{2m}{d}, \quad \ldots, \quad x_0 + \frac{(d-1)m}{d}.$$

The computation

Here $a = 90$, $b = 12$, $m = 102$.

Step 1: the gcd. $102 = 90 + 12$, $90 = 7 \cdot 12 + 6$, $12 = 2 \cdot 6$, so $\gcd(90, 102) = 6$. Since $6 \mid 12$, the congruence is solvable and has exactly 6 incongruent solutions modulo 102.

Step 2: reduce. Divide by 6:

$$15x \equiv 2 \pmod{17}.$$

Note $102/6 = 17$, a prime, which makes the inversion easy.

Step 3: invert. Since $15 \equiv -2 \pmod{17}$, the congruence reads $-2x \equiv 2$, i.e. $2x \equiv -2$, i.e. $x \equiv -1 \equiv 16 \pmod{17}$.

Check: $15 \times 16 = 240 = 14 \times 17 + 2 = 238 + 2$. ✓

Step 4: lift. The six solutions modulo 102 are $x \equiv 16 \pmod{17}$, namely

$$16, \ 33, \ 50, \ 67, \ 84, \ 101.$$

All six are positive and ≤ 102, so all six are counted.

Step 5: sum. These form an arithmetic progression with first term 16, last term 101 and 6 terms:

$$S = \frac{6}{2}(16 + 101) = 3 \times 117 = 351.$$

A shortcut for the sum. The d solutions are $x_0 + k\,\frac{m}{d}$ for $k = 0, \ldots, d-1$, so

$$S = d\,x_0 + \frac{m}{d} \cdot \frac{d(d-1)}{2} = d\,x_0 + \frac{m(d-1)}{2}.$$

Here $S = 6 \times 16 + \frac{102 \times 5}{2} = 96 + 255 = 351$. ✓
Verifying one solution in the original. $90 \times 16 = 1440$, and $1440 = 14 \times 102 + 12 = 1428 + 12$. So $90 \times 16 \equiv 12 \pmod{102}$. ✓

Answer: **(b)** 351.

Question 43. If the given PDE is parabolic on the curve $y = \varphi(x)$, then $\lim_{x\to 0} \varphi(x)$ equals **(a)** 1 **(b)** 1/2 **(c)** 1/4 **(d)** 1/8.

The PDE did not reproduce

Classification of second-order PDEs

For a second-order linear PDE in two variables,

$$A(x,y)u_{xx} + 2B(x,y)u_{xy} + C(x,y)u_{yy} + (\text{lower order}) = 0,$$

the type at a point is determined by the sign of the discriminant

$$\Delta = B^2 - AC.$$

$\Delta > 0$	hyperbolic	two real characteristic families (wave equation)
$\Delta = 0$	parabolic	one real characteristic family (heat equation)
$\Delta < 0$	elliptic	no real characteristics (Laplace equation)

The classification is *pointwise*: a single equation may change type across the plane. The Tricomi equation $u_{xx} + y\,u_{yy} = 0$ has $\Delta = -y$, so it is hyperbolic for $y < 0$, parabolic exactly on the line $y = 0$, and elliptic for $y > 0$. Equations of this mixed kind are precisely the setting of the present question.

The method the item requires

Set $\Delta(x,y) = B^2 - AC = 0$ and solve for y as a function of x: this defines the parabolic curve $y = \varphi(x)$. Then take the limit as $x \to 0$.

Since the coefficients A, B, C are absent, the curve cannot be determined. The option values $1,\ \frac{1}{2},\ \frac{1}{4},\ \frac{1}{8}$ form a geometric progression with ratio $\frac{1}{2}$, which typically arises when the discriminant condition is polynomial in y with the limiting root determined by a coefficient ratio such as $1/2^k$.

Answer: Indeterminate: the PDE is absent from the source text.

Question 44. For the sequence $f_n(x) = \lim_{k\to\infty} \big(\cos(n!\pi x)\big)^{2k}$ on $[0,1]$, the limit function f satisfies **(a)** f is Lebesgue integrable **(b)** $\{f_n\}$ converges uniformly **(c)** all discontinuities of f are removable **(d)** f is differentiable at every rational.

Identifying f_n

Fix n and x, and set $c = \cos(n!\pi x)$, so $|c| \le 1$. Then

$$c^{2k} \longrightarrow \begin{cases} 1 & \text{if } |c| = 1, \\ 0 & \text{if } |c| < 1. \end{cases}$$

Now $|\cos\theta| = 1$ exactly when θ is an integer multiple of π, so $|c| = 1$ exactly when $n!\,x \in \mathbb{Z}$. Hence

$$f_n(x) = \begin{cases} 1 & \text{if } n!\,x \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases}$$

So f_n is the indicator function of the finite set $\{0, \frac{1}{n!}, \frac{2}{n!}, \ldots, 1\} \cap [0,1]$, a set of $n! + 1$ points.

Identifying $f = \lim_n f_n$

If x is rational, write $x = p/q$ in lowest terms. For every $n \geq q$ the integer $n!$ is divisible by q, so $n!\,x = (n!/q)p \in \mathbb{Z}$ and $f_n(x) = 1$. Hence $f(x) = 1$.

If x is irrational, then $n!\,x$ is irrational for every n (a nonzero integer multiple of an irrational is irrational), so $n!\,x \notin \mathbb{Z}$ and $f_n(x) = 0$ for all n. Hence $f(x) = 0$.

Therefore

$$f = \mathbf{1}_{\mathbb{Q}\cap[0,1]},$$

the **Dirichlet function** on $[0,1]$. This double limit is the classical construction of it, due to Baire, and it exhibits the Dirichlet function as a pointwise limit of a sequence of pointwise limits of *continuous* functions — a function of Baire class 2.

Testing the options

(a) True. $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ is measurable, being the indicator of a Borel (indeed countable) set. It is bounded, so it is Lebesgue integrable, and since $\mathbb{Q} \cap [0,1]$ is countable it has Lebesgue measure zero, whence

$$\int_0^1 f \,\mathrm{d}\lambda = \lambda(\mathbb{Q} \cap [0,1]) = 0.$$

(b) False. Uniform convergence would force the limit of continuous functions to be continuous; more directly, each f_n is 0 except at finitely many points while $f = 1$ on a dense set, so $\sup_x |f_n(x) - f(x)| = 1$ for every n. The convergence is pointwise only.

(c) False. A removable discontinuity at a requires $\lim_{x\to a} f(x)$ to exist. But every interval contains both rationals and irrationals, so on every neighbourhood of every point f takes both the values 0 and 1; no limit exists anywhere. The Dirichlet function is discontinuous at *every* point of $[0,1]$, and none of its discontinuities is removable — they are all of the oscillatory (second) kind.

(d) False. Differentiability at a point implies continuity there, and f is continuous nowhere.

Riemann versus Lebesgue. f is *not* Riemann integrable: every upper Darboux sum is 1 and every lower Darboux sum is 0, since every subinterval contains both rationals and irrationals. It *is* Lebesgue integrable with integral 0. This function is the standard exhibit for the strict superiority of the Lebesgue theory, and the present item is testing exactly that awareness — option **(a)** would be false if it read "Riemann integrable".
The Lebesgue criterion explains the failure precisely: a bounded function on $[a,b]$ is Riemann integrable if and only if its set of discontinuities has measure zero. Here that set is all of $[0,1]$, of measure 1.

Answer: **(a)** f is Lebesgue integrable.

Question 45. If $X \sim W_p(n, I_p)$ and A is idempotent of rank r, then $\operatorname{tr}(AX)$ is distributed as **(a)** χ^2_n **(b)** χ^2_r **(c)** χ^2_{pr} **(d)** χ^2_{nr}.

The concept: the Wishart distribution

Definition and basic structure

Let $Z_1, \ldots, Z_n$ be i.i.d. $N_p(\mathbf{0}, \Sigma)$ column vectors. The random matrix

$$X = \sum_{i=1}^{n} Z_i Z_i^{\mathsf{T}}$$

has the *Wishart distribution* $W_p(n, \Sigma)$ with n degrees of freedom and scale matrix Σ. It is the multivariate generalisation of the chi-square: for $p = 1$ and $\Sigma = 1$ it reduces to χ^2_n. The construction as a sum of n independent rank-one terms is the key to every computation with it.

Quadratic forms in normal vectors

If $Z \sim N_p(\mathbf{0}, I_p)$ and A is a symmetric idempotent matrix of rank r, then

$$Z^{\mathsf{T}} A Z \sim \chi^2_r.$$

Proof. Idempotence forces every eigenvalue of A to satisfy $\lambda^2 = \lambda$, hence $\lambda \in \{0, 1\}$; the rank r says there are exactly r eigenvalues equal to 1. Diagonalise $A = Q\Lambda Q^{\mathsf{T}}$ with Q orthogonal. Put $W = Q^{\mathsf{T}} Z$, which is again $N_p(\mathbf{0}, I_p)$ because orthogonal transformations preserve the standard normal. Then

$$Z^{\mathsf{T}} A Z = W^{\mathsf{T}} \Lambda W = \sum_{j=1}^{p} \lambda_j W_j^2 = \sum_{j=1}^{r} W_j^2 \sim \chi^2_r. \qquad \square$$

The computation

Write $X = \sum_{i=1}^{n} Z_i Z_i^{\mathsf{T}}$ with Z_i i.i.d. $N_p(\mathbf{0}, I_p)$, since $\Sigma = I_p$. Then by linearity of the trace,

$$\operatorname{tr}(AX) = \operatorname{tr}\left(A \sum_{i=1}^{n} Z_i Z_i^{\mathsf{T}}\right) = \sum_{i=1}^{n} \operatorname{tr}\big(A Z_i Z_i^{\mathsf{T}}\big).$$

Using the cyclic property $\operatorname{tr}(AZZ^{\mathsf{T}}) = \operatorname{tr}(Z^{\mathsf{T}} A Z)$, and noting that $Z^{\mathsf{T}} A Z$ is a scalar,

$$\operatorname{tr}(AX) = \sum_{i=1}^{n} Z_i^{\mathsf{T}} A Z_i.$$

Each summand is χ^2_r by the lemma, and the summands are independent because the Z_i are. A sum of n independent χ^2_r variables is χ^2_{nr}, since chi-square degrees of freedom add under independent summation. Hence

$$\operatorname{tr}(AX) \sim \chi^2_{nr}.$$

Sanity checks by special cases. Take $A = I_p$, which is idempotent of rank $r = p$. Then $\operatorname{tr}(X) = \sum_i |Z_i|^2$, a sum of np independent squared standard normals, so $\operatorname{tr} X \sim \chi^2_{np}$. The formula gives $\chi^2_{nr} = \chi^2_{np}$. ✓
Take $p = 1$, so A is a 1×1 idempotent, i.e. $A = 0$ or $A = 1$. With $A = 1$, $r = 1$ and $\operatorname{tr}(AX) = X \sim \chi^2_n$; the formula gives $\chi^2_{n\cdot 1} = \chi^2_n$. ✓
The distractors. Option **(a)** χ^2_n ignores the rank; option **(b)** χ^2_r ignores the sample size; option **(c)** χ^2_{pr} confuses the dimension p with the degrees of freedom n. Only the product nr respects both the number of independent Gaussian vectors and the dimension of the subspace onto which A projects.

Answer: **(d)** χ^2_{nr}.

Question 46. If the density of X is symmetric about μ and $Y = 2\mu - X$, then **(a)** only S_1 is true **(b)** only S_2 is true **(c)** both S_1 and S_2 are true **(d)** neither is true.

The statements S_1 and S_2 did not reproduce

Symmetry about a point

A random variable X with density f is *symmetric about* μ if

$$f(\mu + t) = f(\mu - t) \qquad \text{for all } t.$$

Equivalently, $X - \mu$ and $\mu - X$ have the same distribution.
The reflection $Y = 2\mu - X$ is exactly the map $x \mapsto \mu - (x - \mu)$, so the symmetry statement says precisely that

$$Y \overset{d}{=} X.$$

This single identity is the source of every consequence.

The consequences that S_1 and S_2 are most likely to assert

1. **X and Y are identically distributed.** Immediate from the definition, as above. Their common CDF satisfies $F(\mu + t) + F(\mu - t) = 1$ at continuity points.

2. **$\mathbb{E}X = \mu$ whenever the mean exists.** From $Y \overset{d}{=} X$ we get $\mathbb{E}Y = \mathbb{E}X$, and $\mathbb{E}Y = 2\mu - \mathbb{E}X$, so $2\mathbb{E}X = 2\mu$. Note the caveat: the Cauchy density is symmetric about 0 but has no mean, so symmetry alone does not guarantee $\mathbb{E}X = \mu$ — only that if the mean exists it equals μ.

3. **μ is a median.** $\mathbb{P}(X \le \mu) = \mathbb{P}(Y \le \mu) = \mathbb{P}(X \ge \mu)$, and these sum to at least 1, so each is at least $1/2$. Unlike the mean, this requires no integrability.

4. **All odd central moments vanish** when they exist: $\mathbb{E}[(X - \mu)^{2k+1}] = \mathbb{E}[(Y - \mu)^{2k+1}] = \mathbb{E}[-(X - \mu)^{2k+1}]$, forcing the value to be 0. In particular the skewness is zero.

5. **X and Y are *not* independent** (except in degenerate cases), since Y is a deterministic function of X. Indeed $\operatorname{Corr}(X, Y) = -1$ when $\operatorname{Var} X$ is finite and positive, because Y is an

affine function of X with negative slope. This is a frequent trap: identically distributed does not mean independent.

Answer: Indeterminate: the statements S_1 and S_2 are absent from the source text.

Question 47. The number of non-negative integer solutions of $x_1 + x_2 + x_3 = 17$ with the given upper bounds is **(a)** 57 **(b)** 59 **(c)** 60 **(d)** 148.

The upper bounds did not reproduce

Stars and bars

The number of solutions of $x_1 + \cdots + x_k = N$ in non-negative integers is

$$\binom{N+k-1}{k-1}.$$

Reason. Represent a solution as N identical stars separated into k groups by $k-1$ bars. Any arrangement of N stars and $k-1$ bars in a row determines a solution and vice versa, and there are $\binom{N+k-1}{k-1}$ such arrangements.
For strictly positive solutions, substitute $y_i = x_i - 1$ to get $\binom{N-1}{k-1}$.

Here $k = 3$, $N = 17$, so the unrestricted count is

$$\binom{17+2}{2} = \binom{19}{2} = 171.$$

Imposing upper bounds by inclusion–exclusion

Suppose we additionally require $x_i \le u_i$. Let A_i be the set of solutions violating the ith bound, i.e. with $x_i \ge u_i + 1$. Substituting $y_i = x_i - (u_i + 1) \ge 0$ converts the violation into an unrestricted problem with total $N - (u_i + 1)$, so

$$|A_i| = \binom{N - u_i - 1 + k - 1}{k-1},$$

interpreted as 0 when the top is negative. By inclusion–exclusion the admissible count is

$$\binom{N+k-1}{k-1} - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \dots,$$

where $|A_i \cap A_j| = \binom{N-u_i-u_j-2+k-1}{k-1}$, and so on.

What the options reveal

Three of the four options (57, 59, 60) cluster near a third of 171, which is the signature of bounds tight enough to bite in all three variables. The fourth, 148, is close to 171 and would arise from a single loose bound. As a worked illustration of the machinery, suppose the bounds were $x_i \le 9$ for each i. Then

$$|A_i| = \binom{17-10+2}{2} = \binom{9}{2} = 36,$$

and no two bounds can be violated simultaneously since $10+10=20>17$. The count would be $171-3(36)=171-108=63$. Bounds of $x_i \le 10$ would give $171-3\binom{8}{2}=171-84=87$. Neither matches an option, so the actual bounds are unequal and cannot be recovered.

Answer: Indeterminate: the upper bounds are absent from the source text. The unrestricted count is $\binom{19}{2}=171$.

Question 48. If $\varphi : D \to D$ is holomorphic with $\varphi(1/2)=0$ and $\varphi(-1/2)=3/5$, then $\varphi(1/3)$ equals **(a)** 2/7 **(b)** 1/5 **(c)** 2/5 **(d)** 3/5.

The concept: Schwarz and Schwarz–Pick

Schwarz's lemma

Let $g : D \to D$ be holomorphic with $g(0)=0$, where $D=\{z : |z|<1\}$. Then

$$|g(z)| \le |z| \quad \text{for all } z \in D, \qquad \text{and} \qquad \big|g'(0)\big| \le 1.$$

If equality holds at any single point $z \neq 0$, or if $|g'(0)|=1$, then g is a rotation: $g(z)=\lambda z$ with $|\lambda|=1$.

Proof. Since $g(0)=0$, the function $h(z)=g(z)/z$ has a removable singularity at 0 and is holomorphic on D. On $|z|=r<1$ we have $|h| \le 1/r$, so by the maximum modulus principle $|h| \le 1/r$ on $|z| \le r$. Letting $r \to 1$ gives $|h| \le 1$ throughout D. Equality at an interior point forces h to be a unimodular constant by the maximum principle. ☐

Schwarz–Pick, the invariant form

Define the *pseudo-hyperbolic distance* on D by

$$\rho(a,b)=\left|\frac{a-b}{1-\bar{a}b}\right|.$$

Then every holomorphic $\varphi : D \to D$ is a contraction:

$$\rho\big(\varphi(a),\varphi(b)\big) \le \rho(a,b) \qquad \text{for all } a,b \in D,$$

with equality (at one pair, hence at all) if and only if φ is an automorphism of D.
Schwarz's lemma is the special case $b=0$, $\varphi(0)=0$. The general statement follows by pre- and post-composing with the Möbius automorphisms $\sigma_a(z)=\frac{z+a}{1+\bar{a}z}$ that move a to 0.

Applying the bound

We are told $\varphi(1/2)=0$. Apply Schwarz–Pick with $a=1/2$ and $b=1/3$:

$$\rho\big(\varphi(1/3),\varphi(1/2)\big) \le \rho(1/3,1/2).$$

Since $\varphi(1/2)=0$ and $\rho(w,0)=|w|$, the left side is $|\varphi(1/3)|$. The right side is

$$\rho\big(\tfrac{1}{3},\tfrac{1}{2}\big)=\left|\frac{\frac{1}{3}-\frac{1}{2}}{1-\frac{1}{3}\cdot\frac{1}{2}}\right|=\frac{\frac{1}{6}}{\frac{5}{6}}=\frac{1}{5}.$$

Therefore

$$\boxed{|\varphi(1/3)| \le \tfrac{1}{5}}$$

for *every* holomorphic self-map of D vanishing at $1/2$.

Discrimination by the bound

Compare the options against $1/5 = 0.200$:

$$\textbf{(a)}\ \tfrac{2}{7} \approx 0.286 > \tfrac{1}{5}, \qquad \textbf{(b)}\ \tfrac{1}{5} = \tfrac{1}{5}, \qquad \textbf{(c)}\ \tfrac{2}{5} = 0.400 > \tfrac{1}{5}, \qquad \textbf{(d)}\ \tfrac{3}{5} = 0.600 > \tfrac{1}{5}.$$

Options **(a)**, **(c)** and **(d)** exceed the bound and are impossible for any map of the stated kind. Only **(b)** survives, and it is therefore the key.

The item is over-determined. Equality in Schwarz–Pick forces φ to be an automorphism of D, necessarily of the form

$$\varphi(z) = \lambda\,\frac{z-\frac{1}{2}}{1-\frac{1}{2}z} = \lambda\,\frac{2z-1}{2-z}, \qquad |\lambda| = 1.$$

Such a map has $|\varphi(-1/2)| = |\lambda| \cdot \frac{2}{5/2} = \frac{4}{5}$, not $\frac{3}{5}$. So the second datum is incompatible with equality, and the true value of $|\varphi(1/3)|$ must be strictly less than $1/5$.
Indeed, the extremal map consistent with both data is found by conjugating: put $\sigma(w) = \frac{2w+1}{2-\cdot}$, more precisely $\sigma(w) = \frac{w+1/2}{1+w/2} = \frac{2w+1}{2+w}$, so that $\sigma(0) = 1/2$, and set $g = \varphi \circ \sigma$, which fixes 0. Solving $\sigma(w) = -1/2$ gives $w = -4/5$, and $g(-4/5) = 3/5$. The minimal choice $g(w) = cw$ requires $c = -3/4$, giving

$$\varphi(z) = -\tfrac{3}{4}\cdot\frac{2z-1}{2-z}, \qquad \varphi(1/3) = -\tfrac{3}{4}\cdot\frac{-1/3}{5/3} = \tfrac{3}{20} = 0.15.$$

Since $|g(-4/5)| = 3/5 < 4/5$, Schwarz is strict and φ is *not* uniquely determined: many self-maps satisfy both data. The item therefore has no uniquely correct numerical answer, and **(b)** is the keyed answer only by elimination.

Answer: **(b)** $1/5$, by elimination: Schwarz–Pick gives $|\varphi(1/3)| \le 1/5$, ruling out the other three. Note the caveat above.

Question 49. The sum of the series $\sum_{n\ge1} n^2x^n$ for $|x|<1$ is **(a)** $\dfrac{x(1+x)}{(1-x)^3}$ **(b)** $\dfrac{x^2(1+x)}{(1-x)^3}$ **(c)** $\dfrac{x(1+x)}{(1-x)^2}$ **(d)** $\dfrac{x^2(1+x)}{(1-x)^2}$.

The concept: differentiating the geometric series

The operator $x\frac{\mathrm{d}}{\mathrm{d}x}$

Start from the geometric series, valid for $|x|<1$:

$$\sum_{n\ge0} x^n = \frac{1}{1-x}.$$

Power series may be differentiated term by term inside the radius of convergence, and the differentiated series has the same radius. The operator

$$\theta = x\frac{\mathrm{d}}{\mathrm{d}x}$$

acts on the general term by $\theta(x^n) = nx^n$, so applying θ once produces $\sum nx^n$ and applying it twice produces $\sum n^2x^n$. Each application preserves the radius of convergence.

The computation

First application. Differentiate $\sum_{n\geq 0} x^n = (1-x)^{-1}$:

$$\sum_{n\geq 1} nx^{n-1} = \frac{1}{(1-x)^2} \quad\Longrightarrow\quad \sum_{n\geq 1} nx^n = \frac{x}{(1-x)^2}. \tag{49.1}$$

Second application. Differentiate (49.1) with respect to x:

$$\sum_{n\geq 1} n^2x^{n-1} = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{x}{(1-x)^2}\right] = \frac{(1-x)^2\cdot 1 - x\cdot 2(1-x)(-1)}{(1-x)^4}.$$

Factor $(1-x)$ from the numerator:

$$= \frac{(1-x)\big[(1-x)+2x\big]}{(1-x)^4} = \frac{1+x}{(1-x)^3}.$$

Multiplying by x:

$$\sum_{n\geq 1} n^2x^n = \frac{x(1+x)}{(1-x)^3}.$$

Verification

Low-order coefficients. Expand the answer as a power series. Using $(1-x)^{-3} = \sum_{k\geq 0}\binom{k+2}{2}x^k = 1+3x+6x^2+10x^3+\cdots$,

$$x(1+x)(1+3x+6x^2+10x^3+\cdots) = (x+x^2)(1+3x+6x^2+\cdots) = x+4x^2+9x^3+16x^4+\cdots$$

The coefficients are $1, 4, 9, 16$, i.e. n^2. ✓

Numerical spot-check at $x = 1/2$. The formula gives

$$\frac{\frac{1}{2}\cdot\frac{3}{2}}{\left(\frac{1}{2}\right)^3} = \frac{3/4}{1/8} = 6.$$

Direct summation: $\frac{1}{2}+4\cdot\frac{1}{4}+9\cdot\frac{1}{8}+16\cdot\frac{1}{16}+25\cdot\frac{1}{32}+\cdots = 0.5+1+1.125+1+0.781+\cdots \to 6.$ ✓

Dimensional check on the options. As $x \to 0^+$ the sum behaves like its leading term x. Option **(a)** behaves like x; **(b)** and **(d)** behave like x^2 and are eliminated at once. As $x \to 1^-$ the sum diverges like $\sum n^2$ weighted, and the correct rate is $(1-x)^{-3}$; option **(c)** has only $(1-x)^{-2}$, the rate appropriate to $\sum nx^n$, and is eliminated. Two asymptotic checks isolate **(a)** without any differentiation.

The general pattern:

$$\sum_{n\ge 1} n^k x^n = \frac{A_k(x)}{(1-x)^{k+1}},$$

where A_k is the kth Eulerian polynomial: $A_1 = x$, $A_2 = x(1+x)$, $A_3 = x(1+4x+x^2)$.

Answer: **(a)** $\dfrac{x(1+x)}{(1-x)^3}$.

Question 50. For a 2×2 positive definite matrix A with the given trace and determinant conditions, $\det(\operatorname{adj} A)$ equals **(a)** 4 **(b)** 8 **(c)** 16 **(d)** 64.

The trace and determinant conditions did not reproduce

The adjugate and its determinant

For $A \in M_n(\mathbb{R})$ the *adjugate* (classical adjoint) $\operatorname{adj} A$ is the transpose of the cofactor matrix, characterised by

$$A \operatorname{adj} A = \operatorname{adj} A\, A = (\det A)\, I_n. \tag{50.1}$$

Taking determinants in (50.1),

$$\det A \cdot \det(\operatorname{adj} A) = (\det A)^n \quad \Longrightarrow \quad \det(\operatorname{adj} A) = (\det A)^{n-1}$$

whenever $\det A \neq 0$; and by continuity (both sides are polynomials in the entries) the identity holds for all A.

For $n = 2$ this gives $\det(\operatorname{adj} A) = \det A$ — the exponent is $n - 1 = 1$.

The explicit 2×2 case

For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\operatorname{adj} A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

so $\det(\operatorname{adj} A) = da - (-b)(-c) = ad - bc = \det A$. ✓

Note also that in the 2×2 case

$$\operatorname{tr}(\operatorname{adj} A) = a + d = \operatorname{tr} A,$$

so the adjugate has the same trace and determinant as A, hence the same characteristic polynomial. (Indeed $\operatorname{adj} A = (\operatorname{tr} A)I - A$ for 2×2 matrices, by Cayley–Hamilton.)

What positive definiteness supplies

For a 2×2 real symmetric A with eigenvalues λ_1, λ_2:

$$\operatorname{tr} A = \lambda_1 + \lambda_2, \qquad \det A = \lambda_1\lambda_2.$$

Positive definiteness is equivalent to $\lambda_1, \lambda_2 > 0$, equivalently $\operatorname{tr} A > 0$ and $\det A > 0$, equivalently both leading principal minors positive (Sylvester's criterion).

So the answer is simply $\det A$, and the trace condition would be needed only to pin down which of the offered values it is. Note that all four options $(4, 8, 16, 64)$ are positive, consistent with positive definiteness; the option set alone does not discriminate.

The bound worth carrying. By AM–GM applied to the eigenvalues,

$$\det A = \lambda_1\lambda_2 \le \left(\frac{\lambda_1+\lambda_2}{2}\right)^2 = \left(\frac{\operatorname{tr} A}{2}\right)^2,$$

with equality iff $\lambda_1 = \lambda_2$, i.e. iff A is a scalar multiple of the identity. If, for instance, the missing condition were $\operatorname{tr} A = 8$, then $\det A \le 16$ and options **(d)** would be excluded immediately.

Answer: Indeterminate: the conditions are absent. For any 2×2 matrix, $\det(\operatorname{adj} A) = \det A$; in general $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

Question 51. The integral equation corresponding to the IVP $y'' - 5y' + 4y = 0$, $y(0) = 0$, $y'(0) = -1$, for $u = y''$ is **(a)** $u(x) = 4x + 5 + \int_0^x (5 + 4x + 4t)u(t)\,\mathrm{d}t$ **(b)** $u(x) = 4x - 5 + \int_0^x (5 - 4x + 4t)u(t)\,\mathrm{d}t$ **(c)** $u(x) = -4x + 5 + \int_0^x (5 + 4x + 4t)u(t)\,\mathrm{d}t$ **(d)** $u(x) = -4x + 5 + \int_0^x (5 + 4x - 4t)u(t)\,\mathrm{d}t$.

The concept: converting an IVP into a Volterra equation

The reduction formula for repeated integrals

The standard device sets $u = y''$ (the highest derivative) and recovers the lower derivatives by integration, absorbing the initial conditions as they arise:

$$y'(x) = y'(0) + \int_0^x u(t)\,\mathrm{d}t,$$

$$y(x) = y(0) + y'(0)\,x + \int_0^x\int_0^s u(t)\,\mathrm{d}t\,\mathrm{d}s.$$

The double integral collapses to a single one by Cauchy's formula for repeated integration,

$$\int_0^x\int_0^s u(t)\,\mathrm{d}t\,\mathrm{d}s = \int_0^x (x-t)\,u(t)\,\mathrm{d}t,$$

which follows from exchanging the order of integration over the triangle $0 \le t \le s \le x$: the inner variable s ranges over (t, x), contributing a factor $x - t$. More generally the n-fold repeated integral is $\int_0^x \frac{(x-t)^{n-1}}{(n-1)!} u(t)\,\mathrm{d}t$.

The computation

With $y(0) = 0$ and $y'(0) = -1$:

$$y'(x) = -1 + \int_0^x u(t)\,\mathrm{d}t, \tag{51.1}$$

$$y(x) = 0 + (-1)x + \int_0^x (x-t)u(t)\,\mathrm{d}t = -x + \int_0^x (x-t)u(t)\,\mathrm{d}t. \tag{51.2}$$

Now rearrange the differential equation as $y'' = 5y' - 4y$, i.e. $u = 5y' - 4y$, and substitute (51.1) and (51.2):

$$u(x) = 5\left[-1 + \int_0^x u(t)\,\mathrm{d}t\right] - 4\left[-x + \int_0^x (x-t)u(t)\,\mathrm{d}t\right].$$

Expand:

$$u(x) = -5 + 4x + \int_0^x \Big[5 - 4(x-t)\Big]u(t)\,\mathrm{d}t.$$

Simplify the kernel: $5 - 4(x-t) = 5 - 4x + 4t$. Hence

$$\boxed{u(x) = 4x - 5 + \int_0^x \big(5 - 4x + 4t\big)u(t)\,\mathrm{d}t}$$

which is exactly option **(b)**.

Discrimination

The four options differ only in signs, so it is worth isolating the two decisive sign choices.

The free term. It equals $5y'(0) - 4y(0) = 5(-1) - 4(0) = -5$ plus the $-4 \cdot (-x)$ contribution $= +4x$. So the free term is $4x - 5$. Options **(a)** $(4x+5)$, **(c)** and **(d)** (both $-4x+5$) all have at least one sign wrong here. Only **(b)** matches.

The kernel. The kernel must be $5 - 4(x-t)$, in which x appears with a minus sign and t with a plus sign. Options **(a)** and **(c)** carry $+4x + 4t$, which corresponds to the impossible kernel $5 + 4(x+t)$; option **(d)** carries $+4x - 4t$, i.e. $5 + 4(x-t)$, which has the sign of the $4y$ term wrong. Again only **(b)** is right.

A verification at $x = 0$. Every candidate must satisfy $u(0) = y''(0)$. From the ODE, $y''(0) = 5y'(0) - 4y(0) = -5$. Setting $x = 0$ collapses the integral to zero, so the free term at $x = 0$ must equal -5:

Option	free term at $x = 0$	verdict
(a) $4x + 5$	$+5$	fails
(b) $4x - 5$	-5	passes
(c),**(d)** $-4x + 5$	$+5$	fail

This single evaluation settles the item in seconds and is the check to perform first under time pressure.

Why Volterra and not Fredholm. The upper limit of integration is the variable x, not a fixed constant; equations of this shape are *Volterra* equations of the second kind. They are always uniquely solvable for continuous kernels, by Picard iteration, because the Neumann series converges for every value of the parameter — unlike the Fredholm case, where the eigenvalues of the kernel obstruct solvability. (Compare Question 86.)

Answer: **(b)** $u(x) = 4x - 5 + \int_0^x (5 - 4x + 4t)u(t)\,\mathrm{d}t$.

Question 52. For the given discontinuous function f on $[0,1]$, **(a)** only S_1 is true **(b)** only S_2 is true **(c)** both S_1 and S_2 are true **(d)** neither is true.

The function and the statements did not reproduce

The classification of discontinuities

Let f be defined near a. The point a is a discontinuity of
the first kind (jump) if both one-sided limits $f(a^-)$ and $f(a^+)$ exist and are finite. If they are equal but differ from $f(a)$, the discontinuity is *removable*; if they differ from each other, it is a *jump*.
the second kind if at least one one-sided limit fails to exist (finitely). Oscillatory discontinuities, such as that of $\sin(1/x)$ at 0, are of this kind.
Two facts worth carrying. A monotone function on an interval has only jump discontinuities, and at most countably many of them. A function of bounded variation, being a difference of two monotone functions, likewise has only countably many discontinuities, all of the first kind (see Question 101).

Integrability of discontinuous functions

Lebesgue's criterion for Riemann integrability. A bounded $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if its set of discontinuities has Lebesgue measure zero.
Consequences worth having ready:

- A function with finitely or countably many discontinuities, if bounded, is Riemann integrable.
- Thomae's function (equal to $1/q$ at p/q in lowest terms and 0 at irrationals) is discontinuous exactly on $\mathbb{Q}$, a null set, hence Riemann integrable with integral 0 — despite being discontinuous on a dense set.
- The Dirichlet function $\mathbf{1}_{\mathbb{Q}}$ is discontinuous everywhere, hence not Riemann integrable, though it is Lebesgue integrable (Question 44).

Answer: Indeterminate: the function and the statements S_1, S_2 are absent from the source text.

Question 53. Which of the following statements about the mixture of bivariate normals is FALSE? **(a)** marginals under g are $N(0,1)$ **(b)** conditional under f_ρ is $N(\rho x_2,\, 1-\rho^2)$ **(c)** g is bivariate normal for all ρ_1, ρ_2 **(d)** h is not a pdf.

The set-up

The item concerns the standard construction: let f_ρ denote the standard bivariate normal density with correlation ρ,

$$f_\rho(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right),$$

and set

$$g(x_1, x_2) = \tfrac{1}{2} f_{\rho_1}(x_1, x_2) + \tfrac{1}{2} f_{\rho_2}(x_1, x_2),$$

an equal mixture of two such densities with different correlations.

Option (b): the conditional distribution — TRUE

For a standard bivariate normal with correlation ρ, complete the square in the exponent in x_1:

$$x_1^2 - 2\rho x_1 x_2 + x_2^2 = (x_1 - \rho x_2)^2 + (1 - \rho^2) x_2^2.$$

Hence

$$f_\rho(x_1, x_2) = \underbrace{\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x_2^2/2}}_{\text{marginal of } X_2} \cdot \underbrace{\frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(x_1 - \rho x_2)^2}{2(1-\rho^2)}\right)}_{\text{conditional of } X_1 \text{ given } X_2 = x_2},$$

so

$$X_1 \mid X_2 = x_2 \;\sim\; N\big(\rho x_2,\; 1 - \rho^2\big).$$

This is the fundamental regression property of the bivariate normal: the conditional mean is linear in the conditioning variable, and the conditional variance is constant (homoscedastic). **True.**

Option (a): the marginals of the mixture — TRUE

Marginalisation is a linear operation, so it commutes with mixing:

$$g_1(x_1) = \int_{\mathbb{R}} g(x_1, x_2)\, \mathrm{d}x_2 = \tfrac{1}{2} \int f_{\rho_1}\, \mathrm{d}x_2 + \tfrac{1}{2} \int f_{\rho_2}\, \mathrm{d}x_2 = \tfrac{1}{2}\phi(x_1) + \tfrac{1}{2}\phi(x_1) = \phi(x_1),$$

where ϕ is the standard normal density, since *every* standard bivariate normal has $N(0,1)$ marginals regardless of ρ. Identically for g_2. **True.**

Option (c): is the mixture bivariate normal? — FALSE

This is the false statement, and it is the classical counterexample showing that *normal marginals do not imply joint normality.*

Suppose, for contradiction, that g were bivariate normal. Its marginals are $N(0,1)$ by the previous paragraph, so it would have to be f_ρ for some ρ. Compute the correlation under g:

$$\mathbb{E}_g[X_1 X_2] = \tfrac{1}{2}\mathbb{E}_{f_{\rho_1}}[X_1 X_2] + \tfrac{1}{2}\mathbb{E}_{f_{\rho_2}}[X_1 X_2] = \tfrac{1}{2}\rho_1 + \tfrac{1}{2}\rho_2 = \bar{\rho},$$

so necessarily $g = f_{\bar{\rho}}$ with $\bar{\rho} = (\rho_1 + \rho_2)/2$. But a mixture of two *distinct* densities is never equal to a single member of the family unless the family is closed under mixing, and the normal family is not. Concretely, compare fourth moments: for f_ρ,

$$\mathbb{E}[X_1^2 X_2^2] = 1 + 2\rho^2.$$

Under the mixture,

$$\mathbb{E}_g[X_1^2 X_2^2] = \tfrac{1}{2}(1 + 2\rho_1^2) + \tfrac{1}{2}(1 + 2\rho_2^2) = 1 + \rho_1^2 + \rho_2^2,$$

whereas $f_{\bar{\rho}}$ would give $1 + 2\bar{\rho}^2 = 1 + \tfrac{1}{2}(\rho_1 + \rho_2)^2$. These agree if and only if

$$\rho_1^2 + \rho_2^2 = \tfrac{1}{2}(\rho_1 + \rho_2)^2 \iff \tfrac{1}{2}(\rho_1 - \rho_2)^2 = 0 \iff \rho_1 = \rho_2.$$

So g is bivariate normal *only* in the degenerate case $\rho_1 = \rho_2$, and the claim "for all ρ_1, ρ_2" is **false**.

Option (d)

The function h is not specified in the available text. In the usual version of this problem h is constructed so as to integrate to something other than 1 (or to take negative values), making "h is not a pdf" a true statement. Since option **(c)** is demonstrably false and the item asks for the false statement, the key is **(c)** regardless.

> **Answer:** **(c)** — the mixture is bivariate normal only when $\rho_1 = \rho_2$. This is the standard example of normal marginals without joint normality.

> **Question 54.** For the linear map $T(A) = A^{\mathsf{T}}$ on $M_3(\mathbb{R})$, **(a)** only S_1 is true **(b)** only S_2 is true **(c)** both are true **(d)** neither is true.

The statements did not reproduce; the operator is fully analysable

Although S_1 and S_2 are missing, the transpose operator on $M_3(\mathbb{R})$ can be analysed completely, and doing so will decide almost any pair of statements the setter might have posed.

The complete analysis of T

T is linear. $(A + \lambda B)^{\mathsf{T}} = A^{\mathsf{T}} + \lambda B^{\mathsf{T}}$.

T is an involution. $T^2(A) = (A^{\mathsf{T}})^{\mathsf{T}} = A$, so $T^2 = I$. Consequently the minimal polynomial of T divides $x^2 - 1 = (x-1)(x+1)$, which has distinct roots. Hence:

- T is **diagonalisable**;
- every eigenvalue of T lies in $\{+1, -1\}$;
- T is invertible, with $T^{-1} = T$.

The eigenspaces. $T(A) = A$ means $A^{\mathsf{T}} = A$: the symmetric matrices. $T(A) = -A$ means $A^{\mathsf{T}} = -A$: the skew-symmetric matrices. For $n = 3$:

$$\dim\{\text{symmetric}\} = \frac{n(n+1)}{2} = 6, \qquad \dim\{\text{skew}\} = \frac{n(n-1)}{2} = 3,$$

and $6 + 3 = 9 = \dim M_3(\mathbb{R})$. ✓ The decomposition

$$A = \underbrace{\tfrac{1}{2}(A + A^{\mathsf{T}})}_{\text{symmetric}} + \underbrace{\tfrac{1}{2}(A - A^{\mathsf{T}})}_{\text{skew}}$$

exhibits it explicitly, and confirms directly that $M_3(\mathbb{R})$ is the direct sum of the two eigenspaces.

Consequent invariants.

$$\operatorname{tr} T = 6(+1) + 3(-1) = 3, \qquad \det T = (+1)^6(-1)^3 = -1.$$

The characteristic polynomial is $(x-1)^6(x+1)^3$ and the minimal polynomial is $x^2 - 1$.

Further properties. T is an orthogonal (indeed self-adjoint) operator with respect to the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(A^{\mathsf{T}}B)$, since

$$\langle T(A), B\rangle = \operatorname{tr}(AB) = \operatorname{tr}(BA) = \langle A, T(B)\rangle,$$

and the two eigenspaces are orthogonal complements. T is not a ring homomorphism on $M_3(\mathbb{R})$, since $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}} \neq A^{\mathsf{T}}B^{\mathsf{T}}$ in general — it is an *anti*-homomorphism.

Summary table for reference.

Property	Value for $T(A) = A^{\mathsf{T}}$ on $M_n(\mathbb{R})$
Minimal polynomial	$x^2 - 1$
Eigenvalues	$+1$ (multiplicity $n(n+1)/2$), -1 (multiplicity $n(n-1)/2$)
Diagonalisable	Yes
Trace	$n(n+1)/2 - n(n-1)/2 = n$
Determinant	$(-1)^{n(n-1)/2}$
T^{-1}	T

For $n = 3$: trace 3, determinant $(-1)^3 = -1$.

Answer: Indeterminate: S_1 and S_2 are absent. The complete spectral data for T are given above and will settle any standard pair of statements.

Question 55. For the topologies generated by the given collections of intervals, **(a)** $\mathbb{R}$ is not Hausdorff in τ_3 **(b)** $\tau_2 = \tau_3$ **(c)** $\tau_1 \subseteq \tau_3$ **(d)** $\mathbb{R}$ is connected in τ_3.

The generating collections did not reproduce

The three standard interval topologies on $\mathbb{R}$ are these, and virtually all examination items on the topic use them.

The three interval topologies on $\mathbb{R}$

τ_{std}**, the standard topology**, generated by the open intervals (a, b).
τ_ℓ**, the lower-limit (Sorgenfrey) topology**, generated by the half-open intervals $[a, b)$.
τ_u**, the upper-limit topology**, generated by the half-open intervals $(a, b]$.
Comparison. τ_ℓ and τ_u are both *strictly finer* than τ_{std}:

$$(a,b) = \bigcup_n \left[a + \tfrac{1}{n},\, b\right) \in \tau_\ell, \qquad (a,b) = \bigcup_n \left(a,\, b - \tfrac{1}{n}\right] \in \tau_u,$$

so $\tau_{\mathrm{std}} \subseteq \tau_\ell$ and $\tau_{\mathrm{std}} \subseteq \tau_u$; and the inclusions are strict since $[a, b)$ is not standard-open. Moreover τ_ℓ and τ_u are *incomparable* with each other: $[0, 1)$ is τ_ℓ-open but not τ_u-open, and $(0, 1]$ is τ_u-open but not τ_ℓ-open.

The properties in question

Hausdorff. All three are Hausdorff. A topology finer than a Hausdorff topology is automatically Hausdorff, since the separating open sets remain open. So option **(a)** is false under every reading.
Connectedness. $(\mathbb{R}, \tau_{\mathrm{std}})$ is connected. But $\mathbb{R}$ in the Sorgenfrey topology is *totally disconnected*: for any a, the set $[a, \infty)$ is open (a union of sets $[a, n)$) and its complement $(-\infty, a)$ is also open (a union of sets $[a - n, a)$), so $\{[a, \infty), (-\infty, a)\}$ is a separation. The same argument with (a, ∞) and $(-\infty, a]$ works in τ_u.

Further contrasts. Sorgenfrey $\mathbb{R}$ is first countable, separable, Lindelöf and normal, but *not* second countable and *not* metrisable; the Sorgenfrey plane $\mathbb{R}_\ell \times \mathbb{R}_\ell$ is the classical example of a product of normal spaces that is not normal.

Evaluating the options under the standard labelling

Take the most common convention, $\tau_1 = \tau_{\text{std}}$ (open intervals), $\tau_2 = \tau_\ell$ (lower-limit), $\tau_3 = \tau_u$ (upper-limit):

Option	Statement	Verdict
(a)	$\mathbb{R}$ not Hausdorff in τ_3	False — finer than standard
(b)	$\tau_2 = \tau_3$	False — incomparable
(c)	$\tau_1 \subseteq \tau_3$	**True** — standard is coarsest
(d)	$\mathbb{R}$ connected in τ_3	False — totally disconnected

Answer: **(c)** $\tau_1 \subseteq \tau_3$, under the standard labelling $\tau_1 =$ open intervals, $\tau_2 =$ lower-limit, $\tau_3 =$ upper-limit. The labelling could not be verified against the source; the reasoning is given so that any labelling can be evaluated.

Question 56. Concerning the two statements about Rolle-type theorems and mean-value expansions, **(a)** only S_1 is true **(b)** only S_2 is true **(c)** both are true **(d)** neither is true.

The statements did not reproduce

Rolle's theorem

If f is continuous on $[a, b]$, differentiable on (a, b), and $f(a) = f(b)$, then there exists $c \in (a, b)$ with $f'(c) = 0$.
All three hypotheses are essential and each has a standard counterexample:

- Continuity on the *closed* interval: $f(x) = x$ on $[0, 1)$ with $f(1) = 0$ has $f(0) = f(1)$ but no interior critical point.
- Differentiability on the open interval: $f(x) = |x|$ on $[-1, 1]$ has $f(-1) = f(1)$ but f' never vanishes.
- Equal endpoint values: obvious.

Note that differentiability is required only on the *open* interval, so $f(x) = \sqrt{1 - x^2}$ on $[-1, 1]$ satisfies Rolle despite infinite one-sided derivatives at the endpoints.

The mean value theorems

Lagrange MVT. Under the same hypotheses without $f(a) = f(b)$, there is $c \in (a, b)$ with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof: apply Rolle to $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$.
Cauchy MVT. If f, g are continuous on $[a,b]$ and differentiable on (a,b), there is c with

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c).$$

Proof: apply Rolle to $h = \big(f(b) - f(a)\big)g - \big(g(b) - g(a)\big)f$. This is the engine behind L'Hôpital's rule.
Taylor with Lagrange remainder. If f is n times differentiable on (a,b) with $f^{(n-1)}$ continuous on $[a,b]$, then for $x \in (a,b)$

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}(\xi)}{n!}(x-a)^n$$

for some ξ strictly between a and x. The case $n = 1$ is the Lagrange MVT.

Two traps that appear repeatedly in items of this shape.
1. The point c is not unique and not continuous in the data. Statements asserting that "c is unique" or that "$c \to (a+b)/2$" are false in general, though the latter does hold asymptotically for f with $f''(a) \neq 0$ as $b \to a$.
2. Rolle's theorem fails for vector-valued functions. Take $f : [0, 2\pi] \to \mathbb{R}^2$, $f(t) = (\cos t, \sin t)$. Then $f(0) = f(2\pi)$ but $|f'(t)| = 1$ never vanishes. The mean value *theorem* thus has no exact vector analogue; only the mean value *inequality* $|f(b) - f(a)| \leq (b-a) \sup |f'|$ survives.

Answer: Indeterminate: S_1 and S_2 are absent from the source text.

Question 57. The power of the UMP test of size 0.1 for the given densities is **(a)** $1 - (9/10)^{3/5}$ **(b)** $(9/10)^{5/3}$ **(c)** $1 - (9/10)^{5/3}$ **(d)** $(9/10)^{3/5}$.

The densities did not reproduce

The Neyman–Pearson lemma

To test a simple null $H_0 : \theta = \theta_0$ against a simple alternative $H_1 : \theta = \theta_1$ at level α, the most powerful test rejects when the likelihood ratio is large:

$$\text{reject } H_0 \iff \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} > k,$$

where k is chosen so that $\mathbb{P}_{\theta_0}(\text{reject}) = \alpha$. Any other test of size $\leq \alpha$ has power no greater at θ_1.
If the likelihood ratio is monotone in a statistic T and that monotonicity holds for *every* $\theta_1 > \theta_0$ (the *monotone likelihood ratio* property), the same test is uniformly most powerful for the one-sided composite alternative $H_1 : \theta > \theta_0$.

The standard exponential-family computation

Take $X \sim \mathrm{Exp}(\theta)$, i.e. $f(x;\theta) = \theta e^{-\theta x}$ on $x > 0$, and test $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1 < \theta_0$ (smaller rate, larger mean). The likelihood ratio

$$\frac{f_1}{f_0} = \frac{\theta_1}{\theta_0} e^{(\theta_0 - \theta_1)x}$$

is increasing in x, so the UMP test rejects for $X > c$.
Size: $\mathbb{P}_{\theta_0}(X > c) = e^{-\theta_0 c} = \alpha$, so $c = -\ln\alpha/\theta_0$.
Power: $\mathbb{P}_{\theta_1}(X > c) = e^{-\theta_1 c} = e^{(\theta_1/\theta_0)\ln\alpha} = \alpha^{\theta_1/\theta_0}$.
So the power has the closed form $\alpha^{\theta_1/\theta_0}$, and if the rejection region is instead in the lower tail the power becomes $1 - (1-\alpha)^{\theta_1/\theta_0}$.

Reading the options. Every option is built from $9/10 = 1 - \alpha$ raised to $3/5$ or $5/3$. This is the signature of the second form above, with $1 - \alpha = 0.9$ and a parameter ratio of $3/5$ or its reciprocal. Options **(a)** and **(c)** have the shape $1 - (1-\alpha)^\lambda$ (rejection region in the lower tail); options **(b)** and **(d)** have the shape $(1-\alpha)^\lambda$.
Without the densities the exponent cannot be fixed. Note as a consistency check that power must exceed size: 0.1. Evaluating, $1 - (0.9)^{3/5} = 0.0619$ — *below* the size, so option **(a)** would describe a biased test; $1 - (0.9)^{5/3} = 0.1652$; $(0.9)^{5/3} = 0.8348$; $(0.9)^{3/5} = 0.9381$. Option **(a)** is thus eliminable on the grounds that a UMP test cannot be biased in a one-parameter exponential family.

Answer: Indeterminate: the densities are absent. Option **(a)** is eliminable because its value (0.062) falls below the size (0.1), which a UMP test in this setting cannot do.

Question 58. For the matrices P, Q, R, the similarity relations are **(a)** $P \sim Q$ and $Q \sim R$ **(b)** $P \sim Q$ and $Q \nsim R$ **(c)** $P \nsim Q$ and $Q \sim R$ **(d)** $P \nsim Q$ and $Q \nsim R$.

The matrices did not reproduce

Similarity and its complete invariant

$A \sim B$ means $B = S^{-1}AS$ for some invertible S; equivalently, A and B represent the same linear operator in different bases.
The complete invariant over an algebraically closed field is the Jordan canonical form, up to permutation of blocks. Over a general field it is the *rational canonical form*, equivalently the list of invariant factors.

The checklist, in order of cost

Test in this order; the first failure settles the matter.
1. Cheap necessary conditions. Similar matrices share: trace, determinant, rank, characteristic polynomial, minimal polynomial, eigenvalues with algebraic *and* geometric multiplicities, and $\mathrm{rank}(A - \lambda I)^k$ for every λ and k.

2. Characteristic polynomial. If these differ, not similar. If they agree, continue — equality is necessary but not sufficient.
3. Minimal polynomial. If these differ, not similar. Together with the characteristic polynomial this settles all cases up to size 3×3, and all cases where no eigenvalue has algebraic multiplicity exceeding 3.
4. Geometric multiplicities. $\dim\ker(A-\lambda I)$ is the *number* of Jordan blocks for λ. Combined with the algebraic multiplicity (total size) and the minimal polynomial (largest block), this determines the Jordan form in most cases.
5. Full block structure. The ranks $\operatorname{rank}(A-\lambda I)^k$ for $k = 1, 2, \ldots$ determine the block sizes completely, via

$$\#\{\text{blocks of size } \geq k\} = \operatorname{rank}(A-\lambda I)^{k-1} - \operatorname{rank}(A-\lambda I)^k.$$

The classic trap. The matrices

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

have the same characteristic polynomial x^4 and the same rank 2, but different minimal polynomials (x^2 versus x^3), so they are not similar. Equality of characteristic polynomial and rank is not enough — the smallest size at which characteristic and minimal polynomials together fail to determine similarity is 4×4, with λ of algebraic multiplicity 4, minimal polynomial $(x-\lambda)^2$, and block structures $2+2$ versus $2+1+1$.

Answer: Indeterminate: the matrices are absent from the source text.

Question 59. If $\mathbb{E}[\hat{y}_R] = 7/8$ for the given ratio estimator, the value of c is **(a)** 0 **(b)** 3 **(c)** 2 **(d)** 1.

The estimator did not reproduce

The ratio estimator

In survey sampling, suppose an auxiliary variable x is measured on every unit of the population (so the population mean $\bar{X}$ is known) and is correlated with the study variable y, which is measured only on the sample. The *ratio estimator* of the population mean $\bar{Y}$ is

$$\hat{\bar{y}}_R = \frac{\bar{y}}{\bar{x}}\,\bar{X},$$

where $\bar{y}, \bar{x}$ are the sample means.
It is biased. Because $\bar{y}/\bar{x}$ is a nonlinear function of the sample, $\mathbb{E}[\bar{y}/\bar{x}] \neq \mathbb{E}[\bar{y}]/\mathbb{E}[\bar{x}]$ in general. The bias is $O(1/n)$ and to first order

$$\operatorname{Bias}(\hat{\bar{y}}_R) \approx \frac{1}{\bar{X}}\Big(R\operatorname{Var}(\bar{x}) - \operatorname{Cov}(\bar{y},\bar{x})\Big), \qquad R = \frac{\bar{Y}}{\bar{X}},$$

which vanishes when the regression of y on x passes through the origin with slope R — the condition under which the ratio estimator is exactly right.
Its virtue is variance reduction: it beats the simple sample mean whenever

$$\rho > \frac{1}{2} \cdot \frac{C_x}{C_y},$$

where C_x, C_y are the coefficients of variation and ρ the correlation.

How items of this shape are solved. The setter typically defines an estimator containing a free constant c, computes $\mathbb{E}[\hat{y}_R]$ exactly over a small population (usually by enumerating all $\binom{N}{n}$ samples), and sets the result equal to a target. The candidate must:

1. Enumerate the samples and their probabilities under the stated design.
2. Compute the estimator's value on each.
3. Form the expectation as a function of c.
4. Solve the resulting linear equation.

The value $7/8$ with denominators 8 suggests a population of size $N = 4$ with samples of size $n = 2$ (so $\binom{4}{2} = 6$) or a design with 8 equally likely outcomes. Without the estimator's definition the equation cannot be formed.

Answer: Indeterminate: the definition of the ratio estimator and the population are absent from the source text.

Question 60. The dimension of $V + W$ inside $P_4(\mathbb{R})$ is **(a)** 5 **(b)** 6 **(c)** 4 **(d)** 3.

The subspaces V and W did not reproduce

The dimension formula for sums

For finite-dimensional subspaces V, W of a vector space,

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W).$$

Proof. Extend a basis $\{u_1, \dots, u_k\}$ of $V \cap W$ to a basis $\{u_i\} \cup \{v_1, \dots, v_p\}$ of V and to a basis $\{u_i\} \cup \{w_1, \dots, w_q\}$ of W. Then $\{u_i\} \cup \{v_j\} \cup \{w_l\}$ spans $V + W$, and one checks it is independent. Counting gives $k + p + q = (k + p) + (k + q) - k$. □
The sum is *direct*, written $V \oplus W$, exactly when $V \cap W = \{0\}$, in which case $\dim(V \oplus W) = \dim V + \dim W$.

What the option set forces

The ambient space is $P_4(\mathbb{R})$, real polynomials of degree at most 4, with basis $1, x, x^2, x^3, x^4$; hence

$$\dim P_4(\mathbb{R}) = 5.$$

Since $V + W$ is a subspace of $P_4(\mathbb{R})$,

$$\dim(V + W) \leq 5.$$

This immediately eliminates option **(b)**, 6: no subspace of a five-dimensional space can have dimension six. Options **(a)**, **(c)** and **(d)** — the values 5, 4 and 3 — all remain possible, and choosing between them requires the definitions of V and W.

The off-by-one to guard against. $P_n(\mathbb{R})$ denotes polynomials of degree *at most* n and has dimension $n + 1$; the set of polynomials of degree *exactly* n is not a subspace at all (it is not closed under addition, since $x^4 + (-x^4 + 1) = 1$). A candidate who takes $\dim P_4(\mathbb{R}) = 4$ would wrongly eliminate option **(a)** as well.
Compare Question 26, where the same convention gives $\dim P_{30}(\mathbb{R}) = 31$ and changes the answer.

Answer: Indeterminate: V and W are absent. Option **(b)** (6) is impossible, since $\dim P_4(\mathbb{R}) = 5$.

CHAPTER 3

Part C — Advanced Mathematical Sciences

Part C carries sixty items of which twenty are to be attempted. Each item may have one, two, three or four correct options, and — in the standard CSIR scheme — a response scores only if *every* correct option is marked and no incorrect one is. The consequence for strategy is severe: a partially understood item is worth nothing, and it is better to answer fifteen items completely than twenty-five partially.

The commentary below therefore adjudicates every option separately, marking each TRUE or FALSE with its own argument.

> **Question 61.** Which of the following statements is/are true? **(a)** If G is a non-abelian group of order 243, then $|\mathrm{Aut}(G)| \geq 9$ **(b)** $|\mathrm{Aut}(S_3)| = 6$ **(c)** $Z(S_3) \neq \{I\}$ **(d)** The order of 13 in U_{18} is 6.

(a) — TRUE

Inner automorphisms and the centre

For any group G the map $g \mapsto \varphi_g$, where $\varphi_g(x) = gxg^{-1}$, is a homomorphism $G \to \mathrm{Aut}(G)$ with kernel exactly $Z(G)$. Hence

$$\mathrm{Inn}(G) \cong G/Z(G) \ \leq \ \mathrm{Aut}(G), \qquad \text{so} \qquad |\mathrm{Aut}(G)| \ \geq \ [G : Z(G)].$$

Lemma 3.1. If $G/Z(G)$ is cyclic then G is abelian.

Proof. Suppose $G/Z(G) = \langle gZ \rangle$. Every $x \in G$ then has the form $x = g^a z$ with $z \in Z(G)$. For $x = g^a z_1$ and $y = g^b z_2$,

$$xy = g^a z_1 g^b z_2 = g^{a+b} z_1 z_2 = g^b z_2 g^a z_1 = yx,$$

using centrality of z_1, z_2. So G is abelian. □

Now let $|G| = 243 = 3^5$ with G non-abelian. A p-group has non-trivial centre (the class equation $|G| = |Z(G)| + \sum [G : C(x_i)]$ forces $p \mid |Z(G)|$), so $|Z(G)| \in \{3, 9, 27, 81\}$; and $|Z(G)| \neq 243$ since G is non-abelian. Therefore $[G : Z(G)] \geq 3$. If $[G : Z(G)] = 3$ then $G/Z(G)$ has prime order, hence is cyclic, hence G is abelian by the lemma — a contradiction. So

$$[G : Z(G)] \geq 9 \quad \Longrightarrow \quad |\mathrm{Aut}(G)| \geq |\mathrm{Inn}(G)| = [G : Z(G)] \geq 9.$$

True.

(b) — TRUE

$\mathrm{Aut}(S_3) \cong S_3$, so $|\mathrm{Aut}(S_3)| = 6$.

Reason. $Z(S_3) = \{e\}$ (verified below), so $\mathrm{Inn}(S_3) \cong S_3/Z(S_3) \cong S_3$ has order 6, giving $|\mathrm{Aut}(S_3)| \geq 6$. Conversely, any automorphism permutes the three transpositions $(12), (13), (23)$ — these are the only elements of order 2 — and is determined by that permutation, since the transpositions generate S_3. So $|\mathrm{Aut}(S_3)| \leq |S_3| = 6$. Hence equality. **True.**

(More generally $\mathrm{Aut}(S_n) \cong S_n$ for all $n \neq 2, 6$; S_6 is the celebrated exception, possessing outer automorphisms.)

(c) — FALSE

$Z(S_3) = \{e\}$. Indeed a transposition such as (12) does not commute with (13):

$$(12)(13) = (132), \qquad (13)(12) = (123),$$

which differ. And a 3-cycle such as (123) does not commute with (12):

$$(123)(12) = (13), \qquad (12)(123) = (23).$$

So no non-identity element is central. The assertion $Z(S_3) \neq \{I\}$ is **false**. (In general $Z(S_n) = \{e\}$ for $n \geq 3$.)

(d) — FALSE

$U_{18} = (\mathbb{Z}/18\mathbb{Z})^\times$ has order $\varphi(18) = \varphi(2)\varphi(9) = 1 \cdot 6 = 6$, and consists of $\{1, 5, 7, 11, 13, 17\}$. Compute powers of 13 modulo 18:

$$13^1 = 13, \qquad 13^2 = 169 = 9 \cdot 18 + 7 \equiv 7, \qquad 13^3 = 13 \cdot 7 = 91 = 5 \cdot 18 + 1 \equiv 1.$$

So the order of 13 is 3, not 6. **False.**

U_{18} is cyclic of order 6 (since $18 = 2 \cdot 3^2$ is of the form $2p^k$), with generator 5: $5, 7, 17, 13, 11, 1$. The element $13 = 5^4$ has order $6/\gcd(4, 6) = 6/2 = 3$, confirming the computation.

Answer: **(a)** and **(b)**.

Question 62. For the Wilcoxon signed-rank statistic T^+ based on a sample of size 4 from a continuous symmetric distribution, **(a)** $\mathbb{P}(T^+ \text{ even}) = 1/2$ **(b)** $\mathbb{E}(T^+) = 5$ **(c)** $\mathbb{P}(T^+ \text{ odd}) = 7/16$ **(d)** $\mathbb{P}(T^+ = 0) = 1/8$.

The concept: the null distribution of T^+

The Wilcoxon signed-rank statistic

Given $X_1, \ldots, X_n$ i.i.d. from a continuous distribution symmetric about 0 under H_0, rank the absolute values $|X_i|$ from 1 to n and set

$$T^+ = \sum_{i=1}^{n} R_i \, \mathbf{1}_{\{X_i > 0\}},$$

the sum of the ranks attached to positive observations.

The key structural fact. Under H_0, the signs $\operatorname{sgn}(X_i)$ are i.i.d. symmetric Bernoulli and are *independent* of the ranks $|X_i|$. Hence

$$T^+ \stackrel{d}{=} \sum_{r=1}^{n} r\, B_r, \qquad B_r \text{ i.i.d. Bernoulli}(1/2).$$

Equivalently: each of the 2^n subsets of $\{1,2,\ldots,n\}$ is equally likely, and T^+ is the sum of the chosen subset. This is a distribution-free statement — it does not depend on the underlying distribution at all, which is the whole point of the rank test.

The full distribution for $n = 4$

There are $2^4 = 16$ equally likely subsets of $\{1,2,3,4\}$. Enumerate by subset sum:

Sum	Subsets	Count
0	$\emptyset$	1
1	$\{1\}$	1
2	$\{2\}$	1
3	$\{3\},\{1,2\}$	2
4	$\{4\},\{1,3\}$	2
5	$\{1,4\},\{2,3\}$	2
6	$\{2,4\},\{1,2,3\}$	2
7	$\{3,4\},\{1,2,4\}$	2
8	$\{1,3,4\}$	1
9	$\{2,3,4\}$	1
10	$\{1,2,3,4\}$	1
		16

Note the symmetry of the counts about 5, reflecting the fact that $T^+ \stackrel{d}{=} \frac{n(n+1)}{2} - T^+ = 10 - T^+$ (complementing the subset).

Adjudicating the options

(b) TRUE. By linearity,

$$\mathbb{E}[T^+] = \sum_{r=1}^{4} r \cdot \mathbb{E}[B_r] = \tfrac{1}{2}\sum_{r=1}^{4} r = \tfrac{1}{2}\cdot\frac{4\cdot 5}{2} = 5.$$

The general formula is $\mathbb{E}[T^+] = n(n+1)/4$, here $4\cdot 5/4 = 5$. ✓

(a) TRUE. Sum the counts at even values $0,2,4,6,8,10$:

$$1+1+2+2+1+1 = 8,$$

so $\mathbb{P}(T^+ \text{ even}) = 8/16 = 1/2$.

A cleaner argument. Write $T^+ = B_1 + 2B_2 + 3B_3 + 4B_4$. Modulo 2 this is $B_1 + B_3$, a sum of two independent fair coin flips, which is even with probability $1/2$. The even terms $2B_2$ and $4B_4$ are irrelevant to parity. This argument shows more generally that $\mathbb{P}(T^+ \text{ even}) = 1/2$ whenever $n \ge 1$, since the odd ranks contribute at least one independent fair bit.

(c) FALSE. By complementarity with **(a)**, $\mathbb{P}(T^+ \text{ odd}) = 1 - 1/2 = 1/2$, not $7/16$. (Direct count: $1+2+2+2+1 = 8$ at values $1, 3, 5, 7, 9$. ✓)

(d) FALSE. $T^+ = 0$ occurs only for the empty subset — all four observations negative — so

$$\mathbb{P}(T^+ = 0) = \frac{1}{16},$$

not $1/8$. The value $1/8 = 2/16$ would be $\mathbb{P}(T^+ \in \{0,1\})$ or $\mathbb{P}(T^+ = 0)$ for $n = 3$.

Variance, for completeness.

$$\operatorname{Var}(T^+) = \sum_{r=1}^{n} r^2 \operatorname{Var}(B_r) = \tfrac{1}{4} \sum r^2 = \frac{n(n+1)(2n+1)}{24},$$

which for $n = 4$ is $4 \cdot 5 \cdot 9/24 = 7.5$. The normal approximation $T^+ \approx N(5, 7.5)$ is poor at $n = 4$ but becomes serviceable by $n \approx 20$.

Answer: **(a)** and **(b)**.

Question 63. For the function $f(x,y) = \dfrac{x^6 - y^3}{|x| + y^2}$ (with $f = 0$ at the origin), **(a)** f is differentiable at $(0,0)$ **(b)** f is unbounded near $(0,0)$ **(c)** $f_x(1,1) + f_y(1,1) > 0$ **(d)** f increases at $(1,1)$ along $\frac{\mathbf{i}+\mathbf{j}}{\sqrt{2}}$.

Behaviour near the origin

Split the quotient and bound each piece separately. For $(x,y) \neq (0,0)$,

$$\frac{|x|^6}{|x| + y^2} \le \frac{|x|^6}{|x|} = |x|^5, \qquad \frac{|y|^3}{|x| + y^2} \le \frac{|y|^3}{y^2} = |y|.$$

Hence

$$|f(x,y)| \le |x|^5 + |y| \longrightarrow 0 \quad \text{as } (x,y) \to (0,0). \tag{63.1}$$

(b) FALSE. Estimate (63.1) shows f is bounded — indeed $f \to 0$ — near the origin. So f is continuous at $(0,0)$ and certainly not unbounded.

(a) — FALSE

First compute the partial derivatives at the origin from the definition.

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - 0}{h} = \lim_{h\to 0} \frac{1}{h} \cdot \frac{h^6}{|h|} = \lim_{h\to 0} \frac{h^6}{h\,|h|} = \lim_{h\to 0} \pm |h|^4 = 0.$$

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - 0}{k} = \lim_{k\to 0} \frac{1}{k} \cdot \frac{-k^3}{k^2} = \lim_{k\to 0} \frac{-k}{k} = -1.$$

So both partials exist, with $\nabla f(0,0) = (0,-1)$ — but existence of partials is far from differentiability.

Differentiability at the origin requires

$$\frac{f(h,k) - f(0,0) - \big[0 \cdot h + (-1)k\big]}{\sqrt{h^2 + k^2}} \longrightarrow 0 \quad \text{as } (h,k) \to (0,0).$$

Compute the numerator:

$$f(h,k)+k=\frac{h^6-k^3}{|h|+k^2}+k=\frac{h^6-k^3+k\,|h|+k^3}{|h|+k^2}=\frac{h^6+k\,|h|}{|h|+k^2}.$$

Approach along $h=k=t$ with $t>0$:

$$\frac{t^6+t^2}{t+t^2}=\frac{t^2(t^4+1)}{t(1+t)}=\frac{t(t^4+1)}{1+t}\sim t,\qquad \sqrt{h^2+k^2}=t\sqrt{2}.$$

The ratio therefore tends to $1/\sqrt{2}\neq 0$. The linear approximation fails, so f is **not differentiable** at the origin.

The moral. Continuity plus existence of both partials does *not* imply differentiability. What does suffice is continuity of the partials in a neighbourhood (C^1 implies differentiable), and here f_x is not continuous at the origin: from the computation below, $f_x(0,y)=y^{-1}$-type behaviour which does not tend to $f_x(0,0)=0$.

At the point $(1,1)$

Near $(1,1)$ we have $x>0$, so $|x|=x$ and

$$f(x,y)=\frac{x^6-y^3}{x+y^2}$$

is a smooth rational function there. Note $f(1,1)=(1-1)/2=0$.

By the quotient rule, with $N=x^6-y^3$ and $D=x+y^2$:

$$f_x=\frac{6x^5D-N\cdot 1}{D^2},\qquad f_y=\frac{-3y^2D-N\cdot 2y}{D^2}.$$

At $(1,1)$: $N=0$, $D=2$, $D^2=4$. Hence

$$f_x(1,1)=\frac{6\cdot 1\cdot 2-0}{4}=\frac{12}{4}=3,\qquad f_y(1,1)=\frac{-3\cdot 1\cdot 2-0}{4}=\frac{-6}{4}=-\frac{3}{2}.$$

The vanishing of N at $(1,1)$ simplifies both enormously — worth noticing before differentiating.

(c) TRUE.

$$f_x(1,1)+f_y(1,1)=3-\tfrac{3}{2}=\tfrac{3}{2}>0.$$

(d) TRUE. The directional derivative in the unit direction $\mathbf{u}=\frac{1}{\sqrt{2}}(1,1)$ is

$$D_{\mathbf{u}}f(1,1)=\nabla f(1,1)\cdot\mathbf{u}=\left(3,-\tfrac{3}{2}\right)\cdot\left(\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right)=\frac{3-\frac{3}{2}}{\sqrt{2}}=\frac{3}{2\sqrt{2}}\approx 1.06>0.$$

A positive directional derivative means f increases in that direction.

Observe that **(c)** and **(d)** are logically equivalent here: the directional derivative along $(1,1)/\sqrt{2}$ is exactly $(f_x+f_y)/\sqrt{2}$, so one is positive if and only if the other is. Recognising this saves a computation.

Answer: **(c)** and **(d)**.

Question 64. Which of the following statements about Lebesgue measure is/are true? **(a)** Every uncountable set has positive measure **(b)** The set G of numbers with only digits 0 and 4 in base 5 has measure $2/5$ **(c)** The set of numbers in $[0, 1]$ whose decimal expansion avoids the digit 7 has measure zero **(d)** The set S constructed from an enumeration of $\mathbb{Q}$ has the stated property.

(a) — FALSE

The Cantor middle-thirds set $C \subseteq [0, 1]$ is uncountable yet has Lebesgue measure zero.

Uncountability. Every point of C has a unique ternary expansion using only the digits 0 and 2, so C is in bijection with $\{0, 2\}^{\mathbb{N}}$, of cardinality $2^{\aleph_0} = \mathfrak{c}$.

Measure zero. At stage n the construction retains 2^n intervals each of length 3^{-n}, so

$$\lambda(C) \le \left(\frac{2}{3}\right)^n \longrightarrow 0.$$

Hence $\lambda(C) = 0$. So cardinality and measure are entirely independent notions: an uncountable set may be null, and (trivially) a countable set always is.

(b) — FALSE

The set G consists of those $x \in [0, 1]$ whose base-5 expansion uses only the digits 0 and 4. This is a Cantor-type construction in base 5: at each stage, of the 5 available sub-intervals we retain 2.

Stage 1: retain 2 intervals of length $1/5$; total measure $2/5$.
Stage 2: within each, retain 2 of 5; total measure $(2/5)^2 = 4/25$.
Stage n: total measure $(2/5)^n$.

Therefore

$$\lambda(G) = \lim_{n\to\infty} \left(\frac{2}{5}\right)^n = 0.$$

The value $2/5$ is the measure after *one* stage only; the claim confuses the first approximation with the limit. **False.**

(c) — TRUE

Let A be the set of $x \in [0, 1]$ whose decimal expansion contains no digit 7. The same computation applies with base 10 and 9 retained digits:

$$\lambda(A) = \lim_{n\to\infty} \left(\frac{9}{10}\right)^n = 0,$$

since $(9/10)^n \to 0$. **True.**

The probabilistic reading. If the digits were chosen independently and uniformly, the probability that none of the first n is a 7 is $(9/10)^n$, and the event "no 7 ever appears" has probability 0. This is a special case of Borel's normal number theorem: almost every real number in $[0, 1]$ is normal to every base, meaning every finite digit string appears with its expected asymptotic frequency. The exceptional set — which contains all rationals, all numbers with a missing digit, and the whole Cantor set — has measure zero but cardinality $\mathfrak{c}$. **Note also** that A is uncountable (bijective with $\{0, \ldots, 9\} \setminus \{7\}$ strings, i.e. $9^{\mathbb{N}}$), giving a second counterexample to option **(a)**.

(d)

The construction of S from an enumeration of $\mathbb{Q}$, and the property asserted, are not present in the available text. The standard construction is: enumerate $\mathbb{Q} \cap [0,1] = \{q_1, q_2, \dots\}$, fix $\varepsilon > 0$, and set

$$S = \bigcup_{n \geq 1} \left(q_n - \frac{\varepsilon}{2^{n+1}},\ q_n + \frac{\varepsilon}{2^{n+1}} \right).$$

Then S is open and dense in $[0,1]$, yet by countable subadditivity $\lambda(S) \leq \sum_n \varepsilon 2^{-n} = \varepsilon$, which can be made arbitrarily small. Its complement in $[0,1]$ is closed, nowhere dense, and of measure at least $1 - \varepsilon$ — a "fat Cantor set". This example shows that *dense open* does not mean *large in measure*, and that *nowhere dense* does not mean *null*: topological and measure-theoretic smallness are incomparable.

Answer: **(c)**. (**(a)** and **(b)** are false; **(d)** is indeterminate as printed.)

Question 65. For the wave equation $u_{tt} = 4u_{xx}$ with Neumann boundary conditions and the given initial data, the constants in the Fourier solution satisfy **(a)** $A + B = 6$ **(b)** $A^2 + B^2 = 20$ **(c)** $A^2 - B^2 = 12$ **(d)** $A - B = 10$.

The initial data did not reproduce — but the options determine each other

The four options are four algebraic conditions on the same pair (A, B). Since a consistent item must have all its true options simultaneously satisfiable by one pair, the option set can be tested for internal consistency, and this is decisive here.

Testing consistency

*Suppose **(a)** and **(c)** both hold.* From $A + B = 6$ and $A^2 - B^2 = (A+B)(A-B) = 12$ we get

$$6(A - B) = 12 \quad \Longrightarrow \quad A - B = 2.$$

Combined with $A + B = 6$:

$$A = 4, \qquad B = 2.$$

Check **(b)**: $A^2 + B^2 = 16 + 4 = 20$. ✓ **So (a), (b), (c) are mutually consistent and all hold for** $(A, B) = (4, 2)$**.**

Check **(d)**: $A - B = 4 - 2 = 2 \neq 10$. **False.**

*Suppose instead **(a)** and **(d)** both hold.* Then $A + B = 6$ and $A - B = 10$ give $A = 8$, $B = -2$. Then $A^2 + B^2 = 68 \neq 20$ and $A^2 - B^2 = 60 \neq 12$, so **(b)** and **(c)** both fail. This alternative makes only two options true, whereas the first makes three; and any two of **(a)**,**(b)**,**(c)** already imply the third, so they stand or fall together.

Hence the consistent reading is

$$\boxed{A = 4, \quad B = 2}$$

with **(a)**, **(b)**, **(c)** true and **(d)** false.

The underlying theory, for completeness

The wave equation with Neumann data

Consider $u_{tt} = c^2 u_{xx}$ on $0 < x < L$ with Neumann conditions $u_x(0,t) = u_x(L,t) = 0$ (free ends). Separating variables $u = X(x)T(t)$ gives $X'' + \lambda X = 0$ with $X'(0) = X'(L) = 0$, whose eigenpairs are

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \qquad X_n(x) = \cos\frac{n\pi x}{L}, \qquad n = 0, 1, 2, \ldots$$

Note that $n = 0$ is admissible for Neumann (constant eigenfunction, $\lambda = 0$), unlike the Dirichlet case — this is the single most common oversight.
The general solution is therefore

$$u(x,t) = \frac{a_0 + b_0 t}{2} + \sum_{n\geq 1}\left(a_n \cos\frac{n\pi ct}{L} + b_n \sin\frac{n\pi ct}{L}\right)\cos\frac{n\pi x}{L},$$

the term $b_0 t$ arising because $\lambda_0 = 0$ gives $T'' = 0$, i.e. T linear in t. Physically this is the free translation of the whole string, possible only because the ends are unconstrained.
Here $c^2 = 4$, so $c = 2$ and the temporal frequencies are $2n\pi/L$.

The coefficients a_n come from the initial displacement and b_n from the initial velocity, via the cosine-series (half-range) expansions

$$a_n = \frac{2}{L}\int_0^L u(x,0)\cos\frac{n\pi x}{L}\,\mathrm{d}x, \qquad b_n = \frac{2}{n\pi c}\int_0^L u_t(x,0)\cos\frac{n\pi x}{L}\,\mathrm{d}x \quad (n \geq 1).$$

The constants A and B of the item are two such coefficients; the analysis above identifies them as 4 and 2.

Answer: **(a)**, **(b)** and **(c)** (consistent with $A = 4$, $B = 2$). **(d)** is inconsistent with the rest.

Question 66. Which of the following identities involving conditional expectation and variance hold? **(a)** $\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X]$ **(b)** $\mathbb{E}[\mathbb{E}[X \mid Y] \mid Y] = \mathbb{E}[X \mid Y]$ **(c)** $\mathrm{Var}(X \mid Y) = \mathbb{E}[X^2 \mid Y] - (\mathbb{E}[X \mid Y])^2$ **(d)** $\mathrm{Var}(X) = \mathrm{Var}(\mathbb{E}[X \mid Y]) + \mathbb{E}[\mathrm{Var}(X \mid Y)]$.

The concept: conditional expectation as a projection

Definition and defining property

Let X be integrable and $\mathcal{G}$ a sub-σ-algebra (here $\mathcal{G} = \sigma(Y)$). Then $\mathbb{E}[X \mid \mathcal{G}]$ is the a.s.-unique random variable Z such that

1. Z is $\mathcal{G}$-measurable, and
2. $\mathbb{E}[Z\mathbf{1}_G] = \mathbb{E}[X\mathbf{1}_G]$ for every $G \in \mathcal{G}$.

In L^2 this is exactly the orthogonal projection of X onto the closed subspace of $\mathcal{G}$-measurable square-integrable variables — the best $\mathcal{G}$-measurable predictor of X in mean square. Every property below is a consequence of that geometric picture.

(a) — TRUE (the tower property)

Take $G = \Omega$ in the defining property (2):

$$\mathbb{E}\big[\mathbb{E}[X \mid \mathcal{G}]\big] = \mathbb{E}[X\mathbf{1}_\Omega] = \mathbb{E}[X].$$

This is also called the *law of iterated expectations* or *smoothing*, and it is the workhorse of the whole subject: it computes an unconditional expectation by conditioning on whatever makes the inner expectation easy.

(b) — TRUE (idempotence)

$\mathbb{E}[X \mid Y]$ is by construction $\sigma(Y)$-measurable. For any $\mathcal{G}$-measurable integrable Z we have $\mathbb{E}[Z \mid \mathcal{G}] = Z$, since Z itself satisfies both defining conditions. Applying this with $Z = \mathbb{E}[X \mid Y]$ gives

$$\mathbb{E}\big[\mathbb{E}[X \mid Y] \mid Y\big] = \mathbb{E}[X \mid Y].$$

Geometrically: a projection applied twice is the projection. $P^2 = P$.

(c) — TRUE (the definition of conditional variance)

Conditional variance is *defined* by

$$\mathrm{Var}(X \mid Y) = \mathbb{E}\big[(X - \mathbb{E}[X \mid Y])^2 \mid Y\big].$$

Expanding the square inside and using linearity of conditional expectation together with the fact that $\mathbb{E}[X \mid Y]$ is $\sigma(Y)$-measurable (so it may be pulled out of the inner conditional expectation):

$$\mathrm{Var}(X \mid Y) = \mathbb{E}[X^2 \mid Y] - 2\mathbb{E}[X \mid Y]\,\mathbb{E}[X \mid Y] + (\mathbb{E}[X \mid Y])^2 = \mathbb{E}[X^2 \mid Y] - (\mathbb{E}[X \mid Y])^2.$$

(d) — TRUE (the law of total variance)

Write $\mu = \mathbb{E}X$ and $m(Y) = \mathbb{E}[X \mid Y]$. Then

$$\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}\big[(X - m(Y))^2\big] + 2\,\mathbb{E}\big[(X - m(Y))(m(Y) - \mu)\big] + \mathbb{E}\big[(m(Y) - \mu)^2\big].$$

The cross term vanishes: conditioning on Y and using the tower property,

$$\mathbb{E}\big[(X - m(Y))(m(Y) - \mu)\big] = \mathbb{E}\Big[(m(Y) - \mu)\,\mathbb{E}\big[X - m(Y) \mid Y\big]\Big] = \mathbb{E}\big[(m(Y) - \mu)\cdot 0\big] = 0,$$

since $\mathbb{E}[X \mid Y] - m(Y) = 0$. This is precisely the orthogonality of the residual to the projection. The first term is $\mathbb{E}[\mathrm{Var}(X \mid Y)]$ by the tower property, and the third is $\mathrm{Var}(m(Y)) = \mathrm{Var}(\mathbb{E}[X \mid Y])$ because $\mathbb{E}[m(Y)] = \mu$ by **(a)**. Hence

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(\mathbb{E}[X \mid Y]).$$

Reading the decomposition. $\mathrm{Var}(\mathbb{E}[X \mid Y])$ is the "explained" variance — the variability of the conditional mean as Y varies — and $\mathbb{E}[\mathrm{Var}(X \mid Y)]$ is the "unexplained" or residual variance. Their sum being the total is the population version of the ANOVA identity of Question 33: the between-group sum of squares and within-group sum of squares add to the total.
An immediate corollary, used constantly: $\mathrm{Var}(X) \geq \mathrm{Var}(\mathbb{E}[X \mid Y])$, with equality iff $\mathrm{Var}(X \mid Y) = 0$ a.s., i.e. iff X is a function of Y.

Answer: **(a)**, **(b)**, **(c)** and **(d)** — all four.

Question 67. Which of the following field-extension statements is/are true? **(a)** If α is a real root of $x^3 - 3x - 1$, then $\sqrt{2} \notin \mathbb{Q}(\alpha)$ **(b)** $x^4 - \sqrt{2}$ is irreducible over $\mathbb{Q}(\sqrt{2})$ **(c)** $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 8$ **(d)** $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}(\sqrt{2})] = 2$.

(a) — TRUE

Step 1: the cubic is irreducible over $\mathbb{Q}$. A cubic is reducible over $\mathbb{Q}$ iff it has a rational root. By the rational root theorem any such root divides 1, so is ± 1:

$$f(1) = 1 - 3 - 1 = -3 \neq 0, \qquad f(-1) = -1 + 3 - 1 = 1 \neq 0.$$

Hence $f(x) = x^3 - 3x - 1$ is irreducible and is the minimal polynomial of α, so

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3.$$

Step 2: the tower law. If $\sqrt{2} \in \mathbb{Q}(\alpha)$, then $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\alpha)$ and the tower law gives

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}(\sqrt{2})] \cdot 2.$$

So $2 \mid 3$, which is false. Hence $\sqrt{2} \notin \mathbb{Q}(\alpha)$. **True.**

The same argument shows that a degree-3 extension of $\mathbb{Q}$ contains no subfield strictly between $\mathbb{Q}$ and itself, since 3 is prime. More generally, *the degree of any intermediate field divides the degree of the whole extension* — the single most useful corollary of the tower law and the reason it is worth stating separately.

(b) — TRUE

We apply Eisenstein's criterion in the ring $\mathbb{Z}[\sqrt{2}]$, the ring of integers of $\mathbb{Q}(\sqrt{2})$.

Eisenstein's criterion over a general domain

Let R be an integral domain with fraction field K, and let $f(x) = a_n x^n + \cdots + a_0 \in R[x]$. If there is a prime ideal $\mathfrak{p}$ of R with

$$a_n \notin \mathfrak{p}, \qquad a_i \in \mathfrak{p} \ (0 \le i < n), \qquad a_0 \notin \mathfrak{p}^2,$$

then f is irreducible over K (given R a UFD, or more generally after Gauss's lemma).

Take $R = \mathbb{Z}[\sqrt{2}]$, which is a Euclidean domain (hence a PID and a UFD) under the norm $N(a + b\sqrt{2}) = |a^2 - 2b^2|$. The element $\sqrt{2}$ has norm 2, a rational prime, so $\sqrt{2}$ is irreducible and — in a PID — prime. Let $\mathfrak{p} = (\sqrt{2})$.

For $f(x) = x^4 - \sqrt{2}$: the leading coefficient $1 \notin \mathfrak{p}$; the middle coefficients are $0 \in \mathfrak{p}$; the constant term $-\sqrt{2} \in \mathfrak{p}$ but $-\sqrt{2} \notin \mathfrak{p}^2 = (2)$, since $\sqrt{2}/2 = 1/\sqrt{2} \notin \mathbb{Z}[\sqrt{2}]$. Eisenstein applies, and $x^4 - \sqrt{2}$ is irreducible over $\mathbb{Q}(\sqrt{2})$. **True.**

(c) — FALSE

$x^3 - 2$ is irreducible over $\mathbb{Q}$ (Eisenstein at 2, or: no rational root since $\sqrt[3]{2}$ is irrational). Hence

$$[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3 \neq 8.$$

False. The value 8 has no relevance here; note also that a degree of 8 would be impossible for a root of a cubic under any circumstances.

(d) — FALSE

Two independent objections.

First, $\mathbb{Q}(\sqrt{2})$ *is not a subfield of* $\mathbb{Q}(\sqrt[3]{2})$, so the expression $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}(\sqrt{2})]$ is not even defined as a degree over a subfield. Indeed $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, and by the tower law any subfield would have degree dividing 3, so the only subfields are $\mathbb{Q}$ and $\mathbb{Q}(\sqrt[3]{2})$ itself; and $\mathbb{Q}(\sqrt{2})$ has degree 2.

Second, if one reads the expression as the degree of the compositum $\mathbb{Q}(\sqrt{2}, \sqrt[3]{2})$ *over* $\mathbb{Q}(\sqrt{2})$, the answer is 3, not 2. For

$$[\mathbb{Q}(\sqrt{2}, \sqrt[3]{2}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt[3]{2}) : \mathbb{Q}(\sqrt{2})] \cdot 2$$

must also be divisible by $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, hence by 6; and it is at most $2 \cdot 3 = 6$; so it is exactly 6 and the required degree is 3.

False under either reading.

Answer: **(a)** and **(b)**.

Question 68. For the sequence of tent functions of height n^2 supported on $[0, 3/n]$, **(a)** each f_n is Riemann integrable **(b)** $\int_0^1 f_n = 2$ for all n **(c)** $f_n \to 0$ pointwise **(d)** the integral of the limit equals the limit of the integrals.

The construction

A "tent" of height h on $[0, b]$ is the continuous piecewise-linear function rising from 0 at $x = 0$ to h at $x = b/2$ and falling back to 0 at $x = b$, and equal to 0 outside $[0, b]$. Here $h = n^2$ and $b = 3/n$.

(a) — TRUE

Each f_n is continuous on $[0, 1]$: it is piecewise linear with matching values at the joins, and it is 0 at both ends of its support. Continuous functions on a compact interval are Riemann integrable. **True.**

(Even without continuity: f_n is bounded and has at most three points of non-differentiability, so its set of discontinuities is empty and Lebesgue's criterion applies trivially.)

(c) — TRUE

Fix $x \in [0, 1]$.

If $x > 0$*:* choose N with $3/N < x$, i.e. $N > 3/x$. For every $n \geq N$ the support $[0, 3/n]$ does not contain x, so $f_n(x) = 0$. Hence $f_n(x) \to 0$.

If $x = 0$*:* $f_n(0) = 0$ for every n by construction, so again $f_n(0) \to 0$.

Therefore $f_n \to 0$ pointwise on $[0, 1]$. **True.** (The convergence is *not* uniform: $\sup f_n = n^2 \to \infty$.)

(b) — FALSE as printed

The area of a tent is that of a triangle with base b and height h:

$$\int_0^1 f_n = \tfrac{1}{2}\, b\, h = \tfrac{1}{2} \cdot \frac{3}{n} \cdot n^2 = \frac{3n}{2}.$$

This is $3/2$ at $n = 1$, 3 at $n = 2$, and tends to infinity. It is not 2 for any n except the non-integer $n = 4/3$. **False.**

For the assertion $\int_0^1 f_n = 2$ to hold identically, the height would have to be $h = 4n/3$ rather than n^2: then $\frac{1}{2} \cdot \frac{3}{n} \cdot \frac{4n}{3} = 2$. With the printed height n^2 the integral grows without bound. The item's parameters and its option **(b)** are inconsistent; **(b)** is false as printed, and would be true under the alternative height.

(d) — FALSE

The pointwise limit is $f \equiv 0$, so $\int_0^1 f = 0$. But $\int_0^1 f_n = 3n/2 \to \infty$ (or $= 2$ under the alternative height). Either way,

$$\int_0^1 \lim_n f_n = 0 \neq \lim_n \int_0^1 f_n.$$

False.

Why the dominated convergence theorem does not apply. DCT requires an integrable g with $|f_n| \le g$ for all n. Here $\sup_n f_n(x) = \infty$ near 0 — indeed at $x = 3/(2n)$ we have $f_n = n^2$, so any dominating function would need $g(3/(2n)) \ge n^2$, i.e. $g(y) \ge 9/(4y^2)$ near 0, which is not integrable. The failure is thus exactly the failure of domination, and this family is the standard illustration.
Fatou's lemma still applies and is consistent: $\int \liminf f_n = 0 \le \liminf \int f_n$, with strict inequality — "mass escaping to infinity" in the vertical direction.

Answer: **(a)** and **(c)**.

Question 69. Which of the following functions is uniformly continuous on the indicated set? **(a)** $\dfrac{\cos^2 x}{1 + x^2 + \sin^2 x}$ on $[0, \infty)$ **(b)** $\dfrac{1}{x} \sin \dfrac{1}{x}$ on $[1, \infty)$ **(c)** $x \sin \dfrac{2}{\pi x}$ on $(0, \pi/2)$ **(d)** the distance to the interval $[1, 2]$, on $\mathbb{R}$.

The concept: three sufficient criteria

Tools for establishing uniform continuity

1. Compactness (Heine–Cantor). A continuous function on a compact set is uniformly continuous.
2. Lipschitz. If $|f(x) - f(y)| \le L\,|x - y|$ then f is uniformly continuous with $\delta = \varepsilon/L$. In particular a differentiable function with bounded derivative is uniformly continuous, by the mean value theorem.

3. Continuous extension. If f is continuous on (a,b) and both one-sided limits $f(a^+)$, $f(b^-)$ exist finitely, then f extends continuously to $[a,b]$, which is compact, so f is uniformly continuous on (a,b).
4. Finite limit at infinity. If f is continuous on $[a,\infty)$ and $\lim_{x\to\infty} f(x)$ exists finitely, then f is uniformly continuous on $[a,\infty)$. *Proof.* Given ε, choose M with $|f(x)-L|<\varepsilon/2$ for $x>M$; then any two points beyond M are within ε. On $[a,M+1]$, compactness gives a δ. Take the minimum.

An unbounded derivative does not preclude uniform continuity. $f(x)=\sqrt{x}$ on $[0,1]$ has $f'(x)\to\infty$ as $x\to 0^+$, yet is uniformly continuous by compactness. Conversely a bounded function need not be uniformly continuous: $\sin(x^2)$ on $\mathbb{R}$ is bounded but not uniformly continuous, because its oscillations become arbitrarily rapid.

(a) — TRUE

Set $f(x)=\dfrac{\cos^2 x}{1+x^2+\sin^2 x}$. The denominator satisfies $1+x^2+\sin^2 x\ge 1+x^2$, so

$$0\le f(x)\le \frac{1}{1+x^2}\longrightarrow 0 \quad \text{as } x\to\infty.$$

Thus f is continuous on $[0,\infty)$ with $\lim_{x\to\infty} f(x)=0$, and criterion 4 applies. **Uniformly continuous.**

(b) — TRUE

Set $g(x)=\frac{1}{x}\sin\frac{1}{x}$ on $[1,\infty)$. As $x\to\infty$, $|g(x)|\le 1/x\to 0$, so again the limit at infinity exists (and is 0), and g is continuous on $[1,\infty)$. Criterion 4 applies. **Uniformly continuous.**

Alternatively, by the Lipschitz route:

$$g'(x)=-\frac{1}{x^2}\sin\frac{1}{x}-\frac{1}{x^3}\cos\frac{1}{x}, \qquad \left|g'(x)\right|\le\frac{1}{x^2}+\frac{1}{x^3}\le 2 \ \text{ on } [1,\infty),$$

so g is 2-Lipschitz. Note that the awkward behaviour of $\frac{1}{x}\sin\frac{1}{x}$ occurs near 0, which is excluded from the domain.

(c) — TRUE

Set $h(x)=x\sin\dfrac{2}{\pi x}$ on $(0,\pi/2)$. The only concern is the endpoint $x=0$, where the argument of the sine blows up. But the sine is bounded, so

$$|h(x)|\le|x|\longrightarrow 0 \quad \text{as } x\to 0^+.$$

Hence $\lim_{x\to 0^+} h(x)=0$ exists, and $\lim_{x\to(\pi/2)^-} h(x)=\frac{\pi}{2}\sin\frac{4}{\pi^2}$ exists by continuity. Criterion 3 applies: h extends continuously to $[0,\pi/2]$, which is compact. **Uniformly continuous.**

Note that $h'(x)=\sin\frac{2}{\pi x}-\frac{2}{\pi x}\cos\frac{2}{\pi x}$ is *unbounded* as $x\to 0^+$; this is precisely the situation warned of above, where unbounded derivative coexists with uniform continuity.

(d) — TRUE

Let $d(x) = \operatorname{dist}(x, [1,2]) = \inf_{y\in[1,2]} |x-y|$, explicitly

$$d(x) = \begin{cases} 1-x & x<1, \\ 0 & 1 \le x \le 2, \\ x-2 & x>2. \end{cases}$$

The distance function to any non-empty set $A \subseteq \mathbb{R}$ is 1-Lipschitz: for any $y \in A$,

$$d(x) \le |x-y| \le |x-z| + |z-y|,$$

and taking the infimum over $y \in A$ gives $d(x) \le |x-z| + d(z)$, i.e. $d(x) - d(z) \le |x-z|$; by symmetry $|d(x)-d(z)| \le |x-z|$. Criterion 2 applies. **Uniformly continuous.**

All four options are correct. This is unusual but permissible in Part C, where any number of options from one to four may be true. It is worth resisting the instinct that "all four" must be wrong: the four functions here have been constructed to illustrate four *different* routes to uniform continuity — decay at infinity (**(a)**), bounded derivative (**(b)**), continuous extension across a bad endpoint (**(c)**), and the Lipschitz property of a distance function (**(d)**). The item is a survey of the toolkit rather than a discrimination test.

Answer: **(a)**, **(b)**, **(c)** and **(d)** — all four.

Question 70. Which of the following limit statements is/are true? **(a)** $\lim \int_0^{\pi/4} \dfrac{t^n \sin t}{1+\cos^2 t}\,\mathrm{d}t = 0$ **(b)** $\lim n\int_0^1 (f(t))^n\,\mathrm{d}t \le 1$ for continuous $f : [0,1] \to [0,1]$ **(c)** $\lim \int_0^1 f(t)\sin(nt)\,\mathrm{d}t \le 1$ for C^1 functions **(d)** the integral of $f\cos^2(nt)$ diverges for $f > 1$.

(a) — TRUE

On $[0, \pi/4]$ the factor $\dfrac{\sin t}{1+\cos^2 t}$ is continuous, hence bounded; explicitly, $0 \le \sin t \le \frac{1}{\sqrt{2}}$ and $1+\cos^2 t \ge 1$, so the factor is bounded by $M = 1/\sqrt{2}$. Therefore

$$\left| \int_0^{\pi/4} \frac{t^n \sin t}{1+\cos^2 t}\,\mathrm{d}t \right| \le M \int_0^{\pi/4} t^n\,\mathrm{d}t = M\,\frac{(\pi/4)^{n+1}}{n+1}.$$

Since $\pi/4 \approx 0.785 < 1$, the numerator tends to 0 geometrically and the whole expression tends to 0. **True.**

The essential mechanism is that $t^n \to 0$ uniformly on any interval $[0,c]$ with $c<1$. Had the interval been $[0,1]$ the argument would still succeed — the bound becomes $M/(n+1) \to 0$ — but at $t=1$ the pointwise limit of t^n is 1, so the uniformity is lost and one needs dominated convergence instead.

(b) — FALSE

Take $f \equiv 1$, which is continuous and maps $[0,1]$ into $[0,1]$ as required. Then

$$n\int_0^1 (f(t))^n \, \mathrm{d}t = n\int_0^1 1 \, \mathrm{d}t = n \longrightarrow \infty,$$

so the limit is not ≤ 1; it does not exist finitely at all. **False.**

The statement becomes true under a strict inequality. If $f : [0,1] \to [0,1)$ is continuous, then $M = \max f < 1$ by compactness, and

$$n\int_0^1 f^n \leq nM^n \longrightarrow 0,$$

since geometric decay beats linear growth. The failure is entirely at the single value $f = 1$, and the option is testing whether the candidate notices that the codomain $[0,1]$ is closed. More refined: if f attains the value 1 at an interior point where $f'' < 0$, Laplace's method gives $n\int f^n \sim c\sqrt{n} \to \infty$; if $f \equiv 1$ on a set of positive measure μ, then $n\int f^n \geq n\mu \to \infty$.

(c) — TRUE

The Riemann–Lebesgue lemma

If $f \in L^1[a,b]$ then

$$\lim_{n\to\infty}\int_a^b f(t)\sin(nt)\,\mathrm{d}t = 0, \qquad \lim_{n\to\infty}\int_a^b f(t)\cos(nt)\,\mathrm{d}t = 0.$$

Proof sketch for C^1 f. Integrate by parts:

$$\int_a^b f(t)\sin(nt)\,\mathrm{d}t = \left[-\frac{f(t)\cos(nt)}{n}\right]_a^b + \frac{1}{n}\int_a^b f'(t)\cos(nt)\,\mathrm{d}t,$$

and both terms are $O(1/n)$ since f and f' are bounded on $[a,b]$. The general L^1 case follows by density of C^1 in L^1.

Hence for C^1 functions the limit is 0, and $0 \leq 1$. **True.**

The statement is weak — it asserts an inequality where an equality with 0 holds — but weak statements are still true statements. Do not reject an option merely because it understates what you know.

(d) — FALSE

Use the identity $\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$:

$$\int_0^1 f(t)\cos^2(nt)\,\mathrm{d}t = \frac{1}{2}\int_0^1 f(t)\,\mathrm{d}t + \frac{1}{2}\int_0^1 f(t)\cos(2nt)\,\mathrm{d}t.$$

By Riemann–Lebesgue the second term tends to 0, so

$$\lim_{n\to\infty}\int_0^1 f(t)\cos^2(nt)\,\mathrm{d}t = \frac{1}{2}\int_0^1 f(t)\,\mathrm{d}t,$$

a *finite* number whenever f is integrable. If $f > 1$ then this limit exceeds $1/2$, but it is finite — nothing diverges. **False.**

This limit — that oscillating squares average to one half against any fixed integrable weight — is worth memorising in the form

$$\int_a^b f(t)\cos^2(nt)\,\mathrm{d}t \longrightarrow \frac{1}{2}\int_a^b f, \qquad \int_a^b f(t)\sin^2(nt)\,\mathrm{d}t \longrightarrow \frac{1}{2}\int_a^b f.$$

It is the simplest instance of *weak convergence*: $\cos^2(nt) \rightharpoonup \frac{1}{2}$ in the weak-$*$ topology of L^∞, without converging pointwise anywhere.

Answer: **(a)** and **(c)**.

Question 71. For the set $S = \{(1/n, 1/n) : n \in \mathbb{N}\}$, **(a)** every point of S is isolated **(b)** $(0,0)$ is a limit point of S **(c)** S is not compact **(d)** every open cover of S has a finite subcover.

The geometry

S is a sequence of points on the line $y = x$ in $\mathbb{R}^2$, at parameters $1, \frac{1}{2}, \frac{1}{3}, \ldots$, accumulating at the origin, which does not belong to S.

(a) — TRUE

A point $p \in S$ is *isolated* in S if some open ball about p meets S only in p. Take $p_n = (1/n, 1/n)$. Its nearest neighbours in S are p_{n-1} and p_{n+1}, at distances

$$|p_n - p_{n+1}| = \sqrt{2}\left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{\sqrt{2}}{n(n+1)} > 0.$$

Choosing r_n smaller than this distance (and, for $n \geq 2$, smaller than the distance to p_{n-1}) gives a ball meeting S only in p_n. So every point of S is isolated. **True.**

Equivalently, S carries the discrete subspace topology — a fact used again in **(d)**.

(b) — TRUE

A point q is a limit point of S if every punctured neighbourhood of q meets S. Given $\varepsilon > 0$, choose $n > \sqrt{2}/\varepsilon$; then

$$|p_n - (0,0)| = \frac{\sqrt{2}}{n} < \varepsilon,$$

and $p_n \neq (0,0)$ since $1/n > 0$. So every punctured ball about the origin contains points of S. **True.**

Moreover $(0,0)$ is the *only* limit point of S, by **(a)**: no point of S is a limit point of S, and no point off the closure of the sequence can be.

(c) — TRUE

By Heine–Borel, a subset of $\mathbb{R}^2$ is compact iff it is closed and bounded. S is bounded (contained in the ball of radius $\sqrt{2}$), but it is *not closed*: its limit point $(0,0)$ is not in S, so $\bar{S} = S \cup \{(0,0)\} \neq S$. Hence S is not compact. **True.**

(d) — FALSE

This is the definition of compactness, which we have just refuted; but it is worth exhibiting an explicit open cover with no finite subcover, since that is what the option literally asserts.

For each n, let U_n be an open ball about p_n containing no other point of S (possible by **(a)**). Then $\{U_n : n \in \mathbb{N}\}$ is an open cover of S. Any subfamily covers only the finitely many p_n whose indices it uses, so no finite subfamily covers all of S. **False.**

The instructive contrast. The set

$$S' = S \cup \{(0,0)\}$$

is compact — it is closed and bounded — and it is the standard example of a *convergent sequence together with its limit*, the simplest infinite compact metric space. Adding a single point to S changes the answer to **(c)** and **(d)** in both directions. This is the exact content of the item: the distinction between a sequence and a convergent sequence with its limit.
Note also that in S' every point except $(0,0)$ is isolated, so **(a)** fails for S': compactness and complete isolation are incompatible for infinite sets, since a compact space with the discrete topology must be finite.

Answer: **(a)**, **(b)** and **(c)**.

Question 72. For the topology $\mathbb{R}_K$ generated by open intervals and sets of the form $(a,b) \setminus K$, **(a)** $\mathbb{R}_K$ is Hausdorff **(b)** $\mathbb{R}_K$ is regular **(c)** $\mathbb{R}_K$ is normal **(d)** 0 is not a limit point of K.

The space

Here $K = \{1/n : n \in \mathbb{N}\} = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$, and $\mathbb{R}_K$ is the topology on $\mathbb{R}$ with basis

$$\mathcal{B} = \big\{(a,b)\big\} \cup \big\{(a,b) \setminus K\big\}.$$

This is Munkres' *K-topology*, and it is the standard counterexample separating the Hausdorff property from regularity.

Observe first that $\mathcal{B}$ contains all standard basic open sets, so

$$\tau_{\text{std}} \subsetneq \tau_K :$$

τ_K is strictly finer, the set $(-1,1) \setminus K$ being τ_K-open but not standard-open (any standard neighbourhood of 0 meets K).

(a) — TRUE

If $x \neq y$, the standard topology already separates them by disjoint open intervals, and those intervals remain open in the finer topology τ_K. **Hausdorff.**

The general principle: any topology finer than a Hausdorff topology is Hausdorff. Separation axioms T_0, T_1, T_2 are all preserved under refinement, because the separating sets stay open. Regularity and normality are *not* preserved under refinement — there are more closed sets to separate — which is exactly what happens here.

(d) — TRUE

In the standard topology 0 *is* a limit point of K, since $1/n \to 0$. But in τ_K the set

$$U = (-1, 1) \setminus K$$

is a basic open neighbourhood of 0 (note $0 \notin K$, so $0 \in U$), and $U \cap K = \emptyset$. Hence no punctured neighbourhood condition can be satisfied: 0 is **not** a limit point of K in τ_K. **True.**

An immediate consequence, used next: K is *closed* in τ_K. Its complement is

$$\mathbb{R} \setminus K = \big((-\infty, 1) \setminus K\big) \cup (1, \infty) \cup \bigcup_n \left(\tfrac{1}{n+1}, \tfrac{1}{n}\right),$$

a union of τ_K-open sets. (In the standard topology K is not closed, since $0 \in \bar{K} \setminus K$.)

(b) — FALSE

Regularity

A space is *regular* if it is T_1 and, for every closed set C and point $x \notin C$, there exist disjoint open sets $U \ni x$ and $V \supseteq C$.

Take $C = K$, which is closed by the previous paragraph, and $x = 0 \notin K$. Suppose $U \ni 0$ and $V \supseteq K$ were disjoint open sets.

Since U is open and contains 0, it contains a basic set of the form $(a, b) \setminus K$ with $a < 0 < b$. (It cannot contain a plain interval $(a, b) \ni 0$, for such an interval meets K and hence meets V.)

Choose n large enough that $1/n < b$. Since $1/n \in K \subseteq V$ and V is open, V contains a basic neighbourhood of $1/n$, which necessarily contains an interval $(\frac{1}{n} - \varepsilon, \frac{1}{n} + \varepsilon)$ for some $\varepsilon > 0$ — removing K cannot remove a whole interval. Shrinking ε, we may assume this interval lies inside (a, b) and contains no point of K other than $1/n$ itself.

But then any point $z \in (\frac{1}{n} - \varepsilon, \frac{1}{n})$ with $z \notin K$ satisfies $z \in U$ (being in $(a, b) \setminus K$) and $z \in V$ — contradicting disjointness. Such z exists because the interval is uncountable and K is countable.

Hence no such U, V exist, and $\mathbb{R}_K$ is **not regular. False.**

(c) — FALSE

$\mathbb{R}_K$ is T_1: singletons are closed, since $\{x\}^c$ is a union of intervals. For T_1 spaces,

$$\text{normal} \implies \text{regular},$$

because a singleton $\{x\}$ is a closed set and normality separates the two closed sets $\{x\}$ and C. Since $\mathbb{R}_K$ is T_1 and not regular, it cannot be normal. **False.**

The place of $\mathbb{R}_K$ in the hierarchy. The separation axioms satisfy

$$T_4 \ (\text{normal} + T_1) \Rightarrow T_3 \ (\text{regular} + T_1) \Rightarrow T_2 \Rightarrow T_1 \Rightarrow T_0,$$

and each implication is strict. $\mathbb{R}_K$ is the standard witness for $T_2 \not\Rightarrow T_3$. Note also that $\mathbb{R}_K$ is connected (it is finer than the standard topology but not enough to disconnect $\mathbb{R}$ — unlike the Sorgenfrey topology of Question 55) and is not metrisable, since every metric space is normal.

Answer: **(a)** and **(d)**.

Question 73. For an $M/M/3$ queue with $\lambda = 10$, $\mu = 5$, **(a)** $p_0 = 1/9$ **(b)** $p_0 + p_1 = 1/3$ **(c)** $2p_0 = p_3$ **(d)** the tail probability $p_3 + p_4 + \cdots = 4/9$.

The concept: the $M/M/c$ steady state

Birth–death balance for $M/M/c$

Arrivals are Poisson at rate λ; each of c servers works at rate μ. When $n < c$ customers are present, n servers are busy and the total service rate is $n\mu$; when $n \geq c$, all c servers are busy and the rate is $c\mu$.
Detailed balance between adjacent states gives

$$p_n = \begin{cases} \dfrac{a^n}{n!}\, p_0, & 0 \leq n \leq c, \\[2ex] \dfrac{a^n}{c!\, c^{n-c}}\, p_0, & n \geq c, \end{cases} \qquad a = \frac{\lambda}{\mu}, \quad \rho = \frac{\lambda}{c\mu} = \frac{a}{c}.$$

Normalisation $\sum_n p_n = 1$ yields

$$p_0 = \left[\sum_{n=0}^{c-1} \frac{a^n}{n!} + \frac{a^c}{c!\,(1-\rho)} \right]^{-1},$$

valid (i.e. the chain is positive recurrent) precisely when $\rho < 1$.

The computation

Here $c = 3$, $\lambda = 10$, $\mu = 5$, so

$$a = \frac{\lambda}{\mu} = 2, \qquad \rho = \frac{a}{c} = \frac{2}{3} < 1. \checkmark$$

Compute p_0.

$$\sum_{n=0}^{2} \frac{a^n}{n!} = \frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} = 1 + 2 + 2 = 5,$$

$$\frac{a^3}{3!\,(1-\rho)} = \frac{8}{6 \cdot \frac{1}{3}} = \frac{8}{2} = 4.$$

Hence

$$p_0 = \frac{1}{5+4} = \frac{1}{9}.$$

(a) TRUE.

Compute the low-order probabilities.

$$p_1 = \frac{a^1}{1!}p_0 = 2 \cdot \frac{1}{9} = \frac{2}{9}, \qquad p_2 = \frac{a^2}{2!}p_0 = 2 \cdot \frac{1}{9} = \frac{2}{9}, \qquad p_3 = \frac{a^3}{3!}p_0 = \frac{8}{6} \cdot \frac{1}{9} = \frac{4}{27}.$$

(b) TRUE. $p_0 + p_1 = \frac{1}{9} + \frac{2}{9} = \frac{3}{9} = \frac{1}{3}$.

(c) FALSE. $2p_0 = \frac{2}{9} = \frac{6}{27}$, while $p_3 = \frac{4}{27}$. These are unequal. (The correct relation is $p_3 = \frac{4}{3}p_2 = \frac{4}{3} \cdot \frac{2}{9}$.)

(d) TRUE. Since the states $0, 1, 2, 3, \ldots$ exhaust the space,

$$\sum_{n \geq 3} p_n = 1 - (p_0 + p_1 + p_2) = 1 - \left(\frac{1}{9} + \frac{2}{9} + \frac{2}{9}\right) = 1 - \frac{5}{9} = \frac{4}{9}.$$

Interpretation of the tail. $\sum_{n \geq c} p_n$ is the probability that all servers are busy, so an arriving customer must wait. This quantity is *Erlang's C formula*:

$$C(c, a) = \frac{a^c}{c!\,(1-\rho)}\, p_0 = \frac{4}{9} \text{ here.}$$

It is the central quantity in call-centre staffing. Direct check: $4 \times \frac{1}{9} = \frac{4}{9}$. ✓

Further performance measures, for completeness:

$$L_q = \frac{\rho\, C(c, a)}{1 - \rho} = \frac{(2/3)(4/9)}{1/3} = \frac{8}{9}, \qquad W_q = \frac{L_q}{\lambda} = \frac{8}{90} = \frac{4}{45},$$

$$L = L_q + a = \frac{8}{9} + 2 = \frac{26}{9}, \qquad W = \frac{L}{\lambda} = \frac{26}{90} = \frac{13}{45}.$$

Little's law $L = \lambda W$ holds throughout. ✓

Answer: **(a)**, **(b)** and **(d)**.

Question 74. In the space $(C[0,1], d)$ with the integral metric, for the sequence of tent functions rising on $[1/2,\ 1/2 + 1/n]$, **(a)** $\{f_n\}$ is Cauchy **(b)** the space is incomplete **(c)** the pointwise limit exists everywhere **(d)** the pointwise limit is continuous.

The set-up

The metric is

$$d(f, g) = \int_0^1 |f(t) - g(t)|\, dt,$$

the L^1 metric restricted to continuous functions. The functions are the ramps

$$f_n(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2}, \\ n\left(x - \frac{1}{2}\right), & \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n}, \\ 1, & \frac{1}{2} + \frac{1}{n} \leq x \leq 1, \end{cases}$$

each continuous, rising linearly from 0 to 1 over an interval of width $1/n$.

(a) — TRUE

For $m > n$ the two functions agree outside $\left(\frac{1}{2}, \frac{1}{2} + \frac{1}{n}\right)$ and differ by at most 1 inside it. Hence

$$d(f_n, f_m) = \int_{1/2}^{1/2+1/n} |f_n - f_m| \leq 1 \cdot \frac{1}{n} = \frac{1}{n}.$$

Given $\varepsilon > 0$, choose $N > 1/\varepsilon$; then $d(f_n, f_m) < \varepsilon$ for all $m, n \geq N$. **Cauchy.**

(In fact $d(f_n, f_m) = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{m}\right)$ exactly, by computing the area between two ramps, but the bound suffices.)

(c) — TRUE

Fix $x \in [0, 1]$.

- If $x \leq \frac{1}{2}$: $f_n(x) = 0$ for every n, so the limit is 0.
- If $x > \frac{1}{2}$: choose $N > 1/(x - \frac{1}{2})$; then for $n \geq N$ we have $\frac{1}{2} + \frac{1}{n} \leq x$, so $f_n(x) = 1$. The limit is 1.

Thus the pointwise limit exists at every point and equals

$$f(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2}, \\ 1, & \frac{1}{2} < x \leq 1, \end{cases}$$

the Heaviside step at $\frac{1}{2}$. **True.**

(d) — FALSE

The limit f has a jump of size 1 at $x = \frac{1}{2}$: $f(\frac{1}{2}) = 0$ while $f(\frac{1}{2}^+) = 1$. It is not continuous. **False.**

(b) — TRUE

We must show $(C[0, 1], d)$ is not complete, i.e. that $\{f_n\}$ has no limit *in the space*. Suppose $g \in C[0, 1]$ with $d(f_n, g) \to 0$. Since $d(f_n, f) \leq \frac{1}{2n} \to 0$ as well (where f is the step function), the triangle inequality gives

$$\int_0^1 |f - g| = d(f, g) \leq d(f, f_n) + d(f_n, g) \longrightarrow 0,$$

so $\int_0^1 |f - g| = 0$ and hence $f = g$ almost everywhere. But g is continuous and f is not, and two functions agreeing a.e. cannot have one continuous and the other with a genuine jump: if g were continuous and $g = 0$ a.e. on $[0, \frac{1}{2})$ and $g = 1$ a.e. on $(\frac{1}{2}, 1]$, then by continuity $g(\frac{1}{2}) = \lim_{x \uparrow 1/2} g(x) = 0$ and $g(\frac{1}{2}) = \lim_{x \downarrow 1/2} g(x) = 1$ — a contradiction.

Hence the Cauchy sequence $\{f_n\}$ has no limit in $C[0, 1]$, and the space is **incomplete. True.**

The contrast with the sup metric. $(C[0, 1], \|\cdot\|_\infty)$ *is* complete — it is the archetypal Banach space of continuous functions — because uniform limits of continuous functions are continuous. The present sequence is not Cauchy in the sup metric: $\|f_n - f_m\|_\infty = 1$ for $m \neq n$ large, since at the midpoint of the shorter ramp the two functions differ substantially. So completeness genuinely depends on the metric, not merely on the underlying set.

The completion of $(C[0, 1], d)$ is $L^1[0, 1]$, in which the limit is the equivalence class of the step function. This is the standard motivation for introducing the Lebesgue spaces at all.

Answer: **(a)**, **(b)** and **(c)**.

Question 75. For the equicorrelated trivariate normal, the first principal component satisfies **(a)** $Y_1 = \frac{1}{\sqrt{3}}(X_1 + X_2 + X_3)$ **(b)** $\mathrm{Corr}(Y_1, X_1) = 1/\sqrt{3}$ **(c)** Y_1 explains more than 75% of total variance **(d)** $\mathrm{Var}(Y_1) = 4$.

The numerical values of σ^2 and ρ did not reproduce

The general structure is fully determinable, and one option follows from it without any numbers.

The equicorrelation (intraclass) covariance matrix

Let

$$\Sigma = \sigma^2\big[(1-\rho)I_p + \rho J_p\big] = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix},$$

where J_p is the all-ones matrix. Positive definiteness requires $-\frac{1}{p-1} < \rho < 1$.
Its spectrum is completely explicit. Since $J_p\mathbf{1} = p\mathbf{1}$ and $J_p v = 0$ for every $v \perp \mathbf{1}$:

$$\lambda_1 = \sigma^2\big[1 + (p-1)\rho\big] \quad \text{with eigenvector } \tfrac{1}{\sqrt{p}}\mathbf{1},$$

$$\lambda_2 = \cdots = \lambda_p = \sigma^2(1-\rho) \quad \text{with eigenspace } \mathbf{1}^{\perp}\ (\dim p-1).$$

For $\rho > 0$ we have $\lambda_1 > \lambda_2$, so the leading eigenvector is unambiguously the normalised all-ones direction.

(a) — TRUE (for $\rho > 0$)

The first principal component is the normalised leading eigenvector applied to $\mathbf{X}$:

$$Y_1 = \frac{1}{\sqrt{3}}\mathbf{1}^{\mathsf{T}}\mathbf{X} = \frac{1}{\sqrt{3}}(X_1 + X_2 + X_3).$$

This is the "size" or "common factor" component, and its universality is the signature of equicorrelated data: when all variables correlate equally, the dominant mode of variation is their common level. **True.**

The remaining options depend on σ^2 and ρ

With $p = 3$:

$$\mathrm{Var}(Y_1) = \lambda_1 = \sigma^2(1+2\rho), \qquad \text{total variance} = \mathrm{tr}\,\Sigma = 3\sigma^2.$$

(c) *Proportion explained.*

$$\frac{\lambda_1}{\mathrm{tr}\,\Sigma} = \frac{\sigma^2(1+2\rho)}{3\sigma^2} = \frac{1+2\rho}{3}.$$

This exceeds 0.75 if and only if $1 + 2\rho > 2.25$, i.e. $\rho > 0.625$. So the option is true for strongly correlated data and false otherwise — it cannot be adjudicated without ρ.

(b) *Correlation with a coordinate.* Using $\mathrm{Cov}(Y_1, X_1) = \frac{1}{\sqrt{3}}(\sigma^2 + 2\rho\sigma^2) = \frac{\sigma^2(1+2\rho)}{\sqrt{3}}$ and $\mathrm{sd}(Y_1) = \sigma\sqrt{1+2\rho}$, $\mathrm{sd}(X_1) = \sigma$:

$$\mathrm{Corr}(Y_1, X_1) = \frac{\sigma^2(1+2\rho)/\sqrt{3}}{\sigma\sqrt{1+2\rho}\cdot\sigma} = \sqrt{\frac{1+2\rho}{3}}.$$

This equals $1/\sqrt{3}$ only when $\rho = 0$ — in which case there is no distinguished first component at all, since $\Sigma = \sigma^2 I$ has a triple eigenvalue. So **(b)** is false whenever the item is non-degenerate.

(d) $\mathrm{Var}(Y_1) = 4$ requires $\sigma^2(1+2\rho) = 4$, e.g. $\sigma^2 = 2$, $\rho = \frac{1}{2}$, or $\sigma^2 = \frac{4}{1+2\rho}$ generally. Consistent with many parameter choices but not determinable.

The consistency test worth applying. If **(d)** holds with the common choice $\sigma^2 = 2$, $\rho = \frac{1}{2}$, then the proportion explained is $\frac{1+1}{3} = \frac{2}{3} \approx 66.7\% < 75\%$, so **(c)** would be false, and $\mathrm{Corr}(Y_1, X_1) = \sqrt{2/3} \approx 0.816 \neq 1/\sqrt{3}$, so **(b)** would be false. Under that reading the answer is **(a)** and **(d)**. This is the most likely intended key, but it cannot be confirmed from the printed text.

Answer: **(a)** is true unconditionally (for $\rho > 0$). **(b)** is false except in the degenerate case $\rho = 0$. **(c)** and **(d)** require σ^2 and ρ, which are absent; the likely intended key is **(a)** and **(d)**.

Question 76. For the Sturm–Liouville problem $x(xy')' + \lambda y = 0$, $y(1) = y(\mathrm{e}) = 0$, the eigenpairs are **(a)** $\lambda_n = \frac{(2n-1)^2\pi^2}{\mathrm{e}^2}$ **(b)** $\lambda_n = \frac{(2n-1)^2\pi^2}{4}$ **(c)** $y_n = \cos\frac{(2n-1)\pi x}{\mathrm{e}}$ **(d)** $y_n = \cos\frac{(2n-1)\pi\log x}{2}$.

Reducing to constant coefficients

Expand the operator:

$$x(xy')' = x(xy'' + y') = x^2y'' + xy',$$

so the equation is the *Cauchy–Euler* equation

$$x^2y'' + xy' + \lambda y = 0. \tag{76.1}$$

The logarithmic substitution

For Cauchy–Euler equations the substitution $t = \ln x$ (valid for $x > 0$) converts variable coefficients into constant ones. With $\dfrac{\mathrm{d}}{\mathrm{d}x} = \dfrac{1}{x}\dfrac{\mathrm{d}}{\mathrm{d}t}$:

$$xy' = \dot{y}, \qquad x^2y'' = \ddot{y} - \dot{y},$$

where dots denote $\mathrm{d}/\mathrm{d}t$. Substituting into (76.1):

$$(\ddot{y} - \dot{y}) + \dot{y} + \lambda y = 0 \qquad \Longrightarrow \qquad \boxed{\ddot{y} + \lambda y = 0}$$

— the harmonic oscillator. This is why $\log x$ appears in every eigenfunction.

The domain $x \in [1, \mathrm{e}]$ maps to $t \in [0, 1]$, since $\ln 1 = 0$ and $\ln \mathrm{e} = 1$. The whole problem has been transported to the unit interval.

Solving with the printed boundary conditions

The printed conditions $y(1) = y(\mathrm{e}) = 0$ become $y(t = 0) = y(t = 1) = 0$: a pure Dirichlet problem for $\ddot{y} + \lambda y = 0$ on $[0, 1]$. The standard eigenpairs are

$$\lambda_n = n^2\pi^2, \qquad y_n(t) = \sin(n\pi t), \qquad n = 1, 2, 3, \ldots$$

Transporting back,

$$\boxed{\lambda_n = n^2\pi^2, \qquad y_n(x) = \sin\left(n\pi \ln x\right)}$$

None of the four printed options matches this. Options **(a)** and **(c)** carry an e in the denominator, which cannot occur — the logarithmic substitution has already absorbed the interval length. Options **(b)** and **(d)** carry $(2n - 1)$ and a cosine, the signature of a *mixed* boundary condition.

Identifying the intended problem

Suppose the boundary conditions were instead

$$y'(1) = 0, \qquad y(\mathrm{e}) = 0,$$

i.e. in the t variable $\dot{y}(0) = 0$, $y(1) = 0$. Solving $\ddot{y} + \lambda y = 0$: the condition $\dot{y}(0) = 0$ selects the cosine, $y = \cos(\sqrt{\lambda}\, t)$; the condition $y(1) = 0$ requires $\cos\sqrt{\lambda} = 0$, i.e.

$$\sqrt{\lambda} = \frac{(2n-1)\pi}{2}, \qquad n = 1, 2, \ldots$$

Hence

$$\lambda_n = \frac{(2n-1)^2\pi^2}{4}, \qquad y_n(t) = \cos\frac{(2n-1)\pi t}{2}, \qquad \text{i.e.} \qquad y_n(x) = \cos\frac{(2n-1)\pi \ln x}{2}.$$

These are *exactly* options **(b)** and **(d)**, and they are mutually consistent — the eigenvalue in **(b)** is the square of the frequency in **(d)**. That internal consistency is strong evidence that the boundary condition $y(1) = 0$ in the stem is a misprint for $y'(1) = 0$.

The item as printed is inconsistent. Under the stated conditions $y(1) = y(\mathrm{e}) = 0$ the answer is $\lambda_n = n^2\pi^2$, $y_n = \sin(n\pi \ln x)$, absent from the options. Under the conditions $y'(1) = 0$, $y(\mathrm{e}) = 0$ the answer is exactly **(b)** and **(d)**. The latter is evidently intended.

The Sturm–Liouville framework, for context. Writing (76.1) in self-adjoint form,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}y}{\mathrm{d}x}\right) + \frac{\lambda}{x}y = 0,$$

we identify $p(x) = x$, $q = 0$, and weight $w(x) = 1/x$. The general theory then guarantees: real eigenvalues, a discrete spectrum tending to $+\infty$, and eigenfunctions orthogonal with respect to w:

$$\int_1^{\mathrm{e}} y_m(x) y_n(x)\,\frac{\mathrm{d}x}{x} = 0 \quad (m \neq n).$$

Substituting $t = \ln x$ turns $\mathrm{d}x/x$ into $\mathrm{d}t$, recovering the familiar orthogonality of sines or cosines on $[0, 1]$ — a satisfying confirmation that the substitution is the natural one.

Answer: **(b)** and **(d)**, under the boundary conditions $y'(1) = 0$, $y(\mathrm{e}) = 0$. With the boundary conditions exactly as printed, the correct answer ($\lambda_n = n^2\pi^2$, $y_n = \sin(n\pi \log x)$) is not among the options.

Question 77. For the sequence $X_n = 2^n$ on $(0, 1/n]$ and 0 elsewhere, **(a)** $X_n \to 0$ almost surely **(b)** $\mathbb{P}(|X_n| > \varepsilon \text{ i.o.}) = 0$ for every $\varepsilon > 0$ **(c)** $X_n \to 0$ in probability **(d)** $X_n \to 0$ in L^p for every $p > 0$.

The set-up

Take the probability space $\big((0,1), \mathcal{B}, \lambda\big)$ with Lebesgue measure. Then

$$X_n(\omega) = 2^n \, \mathbf{1}_{(0,1/n]}(\omega).$$

The four modes of convergence

$$X_n \to X \text{ a.s.} : \quad \mathbb{P}\big(\lim_n X_n = X\big) = 1.$$

$$X_n \to X \text{ in probability} : \quad \mathbb{P}(|X_n - X| > \varepsilon) \to 0 \ \ \forall \varepsilon > 0.$$

$$X_n \to X \text{ in } L^p : \quad \mathbb{E}\,|X_n - X|^p \to 0.$$

$$X_n \to X \text{ in distribution} : \quad F_n(x) \to F(x) \text{ at continuity points.}$$

The implications:

$$\text{a.s.} \implies \text{in probability} \implies \text{in distribution}, \qquad L^p \implies \text{in probability}.$$

No other implication holds, and in particular a.s. convergence neither implies nor is implied by L^p convergence. The present item is constructed to exhibit exactly that independence.

(a) — TRUE

Fix $\omega \in (0, 1)$. Choose $N > 1/\omega$; then for every $n \geq N$ we have $1/n < \omega$, so $\omega \notin (0, 1/n]$ and $X_n(\omega) = 0$.

Therefore $X_n(\omega) \to 0$ for *every* $\omega \in (0, 1)$ — convergence is not merely almost sure but everywhere on the sample space. **True.**

(b) — TRUE

The events $A_n = \{|X_n| > \varepsilon\}$ satisfy, for any $0 < \varepsilon < 2^n$,

$$A_n = (0, 1/n], \qquad \mathbb{P}(A_n) = \frac{1}{n}.$$

The event $\{A_n \text{ i.o.}\} = \bigcap_N \bigcup_{n \geq N} A_n$. But $\bigcup_{n \geq N}(0, 1/n] = (0, 1/N]$, so

$$\bigcap_N (0, 1/N] = \emptyset,$$

since no positive ω survives all N. Hence $\mathbb{P}(A_n \text{ i.o.}) = 0$. **True.**

This is equivalent to **(a)**: the standard characterisation states that $X_n \to X$ a.s. if and only if $\mathbb{P}(|X_n - X| > \varepsilon \text{ i.o.}) = 0$ for every $\varepsilon > 0$. So **(a)** and **(b)** stand or fall together and it is enough to verify one.
Note that Borel–Cantelli does *not* apply directly: $\sum \mathbb{P}(A_n) = \sum 1/n = \infty$, so the first lemma gives nothing, and the events are nested (hence far from independent), so the second lemma does not apply either. The direct computation is necessary.

(c) — TRUE

For any $\varepsilon > 0$ and n large enough that $2^n > \varepsilon$,

$$\mathbb{P}(|X_n| > \varepsilon) = \lambda\big((0, 1/n]\big) = \frac{1}{n} \longrightarrow 0.$$

True. (This also follows from **(a)** by the implication chain.)

(d) — FALSE

Compute the L^p norm for $p > 0$:

$$\mathbb{E}\,|X_n|^p = \int_0^{1/n} (2^n)^p \, \mathrm{d}\omega = \frac{2^{np}}{n} = \frac{(2^p)^n}{n}.$$

Since $2^p > 1$ for every $p > 0$, the numerator grows geometrically while the denominator grows only linearly, so

$$\mathbb{E}\,|X_n|^p = \frac{(2^p)^n}{n} \longrightarrow \infty \qquad \text{for every } p > 0.$$

Far from converging to 0 in L^p, the norms diverge. **False.**

The mechanism. The sequence converges pointwise because the support shrinks; it fails in L^p because the height grows faster than the support shrinks. The product $(\text{height})^p \times (\text{width}) = 2^{np}/n$ is the quantity that matters, and here it explodes.
Tuning the example. Replace 2^n by n^α: then $\mathbb{E}\,|X_n|^p = n^{\alpha p - 1}$, which tends to 0 iff $p < 1/\alpha$. So $X_n = n^\alpha \mathbf{1}_{(0,1/n]}$ converges a.s. and in L^p for small p but not for large p — a single family generating the whole range of behaviour.
Why dominated convergence fails here: any dominating g would satisfy $g(\omega) \geq 2^n$ for ω near $1/n$, i.e. $g(\omega) \gtrsim 2^{1/\omega}$, which is not integrable near 0.

Answer: **(a)**, **(b)** and **(c)**.

Question 78. For $f(x, y) = \dfrac{xy^\alpha}{x^2 + y^2}$ (with $f = 0$ at the origin), **(a)** continuous on $\mathbb{R}^2$ for all $\alpha \geq 1$ **(b)** differentiable on $\mathbb{R}^2$ for all $\alpha \geq 3$ **(c)** both partial derivatives vanish at the origin for every natural α **(d)** the mixed second partial vanishes at the origin for $\alpha \geq 3$.

Throughout, α is taken to be a natural number, and away from the origin f is a quotient of polynomials with non-vanishing denominator, hence C^∞. All questions concern the origin.

(a) — FALSE

The claim is "for *all* $\alpha \geq 1$", so a single counterexample suffices.
Take $\alpha = 1$: $f(x,y) = \dfrac{xy}{x^2+y^2}$. Along the line $y = mx$ (with $x \neq 0$),

$$f(x, mx) = \frac{mx^2}{x^2(1+m^2)} = \frac{m}{1+m^2},$$

a constant depending on m. Along $y = x$ the value is $\frac{1}{2}$; along $y = 0$ it is 0. Since different approach directions give different limiting values, no limit exists at the origin and f is discontinuous there. **False.**

Where the threshold actually lies. In polar coordinates $x = r\cos\theta$, $y = r\sin\theta$:

$$f = \frac{r^{1+\alpha}\cos\theta\sin^\alpha\theta}{r^2} = r^{\alpha-1}\cos\theta\sin^\alpha\theta.$$

Since $|\cos\theta\sin^\alpha\theta| \leq 1$, we get $|f| \leq r^{\alpha-1}$, which tends to 0 uniformly in θ precisely when $\alpha > 1$. So f is continuous at the origin for every $\alpha \geq 2$ and fails only at $\alpha = 1$ — the single value that makes the option false.

(c) — TRUE

Compute from the definition, for any natural $\alpha \geq 1$:

$$f_x(0,0) = \lim_{h\to 0}\frac{f(h,0)-0}{h} = \lim_{h\to 0}\frac{1}{h}\cdot\frac{h\cdot 0^\alpha}{h^2} = \lim_{h\to 0}\frac{0}{h} = 0,$$

$$f_y(0,0) = \lim_{k\to 0}\frac{f(0,k)-0}{k} = \lim_{k\to 0}\frac{1}{k}\cdot\frac{0\cdot k^\alpha}{k^2} = 0.$$

Both vanish, because f is identically zero on each coordinate axis — the numerator xy^α contains both x and y as factors. **True.**

Note that this holds even for $\alpha = 1$, where f is not even continuous at the origin — another illustration that existence of partials is a very weak condition.

(b) — TRUE

Take $\alpha \geq 3$. Since $f_x(0,0) = f_y(0,0) = 0$ by **(c)**, differentiability at the origin requires

$$\frac{|f(h,k)|}{\sqrt{h^2+k^2}} \longrightarrow 0.$$

In polar form, using the computation above,

$$\frac{|f|}{r} = \frac{r^{\alpha-1}|\cos\theta\sin^\alpha\theta|}{r} \leq r^{\alpha-2}.$$

For $\alpha \geq 3$ we have $\alpha - 2 \geq 1$, so this is bounded by $r \to 0$, uniformly in θ. Hence f is differentiable at the origin with $\nabla f(0,0) = (0,0)$. Away from the origin f is C^∞. **True.**

(For $\alpha = 2$ the bound is $r^0 = 1$ and one checks along $y = x$ that the ratio tends to $\frac{1}{2\sqrt{2}} \neq 0$: differentiability fails. So $\alpha \geq 3$ is sharp.)

(d) — FALSE

Compute $f_{xy}(0,0) = \dfrac{\partial}{\partial y}\Big[f_x\Big]_{(0,0)}$.

First find f_x at points $(0, y)$ with $y \neq 0$. For $(x, y) \neq (0, 0)$,

$$f_x = \frac{y^\alpha(x^2 + y^2) - xy^\alpha \cdot 2x}{(x^2 + y^2)^2} = \frac{y^\alpha\,(y^2 - x^2)}{(x^2 + y^2)^2}.$$

Setting $x = 0$:

$$f_x(0, y) = \frac{y^\alpha \cdot y^2}{y^4} = y^{\alpha-2}.$$

Therefore

$$f_{xy}(0, 0) = \lim_{k\to 0} \frac{f_x(0, k) - f_x(0, 0)}{k} = \lim_{k\to 0} \frac{k^{\alpha-2} - 0}{k} = \lim_{k\to 0} k^{\alpha-3}.$$

Now:

$$\alpha = 3 : \lim_{k\to 0} k^0 = 1 \neq 0; \qquad \alpha \geq 4 : \lim_{k\to 0} k^{\alpha-3} = 0.$$

So the mixed partial *fails* to vanish at $\alpha = 3$, and the claim "for $\alpha \geq 3$" is **false** — it holds only for $\alpha \geq 4$.

The case $\alpha = 3$ is the classical counterexample to equality of mixed partials. With $f = xy^3/(x^2 + y^2)$ we have just found $f_{xy}(0, 0) = 1$. The symmetric computation gives

$$f_y = \frac{\alpha xy^{\alpha-1}(x^2 + y^2) - xy^\alpha \cdot 2y}{(x^2 + y^2)^2}, \qquad f_y(x, 0) = 0 \text{ for } \alpha \geq 2,$$

so $f_{yx}(0, 0) = \lim_h \frac{f_y(h,0)-0}{h} = 0$. Hence

$$f_{xy}(0, 0) = 1 \;\neq\; 0 = f_{yx}(0, 0).$$

Clairaut's theorem is not violated: it requires the mixed partials to be *continuous* at the point, and here they are not.
This is the same phenomenon as the celebrated $f = xy(x^2 - y^2)/(x^2 + y^2)$, and it is one of the most frequently examined examples in multivariable analysis.

Answer: **(b)** and **(c)**.

Question 79. Which of the following ring-theoretic statements is/are true? **(a)** $\mathbb{Z}_{11}[x]/(x^3 + 8x^2 + 4x + 1)$ is a field **(b)** the principal ideal generated by the cubic is not contained in $(x^2 + x + 8)$ **(c)** $\mathbb{Z}_2[x]/(x^3 + x^2 + x + 2)$ is a field **(d)** $\mathbb{Z}_3[x]/(x^3 + x^2 + x + 2)$ has 27 elements.

Quotients of polynomial rings over a field

Let F be a field and $f \in F[x]$ of degree $n \geq 1$. Then:
1. Size. $F[x]/(f)$ has exactly $|F|^n$ elements when F is finite, because every coset has a unique representative of degree $< n$ (division algorithm), and there are $|F|^n$ such polynomials. *This holds whether or not f is irreducible.*
2. Field. $F[x]/(f)$ is a field if and only if f is irreducible over F. Reason: $F[x]$ is a PID, and in a PID the ideal (f) is maximal iff f is irreducible, and the quotient by a maximal ideal is a field.

3. Divisibility and ideals. $(f) \subseteq (g)$ if and only if $g \mid f$.

(a) — FALSE

Let $f(x) = x^3 + 8x^2 + 4x + 1$ over $\mathbb{F}_{11}$. A cubic is reducible over a field precisely when it has a root there, so test all eleven residues. Working modulo 11:

x	0	1	2	3	4	5	6	7	8	9	10
$f(x)$ mod 11	1	3	5	2	**0**	·	·	·	·	·	·

The root is found at $x = 4$:

$$f(4) = 64 + 8 \cdot 16 + 16 + 1 = 64 + 128 + 16 + 1 = 209 = 11 \times 19 \equiv 0 \pmod{11}. \checkmark$$

Hence f is reducible, (f) is not maximal, and the quotient is not a field. **False.**

The quotient is not even an integral domain: with the factorisation found below, $f = (x^2 + x + 8)(x + 7)$, so in the quotient the nonzero classes of $x^2 + x + 8$ and $x + 7$ multiply to zero. They are zero divisors.

(b) — FALSE

The claim is that $(f) \not\subseteq (x^2 + x + 8)$, equivalently that $x^2 + x + 8 \nmid f$. Perform the division over $\mathbb{F}_{11}$.

Divide $x^3 + 8x^2 + 4x + 1$ by $x^2 + x + 8$:

- First term x: subtract $x(x^2 + x + 8) = x^3 + x^2 + 8x$, leaving $7x^2 - 4x + 1$.
- Next term 7: subtract $7(x^2 + x + 8) = 7x^2 + 7x + 56$. Now $56 \equiv 1 \pmod{11}$, so we subtract $7x^2 + 7x + 1$, leaving

$$(-4 - 7)x + (1 - 1) = -11x = 0 \pmod{11}.$$

The remainder is 0, so

$$x^3 + 8x^2 + 4x + 1 = (x^2 + x + 8)(x + 7) \quad \text{in } \mathbb{F}_{11}[x].$$

Verification. $(x^2 + x + 8)(x + 7) = x^3 + 7x^2 + x^2 + 7x + 8x + 56 = x^3 + 8x^2 + 15x + 56$, and modulo 11, $15 \equiv 4$ and $56 \equiv 1$, giving $x^3 + 8x^2 + 4x + 1$. ✓

Therefore $(f) \subseteq (x^2 + x + 8)$, and the statement that it is *not* contained is **false**.

Consistency: the root $x = 4$ found in **(a)** should be a root of one factor. Testing $x^2 + x + 8$ at 4: $16 + 4 + 8 = 28 \equiv 6 \neq 0$. Testing $x + 7$ at 4: $4 + 7 = 11 \equiv 0$. ✓ So $x = 4$ is the root of the linear factor, and $x^2 + x + 8$ is irreducible over $\mathbb{F}_{11}$ (its discriminant $1 - 32 = -31 \equiv 2$, and one checks 2 is a non-residue mod 11, since the residues are $1, 3, 4, 5, 9$).

(c) — FALSE

Over $\mathbb{Z}_2$ the constant $2 \equiv 0$, so

$$x^3 + x^2 + x + 2 \;\equiv\; x^3 + x^2 + x \;=\; x\,(x^2 + x + 1) \quad \text{in } \mathbb{F}_2[x].$$

This is manifestly reducible — it has the root $x = 0$ — so the quotient is not a field. **False.** (Concretely, $g(0) = 0$ and $g(1) = 1 + 1 + 1 = 1$ in $\mathbb{F}_2$, confirming the single root at 0.)

(d) — TRUE

Over $\mathbb{Z}_3$ the polynomial $x^3 + x^2 + x + 2$ has degree 3, and $|\mathbb{Z}_3| = 3$. By part 1 of the concept box,

$$\left|\mathbb{Z}_3[x]/(x^3 + x^2 + x + 2)\right| = 3^3 = 27.$$

This is true *irrespective of irreducibility*, which is the point of the option. **True.**

Is it a field? Test for roots in $\mathbb{F}_3$:

$$g(0) = 2, \qquad g(1) = 1 + 1 + 1 + 2 = 5 \equiv 2, \qquad g(2) = 8 + 4 + 2 + 2 = 16 \equiv 1.$$

No roots, so the cubic is irreducible over $\mathbb{F}_3$ and the quotient is in fact the field $\mathbb{F}_{27}$. The item did not ask this, but it is worth noting: option **(d)** would have been true even had the polynomial been reducible, whereas the field-hood is an extra fact.

Answer: **(d)** only.

Question 80. In the linear model with three observations, the estimable functions include **(a)** β_0 **(b)** β_2 **(c)** $-\beta_1 - \beta_2$ **(d)** $2\beta_1 + 2\beta_2$.

The design matrix did not reproduce

Estimability

In the linear model $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with $\mathbb{E}\boldsymbol{\varepsilon} = 0$, a linear function $\mathbf{c}^\top\boldsymbol{\beta}$ is *estimable* if there exists a vector $\mathbf{a}$ with

$$\mathbb{E}\left[\mathbf{a}^\top\mathbf{y}\right] = \mathbf{c}^\top\boldsymbol{\beta} \quad \text{for every } \boldsymbol{\beta}.$$

Since $\mathbb{E}[\mathbf{a}^\top\mathbf{y}] = \mathbf{a}^\top X\boldsymbol{\beta}$, this holds iff

$$\mathbf{c}^\top = \mathbf{a}^\top X,$$

i.e. **iff $\mathbf{c}$ lies in the row space of X.**
Equivalently $\mathbf{c} \in \mathcal{R}(X^\top) = \mathcal{R}(X^\top X)$, which gives the practical test: $\mathbf{c}^\top\boldsymbol{\beta}$ is estimable iff $\mathbf{c}^\top(X^\top X)^-(X^\top X) = \mathbf{c}^\top$ for a generalised inverse.
When X has full column rank every function is estimable; the notion has content only in the rank-deficient case, which is the norm in designed experiments with over-parametrised models such as $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$.

The structural fact the option set exploits

Proposition 3.2. The estimable functions form a *vector space* — namely the row space of X. In particular, if $\mathbf{c}^\top\boldsymbol{\beta}$ is estimable, so is $k\,\mathbf{c}^\top\boldsymbol{\beta}$ for every scalar $k \neq 0$.

Apply this to options **(c)** and **(d)**:

$$-\beta_1 - \beta_2 = -1 \cdot (\beta_1 + \beta_2), \qquad 2\beta_1 + 2\beta_2 = 2 \cdot (\beta_1 + \beta_2).$$

Both are nonzero scalar multiples of the same functional $\beta_1 + \beta_2$. Therefore:

$$\boxed{\textbf{(c)} \text{ is estimable } \iff \textbf{(d)} \text{ is estimable}}$$

They must be marked together or not at all. Any purported key containing exactly one of them is wrong.

The typical design behind such an item. In a one-way layout $y_i = \beta_0 + \beta_{t(i)} + \varepsilon_i$ with three observations and two treatments, X is rank-deficient: the columns for β_0, β_1, β_2 satisfy $\mathrm{col}_0 = \mathrm{col}_1 + \mathrm{col}_2$. Consequently:

- β_0 alone is *not* estimable (option **(a)** false);
- β_2 alone is *not* estimable (option **(b)** false);
- contrasts such as $\beta_1 - \beta_2$ *are* estimable;
- $\beta_0 + \beta_i$ (the treatment means) *are* estimable.

Whether $\beta_1 + \beta_2$ is estimable depends on the design; in the balanced one-way layout with an intercept it is not, since it is not a contrast and not of the form $\beta_0 + \beta_i$.

Answer: Indeterminate: the design matrix is absent. Note that **(c)** and **(d)** are proportional, hence necessarily both estimable or both not.

Question 81. The range of α for which the fixed-point iteration converges to the root of $\mathrm{e}^{-x} - \sin x = 0$ on $[0, \pi/2]$ is **(a)** $(0, 0.2)$ **(b)** $(0.2, 0.6)$ **(c)** $(0.5, 0.8)$ **(d)** $(0.9, 1)$.

The iteration formula did not reproduce

Fixed-point iteration and the contraction condition

To solve $F(x) = 0$, rewrite it as $x = g(x)$ and iterate $x_{k+1} = g(x_k)$.
Convergence theorem. If g maps a closed interval I into itself and there is $L < 1$ with $|g'(x)| \le L$ on I, then g has a unique fixed point $r \in I$ and $x_k \to r$ from any starting point in I, with

$$|x_k - r| \le L^k \, |x_0 - r| \, .$$

Convergence is linear with asymptotic rate $|g'(r)|$; when $g'(r) = 0$ the convergence is at least quadratic (this is what Newton's method arranges).
Both hypotheses matter. Self-mapping without contraction, or contraction without self-mapping, is insufficient.

The natural relaxation form

The parameter α almost certainly enters through the relaxation scheme

$$g(x) = x + \alpha\big(\mathrm{e}^{-x} - \sin x\big),$$

whose fixed points are exactly the roots of $\mathrm{e}^{-x} - \sin x = 0$ for any $\alpha \neq 0$. Then

$$g'(x) = 1 + \alpha\big(-\mathrm{e}^{-x} - \cos x\big) = 1 - \alpha\big(\mathrm{e}^{-x} + \cos x\big).$$

Write $\psi(x) = \mathrm{e}^{-x} + \cos x$. On $[0, \pi/2]$ this is strictly decreasing (both terms are), with

$$\psi(0) = 1 + 1 = 2, \qquad \psi(\pi/2) = \mathrm{e}^{-\pi/2} + 0 \approx 0.2079.$$

So $\psi\big([0, \pi/2]\big) = [0.2079,\ 2]$.
The contraction condition $|g'| < 1$ becomes

$$|1 - \alpha\psi(x)| < 1 \quad \iff \quad 0 < \alpha\,\psi(x) < 2 \quad \text{for all } x \in [0, \pi/2].$$

Since $\psi > 0$ throughout, the left inequality gives $\alpha > 0$. The right inequality is most binding where ψ is largest, namely at $x = 0$ with $\psi = 2$:

$$2\alpha < 2 \quad \Longrightarrow \quad \alpha < 1.$$

Hence the uniform contraction condition on the whole interval is

$$\boxed{0 < \alpha < 1}$$

What this implies for the options

Every one of the four printed intervals — $(0, 0.2)$, $(0.2, 0.6)$, $(0.5, 0.8)$, $(0.9, 1)$ — is a subset of $(0, 1)$. Under the natural reading of the iteration, therefore, *all four* options describe ranges of α for which the contraction condition holds and the iteration converges.

The item cannot discriminate as printed. Either the iteration function is different from the relaxation form assumed above, or the item intends the *local* rate condition at the root, or — most likely — the option set is defective.
For reference, the root is where $\mathrm{e}^{-r} = \sin r$; numerically $r \approx 0.5885$ (check: $\mathrm{e}^{-0.5885} = 0.5552$ and $\sin 0.5885 = 0.5553$). At the root $\psi(r) = \mathrm{e}^{-r} + \cos r = \sin r + \cos r \approx 0.5552 + 0.8317 = 1.3869$, so the local condition $|1 - \alpha\psi(r)| < 1$ gives $0 < \alpha < 2/1.3869 \approx 1.442$ — again satisfied by all four options. The fastest local convergence occurs at $\alpha = 1/\psi(r) \approx 0.721$, where $g'(r) = 0$ and the convergence becomes quadratic; that value lies in option **(c)**'s interval $(0.5, 0.8)$, which is the most plausible intended key.

Answer: Indeterminate as printed: the iteration function is absent, and under the natural relaxation reading all four intervals satisfy the convergence condition $0 < \alpha < 1$. If the item seeks the interval containing the optimal α (fastest, locally quadratic convergence at $\alpha = 1/(\sin r + \cos r) \approx 0.72$), the answer is **(c)**.

Question 82. For the extremal of $\int_0^1 (x - y')^2\,\mathrm{d}x$ subject to $y(0) = 0$, $y(1) = 2$, **(a)** $\max y = 3$ **(b)** $\min y = 0$ **(c)** $y(1/2) = 7/8$ **(d)** $y'(1/2) = 2$.

The concept: the Euler–Lagrange equation

Euler–Lagrange

To render $J[y] = \int_a^b F(x, y, y')\,\mathrm{d}x$ stationary among functions with fixed endpoint values, y must satisfy

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0.$$

Two first integrals worth remembering.

- If F does not contain y explicitly, then $\partial F/\partial y = 0$ and the equation reduces to $\partial F/\partial y' = \text{const}$ — the *momentum* integral.
- If F does not contain x explicitly, then $F - y'\,\partial F/\partial y' = \text{const}$ — the Beltrami identity.

Here $F(x, y, y') = (x - y')^2$ contains no y, so the first case applies.

The computation

$$\frac{\partial F}{\partial y'} = 2(x - y')(-1) = -2(x - y') = \text{const},$$

so

$$x - y' = c \qquad \Longrightarrow \qquad y' = x - c.$$

Integrating,

$$y(x) = \frac{x^2}{2} - cx + d.$$

Apply the boundary conditions:

$$y(0) = d = 0,$$
$$y(1) = \tfrac{1}{2} - c = 2 \quad \Longrightarrow \quad c = -\tfrac{3}{2}.$$

Hence

$$\boxed{y(x) = \frac{x^2}{2} + \frac{3x}{2}, \qquad y'(x) = x + \frac{3}{2}}$$

Check: $y(0) = 0$ ✓, $y(1) = \frac{1}{2} + \frac{3}{2} = 2$ ✓.

Adjudicating the options

Note that $y'(x) = x + \frac{3}{2} > 0$ throughout $[0, 1]$, so y is strictly increasing and its extreme values occur at the endpoints.

(a) FALSE. $\max_{[0,1]} y = y(1) = 2$, not 3.

(b) TRUE. $\min_{[0,1]} y = y(0) = 0$. ✓

(c) TRUE.

$$y\left(\tfrac{1}{2}\right) = \frac{(1/2)^2}{2} + \frac{3}{2}\cdot\frac{1}{2} = \frac{1/4}{2} + \frac{3}{4} = \frac{1}{8} + \frac{6}{8} = \frac{7}{8}. \checkmark$$

(d) TRUE.

$$y'\left(\tfrac{1}{2}\right) = \frac{1}{2} + \frac{3}{2} = 2. \checkmark$$

The value of the functional. With $x - y' = c = -\frac{3}{2}$ constant,

$$J[y] = \int_0^1 \left(-\tfrac{3}{2}\right)^2 \, dx = \frac{9}{4}.$$

Since the integrand $(x-y')^2 \geq 0$, the absolute minimum of J over all functions would be 0, attained by $y' = x$, i.e. $y = x^2/2 + \text{const}$; but that curve has $y(1) - y(0) = \frac{1}{2} \neq 2$ and so violates the boundary conditions. The extremal we found is therefore the constrained minimiser, and the value $9/4$ measures exactly how much the endpoint constraint costs.

Why this is a genuine minimum. The functional is convex in y' ($\partial^2 F/\partial y'^2 = 2 > 0$), so the Legendre condition holds strictly and the stationary point is a minimum. For convex integrands the Euler–Lagrange solution is automatically the global minimiser among admissible functions.

Answer: **(b)**, **(c)** and **(d)**.

Question 83. If R is the radius of $\sum a_n z^n$ and $c_k = \max_{n \leq k} |a_n|$, the radius R' of $\sum c_k z^k$ satisfies **(a)** $R' > R$ **(b)** $R' \leq R$ **(c)** $R' \leq 1$ **(d)** $R' \leq \min\{1, R\}$.

Throughout assume some $a_n \neq 0$, so that $c_k > 0$ for all large k (otherwise every $c_k = 0$ and all radii are ∞, a degenerate case).

The tool

By Cauchy–Hadamard,

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n}, \qquad \frac{1}{R'} = \limsup_{k\to\infty} c_k^{1/k}.$$

(b) — TRUE

By definition $c_k = \max_{n\leq k} |a_n| \geq |a_k|$, so

$$c_k^{1/k} \geq |a_k|^{1/k} \quad \text{for every } k.$$

Taking $\limsup$ preserves the inequality:

$$\frac{1}{R'} = \limsup_k c_k^{1/k} \;\geq\; \limsup_k |a_k|^{1/k} = \frac{1}{R}.$$

Hence $R' \leq R$. **True**, and consequently **(a)** ($R' > R$) is **FALSE**.

(c) — TRUE

The sequence (c_k) is non-decreasing by construction. Pick k_0 with $c_{k_0} = M > 0$ (possible since some $a_n \neq 0$). Then $c_k \geq M$ for all $k \geq k_0$, so

$$c_k^{1/k} \geq M^{1/k} \longrightarrow 1 \quad \text{as } k \to \infty,$$

because $M^{1/k} = \exp\left(\frac{\ln M}{k}\right) \to e^0 = 1$ for any fixed $M > 0$. Therefore

$$\frac{1}{R'} = \limsup_k c_k^{1/k} \;\geq\; 1 \qquad \Longrightarrow \qquad R' \leq 1.$$

True.

The mechanism is worth isolating: *a bounded-below non-decreasing coefficient sequence forces the radius to be at most* 1, because kth roots of any fixed positive constant tend to 1. This is why $\sum z^k$ (all coefficients 1) has radius exactly 1, and why no power series with coefficients bounded away from zero can have radius exceeding 1.

(d) — TRUE

Immediate from **(b)** and **(c)**: $R' \leq R$ and $R' \leq 1$ together give

$$R' \leq \min\{1, R\}.$$

True. Note that **(d)** is the strongest of the three true statements and implies both **(b)** and **(c)**; but all three are true, and in a multiple-response item all must be marked.

Is the bound in (d) attained? Yes, in both regimes.
If $R \leq 1$*:* take $a_n = R^{-n}$, so $c_k = \max_{n \leq k} R^{-n} = R^{-k}$ (the sequence is increasing when $R < 1$). Then $R' = R = \min\{1, R\}$.
If $R > 1$*:* take $a_0 = 1$ and $a_n = 0$ for $n \geq 1$, so $R = \infty$ but $c_k = 1$ for all k, giving $R' = 1 = \min\{1, R\}$. Here the passage from a_n to c_k collapses an entire function to one with radius exactly 1.
So $\min\{1, R\}$ is the exact best possible bound, and the operation $a \mapsto c$ never produces a radius exceeding 1 no matter how rapidly the original coefficients decay.

Answer: **(b)**, **(c)** and **(d)**.

Question 84. For the Uniform$(\theta, 2\theta)$ sample, which of the given estimators is unbiased for θ? **(a)** T_1 **(b)** T_2 **(c)** $\frac{1}{3}T_2 + \frac{2}{3}T_3$ **(d)** $\frac{5}{6}T_1 + \frac{1}{6}T_2$.

The definitions of T_1, T_2, T_3 did not reproduce

Moments and order statistics for Uniform$(\theta, 2\theta)$

Let $X_1, \ldots, X_n$ be i.i.d. Uniform$(\theta, 2\theta)$, $\theta > 0$. Then:

$$\mathbb{E}X = \frac{\theta + 2\theta}{2} = \frac{3\theta}{2}, \qquad \operatorname{Var} X = \frac{(2\theta - \theta)^2}{12} = \frac{\theta^2}{12}.$$

Hence the moment estimator is

$$\hat{\theta}_{\text{MOM}} = \frac{2\bar{X}}{3}, \qquad \mathbb{E}\left[\frac{2\bar{X}}{3}\right] = \frac{2}{3} \cdot \frac{3\theta}{2} = \theta. \checkmark$$

For the extreme order statistics, write $X_{(1)} = \min$ and $X_{(n)} = \max$. Setting $U_i = (X_i - \theta)/\theta \sim$ Uniform$(0, 1)$:

$$\mathbb{E}[U_{(1)}] = \frac{1}{n+1}, \qquad \mathbb{E}[U_{(n)}] = \frac{n}{n+1},$$

so

$$\mathbb{E}[X_{(1)}] = \theta\left(1+\frac{1}{n+1}\right) = \frac{(n+2)\theta}{n+1}, \qquad \mathbb{E}[X_{(n)}] = \theta\left(1+\frac{n}{n+1}\right) = \frac{(2n+1)\theta}{n+1}.$$

Consequently $\dfrac{(n+1)X_{(1)}}{n+2}$ and $\dfrac{(n+1)X_{(n)}}{2n+1}$ are each unbiased for θ.

The structural fact the options exploit

Proposition 3.3. Unbiased estimators of the *same* parameter form an affine space: if $\mathbb{E}[T_i] = \theta$ for each i and $\sum_i w_i = 1$, then

$$\mathbb{E}\left[\sum_i w_i T_i\right] = \sum_i w_i \theta = \theta,$$

so every convex (indeed affine) combination is again unbiased.

Both options **(c)** and **(d)** are affine combinations with weights summing to 1:

$$\tfrac{1}{3} + \tfrac{2}{3} = 1, \qquad \tfrac{5}{6} + \tfrac{1}{6} = 1.$$

So each is unbiased *provided its constituents are.* In particular:

- If T_1 and T_2 are both unbiased, then **(a)**, **(b)** and **(d)** are all correct.
- If T_2 and T_3 are both unbiased, **(c)** is correct.
- If exactly one of T_1, T_2 is unbiased, then **(d)** is *not* unbiased, since the bias of the other survives with a nonzero weight.

This last observation is the discriminating one and is presumably why the setter chose the weights $\frac{5}{6}, \frac{1}{6}$ rather than something symmetric.

Choosing among unbiased estimators. Once several unbiased candidates exist, the criterion is variance. For a convex combination $wT_1 + (1-w)T_2$ of two uncorrelated unbiased estimators the variance is minimised at

$$w^* = \frac{\operatorname{Var} T_2}{\operatorname{Var} T_1 + \operatorname{Var} T_2},$$

the inverse-variance weighting. For this model the complete sufficient statistic is the pair $(X_{(1)}, X_{(n)})$ — the family is *not* an exponential family, since the support depends on θ at both ends — and it is not complete, so Lehmann–Scheffé does not directly deliver a UMVUE. This is one of the standard examples where a minimal sufficient statistic fails to be complete.

Answer: Indeterminate: the estimators T_1, T_2, T_3 are absent from the source text. Note that **(c)** and **(d)** have weights summing to 1, so each is unbiased exactly when its constituents are.

Question 85. In the ring $\mathbb{Z}[\sqrt{-7}]$, **(a)** $\sqrt{-7}$ is prime **(b)** 4 divides $3+\sqrt{-7}$ **(c)** 4 is prime **(d)** the only units are ± 1.

The ring and its norm

The norm

$\mathbb{Z}[\sqrt{-7}] = \{a + b\sqrt{-7} : a, b \in \mathbb{Z}\}$, with norm

$$N(a+b\sqrt{-7}) = (a+b\sqrt{-7})(a-b\sqrt{-7}) = a^2 + 7b^2 \ \geq 0.$$

The norm is *multiplicative*: $N(\alpha\beta) = N(\alpha)N(\beta)$. Two immediate consequences:

- α is a unit $\iff$ $N(\alpha) = 1$;
- if $N(\alpha)$ is a rational prime, then α is irreducible (any factorisation would split the prime norm as a product of two integers > 1).

A caution. $\mathbb{Z}[\sqrt{-7}]$ is *not* the full ring of integers of $\mathbb{Q}(\sqrt{-7})$; since $-7 \equiv 1 \pmod 4$, the maximal order is $\mathbb{Z}\left[\frac{1+\sqrt{-7}}{2}\right]$. Thus $\mathbb{Z}[\sqrt{-7}]$ is a non-maximal order: not integrally closed, not Dedekind, and not a UFD.

(d) — TRUE

$\alpha = a + b\sqrt{-7}$ is a unit iff $N(\alpha) = a^2 + 7b^2 = 1$ with $a, b \in \mathbb{Z}$. If $b \neq 0$ then $a^2 + 7b^2 \geq 7 > 1$. So $b = 0$ and $a^2 = 1$, giving $a = \pm 1$. Hence the units are exactly $\{+1, -1\}$. **True.**

(Compare $\mathbb{Z}[\sqrt{2}]$, which has infinitely many units generated by $1+\sqrt{2}$; imaginary quadratic orders always have finite unit groups, and for discriminant < -4 the group is just $\{\pm 1\}$.)

(a) — TRUE

We show $(\sqrt{-7})$ is a prime ideal by identifying the quotient.

$$\mathbb{Z}[\sqrt{-7}] \cong \mathbb{Z}[x]/(x^2+7),$$

with $x \mapsto \sqrt{-7}$. Therefore

$$\frac{\mathbb{Z}[\sqrt{-7}]}{(\sqrt{-7})} \cong \frac{\mathbb{Z}[x]}{(x^2+7,\ x)} \cong \frac{\mathbb{Z}}{(7)} = \mathbb{F}_7,$$

since setting $x = 0$ turns $x^2 + 7$ into 7. As $\mathbb{F}_7$ is a field, the ideal $(\sqrt{-7})$ is maximal, hence prime, hence $\sqrt{-7}$ is a prime element. **True.**

(Consistently, $N(\sqrt{-7}) = 7$ is a rational prime, so $\sqrt{-7}$ is at least irreducible; the quotient computation upgrades this to primality, which in a non-UFD does not follow automatically.)

(b) — FALSE

The claim is that $\dfrac{3+\sqrt{-7}}{4} \in \mathbb{Z}[\sqrt{-7}]$, i.e. that

$$\frac{3}{4} + \frac{1}{4}\sqrt{-7} = a + b\sqrt{-7} \quad \text{with } a, b \in \mathbb{Z}.$$

This requires $a = 3/4$, not an integer. **False.**

A norm check confirms it: $N(3+\sqrt{-7}) = 9+7 = 16$ and $N(4) = 16$, so if 4 divided $3+\sqrt{-7}$ the quotient would have norm 1 and hence be a unit ± 1; but $3+\sqrt{-7} \neq \pm 4$. So $4 \nmid 3+\sqrt{-7}$.

The element $\frac{1+\sqrt{-7}}{2}$ *is* an algebraic integer (root of $x^2 - x + 2$) and lies in the maximal order but not in $\mathbb{Z}[\sqrt{-7}]$. This is precisely why $\mathbb{Z}[\sqrt{-7}]$ fails to be integrally closed, and why divisibility questions in it behave unexpectedly.

(c) — FALSE

$4 = 2 \cdot 2$ is a product of non-units, so it is reducible and cannot be prime. **False.**

Indeed 2 itself is irreducible in $\mathbb{Z}[\sqrt{-7}]$: $N(2) = 4$, and a proper factorisation would need an element of norm 2, i.e. $a^2 + 7b^2 = 2$, which has no integer solutions ($b = 0$ gives $a^2 = 2$; $|b| \geq 1$ gives ≥ 7). But 2 is *not* prime: in the maximal order,

$$2 = \frac{1+\sqrt{-7}}{2} \cdot \frac{1-\sqrt{-7}}{2},$$

and in $\mathbb{Z}[\sqrt{-7}]$ we have $2 \mid (1+\sqrt{-7})(1-\sqrt{-7}) = 8$ while $2 \nmid (1 \pm \sqrt{-7})$ (which would need $1/2 \in \mathbb{Z}$). So 2 is irreducible but not prime — the failure of unique factorisation in this order, exhibited concretely.

Answer: **(a)** and **(d)**.

Question 86. For the integral equation $y(x) = ax^2 + \int_0^1 e^{x+t} y(t)\,dt$, **(a)** a unique solution exists when $a \neq 0$ **(b)** the zero solution exists when $a = 0$ **(c)** $y(2) = \frac{6}{3-e^2}$ when $a = 2$ **(d)** $y(2) = \frac{4}{3-e^2}$ when $a = 1$.

The concept: Fredholm equations with degenerate kernels

Separable (degenerate) kernels

A Fredholm equation of the second kind

$$y(x) = f(x) + \int_a^b K(x,t)\,y(t)\,dt$$

is solvable in closed form whenever the kernel separates:

$$K(x,t) = \sum_{i=1}^{m} u_i(x) v_i(t).$$

Then $y(x) = f(x) + \sum_i C_i u_i(x)$ with unknown constants $C_i = \int_a^b v_i(t) y(t)\,dt$, and substituting back gives an $m \times m$ linear system for the C_i. The infinite-dimensional problem collapses to a finite one.

Here $K(x,t) = e^{x+t} = e^x \cdot e^t$ is separable with $m = 1$.

The computation

Put

$$C = \int_0^1 e^t y(t)\,dt,$$

so the equation reads

$$y(x) = ax^2 + C\,e^x. \tag{86.1}$$

Substitute (86.1) into the definition of C:

$$C = \int_0^1 e^t\left(at^2 + Ce^t\right) dt = a\int_0^1 t^2 e^t\,dt + C\int_0^1 e^{2t}\,dt.$$

The two integrals. By parts twice,

$$\int_0^1 t^2 e^t\,dt = \Big[t^2e^t - 2te^t + 2e^t\Big]_0^1 = (e - 2e + 2e) - 2 = e - 2.$$

$$\int_0^1 e^{2t}\,dt = \left[\frac{e^{2t}}{2}\right]_0^1 = \frac{e^2-1}{2}.$$

Hence

$$C = a(e-2) + C\cdot\frac{e^2-1}{2} \quad\Longrightarrow\quad C\left(1 - \frac{e^2-1}{2}\right) = a(e-2) \quad\Longrightarrow\quad C\cdot\frac{3-e^2}{2} = a(e-2).$$

Since $e^2 \approx 7.389 \neq 3$, the coefficient is nonzero and

$$C = \frac{2a(e-2)}{3-e^2}.$$

Therefore

$$\boxed{y(x) = ax^2 + \frac{2a(e-2)}{3-e^2}\,e^x}$$

Adjudicating the options

(a) TRUE. The linear equation for C has the nonzero coefficient $(3-e^2)/2$, so C — and hence y — is uniquely determined. This holds for *every* a, and in particular for $a \neq 0$. **True as stated.**

(b) TRUE. With $a = 0$ we get $C = 0$ and hence $y \equiv 0$ from (86.1). Substituting back: $0 = 0 + \int_0^1 e^{x+t}\cdot 0\,dt = 0$. ✓ **True.**

(c) and **(d) FALSE.** Evaluate the closed form at $x = 2$:

$$y(2) = 4a + \frac{2a(e-2)e^2}{3-e^2} = a\cdot\frac{4(3-e^2) + 2e^2(e-2)}{3-e^2} = a\cdot\frac{12 - 8e^2 + 2e^3}{3-e^2}.$$

Numerically, with $e^2 = 7.3891$ and $e^3 = 20.0855$:

$$12 - 8(7.3891) + 2(20.0855) = 12 - 59.113 + 40.171 = -6.942, \qquad 3 - e^2 = -4.389.$$

So $y(2) = a \times \dfrac{-6.942}{-4.389} = 1.582\,a$. Thus

$$a = 2:\ y(2) \approx 3.163, \qquad a = 1:\ y(2) \approx 1.582.$$

The printed values are $6/(3-e^2) = -1.367$ and $4/(3-e^2) = -0.911$, neither of which matches — and both are negative while the true values are positive. **Both false.**

The eigenvalue question. Introducing a parameter, consider

$$y(x) = f(x) + \mu \int_0^1 \mathrm{e}^{x+t} y(t)\,\mathrm{d}t.$$

The same computation gives $C\big(1 - \mu\frac{\mathrm{e}^2-1}{2}\big) = \int_0^1 \mathrm{e}^t f(t)\,\mathrm{d}t$, so the homogeneous equation has a nontrivial solution precisely when

$$\mu = \frac{2}{\mathrm{e}^2 - 1} \approx 0.3130,$$

the sole characteristic value of this rank-one kernel, with eigenfunction e^x. At that μ the Fredholm alternative applies: solutions exist only if the forcing is orthogonal to the eigenfunction, and are then non-unique. Our problem has $\mu = 1 \neq 0.313$, which is exactly why the solution is unique.

Contrast with Question 51. There the equation was of *Volterra* type, with variable upper limit, and unique solvability holds for every parameter value. Fredholm equations, with fixed limits, possess a genuine spectrum. This is the essential structural difference between the two classes.

Answer: **(a)** and **(b)**. (**(c)** and **(d)** are numerically incorrect; the correct value is $y(2) = a\,(12 - 8\mathrm{e}^2 + 2\mathrm{e}^3)/(3 - \mathrm{e}^2) \approx 1.582a$.)

Question 87. For the discrete distribution with parameter θ, the MLE is given by **(a)** $\frac{1}{n}\sum I(X_i \neq 0)$ **(b)** $1 - \frac{1}{n}\sum I(X_i = 0)$ **(c)** $\frac{1}{n}\sum I(X_i = 0)$ **(d)** $1 - \frac{1}{n}\sum I(X_i \neq 0)$.

The distribution did not reproduce — but the option set collapses

Options (a) and (b) are the same statistic, and so are **(c)** and **(d)**.
For each observation, exactly one of the indicators $I(X_i = 0)$ and $I(X_i \neq 0)$ equals 1, so

$$I(X_i \neq 0) + I(X_i = 0) = 1 \quad \text{identically.}$$

Summing over i and dividing by n:

$$\frac{1}{n}\sum_i I(X_i \neq 0) = 1 - \frac{1}{n}\sum_i I(X_i = 0),$$

which is precisely the assertion that **(a)** = **(b)**. Identically,

$$\frac{1}{n}\sum_i I(X_i = 0) = 1 - \frac{1}{n}\sum_i I(X_i \neq 0),$$

so **(c)** = **(d)**.

Consequently the item has only *two* distinct candidate answers, and the correct key must be either {**(a)**,**(b)**} or {**(c)**,**(d)**} — never a single option, and never a mixed pair. Recognising this before doing any work halves the problem and rules out two of the sixteen possible response patterns immediately.

Which pair, and why

MLE for a zero-versus-nonzero parameter

Let $p_\theta(0) = 1-\theta$ and $\sum_{k\neq 0} p_\theta(k) = \theta$, so that θ is the probability of a nonzero observation — the standard structure for zero-inflated and hurdle models. Write $m = \sum_i I(X_i \neq 0)$. The likelihood factors as

$$L(\theta) = (1-\theta)^{n-m}\,\theta^m \times (\text{terms free of } \theta),$$

so the log-likelihood is

$$\ell(\theta) = (n-m)\ln(1-\theta) + m\ln\theta + \text{const}, \qquad \ell'(\theta) = \frac{m}{\theta} - \frac{n-m}{1-\theta}.$$

Setting $\ell' = 0$ gives $m(1-\theta) = (n-m)\theta$, i.e. $m = n\theta$, so

$$\hat\theta = \frac{m}{n} = \frac{1}{n}\sum_i I(X_i \neq 0).$$

The second derivative $-m/\theta^2 - (n-m)/(1-\theta)^2 < 0$ confirms a maximum.

If instead θ is defined as the probability of a *zero* observation, the same computation with the roles exchanged gives $\hat\theta = \frac{1}{n}\sum I(X_i = 0)$, i.e. the pair $\{$**(c)**,**(d)**$\}$.

The more common parametrisation — and the one for which the estimator is usually written — is the first, giving $\{$**(a)**,**(b)**$\}$.

Sanity properties of the estimator (either version). It is the sample proportion, hence: unbiased, with variance $\theta(1-\theta)/n$; consistent by the law of large numbers; asymptotically normal by the CLT; and equal to the method of moments estimator, since $\mathbb{E}[I(X \neq 0)] = \theta$. It also attains the Cramér–Rao bound, the Fisher information for a Bernoulli observation being $1/\big(\theta(1-\theta)\big)$.

Answer: The distribution is absent, but **(a)** $\equiv$ **(b)** and **(c)** $\equiv$ **(d)**, so the key is one of these two pairs. Under the usual parametrisation $\theta = \mathbb{P}(X \neq 0)$, the answer is **(a)** and **(b)**.

Question 88. In the simple linear regression with the given design, the covariance condition implies **(a)** $\mathrm{Var}(\hat\beta_0) = \frac{2(2n+1)}{n(n-1)}$ **(b)** $\mathrm{Var}(\hat\beta_0) = \frac{2(n+2)}{n(n+1)}$ **(c)** $c = \frac{6}{n(1-n)}$ **(d)** $c = \frac{6}{n(n+1)}$.

Reconstructing the design from the option set

The design matrix is not printed, but the numerical forms of the options determine it. We test the canonical choice $x_i = i$ for $i = 1, \ldots, n$ with unit error variance $\sigma^2 = 1$, and find that two options match exactly — confirming the reconstruction.

Least squares for simple linear regression

Model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ with ε_i uncorrelated, mean 0, variance σ^2. Write

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

The least-squares estimators satisfy

$$\operatorname{Var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}, \qquad \operatorname{Var}(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^{\,2}}{S_{xx}} \right), \qquad \operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2 \bar{x}}{S_{xx}}.$$

The negative covariance reflects the geometry: raising the fitted slope forces the intercept down, since the fitted line must pass through $(\bar{x}, \bar{y})$. The covariance vanishes iff $\bar{x} = 0$, which is why centring the predictor orthogonalises the estimators.

The design $x_i = i$

The mean.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

The corrected sum of squares. Using $\sum i^2 = \frac{n(n+1)(2n+1)}{6}$,

$$S_{xx} = \sum i^2 - n\bar{x}^2 = \frac{n(n+1)(2n+1)}{6} - n \cdot \frac{(n+1)^2}{4} = n(n+1) \left[\frac{2n+1}{6} - \frac{n+1}{4} \right].$$

Common denominator 12:

$$\frac{2(2n+1) - 3(n+1)}{12} = \frac{4n+2-3n-3}{12} = \frac{n-1}{12},$$

so

$$S_{xx} = \frac{n(n+1)(n-1)}{12} = \frac{n(n^2-1)}{12}.$$

Verifying option (a)

With $\sigma^2 = 1$,

$$\operatorname{Var}(\hat{\beta}_0) = \frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} = \frac{1}{n} + \frac{(n+1)^2/4}{n(n^2-1)/12} = \frac{1}{n} + \frac{3(n+1)^2}{n(n^2-1)}.$$

Since $n^2 - 1 = (n+1)(n-1)$, the second term simplifies:

$$\frac{3(n+1)^2}{n(n+1)(n-1)} = \frac{3(n+1)}{n(n-1)}.$$

Hence

$$\operatorname{Var}(\hat{\beta}_0) = \frac{1}{n} + \frac{3(n+1)}{n(n-1)} = \frac{(n-1) + 3(n+1)}{n(n-1)} = \frac{4n+2}{n(n-1)} = \frac{2(2n+1)}{n(n-1)}.$$

This is *exactly* option **(a)**. **TRUE.**

Consequently option **(b)**, $\frac{2(n+2)}{n(n+1)}$, is **FALSE** — two different expressions cannot both equal the same variance (they differ already at $n = 2$: **(a)** gives 5, **(b)** gives 4/3).

Verifying option (c)

The covariance is

$$c = \mathrm{Cov}(\hat\beta_0, \hat\beta_1) = -\frac{\bar x}{S_{xx}} = -\frac{(n+1)/2}{n(n+1)(n-1)/12} = -\frac{(n+1)}{2}\cdot\frac{12}{n(n+1)(n-1)} = -\frac{6}{n(n-1)}.$$

Writing the minus sign into the denominator:

$$c = \frac{6}{n(1-n)},$$

which is *exactly* option **(c)**. **TRUE.**

Option **(d)**, $\frac{6}{n(n+1)}$, is **FALSE**: it is positive, whereas the covariance must be negative for a design with $\bar x > 0$ — a sign check that disposes of it without any arithmetic.

Why the reconstruction is reliable. Two of the four printed expressions are reproduced exactly by the single hypothesis $x_i = i$, $\sigma^2 = 1$, including the unusual form $6/\big(n(1-n)\big)$ with its inverted sign convention. The probability of such a coincidence under a different design is negligible.

Consistency check at $n = 3$. Design $x = (1, 2, 3)$: $\bar x = 2$, $S_{xx} = 1 + 0 + 1 = 2$. Formula: $S_{xx} = 3(9-1)/12 = 2$. ✓ $\mathrm{Var}(\hat\beta_0) = \frac{1}{3} + \frac{4}{2} = \frac{7}{3}$; option **(a)** gives $\frac{2\cdot 7}{3\cdot 2} = \frac{7}{3}$. ✓ $c = -\frac{2}{2} = -1$; option **(c)** gives $\frac{6}{3(1-3)} = \frac{6}{-6} = -1$. ✓

Answer: **(a)** and **(c)**.

Question 89. For the Bayes classifier with equal priors and the given densities, **(a)** classify to π_2 on the central interval involving $\sqrt{3}$ **(b)** classify to π_2 on an outer interval **(c)** classify to π_1 on a left interval **(d)** the misclassification probability equals $1/\sqrt{3}$.

The densities did not reproduce

The Bayes rule

With two populations π_1, π_2 having densities f_1, f_2 and prior probabilities p_1, p_2, and with unit misclassification costs, the rule minimising the total probability of error assigns x to the population with the larger posterior:

$$\text{classify to } \pi_1 \iff p_1 f_1(x) > p_2 f_2(x) \iff \frac{f_1(x)}{f_2(x)} > \frac{p_2}{p_1}.$$

With equal priors $p_1 = p_2 = \frac{1}{2}$ this reduces to the simple comparison

$$\text{classify to } \pi_1 \iff f_1(x) > f_2(x).$$

The boundary of the classification regions is the set where the two densities cross.

Why the regions can be intervals or unions of intervals

The shape of the decision region is determined by the number of crossings of f_1 and f_2.

Two normals with equal variances: $f_1 - f_2$ changes sign once, so the regions are two half-lines separated by a single threshold — the classical linear discriminant.
Two normals with unequal variances: the log-ratio is a quadratic in x, so there are up to two crossings and the region for the higher-variance population is the union of two outer intervals, with the lower-variance population claiming a central interval. This is the quadratic discriminant, and it is exactly the structure the option set describes: a "central interval" (option **(a)**) and "outer intervals" (option **(b)**).

Where $\sqrt{3}$ typically comes from

For $N(0,\sigma_1^2)$ versus $N(0,\sigma_2^2)$ with $\sigma_2^2 > \sigma_1^2$ and equal priors, setting $f_1(x) = f_2(x)$ gives

$$\frac{1}{\sigma_1}\mathrm{e}^{-x^2/2\sigma_1^2} = \frac{1}{\sigma_2}\mathrm{e}^{-x^2/2\sigma_2^2} \quad\Longrightarrow\quad x^2 = \frac{2\sigma_1^2\sigma_2^2}{\sigma_2^2-\sigma_1^2}\ln\frac{\sigma_2}{\sigma_1}.$$

The classification boundary is the symmetric pair $\pm x_0$, and radicals such as $\sqrt{3}$ arise naturally from the resulting square root. Uniform densities on nested intervals produce $\sqrt{3}$ even more directly, since a Uniform$(-a,a)$ has standard deviation $a/\sqrt{3}$.

The total error probability for the Bayes rule with equal priors is

$$\mathbb{P}(\text{error}) = \tfrac{1}{2}\mathbb{P}_1(R_2) + \tfrac{1}{2}\mathbb{P}_2(R_1) = \tfrac{1}{2}\int \min\{f_1, f_2\}\,\mathrm{d}x,$$

which is $\frac{1}{2}\big(1 - \frac{1}{2}\left\|f_1 - f_2\right\|_1\big)$ in terms of total variation distance. Note that this is at most $\frac{1}{2}$; option **(d)**'s value $1/\sqrt{3} \approx 0.577$ *exceeds* $\frac{1}{2}$ and is therefore **impossible** for a Bayes rule with equal priors, which can always guarantee error at most $\frac{1}{2}$ by ignoring the data. So **(d)** is eliminable without knowing the densities.

Answer: Indeterminate: the densities are absent. Option **(d)** is eliminable, since a Bayes rule with equal priors has error probability at most $1/2$, whereas $1/\sqrt{3} \approx 0.577 > 1/2$.

Question 90. For a two-component Poisson failure system, **(a)** series reliability is $\mathrm{e}^{-2\lambda t}$ **(b)** series hazard is 2λ **(c)** parallel reliability is $1-(1-\mathrm{e}^{-\lambda t})^2$ **(d)** parallel hazard is $2\lambda(1-\mathrm{e}^{-\lambda t})$.

The concept: reliability, hazard, and system structure

The three functions

For a lifetime T with density f and CDF F:

$$R(t) = \mathbb{P}(T > t) = 1 - F(t) \quad \text{(reliability)},$$

$$h(t) = \frac{f(t)}{R(t)} = -\frac{R'(t)}{R(t)} = -\frac{\mathrm{d}}{\mathrm{d}t}\ln R(t) \quad \text{(hazard)}.$$

The hazard is the instantaneous failure rate given survival to t; the reliability is recovered by $R(t) = \exp\big(-\int_0^t h(s)\,\mathrm{d}s\big)$.

For an exponential lifetime with rate λ: $R(t) = \mathrm{e}^{-\lambda t}$ and $h(t) = \lambda$, constant. Constant hazard is equivalent to the memoryless property and characterises the exponential distribution.

Series and parallel

Assume two independent components each with reliability $R_i(t)$.
Series (system works iff both work):

$$R_s(t) = R_1(t)R_2(t), \qquad h_s(t) = h_1(t) + h_2(t).$$

Hazards add in series — immediate from $\ln R_s = \ln R_1 + \ln R_2$.
Parallel (system works iff at least one works):

$$R_p(t) = 1 - \big(1 - R_1(t)\big)\big(1 - R_2(t)\big).$$

Hazards do *not* add in parallel, and this is the crux of the present item.

(a) — TRUE

With $R_1 = R_2 = \mathrm{e}^{-\lambda t}$,

$$R_s(t) = \mathrm{e}^{-\lambda t} \cdot \mathrm{e}^{-\lambda t} = \mathrm{e}^{-2\lambda t}. \checkmark$$

(b) — TRUE

$$h_s(t) = -\frac{\mathrm{d}}{\mathrm{d}t} \ln \mathrm{e}^{-2\lambda t} = -\frac{\mathrm{d}}{\mathrm{d}t}(-2\lambda t) = 2\lambda. \checkmark$$

Equivalently, hazards add: $\lambda + \lambda = 2\lambda$. The series system is itself exponential with rate 2λ, i.e. its lifetime is $\min(T_1, T_2)$, and the minimum of independent exponentials is exponential with the summed rate.

(c) — TRUE

$$R_p(t) = 1 - \big(1 - \mathrm{e}^{-\lambda t}\big)^2. \checkmark$$

Expanding, $R_p(t) = 2\mathrm{e}^{-\lambda t} - \mathrm{e}^{-2\lambda t}$, which is the inclusion–exclusion form $\mathbb{P}(T_1 > t) + \mathbb{P}(T_2 > t) - \mathbb{P}(\min > t)$.

(d) — FALSE

Compute the parallel hazard properly. With $R_p = 2\mathrm{e}^{-\lambda t} - \mathrm{e}^{-2\lambda t}$,

$$R_p'(t) = -2\lambda\mathrm{e}^{-\lambda t} + 2\lambda\mathrm{e}^{-2\lambda t} = -2\lambda\,\mathrm{e}^{-\lambda t}\big(1 - \mathrm{e}^{-\lambda t}\big).$$

Hence

$$h_p(t) = -\frac{R_p'}{R_p} = \frac{2\lambda\,\mathrm{e}^{-\lambda t}\big(1 - \mathrm{e}^{-\lambda t}\big)}{\mathrm{e}^{-\lambda t}\big(2 - \mathrm{e}^{-\lambda t}\big)} = \frac{2\lambda\big(1 - \mathrm{e}^{-\lambda t}\big)}{2 - \mathrm{e}^{-\lambda t}}.$$

The printed option omits the denominator $2 - \mathrm{e}^{-\lambda t}$. **False.**

Sanity checks on h_p.

$$h_p(0) = \frac{2\lambda(1-1)}{2-1} = 0,$$

which is right: a brand-new parallel system cannot fail instantaneously, since both components must fail. As $t \to \infty$,

$$h_p(t) \longrightarrow \frac{2\lambda(1-0)}{2-0} = \lambda,$$

also right: after a long time one component has almost surely failed, and the system inherits the single-component rate λ. The hazard is increasing from 0 to λ — the parallel system exhibits *wear-in* rather than memorylessness, and is therefore *not* exponential despite being built from exponential parts.

The printed expression $2\lambda(1-e^{-\lambda t})$ has the correct value at $t = 0$ but tends to 2λ, twice the correct asymptote — a plausible-looking but incorrect formula, and precisely the sort of distractor Part C rewards you for checking.

Mean lifetimes. Series: $\mathbb{E}[\min] = 1/(2\lambda)$. Parallel: $\mathbb{E}[\max] = \int_0^\infty R_p = \frac{2}{\lambda} - \frac{1}{2\lambda} = \frac{3}{2\lambda}$, the familiar $\frac{1}{\lambda}(1 + \frac{1}{2})$.

Answer: **(a)**, **(b)** and **(c)**.

Question 91. Let $A = I - 2uu^*$ where $u \in \mathbb{C}^n$ is a unit vector. Then **(a)** A is skew-Hermitian **(b)** A is unitary **(c)** A is normal **(d)** the real part of A is symmetric.

The concept: the Householder reflector

Householder matrices

For a unit vector $u \in \mathbb{C}^n$ (so $u^*u = 1$), the matrix

$$A = I - 2uu^*$$

is the *Householder reflector* in the hyperplane $u^\perp$. Its geometric action is transparent: decompose any x as $x = (u^*x)u + w$ with $w \perp u$. Then

$$Ax = x - 2u(u^*x) = \big[(u^*x)u + w\big] - 2(u^*x)u = -(u^*x)u + w.$$

So A fixes $u^\perp$ pointwise and negates the u-direction — a reflection in the hyperplane orthogonal to u. Hence its eigenvalues are -1 (once, with eigenvector u) and $+1$ ($n-1$ times, with eigenspace $u^\perp$), giving

$$\operatorname{tr} A = n - 2, \qquad \det A = -1.$$

(a) — FALSE

Take adjoints, using $(uu^*)^* = u^{**}u^* = uu^*$:

$$A^* = (I - 2uu^*)^* = I^* - 2(uu^*)^* = I - 2uu^* = A.$$

So A is **Hermitian**, not skew-Hermitian. **False.**

(A skew-Hermitian matrix satisfies $A^* = -A$ and has purely imaginary eigenvalues; here the eigenvalues are ± 1, real. The only matrix that is both Hermitian and skew-Hermitian is 0.)

(b) — TRUE

Compute A^2, using $u^*u = 1$:

$$A^2 = (I - 2uu^*)(I - 2uu^*) = I - 4uu^* + 4u\,(u^*u)\,u^* = I - 4uu^* + 4uu^* = I.$$

So A is an involution. Combined with $A^* = A$ from the previous step,

$$A^*A = A^2 = I,$$

which is exactly the definition of unitarity. **True.**

Consequently A preserves the Euclidean norm: $\|Ax\|^2 = (Ax)^*(Ax) = x^*A^*Ax = x^*x = \|x\|^2$ — as any reflection must.

(c) — TRUE

A matrix is *normal* if $A^*A = AA^*$. Since $A^* = A$,

$$A^*A = A^2 = AA^*,$$

trivially. **True.**

> Every Hermitian matrix is normal, and every unitary matrix is normal; A is both. The spectral theorem for normal matrices therefore applies, and indeed we have already exhibited the orthogonal eigenbasis: $\{u\} \cup (\text{any orthonormal basis of } u^\perp)$.

(d) — TRUE

Write $u = a + ib$ with $a, b \in \mathbb{R}^n$. Then

$$uu^* = (a + \mathrm{i}b)(a^\top - \mathrm{i}b^\top) = \big(aa^\top + bb^\top\big) + \mathrm{i}\big(ba^\top - ab^\top\big),$$

so

$$\operatorname{Re}(A) = I - 2\big(aa^\top + bb^\top\big).$$

Each of $aa^\top$ and $bb^\top$ is a real symmetric rank-one matrix, and I is symmetric, so $\operatorname{Re}(A)$ is symmetric. **True.**

More cheaply: $A^* = A$ means $\bar{A}^\top = A$, i.e. taking real parts, $\operatorname{Re}(A)^\top = \operatorname{Re}(A)$. The real part of any Hermitian matrix is symmetric (and the imaginary part is skew-symmetric).

> **Why Householder matrices matter numerically.** Given any $x \neq 0$, choosing
>
> $$u = \frac{x - \|x\|\,e_1}{\|x - \|x\|\,e_1\|}$$
>
> gives $Ax = \|x\|\,e_1$: a single reflection annihilates all but the first entry of a vector. Applying $n-1$ such reflectors successively produces the QR factorisation. Because each factor is unitary, the transformation has condition number 1 and does not amplify rounding error — the reason Householder QR is preferred to Gram–Schmidt in practice.

> **Answer:** **(b)**, **(c)** and **(d)**.

Question 92. For an $M/M/1$ queue with arrival rate 8 per hour and service rate 10 per hour, **(a)** the expected number in the system is 4 **(b)** the expected queue length is 3.2 **(c)** the expected waiting time in the queue is 24 minutes **(d)** the expected service time is 6 minutes.

The standard formulae

$M/M/1$ steady state

With arrival rate λ, service rate μ, and traffic intensity $\rho = \lambda/\mu < 1$:

$$p_n = (1-\rho)\rho^n,$$

a geometric distribution on $\{0, 1, 2, \ldots\}$. Hence

$$L = \mathbb{E}[N] = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}, \qquad L_q = \mathbb{E}[N_q] = \frac{\rho^2}{1-\rho} = L - \rho.$$

Little's law $L = \lambda W$ and $L_q = \lambda W_q$ then give

$$W = \frac{1}{\mu-\lambda}, \qquad W_q = \frac{\rho}{\mu-\lambda} = W - \frac{1}{\mu}.$$

The relation $L = L_q + \rho$ says the expected number *in service* is ρ, which is also the server utilisation — obvious once seen, since the server holds either 0 or 1 customer.

The computation

Here $\lambda = 8$, $\mu = 10$ per hour, so

$$\rho = \frac{8}{10} = 0.8 < 1 \checkmark \qquad \text{(the queue is stable)}.$$

(a) TRUE.

$$L = \frac{\rho}{1-\rho} = \frac{0.8}{0.2} = 4.$$

Equivalently $L = \lambda/(\mu-\lambda) = 8/2 = 4$.

(b) TRUE.

$$L_q = L - \rho = 4 - 0.8 = 3.2.$$

Equivalently $\rho^2/(1-\rho) = 0.64/0.2 = 3.2$.

(c) TRUE.

$$W_q = \frac{L_q}{\lambda} = \frac{3.2}{8} = 0.4 \text{ hour} = 0.4 \times 60 = 24 \text{ minutes}.$$

(d) TRUE. Service times are exponential with rate $\mu = 10$ per hour, so the mean service time is

$$\frac{1}{\mu} = \frac{1}{10} \text{ hour} = 6 \text{ minutes}.$$

Internal consistency. $W = W_q + 1/\mu = 24 + 6 = 30$ minutes $= 0.5$ hour, and Little's law demands $L = \lambda W = 8 \times 0.5 = 4 \checkmark$— matching option **(a)** independently.

The sensitivity of L to ρ. The formula $L = \rho/(1-\rho)$ blows up as $\rho \uparrow 1$: at $\rho = 0.8$ we have $L = 4$, at $\rho = 0.9$ we get $L = 9$, at $\rho = 0.95$ we get $L = 19$. A ten-percent increase in load near saturation more than doubles the congestion. This convexity is the central practical lesson of queueing theory and the reason systems are rarely designed to run above about 80% utilisation.

Answer: **(a)**, **(b)**, **(c)** and **(d)** — all four.

Question 93. Simpson's rule on $[0, 2]$ with $h = 1$ is written $\int_0^2 f \approx af(0) + bf(1) + cf(2)$. Then **(a)** $a^2 + b^2 + c^2 = 10$ **(b)** the error term involves f''' **(c)** $a = c$ **(d)** the rule is exact for cubics.

Deriving the weights

Simpson's rule

Approximate f on $[x_0, x_0 + 2h]$ by the quadratic interpolating it at $x_0, x_0 + h, x_0 + 2h$, and integrate that quadratic exactly. The result is

$$\int_{x_0}^{x_0+2h} f(x)\,\mathrm{d}x \approx \frac{h}{3}\big[f(x_0) + 4f(x_0 + h) + f(x_0 + 2h)\big].$$

Here $h = 1$ and the interval is $[0, 2]$:

$$\int_0^2 f \approx \frac{1}{3}\big[f(0) + 4f(1) + f(2)\big], \qquad \text{so} \qquad a = \frac{1}{3}, \quad b = \frac{4}{3}, \quad c = \frac{1}{3}.$$

Verification by the method of undetermined coefficients. Require exactness on $1, x, x^2$:

$$\int_0^2 1 = 2 = a + b + c, \qquad \int_0^2 x = 2 = b + 2c, \qquad \int_0^2 x^2 = \frac{8}{3} = b + 4c.$$

Subtracting the second from the third gives $2c = \frac{8}{3} - 2 = \frac{2}{3}$, so $c = \frac{1}{3}$; then $b = 2 - \frac{2}{3} = \frac{4}{3}$ and $a = 2 - \frac{4}{3} - \frac{1}{3} = \frac{1}{3}$. ✓

(c) — TRUE

$a = c = \frac{1}{3}$. The equality is forced by symmetry: the nodes are symmetric about the midpoint of the interval, and the integration functional is invariant under the reflection $x \mapsto 2 - x$, so the weights at reflected nodes must agree. **True.**

(a) — FALSE

$$a^2 + b^2 + c^2 = \frac{1}{9} + \frac{16}{9} + \frac{1}{9} = \frac{18}{9} = 2.$$

This is 2, not 10. **False.**

(d) — TRUE

Simpson's rule is exact for all polynomials of degree ≤ 3, one degree beyond what the construction guarantees.

Proof. By linearity and the verification above, exactness holds on $1, x, x^2$. For x^3 on $[0,2]$:

$$\int_0^2 x^3\,\mathrm{d}x = \frac{2^4}{4} = 4, \qquad \frac{1}{3}\big[0 + 4\cdot 1 + 8\big] = \frac{12}{3} = 4. \checkmark$$

True.

The mechanism is symmetry: the error functional annihilates odd functions about the midpoint. Writing $x = 1 + s$, the cubic s^3 integrates to 0 over $s \in [-1,1]$ and the symmetric weights also give $\frac{1}{3}[(-1)^3 + 0 + 1^3] = 0$. Every Newton–Cotes rule with an even number of subintervals gains this extra degree of exactness.

(b) — FALSE

Because the rule is exact on cubics, the leading error term cannot involve f''' — it must involve the first derivative not annihilated, namely the fourth. The standard result is

$$\int_{x_0}^{x_0+2h} f - \frac{h}{3}\big[f_0 + 4f_1 + f_2\big] = -\frac{h^5}{90} f^{(4)}(\xi), \qquad \xi \in (x_0, x_0 + 2h).$$

False.

The contrast with the trapezoidal rule, whose error is $-\frac{h^3}{12} f''(\xi)$ per panel: Simpson gains two orders. Over n panels the composite errors are $O(h^2)$ for trapezoid and $O(h^4)$ for Simpson.

Degree of precision is the standard vocabulary: trapezoid has degree 1, Simpson degree 3, Simpson's 3/8 rule also degree 3 (with a worse constant), and n-point Gauss–Legendre degree $2n-1$ — the optimum achievable with n function evaluations, which is why Gaussian quadrature dominates when the nodes may be chosen freely.

With $h = 1$ here, the error bound is $|E| \leq \frac{1}{90} \max\big|f^{(4)}\big|$.

Answer: **(c)** and **(d)**.

Question 94. For the two-stage Runge–Kutta method applied to $y' = f(x, y)$, **(a)** the local truncation error is $O(h^2)$ **(b)** the method is of order 2 **(c)** the parameters satisfy the stated relation **(d)** the stability interval is as given.

The specific scheme did not reproduce

The two-stage Runge–Kutta family

The general explicit two-stage method is

$$y_{n+1} = y_n + h\big(b_1 k_1 + b_2 k_2\big), \qquad k_1 = f(x_n, y_n), \quad k_2 = f(x_n + c_2 h,\ y_n + a_{21} h k_1).$$

Expanding both the exact solution and the numerical one in Taylor series about (x_n, y_n) and matching terms through h^2 yields the *order conditions*

$$b_1 + b_2 = 1, \qquad b_2 c_2 = \tfrac{1}{2}, \qquad a_{21} = c_2.$$

Any parameter choice satisfying these gives a second-order method. Familiar members:

$$\begin{array}{ll} c_2 = 1,\ b_1 = b_2 = \frac{1}{2} & \text{(Heun / improved Euler)}, \\ c_2 = \frac{1}{2},\ b_1 = 0,\ b_2 = 1 & \text{(midpoint method)}, \\ c_2 = \frac{2}{3},\ b_1 = \frac{1}{4},\ b_2 = \frac{3}{4} & \text{(Ralston, minimal error constant)}. \end{array}$$

No choice of two-stage explicit parameters attains order 3: the third order conditions require $b_2 c_2^2 = \frac{1}{3}$ and $\sum b_i a_{ij} c_j = \frac{1}{6}$, and the second of these cannot be met with only two stages.

(b) — TRUE

Provided the order conditions above hold — as they do for any scheme presented as "the two-stage Runge–Kutta method" — the method has **global** order 2: the accumulated error after integrating over a fixed interval is $O(h^2)$. **True.**

(a) — FALSE as printed

Local versus global error

The *local truncation error* is the error committed in a single step, starting from exact data:

$$\tau_{n+1} = y(x_{n+1}) - \Big[y(x_n) + h\Phi\big(x_n, y(x_n), h\big)\Big].$$

A method of order p has $\tau = O(h^{p+1})$ locally. The extra power is lost in accumulation: over an interval of fixed length there are $N = O(1/h)$ steps, so

$$\text{global error} = O(1/h) \times O(h^{p+1}) = O(h^p).$$

For a second-order method, therefore,

$$\tau = O(h^3), \qquad \text{global error} = O(h^2).$$

The option asserts a local truncation error of $O(h^2)$, which corresponds to a *first*-order method. **False as printed.**

This is the single most frequently confused pair of definitions in numerical ODEs. If the item intends "error" loosely — meaning the global error — then the statement would be true. But "local truncation error" is a technical term, and under its standard meaning $O(h^2)$ is wrong for a second-order scheme. The options **(a)** and **(b)** as printed are mutually inconsistent, since a method of order 2 has local error $O(h^3)$.

(c) and (d)

Option **(c)** concerns the specific parameter relation, and **(d)** the interval of absolute stability; neither the parameters nor the interval reproduced.

For reference, applying any second-order two-stage method to the test equation $y' = \lambda y$ gives the amplification factor

$$R(z) = 1 + z + \frac{z^2}{2}, \qquad z = h\lambda,$$

identically for every member of the family — the order conditions determine R completely. Absolute stability requires $|R(z)| \le 1$, and for real negative z this gives

$$-2 \le z \le 0,$$

so the stability interval is $h\lambda \in [-2, 0]$, i.e. $0 < h \le 2/|\lambda|$ for $\lambda < 0$. If option **(d)** states $(-2, 0)$, it is correct.

For comparison: explicit Euler has $R(z) = 1 + z$ and stability interval $(-2, 0)$ as well; classical RK4 has $R(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24}$ and interval approximately $(-2.785, 0)$. All explicit methods have bounded stability regions — the reason stiff problems demand implicit schemes, whose stability regions are unbounded.

Answer: **(b)** is true. **(a)** is false as printed (a second-order method has local truncation error $O(h^3)$, global error $O(h^2)$). **(c)** and **(d)** cannot be adjudicated: the scheme's parameters are absent. If **(d)** names the interval $(-2, 0)$, it is correct.

Question 95. For subspaces w_1, w_2 of a 7-dimensional space with $\dim w_1 = 4$ and $\dim w_2 = 3$, **(a)** $\dim(w_1 \cap w_2) \ge 0$ **(b)** $\dim(w_1 + w_2) = 7$ **(c)** $\dim(w_1 \cap w_2) = 1$ **(d)** $\dim(w_1 + w_2) = 5$.

The governing identity

The dimension formula

For subspaces U, W of a finite-dimensional vector space,

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Why. Extend a basis of $U \cap W$ to a basis of U and separately to a basis of W; the union is a basis of $U + W$, and the vectors from $U \cap W$ have been counted once rather than twice.

With $\dim w_1 = 4$, $\dim w_2 = 3$ inside a space of dimension 7:

$$\dim(w_1 + w_2) = 7 - \dim(w_1 \cap w_2). \tag{95.1}$$

The constraints

Two bounds pin the intersection down:

$$\dim(w_1 \cap w_2) \;\ge\; \dim w_1 + \dim w_2 - 7 = 4 + 3 - 7 = 0,$$

$$\dim(w_1 \cap w_2) \;\le\; \min\{\dim w_1, \dim w_2\} = 3.$$

So $\dim(w_1 \cap w_2) \in \{0, 1, 2, 3\}$ and correspondingly $\dim(w_1 + w_2) \in \{7, 6, 5, 4\}$. All four combinations are realisable by choosing suitable coordinate subspaces of $\mathbb{R}^7$.

The mutual exclusivity the option set conceals

By (95.1):

- **(b)** $\dim(w_1 + w_2) = 7 \iff \dim(w_1 \cap w_2) = 0$;
- **(c)** $\dim(w_1 \cap w_2) = 1 \iff \dim(w_1 + w_2) = 6$;
- **(d)** $\dim(w_1 + w_2) = 5 \iff \dim(w_1 \cap w_2) = 2$.

Hence **(b)**, **(c)** and **(d)** are **pairwise incompatible**: at most one of them can hold. In particular **(c)** and **(d)** cannot both be marked, nor **(b)** with either.

(a) TRUE unconditionally: the dimension of any subspace is a non-negative integer, so $\dim(w_1 \cap w_2) \geq 0$ is vacuously true. It is also uninformative — but true statements score.

The sharp version of (a). The genuinely informative bound is

$$\dim(U \cap W) \;\geq\; \dim U + \dim W - \dim V,$$

which here also gives ≥ 0 but has real content when the dimensions are larger: two 4-dimensional subspaces of $\mathbb{R}^7$ must intersect in dimension at least 1, so they cannot be disjoint. Geometrically, two hyperplanes in $\mathbb{R}^n$ always meet in an $(n-2)$-dimensional subspace, and a line and a hyperplane in general position meet in $\{0\}$ only when the line is transverse.

Answer: Indeterminate: the specific subspaces are absent. **(a)** is true unconditionally; **(b)**, **(c)**, **(d)** are mutually exclusive, so at most one of them can be part of the key.

Question 96. The block design described is **(a)** connected **(b)** binary **(c)** symmetric **(d)** proper.

The block layout did not fully reproduce

Vocabulary of block designs

Let v treatments be arranged in b blocks. Write n_{ij} for the number of times treatment i appears in block j, r_i for the replication of treatment i, and k_j for the size of block j.
Binary: every $n_{ij} \in \{0, 1\}$ — no treatment appears twice in the same block.
Proper (equiblock): all blocks have the same size, $k_j \equiv k$.
Equireplicate: all treatments have the same replication, $r_i \equiv r$.
Connected: every elementary contrast $\tau_i - \tau_{i'}$ is estimable; equivalently the treatment-block incidence graph is connected; equivalently the rank of the C-matrix is $v - 1$.
Balanced incomplete block design (BIBD): binary, proper, equireplicate, and every *pair* of treatments occurs together in exactly λ blocks. The parameters then satisfy

$$vr = bk, \qquad \lambda(v-1) = r(k-1), \qquad b \geq v \quad \text{(Fisher's inequality)}.$$

Symmetric BIBD: a BIBD with $b = v$ (equivalently $r = k$).

The cyclic design most likely intended

The layout suggested by the item is the cyclic arrangement on four treatments A, B, C, D in four blocks of size two:

$$\{A, B\}, \quad \{B, C\}, \quad \{C, D\}, \quad \{D, A\}.$$

Adjudicate against the definitions.

(b) Binary — TRUE. No treatment is repeated within any block; every $n_{ij} \in \{0, 1\}$.

(d) Proper — TRUE. Every block has size $k = 2$.

(a) Connected — TRUE. The incidence graph is the 4-cycle $A - B - C - D - A$, which is connected: any treatment can be reached from any other through a chain of blocks. Hence all elementary contrasts are estimable. (This design is also equireplicate with $r = 2$, since each treatment sits in exactly two blocks; note $vr = 4 \cdot 2 = 8 = bk = 4 \cdot 2$ ✓.)

(c) Symmetric — doubtful. The design does satisfy $b = v = 4$ and $r = k = 2$, which is the numerical condition for symmetry. But "symmetric" is properly an attribute of a *BIBD*, and this design is *not balanced*:

$$\lambda_{AB} = \lambda_{BC} = \lambda_{CD} = \lambda_{DA} = 1, \qquad \lambda_{AC} = \lambda_{BD} = 0.$$

The diagonal pairs never appear together, so no common λ exists. Confirm against the parameter relation: a BIBD would need $\lambda(v-1) = r(k-1)$, i.e. $3\lambda = 2$, which has no integer solution. So no BIBD with these parameters exists at all, and the design cannot be a symmetric BIBD.

If "symmetric" is read loosely as $b = v$, the option is true. This is a genuine ambiguity of terminology: some elementary texts call any design with $b = v$ symmetric, while the standard usage reserves the term for BIBDs. Under the standard usage **(c)** is false.
The smallest genuine BIBD with $k = 2$ on four treatments is the complete pairing $\{A, B\}, \{A, C\}, \{A, D\}, \{B, C\}, \{B, D\}, \{C, D\}$, with $b = 6$, $r = 3$, $\lambda = 1$. Here $b = 6 > v = 4$, consistent with Fisher's inequality, and the design is not symmetric.

Answer: **(a)**, **(b)** and **(d)** for the cyclic design; **(c)** is false under the standard definition, since the design is not balanced. The block layout did not reproduce in full, so this reading is provisional.

Question 97. For the harmonic numbers $H_n = \sum_{k=1}^{n} 1/k$, **(a)** $H_{2026} > 6$ **(b)** $H_{2052} < 6.5$ **(c)** $H_{52} < 3.5$ **(d)** $H_{153} < 4.5$.

The tools

Estimating harmonic numbers

Integral bounds. Since $1/x$ is decreasing,

$$\int_1^{n+1} \frac{\mathrm{d}x}{x} \;\le\; H_n \;\le\; 1 + \int_1^{n} \frac{\mathrm{d}x}{x}, \qquad \text{i.e.} \qquad \ln(n+1) \le H_n \le 1 + \ln n.$$

Asymptotics. More precisely,

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O(n^{-4}), \qquad \gamma = 0.5772156649\ldots$$

the Euler–Mascheroni constant. For n in the hundreds this is accurate to five decimal places, and $H_n \approx \ln n + 0.5772$ suffices for all four options.
Oresme's dyadic bound (for rigorous lower estimates without logarithms): grouping terms in blocks of doubling length,

$$H_{2^m} \;\geq\; 1 + \frac{m}{2}.$$

(a) — TRUE

Estimate H_{2026}:

$$\ln 2026 \approx 7.6140, \qquad H_{2026} \approx 7.6140 + 0.5772 + 0.0002 = 8.1914.$$

So $H_{2026} \approx 8.19 > 6$. **True.**

A rigorous argument avoiding decimals. Since $2026 > 2^{10} = 1024$ and H_n is increasing,

$$H_{2026} > H_{1024} = H_{2^{10}} \;\geq\; 1 + \frac{10}{2} = 6.$$

Strictly $H_{2026} > H_{1024} \geq 6$, so the inequality is strict. ✓

(b) — FALSE

$$\ln 2052 \approx 7.6267, \qquad H_{2052} \approx 7.6267 + 0.5772 = 8.2039.$$

This is about 8.20, comfortably *greater* than 6.5. **False.**

Note that H_{2052} exceeds H_{2026} by only $\sum_{k=2027}^{2052} 1/k \approx 26/2039 \approx 0.0128$ — the harmonic series grows so slowly that adding twenty-six terms near $n = 2000$ changes the sum by about one part in six hundred. Any option asserting that H_{2052} is materially smaller than H_{2026}-type values is implausible on its face.

(c) — FALSE

$$\ln 52 \approx 3.9512, \qquad H_{52} \approx 3.9512 + 0.5772 + 0.0096 = 4.5380.$$

So $H_{52} \approx 4.54$, which is *not* less than 3.5. **False.**

Rigorous check. $H_{32} = H_{2^5} \geq 1 + \frac{5}{2} = 3.5$, and $52 > 32$, so $H_{52} > H_{32} \geq 3.5$. The claim fails. ✓

(d) — FALSE

$$\ln 153 \approx 5.0304, \qquad H_{153} \approx 5.0304 + 0.5772 + 0.0033 = 5.6109.$$

So $H_{153} \approx 5.61$, not less than 4.5. **False.**

Rigorous check. $H_{128} = H_{2^7} \geq 1 + \frac{7}{2} = 4.5$, and $153 > 128$, so $H_{153} > 4.5$. ✓

The pattern in the distractors. Options **(c)** and **(d)** both propose upper bounds that the dyadic estimate immediately refutes, because $52 > 2^5$ and $153 > 2^7$ push the sums past 3.5 and 4.5 respectively. The item is testing whether the candidate can locate H_n to within about half a unit without a calculator, and Oresme's bound — a fourteenth-century argument — is entirely sufficient for all four.

Useful reference values:

$$H_{10} \approx 2.929, \quad H_{100} \approx 5.187, \quad H_{1000} \approx 7.485, \quad H_{10^4} \approx 9.788.$$

Each factor of ten adds $\ln 10 \approx 2.303$.

Answer: **(a)** only.

Question 98. For the given second-order differential equation, the solution satisfies the stated asymptotic and boundedness properties.

The equation and its options did not reproduce

The source text for this item lost both the differential equation and the four option statements. What follows sets out the standard framework such items test, so that the reader can apply it once the original is to hand.

Classification of points and the nature of solutions

For $y'' + p(x)y' + q(x)y = 0$:
Ordinary point at x_0: p, q analytic there. Two linearly independent power-series solutions exist, with radius of convergence at least the distance to the nearest singularity of p or q in the complex plane.
Regular singular point at x_0: $(x-x_0)p$ and $(x-x_0)^2q$ are analytic. The Frobenius method applies: solutions of the form $(x-x_0)^\rho \sum a_n(x-x_0)^n$, where ρ solves the indicial equation

$$\rho(\rho-1) + p_0\rho + q_0 = 0.$$

When the roots differ by an integer, the second solution may involve a logarithm.
Irregular singular point: neither of the above. Solutions typically have essential singularities and require asymptotic rather than convergent expansions.

Oscillation and boundedness

For $y'' + q(x)y = 0$ with $q > 0$:
Sturm comparison. If $q_1(x) \le q_2(x)$ then between consecutive zeros of a solution of $y'' + q_1y = 0$ there lies a zero of every solution of $y'' + q_2y = 0$. Larger q means faster oscillation.
Oscillation criterion. If $q(x) \ge m > 0$ for all large x, every solution oscillates with zeros spaced at most $\pi/\sqrt{m}$ apart. If $q(x) \to 0$ fast enough — specifically if $\int^\infty xq(x)\,\mathrm{d}x < \infty$ — solutions are non-oscillatory.
Boundedness. For $y'' + q(x)y = 0$ with q positive, increasing and $q \to \infty$, all solutions are bounded and in fact decay like $q^{-1/4}$; this is the WKB or Liouville–Green approximation

$$y \sim q(x)^{-1/4} \exp\left(\pm \mathrm{i} \int^x \sqrt{q(s)}\,\mathrm{d}s\right).$$

Conversely if q decreases to zero the amplitude grows.

The transformation that removes the first-derivative term is worth having at hand, since it converts any second-order linear equation into the normal form to which the criteria above apply. Setting

$$y = u(x)\exp\left(-\tfrac{1}{2}\int p(x)\,\mathrm{d}x\right)$$

in $y'' + py' + qy = 0$ yields

$$u'' + Q(x)u = 0, \qquad Q = q - \frac{p^2}{4} - \frac{p'}{2}.$$

Question 102 of this paper is precisely such a reduction, with $p = -4x$ and $q = 4x^2 - 1$ giving $Q = 1$. Boundedness and oscillation of y then follow from those of u together with the behaviour of the exponential factor — and it is exactly this factor that makes the solutions of Question 102 unbounded despite u being purely trigonometric.

Answer: Indeterminate: the differential equation and the option statements are absent from the available text.

Question 99. Let A be a matrix with the stated minimal polynomial. Which conclusions follow? **(a)** $\mu_A(x) = x^2 + 1 \Rightarrow A^4 = I$ **(b)** $\mu_A(x) = x^2 + x \Rightarrow A^8 = I$ **(c)** $\mu_A(x) = x^2 - x + 1 \Rightarrow A^6 = I$ **(d)** $\mu_A(x) = x^2 + x + 1 \Rightarrow A^3 = I$.

The governing principle

Minimal polynomials and annihilation

The minimal polynomial μ_A is the monic polynomial of least degree with $\mu_A(A) = 0$, and it has the fundamental divisibility property:

$$p(A) = 0 \iff \mu_A \mid p.$$

Why. Divide $p = q\mu_A + r$ with $\deg r < \deg \mu_A$. Then $r(A) = p(A) = 0$, and minimality forces $r = 0$.

So to test whether $A^m = I$, one asks simply: **does $\mu_A(x)$ divide $x^m - 1$?** This converts a matrix question into a polynomial divisibility question, requiring no knowledge of A beyond its minimal polynomial.

A second fact is often decisive: the roots of μ_A are exactly the eigenvalues of A. In particular A is invertible iff 0 is not a root, i.e. iff the constant term of μ_A is nonzero.

(a) — TRUE

If $\mu_A = x^2 + 1$ then $A^2 = -I$, so

$$A^4 = (A^2)^2 = (-I)^2 = I. \checkmark$$

Equivalently, $x^2 + 1 \mid x^4 - 1$ since $x^4 - 1 = (x^2 - 1)(x^2 + 1)$. **True.**

(The eigenvalues are $\pm\mathrm{i}$, primitive fourth roots of unity, so 4 is exactly the order of A.)

(b) — FALSE

If $\mu_A = x^2 + x = x(x+1)$, then 0 is a root, so 0 is an eigenvalue and A is **singular**. A singular matrix satisfies $\det A = 0$, hence

$$\det(A^8) = (\det A)^8 = 0 \neq 1 = \det I,$$

so $A^8 \neq I$ for any such A. **False.**

Equivalently, $x^2 + x \nmid x^8 - 1$ because $x \mid x^2 + x$ but $x \nmid x^8 - 1$ (the constant term of $x^8 - 1$ is $-1 \neq 0$).

> Such an A is not periodic but *idempotent-like*: from $A^2 = -A$ we get $(-A)^2 = A^2 = -A$, so $P = -A$ satisfies $P^2 = P$. Thus $A = -P$ for a projection P, and $A^n = (-1)^n P$ for $n \geq 1$, which cycles between P and $-P$ but never reaches I.

(c) — TRUE

We ask whether $x^2 - x + 1$ divides $x^6 - 1$. Factor the latter into cyclotomics:

$$x^6 - 1 = (x-1)(x+1)(x^2+x+1)(x^2-x+1) = \Phi_1\Phi_2\Phi_3\Phi_6.$$

The factor $x^2 - x + 1 = \Phi_6(x)$ appears explicitly. **True.**

Direct verification. $\mu_A = x^2 - x + 1$ gives $A^2 = A - I$. Then

$$A^3 = A \cdot A^2 = A^2 - A = (A - I) - A = -I,$$

and hence

$$A^6 = (A^3)^2 = (-I)^2 = I. \checkmark$$

The eigenvalues are the primitive sixth roots of unity $\mathrm{e}^{\pm i\pi/3}$, so A has order exactly 6.

(d) — TRUE

Here $\mu_A = x^2 + x + 1 = \Phi_3(x)$, and

$$x^3 - 1 = (x-1)(x^2+x+1),$$

so $\mu_A \mid x^3 - 1$. **True.**

Direct verification. $A^2 = -A - I$, so

$$A^3 = A \cdot A^2 = -A^2 - A = -(-A - I) - A = A + I - A = I. \checkmark$$

The eigenvalues are the primitive cube roots of unity ω, ω^2, and A has order exactly 3.

> **The unifying view.** $A^m = I$ holds iff $\mu_A \mid x^m - 1$, iff μ_A is a product of distinct cyclotomic polynomials Φ_d with $d \mid m$. The requirement that the factors be *distinct* is the statement that A is diagonalisable, since $x^m - 1$ is squarefree over any field of characteristic prime to m. Reading the four options through this lens:
>
> $$x^2 + 1 = \Phi_4 \ (m=4), \quad x^2 - x + 1 = \Phi_6 \ (m=6), \quad x^2 + x + 1 = \Phi_3 \ (m=3),$$

all cyclotomic, hence all giving finite-order matrices; while $x^2+x = x\Phi_1\cdot$ is not cyclotomic and contains the factor x, killing invertibility outright. Recognising the cyclotomic polynomials on sight resolves this item in seconds.

Answer: **(a)**, **(c)** and **(d)**.

Question 100. On $\mathbb{Z}$ give the topology generated by the partition into $\{0\}$ and the pairs $\{-n, n\}$, $n \geq 1$. Then **(a)** $\mathbb{N} \cup \{0\}$ is closed **(b)** $\mathbb{N} \cup \{0\}$ is dense **(c)** every closed subset of $\mathbb{Z} \setminus \{0\}$ has even cardinality **(d)** there are infinitely many closed sets of cardinality 5.

The concept: partition topologies

The topology generated by a partition

Let $\mathcal{P}$ be a partition of a set X. Taking $\mathcal{P}$ as a basis generates a topology in which

$$U \text{ is open} \iff U \text{ is a union of blocks} \iff U \text{ is closed.}$$

The last equivalence holds because the complement of a union of blocks is again a union of blocks. So in a partition topology the open sets and the closed sets coincide — every set is clopen, and the space is as disconnected as its partition allows.
Consequently, for any $S \subseteq X$:

$$\bar{S} = \bigcup\{B \in \mathcal{P} : B \cap S \neq \emptyset\}, \qquad \operatorname{int} S = \bigcup\{B \in \mathcal{P} : B \subseteq S\}.$$

The closure *swells* S to the union of all blocks it touches; the interior *shrinks* it to the blocks it wholly contains.

Here $X = \mathbb{Z}$ and

$$\mathcal{P} = \big\{\{0\}\big\} \cup \big\{\{-n, n\} : n \geq 1\big\}.$$

Blocks are the singleton $\{0\}$ and the symmetric pairs.

(a) — FALSE

Let $S = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$. Is S a union of blocks? It contains $\{0\}$, a block. But it contains 1 without containing -1, so it does not contain the block $\{-1, 1\}$. Hence S is not a union of blocks, and is therefore neither open nor closed. **False.**

Explicitly,

$$\bar{S} = \{0\} \cup \bigcup_{n \geq 1} \{-n, n\} = \mathbb{Z} \neq S.$$

(b) — TRUE

From the computation just made, $\bar{S} = \mathbb{Z}$: every block meets S, since $\{0\}$ does and each pair $\{-n, n\}$ contains the positive element $n \in S$. Therefore S is **dense**. **True.**

The same argument shows $\mathbb{N}$ alone (excluding 0) has closure $\mathbb{Z} \setminus \{0\}$, not all of $\mathbb{Z}$ — so the inclusion of 0 genuinely matters for density.

(c) — TRUE

A subset $C \subseteq \mathbb{Z} \setminus \{0\}$ is closed in the subspace $\mathbb{Z} \setminus \{0\}$ iff it is a union of the blocks contained in $\mathbb{Z} \setminus \{0\}$, which are exactly the pairs $\{-n, n\}$ with $n \geq 1$. Every such block has cardinality 2, and the blocks are disjoint, so

$$|C| = 2 \cdot \#\{\text{blocks used}\},$$

which is even (or infinite, or zero). **True.**

Note the essential role of excluding 0: the block $\{0\}$ is the only odd-sized block, and removing it makes every remaining block even.

(d) — TRUE

A closed subset $C \subseteq \mathbb{Z}$ is a union of blocks. If $|C| = 5$, an odd number, then C must include the odd-sized block $\{0\}$, together with two pair-blocks:

$$C = \{0\} \cup \{-m, m\} \cup \{-n, n\}, \qquad 1 \leq m < n.$$

Then $|C| = 1 + 2 + 2 = 5$ ✓. There are infinitely many such choices — one for each pair $\{m, n\}$ with $m < n$, and there are countably infinitely many. **True.**

For instance $\{0, \pm 1, \pm 2\}$, $\{0, \pm 1, \pm 3\}$, $\{0, \pm 2, \pm 7\}$, and so on without end.

Separation properties of this space. It is not T_1: the point 1 is not closed, since the smallest closed set containing it is $\{-1, 1\}$. It is not even T_0, since 1 and -1 have exactly the same neighbourhoods and cannot be distinguished topologically. The quotient by the partition — collapsing each block to a point — is a countable discrete space, and the original is homeomorphic in the appropriate sense to that discrete space with points doubled. **Compactness.** $\mathbb{Z}$ with this topology is not compact: the blocks themselves form an open cover with no finite subcover.

Answer: **(b)**, **(c)** and **(d)**.

Question 101. Which of the following statements about functions of bounded variation on $[a, b]$ is FALSE? **(a)** every BV function is bounded **(b)** every BV function is a difference of two increasing functions **(c)** every BV function has at most countably many discontinuities **(d)** every BV function is continuous.

The definition

Bounded variation

For $f : [a, b] \to \mathbb{R}$ and a partition $P : a = x_0 < x_1 < \cdots < x_n = b$, set

$$V(f, P) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|, \qquad V_a^b(f) = \sup_P V(f, P).$$

f is of *bounded variation* if $V_a^b(f) < \infty$. The total variation measures the cumulative rise and fall of f — the distance travelled by a point moving along the graph's vertical axis.

(a) — TRUE

For any $x \in [a,b]$, use the partition $\{a, x, b\}$:

$$|f(x)| \le |f(a)| + |f(x) - f(a)| \le |f(a)| + V_a^b(f) < \infty.$$

So f is bounded by $|f(a)| + V_a^b(f)$. **True.**

(b) — TRUE (the Jordan decomposition)

Theorem 3.4 (Jordan)**.** f is of bounded variation on $[a,b]$ if and only if $f = g - h$ with g, h increasing.

Proof. ($\Leftarrow$) An increasing function has $V_a^b(g) = g(b) - g(a) < \infty$, and variation is subadditive: $V(f) \le V(g) + V(h)$.

($\Rightarrow$) Set $g(x) = V_a^x(f)$, the variation function, and $h = g - f$. Then g is increasing since variation is additive over adjacent intervals. And h is increasing: for $x < y$,

$$h(y) - h(x) = V_x^y(f) - \big(f(y) - f(x)\big) \;\ge\; V_x^y(f) - |f(y) - f(x)| \;\ge\; 0,$$

because the single-interval partition already contributes $|f(y) - f(x)|$ to $V_x^y(f)$. Then $f = g - h$. $\square$

True.

(c) — TRUE

An increasing function on $[a,b]$ has only jump discontinuities, and the jumps have positive size summing to at most $g(b) - g(a)$; hence for each k there are at most finitely many jumps of size $> 1/k$, and the total set of discontinuities is a countable union of finite sets, hence countable.

By the Jordan decomposition, a BV function is a difference of two such, so its discontinuity set is contained in the union of two countable sets — countable. **True.**

(d) — FALSE

This is the false statement the item seeks.

Counterexample. The Heaviside step

$$f(x) = \begin{cases} 0, & 0 \le x < \frac{1}{2}, \\ 1, & \frac{1}{2} \le x \le 1 \end{cases}$$

is increasing, hence of bounded variation with $V_0^1(f) = 1$, yet it is discontinuous at $x = \frac{1}{2}$. **False.**

The four-way independence worth fixing in mind.

- *BV does not imply continuous*: the step function above.
- *Continuous does not imply BV*: $f(x) = x\sin(1/x)$ on $(0,1]$, $f(0) = 0$, is continuous but has infinite variation, since the sum of $|f|$ at the extrema behaves like the harmonic series.
- *BV does imply Riemann integrable*: by **(a)** it is bounded and by **(c)** its discontinuity set is countable hence null, so Lebesgue's criterion applies.

- *Continuous plus BV does not imply absolutely continuous*: the Cantor function is continuous and increasing (so BV, with variation 1) yet has derivative 0 almost everywhere, so it is not the integral of its derivative.

The correct hierarchy on $[a, b]$ is

$$C^1 \subset \text{Lipschitz} \subset AC \subset BV \cap C \subset BV \subset \text{Riemann integrable},$$

each inclusion strict.

Answer: **(d)** is the false statement.

Question 102. For the equation $y'' - 4xy' + (4x^2 - 1)y = 0$ (Wronskian-type item), which statements hold? **(a)** $\mathrm{e}^{-x^2} y(x)$ is periodic for every solution **(b)** every solution is bounded **(c)** the solutions are unbounded **(d)** the equation has a polynomial solution.

Reducing the equation

The coefficients suggest a perfect-square structure. Try the substitution

$$y(x) = \mathrm{e}^{x^2} v(x).$$

Then

$$y' = \mathrm{e}^{x^2}\big(v' + 2xv\big), \qquad y'' = \mathrm{e}^{x^2}\big(v'' + 4xv' + (4x^2 + 2)v\big).$$

Substituting into the equation and dividing by $\mathrm{e}^{x^2} \neq 0$:

$$\big[v'' + 4xv' + (4x^2 + 2)v\big] - 4x\big[v' + 2xv\big] + \big(4x^2 - 1\big)v = 0.$$

Collect terms:

$$v'' + \underbrace{(4x - 4x)}_{=0} v' + \underbrace{\big(4x^2 + 2 - 8x^2 + 4x^2 - 1\big)}_{=1} v = 0,$$

that is,

$$\boxed{v'' + v = 0}$$

— the harmonic oscillator. Hence

$$v(x) = A\cos x + B\sin x, \qquad y(x) = \mathrm{e}^{x^2}\big(A\cos x + B\sin x\big),$$

and a fundamental pair is

$$y_1 = \mathrm{e}^{x^2}\cos x, \qquad y_2 = \mathrm{e}^{x^2}\sin x.$$

Why the substitution works. For $y'' + py' + qy = 0$, putting $y = u\exp\big(-\frac{1}{2}\int p\big)$ removes the first-derivative term and leaves $u'' + Qu = 0$ with $Q = q - \frac{p^2}{4} - \frac{p'}{2}$. Here $p = -4x$, $q = 4x^2 - 1$, and

$$Q = (4x^2 - 1) - \frac{16x^2}{4} - \frac{-4}{2} = 4x^2 - 1 - 4x^2 + 2 = 1.$$

The exponential factor is $\exp\big(-\frac{1}{2}\int(-4x)\,dx\big) = e^{x^2}$, exactly the substitution used. The constancy of Q is what makes this equation solvable in closed form.

(a) — TRUE

For any solution, $e^{-x^2}y(x) = A\cos x + B\sin x$, which can be written as $R\cos(x-\phi)$ with $R = \sqrt{A^2+B^2}$. This is periodic with period 2π (for $R \neq 0$; the zero solution is trivially periodic). **True.**

(c) — TRUE, and (b) — FALSE

Consider $y = e^{x^2}R\cos(x-\phi)$ with $R > 0$. At the points $x_k = \phi + 2k\pi$ we have $\cos(x_k - \phi) = 1$, so

$$|y(x_k)| = R\,e^{x_k^2} = R\,e^{(\phi+2k\pi)^2} \longrightarrow \infty \quad \text{as } k \to \infty.$$

So every non-trivial solution is unbounded: the trigonometric factor oscillates between $\pm R$ while the envelope e^{x^2} grows super-exponentially. **(c) true, (b) false.**

Note that the solutions also have infinitely many zeros — at $x = \phi + \frac{\pi}{2} + k\pi$ — so they oscillate with unboundedly growing amplitude. This is consistent with the normal form $v'' + v = 0$ having $Q = 1 > 0$: Sturm's criterion guarantees oscillation with zeros spaced exactly π apart, and the spacing is unaffected by the exponential factor.

(d) — FALSE

Suppose $y = P(x)$ were a non-zero polynomial solution. Then $e^{-x^2}P(x) = A\cos x + B\sin x$, so

$$P(x) = e^{x^2}\big(A\cos x + B\sin x\big).$$

The right side is not a polynomial for any $(A, B) \neq (0, 0)$: it has infinitely many zeros (at $x = \phi + \frac{\pi}{2} + k\pi$ for every integer k), whereas a non-zero polynomial has finitely many. Hence no non-trivial polynomial solution exists. **False.**

The Wronskian, as a check. Abel's formula gives

$$W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} p(s)\,ds\right) = W(x_0)\,e^{2x^2-2x_0^2}.$$

Direct computation from our pair:

$$\begin{aligned} W = y_1y_2' - y_2y_1' &= e^{x^2}\cos x\cdot e^{x^2}(\cos x + 2x\sin x) - e^{x^2}\sin x\cdot e^{x^2}(-\sin x + 2x\cos x)\\ &= e^{2x^2}\big[\cos^2 x + 2x\sin x\cos x + \sin^2 x - 2x\sin x\cos x\big] = e^{2x^2}. \end{aligned}$$

This matches Abel's formula with $W(0) = 1$, and being everywhere nonzero confirms that y_1, y_2 are linearly independent. ✓

Answer: **(a)** and **(c)**.

Question 103. Let m be the number of ring homomorphisms $\mathbb{Z}_{12} \to \mathbb{Z}_{20}$ and n the number of ring homomorphisms $\mathbb{Z} \to \mathbb{Z}$. Then **(a)** $m = 2$ **(b)** $m = 4$ **(c)** $n = 1$ **(d)** $n = 2$.

A convention to settle first. Some texts require ring homomorphisms to send $1 \mapsto 1$ (*unital*); others do not. The counts differ, and the option set here ($m = 2$, $n = 2$) is consistent with the *non-unital* convention, which is the one used below. Under the unital convention one would get $m = 0$ (since no unital map $\mathbb{Z}_{12} \to \mathbb{Z}_{20}$ exists, as $12 \nmid$ the relevant order) and $n = 1$. The item's inclusion of $n = 2$ as an option indicates the non-unital reading.

Counting homomorphisms $\mathbb{Z}_{12} \to \mathbb{Z}_{20}$

Homomorphisms out of $\mathbb{Z}_k$

A ring homomorphism $f : \mathbb{Z}_k \to R$ is determined by $e = f(1)$, since $f(a) = f(a \cdot 1) = a\,e$ for every a. Two conditions constrain e:
1. Multiplicativity. $e = f(1) = f(1 \cdot 1) = f(1)f(1) = e^2$, so e must be an *idempotent* of R.
2. Additivity / well-definedness. Since $k \cdot 1 = 0$ in $\mathbb{Z}_k$, we need $k\,e = 0$ in R.
Conversely, any e satisfying both defines a homomorphism. So

$$\#\,\mathrm{Hom}(\mathbb{Z}_k, R) = \#\{e \in R : e^2 = e \text{ and } ke = 0\}.$$

Step 1: find the idempotents of $\mathbb{Z}_{20}$. By the Chinese Remainder Theorem, $\mathbb{Z}_{20} \cong \mathbb{Z}_4 \times \mathbb{Z}_5$. Since $\mathbb{Z}_4$ has idempotents $\{0, 1\}$ (as $2^2 = 0$, $3^2 = 1$) and $\mathbb{Z}_5$ is a field with idempotents $\{0, 1\}$, the product has $2 \times 2 = 4$ idempotents. Translating back:

$$(0,0) \mapsto 0, \quad (1,1) \mapsto 1, \quad (0,1) \mapsto 16, \quad (1,0) \mapsto 5.$$

Check: $16^2 = 256 = 12 \cdot 20 + 16 \equiv 16$ ✓; $5^2 = 25 \equiv 5$ ✓.

Step 2: impose $12e \equiv 0 \pmod{20}$. Test each idempotent:

e	$12e$ mod 20	admissible?
0	0	yes
1	12	no
5	$60 \equiv 0$	yes
16	$192 = 9 \cdot 20 + 12 \equiv 12$	no

So exactly two idempotents survive: $e = 0$ and $e = 5$. Therefore

$$m = 2.$$

(a) TRUE, **(b) FALSE**.

The two maps are the zero map and $f(a) = 5a$ mod 20. The latter is a genuine homomorphism: $f(a)f(b) = 25ab \equiv 5ab = f(ab)$ since $25 \equiv 5 \pmod{20}$.

Counting homomorphisms $\mathbb{Z} \to \mathbb{Z}$

Again f is determined by $e = f(1)$, which must satisfy $e^2 = e$. In $\mathbb{Z}$ this gives $e(e-1) = 0$, so $e = 0$ or $e = 1$; there is no condition of the second type since $\mathbb{Z}$ has no torsion. Hence

$$n = 2,$$

the two maps being the zero map and the identity. **(d) TRUE**, **(c) FALSE**.

The general count. For $\mathrm{Hom}(\mathbb{Z}_k, \mathbb{Z}_\ell)$ under the non-unital convention, one finds

$$\#\,\mathrm{Hom}(\mathbb{Z}_k, \mathbb{Z}_\ell) = \#\{d \text{ idempotent mod } \ell \text{ with } \ell \mid kd\},$$

and the number of idempotents in $\mathbb{Z}_\ell$ is $2^{\omega(\ell)}$, where $\omega(\ell)$ counts distinct prime factors. Here $\ell = 20 = 2^2 \cdot 5$ has $\omega = 2$, giving 4 idempotents — confirming Step 1 without enumeration.
The contrast with group homomorphisms. $\#\,\mathrm{Hom}_{\mathrm{grp}}(\mathbb{Z}_{12}, \mathbb{Z}_{20}) = \gcd(12, 20) = 4$, since only the additive condition $12e \equiv 0$ applies. The idempotency requirement of ring homomorphisms cuts this from 4 down to 2 — the multiplicative structure is genuinely restrictive.

Answer: **(a)** and **(d)**.

Question 104. For the joint density $f(x, y) = x(y - x)e^{-y}$ on $0 \le x \le y < \infty$, **(a)** $X \sim \mathrm{Gamma}(2, 1)$ **(b)** $Y - X$ has the stated density **(c)** $\mathbb{E}[Y \mid X] = X + 1$ **(d)** $X/Y \sim \mathrm{Beta}(2, 2)$.

The source text prints the density with a factor e^{-xy}, which does not integrate to 1 over the region and is not a probability density. The intended density is

$$f(x, y) = x\,(y - x)\,e^{-y}, \qquad 0 \le x \le y < \infty,$$

which does. All computations below use this corrected form; the check is given immediately.

Verifying normalisation — and the change of variables that makes everything easy

Substitute

$$u = x, \qquad v = y - x,$$

so that $x = u$, $y = u + v$, the Jacobian is 1, and the region $0 \le x \le y$ becomes the quadrant $u, v \ge 0$. Then $e^{-y} = e^{-(u+v)}$ and

$$f(x, y)\,dx\,dy = u\,v\,e^{-u}e^{-v}\,du\,dv = \underbrace{\left(ue^{-u}\right)}_{\mathrm{Gamma}(2,1)}\ \underbrace{\left(ve^{-v}\right)}_{\mathrm{Gamma}(2,1)}\ du\,dv.$$

The structural discovery

The joint density *factorises* in the variables $(U, V) = (X, Y - X)$:

$$U = X \text{ and } V = Y - X \text{ are } \textbf{independent}, \text{ each } \sim \mathrm{Gamma}(2, 1).$$

Every option now follows from this single observation. In particular $\int\!\int f = 1 \times 1 = 1$ ✓, confirming the corrected density.

(a) — TRUE

$X = U \sim \mathrm{Gamma}(2, 1)$, with density xe^{-x} on $x > 0$. **True.**

Direct check. Integrating y out:

$$f_X(x) = \int_x^\infty x(y-x)\mathrm{e}^{-y}\,\mathrm{d}y \overset{v=y-x}{=} x\mathrm{e}^{-x}\int_0^\infty v\mathrm{e}^{-v}\,\mathrm{d}v = x\mathrm{e}^{-x}\cdot\Gamma(2) = x\mathrm{e}^{-x}.\ \checkmark$$

(b) — TRUE (under the natural reading)

$Y - X = V \sim \mathrm{Gamma}(2,1)$, with density $(y-x)\mathrm{e}^{-(y-x)}$ in the obvious notation — i.e. $v\mathrm{e}^{-v}$ on $v > 0$. If option **(b)** states this density, it is **true**. Note that X and $Y - X$ are identically distributed, a pleasing symmetry of this model.

(c) — FALSE

By independence of X and $V = Y - X$,

$$\mathbb{E}[Y \mid X] = \mathbb{E}[X + V \mid X] = X + \mathbb{E}[V] = X + 2,$$

since a $\mathrm{Gamma}(2,1)$ variable has mean 2. The option asserts $X + 1$. **False.**

> The value 1 would be correct if V were $\mathrm{Exponential}(1)$, i.e. $\mathrm{Gamma}(1,1)$. The shape parameter 2 — visible as the factor $(y-x)$ in the density — doubles the conditional increment. This is exactly the sort of distractor that rewards computing $\mathbb{E}[V]$ rather than assuming exponentiality.

(d) — TRUE

Theorem 3.5 (The beta–gamma algebra)**.** If $A \sim \mathrm{Gamma}(\alpha,\theta)$ and $B \sim \mathrm{Gamma}(\beta,\theta)$ are independent (same scale θ), then

$$\frac{A}{A+B} \sim \mathrm{Beta}(\alpha,\beta),$$

and moreover $\dfrac{A}{A+B}$ is *independent* of $A + B \sim \mathrm{Gamma}(\alpha+\beta,\theta)$.

Apply with $A = U = X$ and $B = V = Y - X$, both $\mathrm{Gamma}(2,1)$ and independent. Since $A + B = X + (Y - X) = Y$,

$$\frac{X}{Y} = \frac{A}{A+B} \sim \mathrm{Beta}(2,2).$$

True.

> **Consequences worth recording.**
>
> $$Y = U + V \sim \mathrm{Gamma}(4,1), \qquad \mathbb{E}Y = 4, \quad \mathrm{Var}\,Y = 4.$$
>
> $$\mathbb{E}\left[\frac{X}{Y}\right] = \frac{2}{2+2} = \frac{1}{2}, \qquad \mathrm{Var}\left(\frac{X}{Y}\right) = \frac{2\cdot 2}{4^2\cdot 5} = \frac{1}{20}.$$
>
> The $\mathrm{Beta}(2,2)$ density is $6t(1-t)$ on $[0,1]$ — symmetric about $\frac{1}{2}$, as it must be, since X and $Y - X$ are exchangeable.
> Also $\mathbb{E}[X \mid Y] = Y\,\mathbb{E}[X/Y] = Y/2$ by the independence of the ratio from the sum — a clean answer that follows from the beta–gamma theorem and would be painful to obtain by direct integration.

Answer: **(a)** and **(d)** certainly; **(b)** also, under the reading that $Y - X$ has density $(y-x)\mathrm{e}^{-(y-x)}$. **(c)** is false: $\mathbb{E}[Y \mid X] = X + 2$.

Question 105. For a linear operator T on a finite-dimensional space with the given nilpotency structure, **(a)** $\operatorname{rank} T > \operatorname{rank} T^2$ **(b)** the stated numerical value holds **(c)** T is not diagonalisable **(d)** $\operatorname{rank} T^2 = 0$.

The concept: the rank sequence of a nilpotent operator

Strictly decreasing then constant

For any operator T on a finite-dimensional space,

$$\ker T \subseteq \ker T^2 \subseteq \ker T^3 \subseteq \cdots, \qquad \operatorname{im} T \supseteq \operatorname{im} T^2 \supseteq \operatorname{im} T^3 \supseteq \cdots$$

and the sequences are *strictly* monotone until they stabilise, after which they are constant. Precisely: if $\ker T^k = \ker T^{k+1}$ for some k, then $\ker T^j = \ker T^k$ for all $j \geq k$. Consequently the rank sequence

$$\operatorname{rank} T > \operatorname{rank} T^2 > \cdots > \operatorname{rank} T^p = \operatorname{rank} T^{p+1} = \cdots$$

drops strictly at each step until stabilisation. For a *nilpotent* operator of index p (meaning $T^p = 0$ but $T^{p-1} \neq 0$), the sequence reaches 0 exactly at step p.

(a) — TRUE for any non-zero nilpotent T

If T is nilpotent and $T \neq 0$, its index p satisfies $p \geq 2$, so $\operatorname{rank} T^{p-1} > 0$ while $\operatorname{rank} T^p = 0$; and by strict decrease,

$$\operatorname{rank} T > \operatorname{rank} T^2.$$

Equality $\operatorname{rank} T = \operatorname{rank} T^2$ would force stabilisation at step 1, hence $\operatorname{rank} T = \operatorname{rank} T^j$ for all j, hence $\operatorname{rank} T = 0$ and $T = 0$. **True** for any nonzero nilpotent operator.

(c) — TRUE for any non-zero nilpotent T

A nilpotent operator has all eigenvalues 0: if $Tv = \lambda v$ with $v \neq 0$, then $0 = T^p v = \lambda^p v$, so $\lambda = 0$. Were T diagonalisable, it would be similar to the zero matrix, hence equal to 0. Since $T \neq 0$, it is **not diagonalisable**. **True.**

Equivalently, the minimal polynomial is x^p with $p \geq 2$, which is not squarefree — the standard criterion for non-diagonalisability.

(d) — FALSE in general

The claim $\operatorname{rank} T^2 = 0$ says $T^2 = 0$, i.e. the index is at most 2. This is a genuine restriction, and it fails whenever the index exceeds 2.

Counterexample. On $\mathbb{R}^3$ take the single Jordan block

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad N^3 = 0.$$

Here rank $N = 2$, rank $N^2 = 1 \neq 0$, rank $N^3 = 0$. **False.**

Reading the rank sequence off the Jordan structure. If T is nilpotent with Jordan blocks of sizes $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ (a partition of n), then

$$\operatorname{rank} T^j = \sum_i \max\{\lambda_i - j,\ 0\}, \qquad \dim\ker T^j = \sum_i \min\{\lambda_i,\ j\}.$$

The index of nilpotency is λ_1, the largest block size, and the number of blocks is $s = \dim\ker T = n - \operatorname{rank} T$.
For the example above: partition (3), so rank $N = 3-1 = 2$, rank $N^2 = 3-2 = 1$, rank $N^3 = 0$ ✓. Contrast the partition $(2, 1)$ on $\mathbb{R}^3$: then rank $T = 1$ and rank $T^2 = 0$, so option **(d)** would hold. The truth of **(d)** depends entirely on the partition, which is what the missing data specified.

Answer: **(a)** and **(c)** hold for every nonzero nilpotent operator. **(d)** is false in general (it requires index ≤ 2). **(b)** cannot be adjudicated: the numerical data are absent.

Question 106. Let V be the space of entire functions satisfying $f(\omega z) = f(z)$ where $\omega = e^{2\pi i/m}$. Then **(a)** V contains all polynomials **(b)** V is finite-dimensional **(c)** $z^n \in V$ iff $m \mid n$ **(d)** $f \in V$ iff $f(z) = \frac{1}{m}\sum_{k=0}^{m-1} h(\omega^k z)$ for some entire h.

The structure theorem

Theorem 3.6. An entire function f satisfies $f(\omega z) = f(z)$ for $\omega = e^{2\pi i/m}$ if and only if

$$f(z) = g(z^m)$$

for some entire function g.

Proof. ($\Leftarrow$) $g\big((\omega z)^m\big) = g(\omega^m z^m) = g(z^m)$ since $\omega^m = 1$.
($\Rightarrow$) Expand $f(z) = \sum_{n\geq 0} a_n z^n$, convergent on all of $\mathbb{C}$. Then

$$f(\omega z) = \sum_{n\geq 0} a_n \omega^n z^n,$$

and equality of power series forces $a_n\omega^n = a_n$ for every n, i.e.

$$a_n\big(\omega^n - 1\big) = 0.$$

Now $\omega^n = 1$ precisely when $m \mid n$. Hence $a_n = 0$ whenever $m \nmid n$, and

$$f(z) = \sum_{j\geq 0} a_{jm} z^{jm} = g(z^m), \qquad g(w) = \sum_{j\geq 0} a_{jm} w^j,$$

with g entire since its coefficients are a subsequence of a coefficient sequence with infinite radius of convergence. ∎

(c) — TRUE

Immediate from the proof: $z^n \in V \iff \omega^n = 1 \iff m \mid n$. **True.** (For $m \geq 2$ this already shows $z \notin V$.)

(a) — FALSE

By **(c)**, the polynomial z does not lie in V for $m \geq 2$, since $m \nmid 1$. More generally the only polynomials in V are those whose exponents are all multiples of m. **False.**

(For $m = 1$ the condition is vacuous and V is everything, but $\omega = 1$ makes the item trivial; the intended case is $m \geq 2$.)

(b) — FALSE

The functions $1, z^m, z^{2m}, z^{3m}, \ldots$ all lie in V and are linearly independent, being monomials of distinct degrees. So V contains an infinite independent set and is **infinite-dimensional**. **False.**

Indeed $V \cong \{\text{entire functions}\}$ via $g \mapsto g(z^m)$, an isomorphism of vector spaces, so V is "as large as" the whole space of entire functions.

(d) — TRUE

Define the *averaging* (Reynolds) operator

$$(\mathcal{A}h)(z) = \frac{1}{m}\sum_{k=0}^{m-1} h\big(\omega^k z\big).$$

$\mathcal{A}h \in V$ always. Replace z by ωz:

$$(\mathcal{A}h)(\omega z) = \frac{1}{m}\sum_{k=0}^{m-1} h\big(\omega^{k+1} z\big) = \frac{1}{m}\sum_{j=1}^{m} h\big(\omega^j z\big) = (\mathcal{A}h)(z),$$

the last step because $\omega^m = \omega^0 = 1$, so the sum is over the same set of terms, merely cyclically permuted. And $\mathcal{A}h$ is entire, being a finite sum of entire functions.

Every $f \in V$ arises this way. Take $h = f$. Then each term of the average equals $f(z)$ by the invariance, so $\mathcal{A}f = f$.

Hence $V = \operatorname{im}\mathcal{A}$ exactly, which is the assertion. **True.**

$\mathcal{A}$ is a projection onto V. We have just shown $\mathcal{A}^2 = \mathcal{A}$ (apply $\mathcal{A}$ to something already in V and nothing changes) and $\operatorname{im}\mathcal{A} = V$. This is the analytic shadow of a general principle from representation theory: averaging over a finite group G acting on a vector space produces the projection onto the subspace of G-invariants,

$$\mathcal{A} = \frac{1}{|G|}\sum_{g\in G} g,$$

valid whenever $|G|$ is invertible in the ground field. Here $G = \mathbb{Z}/m\mathbb{Z}$ acts by rotation through ω.

Coefficient interpretation. In terms of Taylor coefficients, $\mathcal{A}$ kills every a_n with $m \nmid n$ and fixes the rest, since $\frac{1}{m}\sum_k \omega^{kn} = \mathbf{1}_{m\mid n}$ — the orthogonality relation for characters of the cyclic group, and the discrete Fourier transform in its simplest guise.

Answer: **(c)** and **(d)**.

Question 107. For a matrix with the stated characteristic and minimal polynomials, which conclusions about its Jordan structure hold?

The polynomials did not reproduce

What each polynomial tells you

Let A be $n \times n$ over $\mathbb{C}$, with characteristic polynomial χ_A and minimal polynomial μ_A.
Shared roots. χ_A and μ_A have the same roots (the eigenvalues), differing only in multiplicities. Cayley–Hamilton gives $\mu_A \mid \chi_A$.
From χ_A: the multiplicity of λ in χ_A is the *algebraic* multiplicity, which equals the **total size** of all Jordan blocks for λ.
From μ_A: the multiplicity of λ in μ_A equals the **largest** Jordan block size for λ.
From ranks: the number of Jordan blocks for λ is the *geometric* multiplicity $\dim\ker(A - \lambda I) = n - \operatorname{rank}(A - \lambda I)$. More finely,

$$\#\{\text{blocks for } \lambda \text{ of size} \geq j\} = \operatorname{rank}(A - \lambda I)^{j-1} - \operatorname{rank}(A - \lambda I)^j.$$

When the two polynomials determine the Jordan form completely. For each eigenvalue, χ_A gives the total size a_λ and μ_A gives the largest block p_λ. These determine the partition of a_λ uniquely precisely when the partition is forced — for instance when $a_\lambda \leq 3$, or when $p_\lambda = a_\lambda$ (a single block), or when $p_\lambda = 1$ (all blocks of size 1, i.e. diagonalisable at λ). The smallest ambiguous case is $a_\lambda = 4$ with $p_\lambda = 2$: the partition could be $(2, 2)$ or $(2, 1, 1)$, and the two are distinguished only by the rank of $(A - \lambda I)$. This is the standard example showing that χ_A and μ_A together do not always suffice.
Diagonalisability has the clean criterion: A is diagonalisable over $\mathbb{C}$ iff μ_A is squarefree, i.e. has no repeated roots. Over a general field F, add the requirement that μ_A split into linear factors over F.

Answer: Indeterminate: the characteristic and minimal polynomials are absent from the available text.

Question 108. For the linear transformation T defined by the given values on a basis, which of the stated properties hold?

The defining values of T did not reproduce

A transformation is determined by its values on a basis

If $\{v_1, \dots, v_n\}$ is a basis of V and $w_1, \dots, w_n$ are arbitrary vectors in W, there is exactly one linear map $T : V \to W$ with $Tv_i = w_i$, namely

$$T\left(\sum_i c_i v_i\right) = \sum_i c_i w_i.$$

Existence uses that every vector has a representation; uniqueness uses that the representation is unique. Consequently:

$$\operatorname{im} T = \operatorname{span}\{w_1, \dots, w_n\}, \qquad \operatorname{rank} T = \dim \operatorname{span}\{w_i\},$$

and by rank–nullity $\dim \ker T = n - \operatorname{rank} T$.
If the specified values are given on a *spanning but dependent* set, or on a *linearly independent but non-spanning* set, the situation changes: in the first case T may fail to exist (the prescribed values must respect every dependency among the v_i); in the second, T exists but is not unique.

The consistency test that often decides such items. If the stem prescribes T on vectors satisfying a linear relation $\sum \alpha_i v_i = 0$, then linearity forces $\sum \alpha_i w_i = 0$. When the prescribed w_i violate this, *no such T exists* — and options asserting properties of T are all vacuous or false. Checking for dependencies among the listed v_i before computing anything is therefore the first move.
Standard follow-up questions and how each is answered:

- *Is T injective?* Iff $\{w_i\}$ is independent, iff $\operatorname{rank} T = n$.
- *Is T surjective?* Iff $\operatorname{span}\{w_i\} = W$.
- *Eigenvalues?* Assemble the matrix $[T]$ in the given basis and compute $\det([T] - \lambda I)$.
- *Is T nilpotent?* Iff all eigenvalues are 0, iff $[T]^n = 0$.
- *Trace and determinant?* Basis-independent, computable from $[T]$ in any basis.

Answer: Indeterminate: the values defining T are absent from the available text.

Question 109. For the first-order partial differential equation with the given initial curves, which solutions or existence statements hold?

The equation and initial curves did not reproduce

The method of characteristics

For the quasi-linear equation

$$a(x, y, u)\, u_x + b(x, y, u)\, u_y = c(x, y, u),$$

the characteristic system is
$$\frac{\mathrm{d}x}{a} = \frac{\mathrm{d}y}{b} = \frac{\mathrm{d}u}{c} \quad (= \mathrm{d}t).$$
Solving gives two independent first integrals $\phi(x, y, u) = c_1$ and $\psi(x, y, u) = c_2$, and the general solution is $F(\phi, \psi) = 0$ for arbitrary F, or equivalently $\phi = G(\psi)$.
Lagrange multipliers. When the ratios are not directly integrable, use

$$\frac{\mathrm{d}x}{a} = \frac{\mathrm{d}y}{b} = \frac{\mathrm{d}u}{c} = \frac{\ell\ \mathrm{d}x + m\ \mathrm{d}y + n\ \mathrm{d}u}{\ell a + mb + nc}$$

and choose (ℓ, m, n) making either the numerator an exact differential or the denominator zero.

The transversality condition

Given initial data prescribed along a curve Γ parametrised by s as $\big(x_0(s), y_0(s), u_0(s)\big)$, the Cauchy problem has a unique local solution near s_0 provided

$$\det\begin{pmatrix} x_0'(s) & y_0'(s) \\ a & b \end{pmatrix} = x_0'(s)\, b - y_0'(s)\, a \ \neq\ 0,$$

i.e. the initial curve is *not* tangent to the characteristic direction.
When the determinant vanishes identically, Γ is itself a characteristic curve, and then:

- if the data are *compatible* with the characteristic ODE, there are **infinitely many** solutions — any surface containing that characteristic will do;
- if the data are *incompatible*, there is **no** solution.

This trichotomy — unique, none, or infinitely many — is the standard subject of multiple-select items of this type, and it is worth checking the transversality determinant before attempting to solve anything.

Two initial curves, two verdicts. Items in this family typically give one curve transverse to the characteristics (unique solution) and one lying along a characteristic (no solution, or infinitely many). Recognising which is which resolves the item without integrating.

Answer: Indeterminate: the partial differential equation and the initial curves are absent from the available text.

Question 110. Let (X_1, X_2) be standard bivariate normal with correlation $\rho = 0$. Then **(a)** $\mathrm{Cov}(X_1 + X_2,\ X_1 - X_2) = 0$ **(b)** $X_1 + X_2$ and $X_1 - X_2$ are independent **(c)** $\mathbb{E}[X_1^3 X_2] = 0$ **(d)** $\mathbb{E}[X_1 X_2 \mid X_1 + X_2] = 0$.

Throughout, $X_1, X_2 \sim N(0, 1)$ independently (uncorrelated jointly normal variables are independent).

(a) — TRUE

By bilinearity of covariance,

$$\operatorname{Cov}(X_1 + X_2,\ X_1 - X_2) = \operatorname{Var} X_1 - \operatorname{Cov}(X_1, X_2) + \operatorname{Cov}(X_2, X_1) - \operatorname{Var} X_2 = \operatorname{Var} X_1 - \operatorname{Var} X_2.$$

The cross terms cancel identically — note this requires no assumption on ρ at all. Since both variances equal 1,

$$\operatorname{Cov}(X_1 + X_2, X_1 - X_2) = 1 - 1 = 0.$$

True.

The identity $\operatorname{Cov}(X + Y, X - Y) = \operatorname{Var} X - \operatorname{Var} Y$ is worth memorising: sum and difference are uncorrelated *exactly when the two variances are equal*, regardless of their correlation. This underlies the paired t-test and the rotation to principal axes for equicorrelated data (Question 75).

(b) — TRUE

Let $S = X_1 + X_2$ and $D = X_1 - X_2$. The vector $(S, D)^\intercal = M(X_1, X_2)^\intercal$ with

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

a linear map. A linear image of a multivariate normal vector is multivariate normal, so (S, D) is bivariate normal. For *jointly* normal variables,

$$\text{uncorrelated} \iff \text{independent},$$

and **(a)** established uncorrelatedness. Hence S and D are independent. **True.**

Explicitly $S \sim N(0, 2)$ and $D \sim N(0, 2)$, independent.

The implication "uncorrelated $\Rightarrow$ independent" requires *joint* normality, not merely normal marginals. The standard counterexample: let $X \sim N(0, 1)$ and $Y = \varepsilon X$ with $\varepsilon = \pm 1$ a fair coin independent of X. Then $Y \sim N(0, 1)$ and $\operatorname{Cov}(X, Y) = 0$, yet $|X| = |Y|$ always — about as dependent as two variables can be. The pair (X, Y) is not bivariate normal.

(c) — TRUE

With $\rho = 0$ the variables are independent, so

$$\mathbb{E}[X_1^3 X_2] = \mathbb{E}[X_1^3]\,\mathbb{E}[X_2] = 0 \cdot 0 = 0.$$

True. (Both factors vanish: odd moments of a centred normal are zero.)

For general ρ the answer is $\mathbb{E}[X_1^3 X_2] = 3\rho$, obtainable from Isserlis' theorem (Wick's formula): for centred jointly Gaussian variables,

$$\mathbb{E}[Y_1 Y_2 Y_3 Y_4] = \mathbb{E}[Y_1 Y_2]\mathbb{E}[Y_3 Y_4] + \mathbb{E}[Y_1 Y_3]\mathbb{E}[Y_2 Y_4] + \mathbb{E}[Y_1 Y_4]\mathbb{E}[Y_2 Y_3].$$

With $Y_1 = Y_2 = Y_3 = X_1$ and $Y_4 = X_2$ this gives $1 \cdot \rho + 1 \cdot \rho + 1 \cdot \rho = 3\rho$, which is 0 at $\rho = 0$. ✓

(d) — FALSE

This is the discriminating option. Express the product in terms of S and D:

$$S^2 - D^2 = (X_1 + X_2)^2 - (X_1 - X_2)^2 = 4X_1X_2 \qquad \Longrightarrow \qquad X_1X_2 = \frac{S^2 - D^2}{4}.$$

Now condition on S. Since D is independent of S with $D \sim N(0,2)$, we have $\mathbb{E}[D^2 \mid S] = \mathbb{E}[D^2] = 2$, and S^2 is a function of S. Hence

$$\mathbb{E}[X_1X_2 \mid S] = \frac{S^2 - \mathbb{E}[D^2 \mid S]}{4} = \frac{S^2 - 2}{4},$$

which is a non-constant function of S, not identically zero. **False.**

Sanity check by the tower property. $\mathbb{E}[S^2] = \operatorname{Var} S = 2$, so

$$\mathbb{E}\big[\mathbb{E}[X_1X_2 \mid S]\big] = \frac{\mathbb{E}[S^2] - 2}{4} = \frac{2-2}{4} = 0 = \mathbb{E}[X_1X_2]. \checkmark$$

The conditional expectation averages to zero, as it must, but is not zero pointwise — knowing S is large in absolute value makes a large positive product likely.

The distinction the item is built on. Options **(c)** and **(d)** both concern the product X_1X_2, but **(c)** is an unconditional moment (zero by symmetry) and **(d)** is a conditional expectation (non-zero, because conditioning on S destroys the symmetry that made the unconditional moment vanish). Independence of X_1 and X_2 does *not* imply that X_1X_2 is independent of $X_1 + X_2$ — indeed, given S, the product is determined up to the value of D^2.

Answer: **(a)**, **(b)** and **(c)**.

Question 111. Regarding the three statements S_1, S_2, S_3, which are correct?

The statements did not reproduce

The source text for this item lost the content of S_1, S_2 and S_3, so no adjudication is possible. The structural remark below is nonetheless worth making, since it applies to every item of this shape.

Items of the form "which of S_1, S_2, S_3 are correct" with options such as "S_1 and S_2 only", "S_2 and S_3 only", "all three", "S_1 only" are *single-answer items in disguise*, since the four options are mutually exclusive descriptions of the same underlying truth-assignment. In such cases:

- exactly one option can be correct, so the usual Part C caution about multiple keys does not apply;
- establishing the truth value of a *single* statement often eliminates two or three options at once — so attack whichever of S_1, S_2, S_3 is easiest, not whichever comes first;
- a statement appearing in three of the four options is more likely true than false, purely as a matter of how such option sets are usually constructed. This is a heuristic of last resort, not an argument.

If instead the four options are the statements themselves, the item is an ordinary multiple-select and each must be adjudicated separately.

Answer: Indeterminate: the statements S_1, S_2, S_3 are absent from the available text.

Question 112. Which of the following algebraic numbers has degree 4 over $\mathbb{Q}$? **(a)** $\sqrt{2}+\sqrt{3}$ **(b)** $\sqrt{3+2\sqrt{2}}$ **(c)** $\sqrt[4]{2}$ **(d)** $\mathrm{i}+\sqrt{2}$.

(a) — TRUE, degree 4

Let $\alpha = \sqrt{2}+\sqrt{3}$. Then

$$\alpha^2 = 5 + 2\sqrt{6} \quad\Longrightarrow\quad \alpha^2 - 5 = 2\sqrt{6} \quad\Longrightarrow\quad (\alpha^2-5)^2 = 24,$$

so α satisfies

$$x^4 - 10x^2 + 1 = 0.$$

Irreducibility. The quartic has no rational roots (by the rational root theorem, only ± 1 are candidates, and $1 - 10 + 1 = -8 \neq 0$). Nor does it factor into two rational quadratics: the four roots are $\pm\sqrt{2}\pm\sqrt{3}$, and any monic quadratic factor with rational coefficients would have to pair two of these with rational sum and product, which one checks fails in every case.

Cleaner argument via the tower law. $\mathbb{Q}(\sqrt{2},\sqrt{3})$ has degree 4 over $\mathbb{Q}$, since $[\mathbb{Q}(\sqrt{2}):\mathbb{Q}] = 2$ and $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$ (else $\sqrt{3} = a + b\sqrt{2}$ gives $3 = a^2 + 2b^2 + 2ab\sqrt{2}$, forcing $ab = 0$ and then either $\sqrt{3} \in \mathbb{Q}$ or $\sqrt{3/2} \in \mathbb{Q}$, both false). And $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{2},\sqrt{3})$, because

$$\alpha^3 = 11\sqrt{2} + 9\sqrt{3} \quad\Longrightarrow\quad \alpha^3 - 9\alpha = 2\sqrt{2},$$

so $\sqrt{2} \in \mathbb{Q}(\alpha)$ and hence $\sqrt{3} = \alpha - \sqrt{2} \in \mathbb{Q}(\alpha)$. Therefore $[\mathbb{Q}(\alpha):\mathbb{Q}] = 4$. **Degree 4.**

(b) — FALSE, degree 2

The key observation is that the radicand is a perfect square:

$$3 + 2\sqrt{2} = 2 + 2\sqrt{2} + 1 = \left(\sqrt{2}+1\right)^2.$$

Hence

$$\sqrt{3+2\sqrt{2}} = \sqrt{2} + 1,$$

taking the positive root. This lies in $\mathbb{Q}(\sqrt{2})$, so its degree over $\mathbb{Q}$ is 2, not 4. Its minimal polynomial is

$$(x-1)^2 = 2, \qquad \text{i.e.} \qquad x^2 - 2x - 1 = 0.$$

Degree 2 — false.

Denesting radicals. An expression $\sqrt{a + b\sqrt{c}}$ simplifies to $\sqrt{m}+\sqrt{n}$ exactly when $a^2 - b^2c$ is a perfect square, via

$$\sqrt{a \pm b\sqrt{c}} = \sqrt{\frac{a + \sqrt{a^2 - b^2c}}{2}} \pm \sqrt{\frac{a - \sqrt{a^2-b^2c}}{2}}.$$

Here $a=3$, $b=2$, $c=2$ give $a^2-b^2c=9-8=1$, a perfect square, so denesting succeeds: $\sqrt{(3+1)/2}+\sqrt{(3-1)/2}=\sqrt{2}+1$. ✓
Always test for denesting before assuming a nested radical has degree 4. This option exists precisely to catch the candidate who does not.

(c) — TRUE, degree 4

$\sqrt[4]{2}$ is a root of x^4-2, which is irreducible over $\mathbb{Q}$ by Eisenstein's criterion at the prime 2: the leading coefficient 1 is not divisible by 2, all lower coefficients are (0 and -2), and the constant -2 is not divisible by 4. Hence

$$[\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}]=4.$$

Degree 4.

(d) — TRUE, degree 4

Let $\beta=\mathrm{i}+\sqrt{2}$. Then

$$\beta-\sqrt{2}=\mathrm{i}\quad\Longrightarrow\quad(\beta-\sqrt{2})^2=-1\quad\Longrightarrow\quad\beta^2-2\sqrt{2}\beta+2=-1,$$

so $\beta^2+3=2\sqrt{2}\beta$ and squaring,

$$(\beta^2+3)^2=8\beta^2\quad\Longrightarrow\quad\beta^4+6\beta^2+9=8\beta^2\quad\Longrightarrow\quad\beta^4-2\beta^2+9=0.$$

Degree via the tower law. $\mathbb{Q}(\mathrm{i},\sqrt{2})$ has degree 4 over $\mathbb{Q}$, since $\mathrm{i}\notin\mathbb{Q}(\sqrt{2})\subset\mathbb{R}$. And $\mathbb{Q}(\beta)=\mathbb{Q}(\mathrm{i},\sqrt{2})$: from $\beta^2=1+2\sqrt{2}\,\mathrm{i}$ we get $\sqrt{2}\,\mathrm{i}\in\mathbb{Q}(\beta)$, and then

$$\beta\cdot\sqrt{2}\,\mathrm{i}=(\mathrm{i}+\sqrt{2})\sqrt{2}\,\mathrm{i}=\sqrt{2}\,\mathrm{i}^2+2\mathrm{i}=-\sqrt{2}+2\mathrm{i},$$

which combined with $\beta=\mathrm{i}+\sqrt{2}$ lets us solve for i and $\sqrt{2}$ separately. Hence $[\mathbb{Q}(\beta):\mathbb{Q}]=4$.
Degree 4.

The general principle behind (a) and (d). If $[\mathbb{Q}(\gamma):\mathbb{Q}]=m$ and $[\mathbb{Q}(\delta):\mathbb{Q}]=n$ with $\gcd(m,n)=1$, then $[\mathbb{Q}(\gamma,\delta):\mathbb{Q}]=mn$. Here both degrees are 2, so the gcd argument does not apply directly — one must check that neither square root lies in the other's field, which is what the explicit computations above accomplish. When $\gcd(m,n)=1$ the conclusion is automatic; when it is not, the compositum may collapse, and option **(b)** is exactly such a collapse in disguise.

Answer: **(a)**, **(c)** and **(d)**. (**(b)** denests to $1+\sqrt{2}$ and has degree 2.)

Question 113. For the given chain of subgroups or modules, which containments and indices hold?

The chain did not reproduce

Multiplicativity of the index

For subgroups $K \le H \le G$ with $[G : K]$ finite,

$$[G : K] = [G : H] \cdot [H : K].$$

The proof is a counting argument: each coset of H in G splits into exactly $[H : K]$ cosets of K. Two consequences used constantly:

- $[G : H]$ divides $[G : K]$ for any intermediate H;
- Lagrange's theorem is the case $K = \{e\}$.

The same formula holds for field extensions (the tower law, used in Questions 67 and 112) and for modules, with dimension or length in place of index.

The usual traps in chain items.
Normality is not transitive. $K \trianglelefteq H$ and $H \trianglelefteq G$ do *not* imply $K \trianglelefteq G$. The standard counterexample sits inside the dihedral group of order 8: with $G = D_4$, $H = \{e, r^2, s, r^2 s\}$ (a Klein four-subgroup, normal of index 2) and $K = \{e, s\}$ (normal in H since H is abelian), one finds $rsr^{-1} = r^2 s \notin K$, so K is not normal in G. A subgroup normal in every intermediate step of a chain is called *subnormal*, and the distinction from normality is precisely what makes solvability a subtler notion than abelianness.
Index and order move oppositely. A larger subgroup has smaller index. When an item lists both, check the product against $|G|$ before anything else.
Composition series. By Jordan–Hölder, any two composition series of a finite group have the same length and the same multiset of composition factors up to isomorphism — so a chain's factors are an invariant of G, not of the chosen chain.

Answer: Indeterminate: the chain of subgroups is absent from the available text.

Question 114. Let f be entire with $f(z) = f(1/z)$ for all $z \neq 0$. Then **(a)** $\oint_{|z|=1} \frac{f'(z)}{z^2}\,\mathrm{d}z = 0$ **(b)** f is an open map **(c)** $f(\mathbb{C})$ is open **(d)** $\lim_{n\to\infty} n!\,a_n = 0$ where a_n are the Taylor coefficients.

The decisive structural fact

Theorem 3.7. If f is entire and $f(z) = f(1/z)$ for all $z \neq 0$, then f is constant.

Proof. Consider the behaviour of f as $|z| \to \infty$. For $|z| > 1$ we have $|1/z| < 1$, so

$$|f(z)| = |f(1/z)| \;\le\; \max_{|w|\le 1} |f(w)| =: M < \infty,$$

the maximum being finite because f is continuous on the compact closed unit disc. Hence $|f| \le M$ on $\{|z| > 1\}$, and $|f| \le M$ on $\{|z| \le 1\}$ as well. So f is a bounded entire function, and by Liouville's theorem f is constant. □

An alternative route via the coefficients. Writing $f(z) = \sum_{n\geq 0} a_n z^n$ and noting that $f(1/z) = \sum_{n\geq 0} a_n z^{-n}$ has a pole or essential singularity at 0 unless $a_n = 0$ for all $n \geq 1$: the identity $f(z) = f(1/z)$ forces the Laurent expansion of f at 0 to have no negative powers *and* to equal one with no positive powers. Hence $a_n = 0$ for every $n \geq 1$ and $f = a_0$. This is the same conclusion by bookkeeping rather than by Liouville.

So $f \equiv c$ for some constant $c \in \mathbb{C}$, and every option is adjudicated against this.

(a) — TRUE

Since f is constant, $f' \equiv 0$, so the integrand vanishes identically and

$$\oint_{|z|=1} \frac{f'(z)}{z^2}\,\mathrm{d}z = 0.$$

True.

(Even without constancy the integral would vanish: by the residue theorem it equals $2\pi\mathrm{i} \cdot \mathrm{Res}_{z=0} \frac{f'(z)}{z^2} = 2\pi\mathrm{i}\, f''(0)$, and one could argue separately that $f''(0) = 0$. But constancy makes it immediate.)

(b) — FALSE

The open mapping theorem

A *non-constant* holomorphic function on a domain is an open map: it carries open sets to open sets. The hypothesis of non-constancy is essential — indeed the theorem is one of the sharpest expressions of the rigidity of holomorphic maps, and it immediately yields the maximum modulus principle.

Here f is constant, so it maps every non-empty open set to the single point $\{c\}$, which is not open in $\mathbb{C}$. **False.**

(c) — FALSE

$f(\mathbb{C}) = \{c\}$, a singleton. Singletons are closed, not open, in $\mathbb{C}$. **False.**

(d) — TRUE

With $f \equiv c$ the Taylor coefficients are $a_0 = c$ and $a_n = 0$ for all $n \geq 1$. Hence

$$n!\,a_n = 0 \quad \text{for every } n \geq 1, \qquad \text{so} \qquad \lim_{n\to\infty} n!\,a_n = 0.$$

True.

Why the factorial appears. By Cauchy's formula $a_n = f^{(n)}(0)/n!$, so $n!\,a_n = f^{(n)}(0)$. The option is therefore asserting that the derivatives at the origin tend to zero — true here because they are all zero from the first onward.

The family of related traps. For a general entire function of finite order, Cauchy's estimates give $|a_n| \leq M(r)/r^n$ for every r, so rapid decay of a_n is the norm; but $n!a_n = f^{(n)}(0)$ can still grow without bound (take $f(z) = \mathrm{e}^{z^2}$). It is only the constancy forced by the functional equation that makes **(d)** true here.

Answer: **(a)** and **(d)**.

Question 115. A bead of unit mass slides without friction on the parabola $y = x^2$ in a uniform gravitational field g (with y vertical). Then **(a)** the Lagrangian is $\frac{1}{2}\dot{x}^2 - gx^2$ **(b)** the kinetic energy is $\frac{1}{2}(1+4x^2)\dot{x}^2$ **(c)** the Lagrangian is $\frac{1}{2}(1+4x^2)\dot{x}^2 - gx^2$ **(d)** the equation of motion is $(1+4x^2)\ddot{x} + 4x\dot{x}^2 + 2gx = 0$.

Setting up the constrained system

The bead is confined to the curve $y = x^2$, so the system has one degree of freedom; take x as the generalised coordinate. Differentiating the constraint,

$$\dot{y} = \frac{\mathrm{d}}{\mathrm{d}t}(x^2) = 2x\dot{x}.$$

(b) — TRUE

For unit mass,

$$T = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}(\dot{x}^2 + 4x^2\dot{x}^2) = \tfrac{1}{2}(1+4x^2)\dot{x}^2.$$

True.

The factor $(1+4x^2)$ is the square of the arc-length element: $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = (1+4x^2)\,\mathrm{d}x^2$, so $T = \frac{1}{2}\dot{s}^2$ as it must be for a bead moving along a curve. Geometrically, $(1+4x^2)$ is the induced metric on the parabola — the “effective mass” the bead feels, which grows as the curve steepens.

(c) — TRUE, (a) — FALSE

The potential energy is $V = mgy = g\,y = gx^2$ for unit mass. Hence

$$L = T - V = \tfrac{1}{2}(1+4x^2)\dot{x}^2 - gx^2,$$

which is exactly option **(c)**. **True.**

Option **(a)** omits the factor $(1+4x^2)$ from the kinetic term — it is the Lagrangian of a *free* particle in a quadratic potential, i.e. a simple harmonic oscillator, not of a bead on a parabola. **False.**

The distinction between **(a)** and **(c)** is the entire content of the item. It is tempting to write $T = \frac{1}{2}\dot{x}^2$ because x is the chosen coordinate, but x is not arc length: the bead's speed involves the vertical motion too. The constraint must be differentiated, not merely substituted into the potential.

(d) — TRUE

Apply the Euler–Lagrange equation $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial \dot{x}}\right) - \dfrac{\partial L}{\partial x} = 0.$

The momentum term.

$$\frac{\partial L}{\partial \dot{x}} = (1+4x^2)\dot{x},$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[(1+4x^2)\dot{x}\Big] = (1+4x^2)\ddot{x} + 8x\dot{x}\cdot\dot{x} = (1+4x^2)\ddot{x} + 8x\dot{x}^2.$$

The force term.

$$\frac{\partial L}{\partial x} = \tfrac{1}{2} \cdot 8x\,\dot{x}^2 - 2gx = 4x\dot{x}^2 - 2gx.$$

Combining.

$$\left(1 + 4x^2\right)\ddot{x} + 8x\dot{x}^2 - 4x\dot{x}^2 + 2gx = 0,$$

that is,

$$\left(1 + 4x^2\right)\ddot{x} + 4x\dot{x}^2 + 2gx = 0.$$

This is exactly option **(d)**. **True.**

Reading the three terms. $(1 + 4x^2)\ddot{x}$ is the inertial term with the position-dependent effective mass; $4x\dot{x}^2$ is a centrifugal-type term arising from the curvature of the constraint (it vanishes at the vertex $x = 0$ where the curve is locally flat); $2gx$ is the restoring gravitational force.
Small-oscillation limit. For $|x| \ll 1$ the equation linearises to $\ddot{x} + 2gx = 0$, simple harmonic motion with

$$\omega = \sqrt{2g}, \qquad T_{\text{period}} = \frac{2\pi}{\sqrt{2g}}.$$

Note this is *independent of amplitude* only in the linear approximation; the exact motion on a parabola is anharmonic. (The curve for which oscillations are exactly isochronous is the cycloid — Huygens' tautochrone — not the parabola.)
Conserved energy. Since L has no explicit t-dependence, the Hamiltonian

$$H = \dot{x}\frac{\partial L}{\partial \dot{x}} - L = \tfrac{1}{2}(1 + 4x^2)\dot{x}^2 + gx^2 = T + V$$

is conserved, providing a first integral and an independent check on **(d)**: differentiating H and setting $\dot{H} = 0$ reproduces the equation of motion.

Answer: **(b)**, **(c)** and **(d)**.

Question 116. For the multivariate normal vector with the given mean and covariance, which distributional statements hold?

The mean vector and covariance matrix did not reproduce

Linear transformations of a multivariate normal

If $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ then for any $q \times p$ matrix A and vector $\mathbf{b}$,

$$A\mathbf{X} + \mathbf{b} \;\sim\; N_q\left(A\boldsymbol{\mu} + \mathbf{b},\; A\Sigma A^{\mathsf{T}}\right).$$

Every marginal and every linear combination is therefore normal, with parameters read off by this single rule. In particular $\mathbf{a}^{\mathsf{T}}\mathbf{X} \sim N\left(\mathbf{a}^{\mathsf{T}}\boldsymbol{\mu},\; \mathbf{a}^{\mathsf{T}}\Sigma\mathbf{a}\right)$.

Conditioning and independence

Partition $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ conformably with

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Then

$$\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{x}_2 \ \sim \ N\Big(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \ \ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Big).$$

Two features are worth emphasising: the conditional mean is *linear* in $\mathbf{x}_2$, and the conditional covariance — the Schur complement — does *not* depend on $\mathbf{x}_2$ at all. Independence of $\mathbf{X}_1$ and $\mathbf{X}_2$ holds iff $\Sigma_{12} = 0$.

Quadratic forms

If $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ with Σ non-singular, then

$$(\mathbf{X} - \boldsymbol{\mu})^{\mathsf{T}}\Sigma^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_p.$$

More generally $\mathbf{X}^{\mathsf{T}}A\mathbf{X} \sim \chi^2_r$ (central, when $\boldsymbol{\mu} = 0$) iff $A\Sigma$ is idempotent of rank r — the criterion underlying Cochran's theorem and the distribution theory of ANOVA.

The pairing structure of the option set. Items of this kind typically pair options **(a)** with **(c)** and **(b)** with **(d)**, offering the same distributional claim with two different parameter values — so that at most one of each pair can be correct. Identifying such pairs before computing halves the work and rules out most of the sixteen possible response patterns.

Answer: Indeterminate: the mean vector and covariance matrix are absent from the available text.

Question 117. For a symmetric positive-definite 2×2 matrix A with the given entries, **(a)** $\max \frac{x^{\mathsf{T}}A^2x}{x^{\mathsf{T}}x} = 2(3+\sqrt{2})$ **(b)** $\min \frac{x^{\mathsf{T}}A^2x}{x^{\mathsf{T}}x} = 2(3-2\sqrt{2})$ **(c)** $\max \frac{x^{\mathsf{T}}A^{-1}x}{x^{\mathsf{T}}x} = \frac{2+\sqrt{2}}{2}$ **(d)** $\min \frac{x^{\mathsf{T}}A^{-1}x}{x^{\mathsf{T}}x} = \frac{1+\sqrt{2}}{2}$.

The concept: the Rayleigh quotient

Rayleigh–Ritz

For a real symmetric A with eigenvalues $\lambda_{\min} = \lambda_1 \le \cdots \le \lambda_n = \lambda_{\max}$,

$$\lambda_{\min} = \min_{x \neq 0} \frac{x^{\mathsf{T}}Ax}{x^{\mathsf{T}}x}, \qquad \lambda_{\max} = \max_{x \neq 0} \frac{x^{\mathsf{T}}Ax}{x^{\mathsf{T}}x},$$

the extrema being attained at the corresponding eigenvectors.

The functional calculus. If A has eigenvalues λ_i with orthonormal eigenvectors v_i, then for any function φ defined on the spectrum, $\varphi(A)$ has eigenvalues $\varphi(\lambda_i)$ with the *same* eigenvectors. In particular

$$\operatorname{eig}(A^2) = \{\lambda_i^2\}, \qquad \operatorname{eig}(A^{-1}) = \{1/\lambda_i\}.$$

Note the order reversal for the inverse: the largest eigenvalue of A^{-1} is $1/\lambda_{\min}$, and the smallest is $1/\lambda_{\max}$.

Reconstructing A from the option set

The entries did not reproduce, but the repeated appearance of $\sqrt{2}$ and the combinations $3 \pm 2\sqrt{2} = (\sqrt{2} \pm 1)^2$ identify the eigenvalues. The natural candidate is

$$A = \begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix},$$

whose eigenvalues are $2 \pm \sqrt{2}$ (for a matrix $\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)$ the eigenvalues are $a \pm b$, with eigenvectors $(1, \pm 1)$). Both are positive ($2 - \sqrt{2} \approx 0.586 > 0$), so A is positive definite as required. Two options match this hypothesis exactly, confirming it.

Adjudicating

For A^2, the eigenvalues are $(2 \pm \sqrt{2})^2$:

$$(2+\sqrt{2})^2 = 4 + 4\sqrt{2} + 2 = 6 + 4\sqrt{2} = 2\big(3 + 2\sqrt{2}\big),$$
$$(2-\sqrt{2})^2 = 4 - 4\sqrt{2} + 2 = 6 - 4\sqrt{2} = 2\big(3 - 2\sqrt{2}\big).$$

(b) TRUE. The minimum of the Rayleigh quotient of A^2 is the smallest eigenvalue, $2(3 - 2\sqrt{2})$. ✓ (Numerically $2(3 - 2.828) = 0.343$, and indeed $(2 - \sqrt{2})^2 = 0.586^2 = 0.343$. ✓)

(a) FALSE. The maximum is $2(3 + 2\sqrt{2}) \approx 11.66$, whereas the option states $2(3 + \sqrt{2}) \approx 8.83$. The printed expression is missing the factor 2 inside the bracket — evidently a typographical slip, since its partner **(b)** carries the 2 correctly.

For A^{-1}, the eigenvalues are $1/(2 \pm \sqrt{2})$. Rationalise:

$$\frac{1}{2-\sqrt{2}} = \frac{2+\sqrt{2}}{(2-\sqrt{2})(2+\sqrt{2})} = \frac{2+\sqrt{2}}{4-2} = \frac{2+\sqrt{2}}{2},$$
$$\frac{1}{2+\sqrt{2}} = \frac{2-\sqrt{2}}{2}.$$

(c) TRUE. The maximum Rayleigh quotient of A^{-1} is $1/\lambda_{\min}(A) = 1/(2 - \sqrt{2}) = \frac{2+\sqrt{2}}{2}$. ✓ (Numerically 1.707, and $1/0.586 = 1.707$. ✓)

(d) FALSE. The minimum is $1/\lambda_{\max}(A) = \frac{2-\sqrt{2}}{2} \approx 0.293$, not $\frac{1+\sqrt{2}}{2} \approx 1.207$.

Consistency checks. The product of the extreme Rayleigh quotients of A^{-1} should be $1/\det A$:

$$\frac{2+\sqrt{2}}{2} \cdot \frac{2-\sqrt{2}}{2} = \frac{4-2}{4} = \frac{1}{2}, \qquad \det A = 4 - 2 = 2, \qquad \frac{1}{\det A} = \frac{1}{2}. \checkmark$$

And their sum should be $\operatorname{tr}(A^{-1}) = \operatorname{tr} A / \det A = 4/2 = 2$:

$$\frac{2+\sqrt{2}}{2} + \frac{2-\sqrt{2}}{2} = \frac{4}{2} = 2. \checkmark$$

Similarly for A^2: $\operatorname{tr}(A^2) = (6 + 4\sqrt{2}) + (6 - 4\sqrt{2}) = 12$, and directly $\operatorname{tr}(A^2) = \sum a_{ij}^2 = 4 + 2 + 2 + 4 = 12$ ✓.

The condition number. $\kappa_2(A) = \lambda_{\max}/\lambda_{\min} = (2+\sqrt{2})/(2-\sqrt{2}) = 3+2\sqrt{2} \approx 5.83$ — modest, so A is well conditioned. This quantity is exactly the ratio appearing in options **(a)** and **(b)**, which is why $3 \pm 2\sqrt{2}$ recurs.

Answer: **(b)** and **(c)**. (**(a)** appears to be a misprint for $2(3+2\sqrt{2})$, which would be correct; **(d)** is false.)

Question 118. For the given parameter sets, which of the stated properties of the family of distributions hold?

The parameter sets did not reproduce

Exponential families

A family $\{f_\theta\}$ is a *k-parameter exponential family* if

$$f_\theta(x) = h(x)\, c(\theta) \exp\left(\sum_{j=1}^{k} w_j(\theta)\, t_j(x)\right),$$

with the crucial requirement that the *support of f_θ does not depend on θ*. Then $T = (t_1, \ldots, t_k)$ is sufficient, and it is complete provided the natural parameter space $\{w(\theta)\}$ contains an open k-dimensional rectangle.

Families that are exponential: normal, Poisson, binomial (fixed n), gamma, beta, negative binomial.

Families that are not: uniform $(0, \theta)$, uniform $(\theta, 2\theta)$, shifted exponential, Cauchy, and any family whose support moves with θ.

Sufficiency, completeness, and their uses

Factorisation criterion. T is sufficient for θ iff $f_\theta(x) = g_\theta\big(T(x)\big)h(x)$.

Completeness. T is complete if $\mathbb{E}_\theta[g(T)] = 0$ for all θ implies $g \equiv 0$ a.s. This is what makes an unbiased function of T *unique*.

Lehmann–Scheffé. If T is complete sufficient and $g(T)$ is unbiased for $\tau(\theta)$, then $g(T)$ is the unique UMVUE.

Basu's theorem. A complete sufficient statistic is independent of every ancillary statistic. The classic application: for $N(\mu, \sigma^2)$ with σ^2 known, $\bar{X}$ is complete sufficient and S^2 is ancillary, so they are independent — a fact otherwise requiring a change-of-variables computation.

A caution on truncated parameter sets. Items of this type often restrict θ to a set such as $\{1, 2, 3\}$ or an interval like $(0, 1)$, and the restriction can destroy completeness while preserving sufficiency. The standard example: for $N(\theta, 1)$ with $\theta \in \mathbb{R}$, $\bar{X}$ is complete; but for θ restricted to an interval of positive length completeness survives, while for θ restricted to a *finite* set it typically fails, since a non-zero polynomial can vanish at finitely many points.

The dimension of the natural parameter space is what matters, and checking it is the first step in any such item.

Answer: Indeterminate: the parameter sets and the associated distributional family are absent from the available text.

Question 119. For the given linear programming problem, which statements about its optimal solution hold?

The programme did not reproduce

The fundamental theorem of linear programming

For $\max\{c^{\mathsf{T}}x : Ax \le b,\ x \ge 0\}$ with a non-empty feasible region:

- the feasible region is a convex polyhedron;
- if an optimum exists, it is attained at an *extreme point* (vertex);
- if two vertices are both optimal, every point of the segment joining them is optimal, so the optimum is non-unique and the solution set is a face;
- the problem may be infeasible (empty region) or unbounded (objective improves without limit along a recession direction).

These four possibilities — unique optimum, multiple optima, infeasible, unbounded — exhaust the cases, and multiple-select items in this area usually test the ability to distinguish them.

Duality

The dual of $\max\{c^{\mathsf{T}}x : Ax \le b,\ x \ge 0\}$ is $\min\{b^{\mathsf{T}}y : A^{\mathsf{T}}y \ge c,\ y \ge 0\}$.
Weak duality: $c^{\mathsf{T}}x \le b^{\mathsf{T}}y$ for any feasible pair.
Strong duality: if either has a finite optimum, both do, and the optimal values coincide.
Complementary slackness: at optimality,

$$y_i\big(b_i - (Ax)_i\big) = 0 \ \ \forall i, \qquad x_j\big((A^{\mathsf{T}}y)_j - c_j\big) = 0 \ \ \forall j.$$

A binding constraint may have a positive dual price; a slack constraint must have dual price zero. This is the most useful tool for verifying a claimed optimum without running the simplex method.
The primal–dual possibility table. If the primal is unbounded the dual is infeasible, and vice versa; both may be infeasible simultaneously; but they cannot both be unbounded.

Checking a candidate optimum quickly. Given a proposed optimal x^*, identify the binding constraints, solve the complementary-slackness equations for y, and verify that $y \ge 0$ and $A^{\mathsf{T}}y \ge c$. If so, weak duality with matching objective values certifies optimality — no simplex iterations required. This is far faster than re-solving, and is the intended route for most examination items that present a candidate solution.

Answer: Indeterminate: the linear programming problem is absent from the available text.

Question 120. For f holomorphic on $\mathbb{C}$ and I_n the number of solutions of $f(z) = 1$ in $|z| < n$ counted with multiplicity, **(a)** if $I_n = 0$ for all n then f has a removable singularity at ∞ **(b)** if f is a polynomial then I_n is eventually constant **(c)** if f has a removable singularity at ∞ then $I_n = 0$ for all n **(d)** there exists f with an essential singularity at ∞ and $I_n = 0$ for all n.

The counting principle

The argument principle

If g is holomorphic inside and on a simple closed contour γ, with no zeros on γ, then

$$\frac{1}{2\pi\mathrm{i}} \oint_\gamma \frac{g'(z)}{g(z)}\,\mathrm{d}z = N,$$

the number of zeros of g inside γ, counted with multiplicity.
Applied with $g = f - 1$ on $\gamma = \{|z| = n\}$, this is exactly I_n: the number of solutions of $f(z) = 1$ in the disc. Since I_n counts zeros in an increasing family of discs, I_n is non-decreasing in n.

(b) — TRUE

If f is a polynomial of degree $d \geq 1$, then $f - 1$ is a polynomial of the same degree, with exactly d roots in $\mathbb{C}$ counted with multiplicity (fundamental theorem of algebra). All lie within some disc $|z| < R$, so

$$I_n = d \qquad \text{for every } n > R.$$

The sequence is eventually constant. **True.**

(If f is a constant polynomial, then $f - 1$ is constant: either identically zero — excluded, since I_n would be undefined — or nowhere zero, giving $I_n = 0$ for all n, again eventually constant.)

(c) — TRUE

Classifying the singularity at infinity

Set $g(w) = f(1/w)$ and classify g's singularity at $w = 0$. Then f has at ∞:

- a **removable** singularity iff f is bounded near ∞;
- a **pole** of order d iff f is a polynomial of degree d;
- an **essential** singularity otherwise (e.g. e^z, $\sin z$).

Suppose f is entire with a removable singularity at ∞. Then f is bounded on $|z| > R$ and, being continuous, bounded on $|z| \le R$; so f is a bounded entire function. By Liouville's theorem $f \equiv c$ is constant.

Now if $c = 1$ then $f - 1 \equiv 0$ and I_n is not defined (every point is a solution); the statement implicitly excludes this. Otherwise $c \neq 1$, so $f - 1$ is a non-zero constant with no zeros at all, and

$$I_n = 0 \quad \text{for every } n.$$

True.

(d) — TRUE

Take

$$f(z) = \mathrm{e}^z + 1.$$

This is entire. Its singularity at ∞ is essential: f is unbounded (along the positive real axis $\mathrm{e}^z \to \infty$) yet is not a polynomial (along the negative real axis $\mathrm{e}^z \to 0$, so $f \to 1$, which no non-constant polynomial does). Formally, $f(1/w) = \mathrm{e}^{1/w} + 1$ has an essential singularity at $w = 0$.

And the equation $f(z) = 1$ reads

$$\mathrm{e}^z + 1 = 1 \quad \Longleftrightarrow \quad \mathrm{e}^z = 0,$$

which has *no* solutions, since the exponential never vanishes. Hence $I_n = 0$ for every n. **True.**

(a) — FALSE

The example just constructed refutes it directly. Take $f(z) = \mathrm{e}^z + 1$: we have $I_n = 0$ for all n, yet f has an *essential* singularity at ∞, not a removable one. **False.**

The logical structure of the item. Options **(a)** and **(c)** are converses of each other. Option **(c)** is true; option **(a)**, its converse, is false; and option **(d)** supplies precisely the counterexample that refutes **(a)**. Recognising this dependency means that establishing **(d)** settles **(a)** at no extra cost — and that marking both **(a)** and **(d)** would be self-contradictory.

Why the exponential evades the argument principle's reach. By Picard's great theorem, a function with an essential singularity takes every complex value, with at most one exception, infinitely often in any neighbourhood of the singularity. For e^z near ∞ the exceptional value is 0; hence $\mathrm{e}^z + 1$ omits exactly the value 1. The item is, at bottom, an invitation to recall which value the exponential omits.

Answer: **(b)**, **(c)** and **(d)**.

CHAPTER 4

Register of Defective and Indeterminate Items

This chapter is an accounting of the paper's own condition. It is included because a solution commentary that silently repairs a broken item teaches the reader nothing about how to behave when an examination misbehaves — and because a candidate who spends four minutes hunting for an answer that does not exist has lost those minutes to the paper, not to the mathematics.

Two categories are distinguished.

The two categories

Defective. The item is fully legible, but the mathematics of the stem and the mathematics of the options do not agree. Either the correct answer is absent from the option list, or two options are identical, or an option contains an arithmetical or typographical error, or the stem's data are internally inconsistent. These are faults *of the paper*.

Indeterminate. The item is presumably sound, but material necessary to answer it — a figure, a matrix, a list of statements, a set of estimators — is not present in the source text available for this commentary. These are faults *of the reproduction*, and the underlying items may be perfectly good. For each, the relevant theory has been developed in full, and what is missing has been stated precisely.

4.1 Defective items

Item	Nature of the defect
A4	Options **(b)** and **(c)** describe the same configuration — the isosceles right triangle of height r and area r^2. Two options are simultaneously correct in a single-answer item.
A6	The stem's worked example COMPUTE → RFUVQPC is corrupt. Under the rule the options encode (reverse the word, then advance each letter by one) the image should be FUVQNPD. The options themselves are consistent, so **(a)** is recoverable.
A8	The three intervals have least common multiple 600 minutes, i.e. ten hours, giving 6:00 PM. This time does not appear among the options.
A16	The displacement is $(3, 3)$, of magnitude $3\sqrt{2} = \sqrt{18}$ km. No option carries this value.
B22	No option satisfies either the printed partial differential equation or the printed initial condition. The Lagrange system and its multipliers are developed in the commentary; the option set does not match any solution branch.
B30	Ambiguity: more than one entire function in the option list meets the stated growth condition, depending on how the normalisation at the origin is read.
B36	Options **(a)**, **(b)** and **(c)** are all entire; only $\lvert z\rvert^2$ in **(d)** fails to be. The item asks which is *not* entire and admits exactly one answer, but the distractor structure suggests a different intent.
B48	Over-determined. Schwarz–Pick gives $\lvert\varphi(1/3)\rvert \le 1/5$, which eliminates three options; but the bound is not attained under the stem's full hypotheses, the true minimal value being $3/20$. The item is answerable by elimination and unanswerable by direct computation.
C68	The tent of height n^2 on $[0, 3/n]$ has integral $3n/2$, not the value 2 asserted in option **(b)**. Option **(b)** would be correct for height $4n/3$. The stem's parameters and its options are inconsistent.
C76	Under the printed conditions $y(1) = y(\mathrm{e}) = 0$ the eigenpairs are $\lambda_n = n^2\pi^2$, $y_n = \sin(n\pi \ln x)$ — absent from the options. Under $y'(1) = 0$, $y(\mathrm{e}) = 0$ the answer is exactly **(b)** and **(d)**. The boundary condition is misprinted.
C81	All four printed intervals lie within $(0, 1)$, the full range over which the natural relaxation iteration contracts. The item cannot discriminate as printed; the iteration function is also absent.
C86	Options **(c)** and **(d)** give numerically incorrect values. The correct result is $y(2) = a(12 - 8\mathrm{e}^2 + 2\mathrm{e}^3)/(3 - \mathrm{e}^2) \approx 1.582a$, positive, whereas both printed values are negative.
C87	Options **(a)** and **(b)** are the same statistic, as are **(c)** and **(d)**, since $I(X_i \neq 0) + I(X_i = 0) = 1$ identically. The item has two distinct candidates, not four.
C93	Option **(a)** states $a^2 + b^2 + c^2 = 10$; the Simpson weights $(\frac{1}{3}, \frac{4}{3}, \frac{1}{3})$ give 2.
C94	Options **(a)** and **(b)** are mutually inconsistent: a method of order 2 has local truncation error $O(h^3)$, not $O(h^2)$.
C104	The printed density carries e^{-xy}, which does not normalise over $0 \le x \le y$. The intended density is $x(y - x)\mathrm{e}^{-y}$, under which the computations close.
C117	Option **(a)** reads $2(3 + \sqrt{2})$; the correct maximum is $2(3 + 2\sqrt{2})$. Its partner **(b)** carries the factor 2 correctly, confirming the slip.

4.2 Indeterminate items

The following items could not be adjudicated because material essential to them is absent from the source text. In each case the commentary develops the governing theory in full and states exactly what is missing; several are partially adjudicable, and those partial verdicts are recorded here.

Item	What is missing, and what survives
A10, A11, A13	Figures, pie chart and bar diagram absent.
B21	The two lists to be matched are absent.
B23	Transition matrix absent.
B25	Matrix absent. For 2×2, $\operatorname{tr} T_{\mathrm{GS}} = bc/(ad)$.
B29, B43, B46, B52, B56, B58, B59	Stem data absent.
B38	Primal problem absent. The number of dual variables equals the number of primal constraints.
B47	Constraint set absent. The unrestricted count is $\binom{19}{2} = 171$.
B50	Matrix absent. For $n \times n$, $\det(\operatorname{adj} A) = (\det A)^{n-1}$, which equals $\det A$ when $n = 2$.
B54	Space absent. The transpose operator on 3×3 matrices has minimal polynomial $x^2 - 1$, eigenvalue $+1$ with multiplicity 6 and -1 with multiplicity 3, trace 3 and determinant -1.
B57	Test absent. Option **(a)** $= 0.062$ is eliminable, falling below the stated size 0.1.
B60	Data absent. Option **(b)** $= 6$ is impossible, since $\dim P_4(\mathbb{R}) = 5$.
C75	σ^2 and ρ absent. Option **(a)** holds unconditionally for $\rho > 0$; **(b)** is false except in the degenerate case; the likely key is **(a)** and **(d)**.
C80	Design matrix absent. Options **(c)** and **(d)** are proportional, hence both estimable or both not.
C84	The estimators T_1, T_2, T_3 are absent. Options **(c)** and **(d)** have weights summing to 1, so each is unbiased exactly when its constituents are.
C89	Densities absent. Option **(d)** is eliminable: a Bayes rule with equal priors has error at most $1/2$, whereas $1/\sqrt{3} \approx 0.577$.
C94	Scheme parameters and stability interval absent; **(b)** adjudicated true, **(a)** false.
C95	Subspaces absent. Option **(a)** is true unconditionally; **(b)**, **(c)**, **(d)** are pairwise incompatible, so at most one can be keyed.
C96	Block layout incomplete. For the cyclic four-block design the verdicts are **(a)**, **(b)**, **(d)** true and **(c)** false.
C98, C107, C108, C109, C111, C113, C116, C118, C119	Equation, polynomials, transformation values, initial curves, statement list, subgroup chain, mean and covariance, parameter sets, and linear programme respectively — all absent.

4.3 Summary

	Part A	Part B	Part C
Items	20	40	60
Fully adjudicated	16	21	44
Defective	4	4	9
Indeterminate	3	19	16

(The categories overlap: an item may be both defective in its options and indeterminate in its data, and C94 is counted in both columns.)

What to do in the hall. The practical lesson of this register is a budgeting one. If an item resists after a minute of honest work, the possibility that the item is at fault deserves

genuine weight — and the correct response is to leave it and return, not to persist. In Part C especially, where a partial response scores nothing, the marginal value of a fifth minute on a doubtful item is negative: the same minutes spent on a clean item are worth a full mark.
Where an item is defective but recoverable — A6 and C76 are the clearest instances — the options themselves are often more reliable than the stem, because option sets are typically generated from a worked solution while stems are retyped. Reading the options for internal consistency, as was done for C65, C87 and C88, is a genuine technique and not merely a last resort.