

ORCHESTRATEUR

Mathematics - Defining, proving, calculating and interpreting

An in-depth guide, available offline

Read, understand, return to a chapter

LAFAYETTE DIGITAL PRODUCTS

Contents

The contents entries are clickable. PDF bookmarks provide chapter and section access. The link at the bottom of each page returns here.

Chapter 1 - From a problem to a proof: domains, implications and counterexamples 7

Chapter 2 - Sequences and induction: understand discrete change 16

Chapter 3 - Limits: establish behaviour without confusing calculation with proof 24

Chapter 4 - Continuity: prove existence, uniqueness and bounds 33

Chapter 5 - Differentiation and optimisation: move from local change to a conclusion 42

Chapter 6 - Convexity: compare a curve, its chords and its tangents 51

Chapter 7 - Exponentials, logarithms and trigonometry: transform without losing the domain 60

Chapter 8 - Antiderivatives and integrals: measure accumulation and check an approximation 69

Chapter 9 - Differential equations: connect change with its rate 78

Chapter 10 - Counting, probability and the binomial distribution 87

Chapter 11 - Geometry in space: vectors, planes and distances 98

Chapter 12 - Statistics, concentration and algorithmic methods 109

Appendices - Extended problems, worked solutions and revision strategies 121

Mathematics - Defining, proving, calculating and interpreting

A progression at French Terminale level

This guide develops the main mathematical tools of the mathematics specialty in the final year of French general secondary education, together with the earlier concepts needed to understand them. It brings together definitions, proofs, methods and original exercises. The official curriculum sets a reference for the scope; its text is not reproduced, and it does not guarantee coverage of every arrangement for a future examination session. Further study is identified where it goes beyond the core progression.

The aim is to recognise the conditions under which formulas can be used, rather than to collect formulas. A correctly calculated derivative does not, by itself, answer an optimisation problem. A binomial distribution is not suitable for every repeated drawing experiment. A diagram in three dimensions does not prove an intersection. Each topic therefore connects the meaning of its objects with a method whose steps can be justified.

Numerical situations are educational and fictional when they do not simply describe a mathematical object. No financial return, medical outcome or actual product behaviour is guaranteed. Applications help explain a line of reasoning and its limitations. They do not replace professional analysis of a sensitive decision.

How to use the twelve chapters

The first chapter sets out mathematical language and reasoning. The next two develop sequences and limits. Continuity, differentiation and convexity then form a coherent set of tools for studying functions. The seventh chapter brings together exponentials, logarithms and trigonometry; the eighth covers integrals, and the ninth differential equations. Probability, geometry in three dimensions, and statistics with algorithms complete the progression.

For a first reading, work through this order and redo the simple examples. For revision, choose a precise question and find the property on which its answer depends. The annexes provide problems that combine several tools. They do more than repeat the calculations from the chapters: they show how to organise a longer solution and discuss what changes when the assumptions change.

Work with paper or a place to take notes, even when consulting the guide on a phone. Read the problem, hide the worked solution and write at least one attempt. An explained mistake is often more instructive than a solution that immediately seems clear. Record the first point at which your reasoning stopped being justified, rather than only the final result that differed.

Writing, exactness and approximation

Symbols are defined in context. An exact equality and a numerical approximation play different roles: keeping an expression such as $\ln 2$ or $\sqrt{3}$ can prevent rounding errors. An interval, a domain condition or an initial index is part of the result. Removing this information can turn a correct answer into a false statement.

A proof establishes a proposition for every value in the stated domain. An example checks only one case; a single

counterexample, however, is enough to refute a universal statement. A simulation explores a model but does not replace a proof. These distinctions run throughout the guide and are a skill as important as technical fluency.

To check a solution, look for an independent method: differentiate to check an antiderivative, substitute to check a differential equation, use a complementary probability to check an event, or use a plane equation to check a calculated point. The check should not simply repeat the operation that could have produced the error.

The length of an explanation guarantees neither mastery nor a school grade. To use this guide, choose a concept, reconstruct its reasoning without looking at the solution, and then check the assumptions and calculations. The aim is to transfer a method to a new problem, rather than simply to have read more pages.

Chapter 1 - From a problem to a proof: domains, implications and counterexamples

Objectives: identify the kind of result being sought, distinguish a check from a proof, and make domains and transformations explicit. The initial module is included here once as a foundation; the following chapters develop each family of problems in greater depth.

Historical perspective - Hilbert: make assumptions visible

In 1899, David Hilbert published Grundlagen der Geometrie, an axiomatic presentation of geometry. This historical perspective reminds us that a familiar diagram does not remove the need to state which relations are assumed and which properties are proved. In this chapter, the same requirement applies to the domain of a variable and the conditions for a transformation: a conclusion holds only under the assumptions actually used. This is a teaching comparison; the guide neither presents Hilbert's system of axioms nor claims to resolve questions about foundations.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Hilbert/

1. Know what you are trying to establish

A calculation can produce a value, a conjecture or a proof: these results do not have the same status. Checking a relation for five values suggests that it might be true; it does not prove it for all numbers. A proof explains why the conclusion follows from the assumptions throughout the stated domain. A single counterexample, on the other hand, is enough to refute a universal statement.

Consider “for every real x, $x^2 > x$”. The values 2 and 3 satisfy the inequality, but $x = 0$ already refutes it. The real number 1/2 even gives $x^2 < x$. The failure does not arise from doing too few calculations: the proposition was too broad. To find a correct statement, factor $x^2 - x = x(x - 1)$, then study its sign on the relevant intervals. This gives $x^2 > x$ when $x < 0$ or $x > 1$.

Before solving a problem, specify the domain, the given information and the kind of answer required: an exact value, an approximation, existence, uniqueness, or a property valid for every index. These distinctions guide the method and prevent a graphical approximation from being presented as a proof.

2. Implication, equivalence and domains of definition

Two propositions are equivalent when each implies the other. Squaring an equality can preserve an implication without preserving equivalence. The equality $x = 2$ implies $x^2 = 4$; the equation $x^2 = 4$ also has the solution $x = -2$. In a sequence of steps, stating the conditions that make transformations reversible allows extra solutions to be detected.

Dividing by an expression requires it to be nonzero. To solve $x(x - 2) = 0$, immediately dividing by x removes the solution $x = 0$. The zero-product property directly gives $x = 0$ or $x = 2$. Similarly, an expression containing $\ln(x)$ requires $x > 0$. The domain is not a precaution added after the calculation: it is part of the problem from the start.

A useful habit is to set out the restrictions at the beginning, then check the solutions obtained in the original equation. This final check does not replace a proof that all solutions have been found, but it eliminates those introduced by an irreversible transformation.

3. A sequence: describing discrete change

A sequence assigns a value u_n to each integer index in its domain. An explicit expression gives u_n directly in terms of n; a recurrence relation calculates a value from earlier values. The two descriptions can represent the same change, but they make different questions easier to answer.

Take u_0 = 1 and u_(n+1) = 0.5 u_n + 2. The first terms are 1; 2.5; 3.25; 3.625. They suggest an increase towards 4 without proving it. Look for a fixed value L satisfying L = 0.5 L + 2: this gives L = 4. Set v_n = u_n − 4. Then v_(n+1) = 0.5 v_n, so v_n = −3 × (0.5)^n and u_n = 4 − 3 × (0.5)^n.

This expression justifies the observations. Since (0.5)^n is strictly positive, u_n < 4. Also, u_(n+1) − u_n = 1.5 × (0.5)^n > 0: the sequence is increasing. Finally, (0.5)^n tends to zero, so u_n tends to 4. Calculating values has become a proof because an underlying structure has been identified.

4. Induction: do not forget the first link

A proof by induction has a base case, an induction step and a conclusion. To prove u_n < 4 in the previous example, begin with u_0 = 1 < 4. If u_n < 4, then u_(n+1) = 0.5 u_n + 2 < 0.5 × 4 + 2 = 4. The property therefore holds for every integer n ≥ 0.

The positive sign of 0.5 is used here: multiplying an inequality by a negative number would reverse its direction. Merely writing “assume the property is true” and arriving at a similar-looking statement is not an induction step. You must show exactly how the assumption about n produces the conclusion about n + 1.

Induction is not circular reasoning. The base case establishes one case without assuming the general conclusion; the induction step

describes a valid transition from any case that has already been established. Removing either element destroys the argument.

5. Differentiating to understand change

The quotient [f(x + h) − f(x)] / h describes an average rate of change when h is nonzero. If its limit exists as h tends to zero, that limit defines f'(x). Geometrically, the derivative gives the slope of the tangent; in a position model, it can represent velocity. The derivative has the units of f divided by the units of x.

On an interval, the sign of the derivative provides information about monotonicity under the usual differentiability assumptions. Finding the zeros of f' is not, however, enough to determine a global maximum. You must examine the sign, the domain and, depending on the problem, the endpoints. A point where f' is zero may also be a minimum or may not correspond to any extremum.

For a composite function g(u(x)), the derivative is g'(u(x)) × u'(x). Thus the derivative of e^(−2x) is −2 e^(−2x), not simply e^(−2x). The factor associated with the inner function is often the part that is forgotten. Checking the expected behaviour helps here: a decreasing exponential must have a negative derivative.

6. Optimising under a constraint

A rectangle has a perimeter of 20 m. If its sides are x and y, then x + y = 10. Its area is therefore not a function of two independent variables: A(x) = x(10 − x), with 0 ≤ x ≤ 10 if the limiting degenerate cases are included in the mathematical study.

The derivative A'(x) = 10 − 2x is positive before 5 and negative after 5. The maximum area is A(5) = 25 m²; the corresponding rectangle is a square. Another proof comes from A(x) = 25 − (x − 5)². These approaches are compatible: one uses monotonicity, the

other an algebraic property. Comparing suitable methods enriches the reasoning without creating an artificial disagreement.

The answer must still be interpreted. If an additional constraint required a width of less than 3 m, the previous maximum would no longer be admissible. Solving the ideal problem and then forgetting the actual constraints would mean answering a different question.

7. Existence and uniqueness are distinct results

Consider f(x) = x³ + x − 1. It is continuous on [0; 1], with f(0) = −1 and f(1) = 1. The intermediate value theorem therefore guarantees at least one solution of f(x) = 0 in this interval. Since f'(x) = 3x² + 1 > 0, the function is strictly increasing: the solution is unique.

Continuity establishes existence; strict monotonicity establishes uniqueness. The properties must not be confused. A continuous function can take the same value several times. A strictly increasing function may fail to reach the required value on a given interval.

The bisection method then gives bounds. Here f(0.5) = −0.375 and f(0.75) = 0.171875: the solution lies in [0.5; 0.75]. Each step preserves a sign change and halves the interval's length. Bounds accompanied by their method are more informative than a long, unjustified decimal value.

8. Logarithms and interpreting a threshold

The exponential and natural logarithm are inverses on the appropriate domains. To solve e^(−2t) = 0.1, take the logarithm of positive quantities: −2t = ln(0.1) = −ln(10), giving t = ln(10)/2, approximately 1.1513 in the model's time unit.

The solution of an equation is not always the same result as an inequality threshold. If the requirement is e^(−2t) ≤ 0.1 for t ≥ 0,

the answer is t ≥ ln(10)/2, because this exponential is decreasing. The expression ln(a + b) = ln(a) + ln(b) is generally false: the logarithm turns a product of positive factors into a sum, not a sum into a sum.

For our sequence, requiring 4 − u_n < 0.01 gives 3 × (0.5)^n < 0.01. The first admissible index is n = 9: the gap at n = 8 is 0.01171875 and at n = 9 it is 0.005859375. Since the index is an integer, a real value obtained using logarithms must still be interpreted correctly.

9. Worked exercises

A - Counterexample. “If $a^2 = b^2$, then a = b.” **Solution:** a = 1 and b = −1 refute the statement. The correct conclusion over the reals is a = b or a = −b, obtained by factoring $a^2 - b^2$.

B - Monotonicity. Study $f(x) = x^3$ over the reals. **Solution:** $f'(x) = 3x^2 \geq 0$; the function is strictly increasing, despite f'(0) = 0. You can also compare two cubes directly. A derivative that is zero at a single point does not imply an extremum.

C - Antiderivative. Find an antiderivative of 2x over the reals, then the integral of 2x between 0 and 3. **Solution:** $F(x) = x^2$ works because F' = 2x. The integral is F(3) − F(0) = 9. Adding a constant to F does not change this difference.

D - Definition. Calculate the median of 2; 3; 8; 11. **Solution:** the list is ordered and contains four values; the usual median is the mean of the two middle values, (3 + 8)/2 = 5.5. The arithmetic mean of the series is 6: the indicators do not answer exactly the same question.

10. Revise a method, not only a result

For each exercise, ask: which domain, which property, which assumptions, which transformation and which conclusion? Repeating the calculation with different values checks that the method does not depend on a memorised example. Explaining aloud why a step is allowed often reveals understanding that is still fragile.

A readable solution distinguishes the given information, the reasoning and the interpretation. Connectives such as “because”, “therefore” and “conversely” must correspond to a real logical link. Keep an error notebook with a correction and a counterexample; a list of formulas alone does not show when to use them.

11. Quantifiers: say exactly what the proof asserts

“For every x, there exists y” and “there exists y such that for every x” do not say the same thing. If each student can find a book that suits them, the book can vary from one student to another. This does not establish the existence of one and the same book that suits everyone. In mathematics, the distinction appears in properties of functions: for every real x, there exists a real y such that $y > x$; there is no real y greater than every real x. The first statement is proved by choosing $y = x + 1$; the second is refuted by taking $x = y + 1$, whatever candidate y is proposed.

To negate “all terms are positive”, it is enough to assert that at least one term is not positive. The negation of $u_n > 0$ is $u_n \leq 0$, not $u_n < 0$. Equality can be the missing counterexample. Conversely, negating “there exists a solution” means that no value in the domain satisfies the equation. That assertion requires a proof covering the entire domain; failing to find a solution on a calculator is not such a proof.

When writing, introduce the arbitrary object: “let x be a strictly positive real number”. This choice imposes no particular value and allows a general property to be proved within the domain. When a constant depends on x, say so. Quantifiers prevent a construction suited to one case from becoming a uniform promise. This precision is useful for convergence thresholds, probabilities and optimisation problems.

12. Choosing between a direct proof, contraposition and contradiction

A direct proof begins with the assumptions and obtains the conclusion. To prove that $x > 1$ implies $x^2 > x$, write $x^2 - x = x(x - 1)$, then observe that both factors are positive. The contrapositive states the same logical relationship in another direction: if $x^2 \leq x$, then $x \leq 1$. You must preserve the exact negation and the domain. The converse, “if $x^2 > x$, then $x > 1$”, is false, because $x = -1$ satisfies the assumption but not the conclusion.

A proof by contradiction temporarily assumes the negation of the conclusion and seeks an inconsistency with the assumptions or an established property. Reaching a surprising statement is not enough. For example, to show that there is no largest natural number, assume that such a number N exists. The integer $N + 1$ is still a natural number and is greater than N: this is the required contradiction. The reasoning also explains the result, because it constructs a larger candidate than any proposed maximum.

A solution can use several methods, but every transition must be identifiable. Beginning with “assume the property is true” without specifying whether you are preparing induction, contraposition or contradiction leaves the reader to guess the assumption's role. A successful proof sets out that role, rather than presenting a sequence of calculations already familiar to its author.

Additional exercise. For real numbers a and b, consider the statement: “if ab = 0, then a = 0”. Find a precise correction.
Solution: a = 2 and b = 0 provide a counterexample. The valid conclusion is a = 0 or b = 0. If b ≠ 0 is added to the assumptions, the conclusion a = 0 becomes correct by division. The property was not completely unreasonable; it had lost an essential assumption or alternative. Learning to repair a statement develops understanding more than simply classifying a sentence as true or false.

Chapter 2 - Sequences and induction: understand discrete change

Historical perspective - Fibonacci: a recurrence comes from assumptions

Leonardo of Pisa, known as Fibonacci, wrote Liber abaci in 1202. It contains a fictional rabbit-breeding problem associated with the sequence that now bears his name. Its relevance to this chapter is the move from assumptions to a recurrence rule, followed by an examination of what that rule allows us to conclude. Such a model does not become a biological forecast merely because its first terms are easy to calculate. This historical example means neither that Fibonacci invented every sequence of this form nor that every natural phenomenon follows this recurrence.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Fibonacci/

1. Identify the index and the rule for change

A sequence is a function defined on a set of integers, often the natural numbers. The index n is not the value u_n: it identifies a stage. When u_n describes a quantity after n operations, u_0 describes its state before the first operation. A mistake at the starting point shifts the whole interpretation without necessarily making the calculations visibly wrong. Begin with a sentence explaining the index and the initial condition.

An explicit definition, such as u_n = 3n + 2, allows the term with index 100 to be calculated directly. A recurrence definition, such as u_(n+1) = 0.8u_n + 5, requires an initial value and builds the change step by step. The second expression does not allow u_n to be replaced by n: u_(n+1) = 0.8n + 5 would define a different

sequence. The rule acts on the current value, not directly on the number of the stage.

A table of the first terms and a graph can suggest a pattern. They help identify a possible sign or bound, but do not constitute a proof for infinitely many terms. Rounding on a calculator can also hide slow divergence or a very small oscillation. Observations therefore guide the investigation; algebraic arguments and theorems establish the result.

2. Arithmetic sequences: add a constant

An arithmetic sequence satisfies u_(n+1) = u_n + r, where r is a constant called the common difference. With first term u_0, its explicit form is u_n = u_0 + nr. This can be proved by induction: the case n = 0 is immediate; if the formula holds at index n, then u_(n+1) = u_0 + nr + r = u_0 + (n + 1)r. The property is thus passed on to the next term.

The sign of r determines the direction of change: a positive common difference produces strict increase; a negative one, strict decrease; a zero one, a constant sequence. A constant increase of 5 units differs from an increase of 5%. In the first case, we add 5; in the second, we multiply by 1.05. This distinction must be understood before seeking a formula, particularly when everyday language uses “regular increase” for both situations.

The sum of N consecutive terms of an arithmetic sequence is N times the average of the first and last terms. There are n + 1 terms from u_0 to u_n inclusive; their sum is (n + 1)(u_0 + u_n)/2. One proof writes the sum in forward and reverse order: each pair of corresponding terms has the same sum u_0 + u_n. This proof explains the factor 1/2 and avoids confusing the last index with the number of terms.

Fictional example: a series of lengths is 4; 7; 10; ...; 34. The common difference is 3. Solving $4 + 3n = 34$ gives $n = 10$, so there are eleven terms, not ten. The sum is $11 \times (4 + 34)/2 = 209$. Checking small cases provides a useful safeguard: for $n = 0$, the formula must give u_0, not zero.

3. Geometric sequences: multiply by a constant

A geometric sequence satisfies $u_{n+1} = qu_n$. If it starts at index zero and $q \neq 0$, its explicit form is $u_n = u_0 q^n$ for $n \geq 0$. When $q = 0$, we keep the given initial term u_0, then $u_n = 0$ for every $n \geq 1$; treating this case separately avoids introducing 0^0. The common ratio q can be positive, zero or negative; its value and the sign of u_0 must be examined before stating how the sequence changes. For $u_0 > 0$ and $q > 1$, the sequence increases. For $u_0 > 0$ and $0 < q < 1$, it decreases towards zero. A negative ratio causes alternating signs when u_0 is nonzero.

A decrease of 20% at each stage corresponds to $q = 0.80$. After two stages, the value has been multiplied by $0.80^2 = 0.64$: the total decrease is 36%, not 40%. Successive percentages apply to different bases. A fictional initial quantity of 60 units becomes 48, then 38.4. This is a calculation exercise, not a forecast of economic, biological or medical change.

For q different from 1, the sum $1 + q + \ldots + q^n$ equals $(1 - q^{n+1})/(1 - q)$. To establish this, multiply the sum S by q, then subtract: $S - qS = 1 - q^{n+1}$. Division is valid because $q \neq 1$. For $q = 1$, the sum is simply $n + 1$. Separating this case is not a formality: the previous formula would have a zero denominator.

The sum of the terms from u_0 to u_n is therefore $u_0(1 - q^{n+1})/(1 - q)$. For $u_0 = 2$ and $q = 1/2$, the first four terms give $2 + 1 + 0.5 + 0.25 = 3.75$. The formula gives $2(1 - (1/2)^4)/(1/2)$

= 3.75. This check confirms both the number of terms and the exponent n + 1.

4. Induction: formulate a property that can be passed on

A proof by induction concerns a proposition P(n) for each integer in a domain. It consists of a base case, an induction step and a conclusion. In the induction step, we fix an arbitrary integer n in the domain, assume P(n) and prove P(n + 1). We do not simultaneously assume the conclusion at every index. The assumption plays a local and conditional role.

The form of P(n) sometimes needs strengthening. To prove only u_n < 5, it may be more useful to prove 0 ≤ u_n < 5, because a lower bound then allows a square root to be taken or a sign to be preserved. Strengthening an induction hypothesis does not mean arbitrarily adding whatever makes the calculation convenient: every component must be proved both at the start and in the step to the next index.

Here is an original proof of an identity: for every n ≥ 0, 1 + 3 + ... + (2n + 1) = (n + 1)². At index zero, both sides equal 1. Assume the identity holds at index n. At index n + 1, we add the term 2n + 3: (n + 1)² + 2n + 3 = n² + 4n + 4 = (n + 2)². The equality is passed on. A diagram of nested squares can illustrate this growth, but the algebraic proof covers every index.

Frequent mistakes include a base case at the wrong index, a calculation that never uses the hypothesis, and a conclusion without a stated domain. A property such as 2^n ≥ n can have a short proof; this does not justify omitting the stages. The writing can be concise while remaining complete: identifying the base case and explicitly showing the step is often enough.

5. Study change without an automatic method

To compare two consecutive terms, we can study $u_{n+1} - u_n$. A positive difference at every index establishes increase. If all the terms are strictly positive, we can also study the ratio u_{n+1}/u_n and compare it with 1. Positivity is essential to this second strategy; a negative ratio or a zero term requires another analysis.

An associated function can also help. For $u_n = f(n)$, if f is increasing on an interval containing the indices under consideration, the sequence is increasing. The converse is not a general rule: a function can vary between integers without preventing its integer values alone from increasing. Do not therefore replace a sequence with a function and assert an equivalence without justification.

Consider $u_n = n/(n + 1)$, for $n \geq 0$. The difference is $(n + 1)/(n + 2) - n/(n + 1) = 1/[(n + 1)(n + 2)] > 0$. The sequence increases and remains strictly below 1. The expression $u_n = 1 - 1/(n + 1)$ gives a second, immediate approach. The two methods reveal different structures: a local comparison between neighbours and an explicit distance from the limit.

6. Affine recurrences: shift the origin of the values

For $u_{n+1} = au_n + b$, with $a \neq 1$, the fixed value L solves $L = aL + b$, so $L = b/(1 - a)$. Setting $v_n = u_n - L$ transforms the relation into $v_{n+1} = av_n$. Thus $u_n = L + (u_0 - L)a^n$. The new variable measures the distance from the fixed point; it makes the geometric structure visible.

Take $u_0 = 10$ and $u_{n+1} = 0.75u_n + 3$. The fixed point is 12 and $u_n = 12 - 2(0.75)^n$. The sequence is increasing, remains below 12 and converges to 12. If we chose $u_0 = 14$ with the same rule, we would obtain $12 + 2(0.75)^n$: a decrease towards the

same limit. The rule alone therefore does not always determine the direction of change; the initial condition also determines it.

With a = −1/2 and L = 4, the distance alternates in sign and its absolute value decreases. A sequence can converge without being monotonic. With a = 2, a nonzero initial distance doubles at each stage and moves away from the fixed point. Yet if u_0 = L, the sequence is constant. Saying “a > 1 always implies divergence” would overlook this exceptional case. The chapter on limits will clarify these distinctions.

7. A sequence defined using a square root

Study u_0 = 0 and u_(n+1) = sqrt(u_n + 2). First, we must ensure that every term is defined. Prove by induction that 0 ≤ u_n ≤ 2. This holds initially. If it holds at index n, then 2 ≤ u_n + 2 ≤ 4, so sqrt(2) ≤ u_(n+1) ≤ 2. The expression remains defined and the next term satisfies the stated bound.

The function f(x) = sqrt(x + 2) is increasing on [0; 2]. Since u_1 = sqrt(2) > u_0, induction gives u_(n+1) ≥ u_n: if u_n ≥ u_(n−1), applying f preserves the inequality. We can also compare f(x) and x for 0 ≤ x ≤ 2. Both quantities are nonnegative; f(x) ≥ x is then equivalent to x + 2 ≥ x², or (x − 2)(x + 1) ≤ 0, which holds on this interval.

The sequence is increasing and bounded above: the monotone convergence theorem ensures that it has a finite limit. Using the continuity of the square root, this limit L satisfies L = sqrt(L + 2). Since L lies in [0; 2], we can square both sides and obtain L² − L − 2 = 0. The candidates are 2 and −1; only 2 belongs to the domain. The fixed-point equation alone was not enough to prove convergence, and its domain eliminates an extraneous algebraic solution.

8. Thresholds, integers and algorithms

For the fictional quantity u_n = 60(0.80)^n, find the first index at which u_n < 5. We obtain (0.80)^n < 1/12. The logarithm is increasing on positive numbers, so n ln(0.80) < ln(1/12). Dividing by ln(0.80), which is negative, reverses the inequality: n > ln(1/12)/ln(0.80), approximately 11.136. The first admissible integer is 12.

The final check establishes minimality: u_11 is approximately 5.154, so it does not satisfy the strict threshold; u_12 is approximately 4.123, so it does. Rounding n to the nearest integer would be the wrong method here. The question requires an integer satisfying an inequality, not a symmetric approximation to a real number.

An algorithm can proceed without a logarithm: initialise n = 0 and u = 60, then, while u ≥ 5, replace u by 0.80u and n by n + 1. The state u = 60(0.80)^n is a loop invariant. Convergence to zero ensures here that the threshold will eventually be reached. For another sequence, a similar loop might never stop; its termination must therefore be discussed, not merely its syntax.

9. Worked exercises

A - Shifted index. An arithmetic sequence starts with u_1 = 7 and has common difference 4. **Solution:** u_n = 7 + 4(n − 1), for n ≥ 1. The term u_6 is 27. Writing 7 + 4n would count an extra increase. The sum of the first six terms is 6(7 + 27)/2 = 102.

B - Ratio. Is the sequence u_n = (−1/2)^n decreasing? **Solution:** no; its terms alternate about zero. Its absolute value (1/2)^n decreases and tends to zero. The ratio of consecutive terms is −1/2 and does not permit the criterion based on comparison with 1 for positive terms.

C - Affine formula. u_0 = 3 and u_(n+1) = 2u_n − 1. **Solution:** L = 1, then u_n = 1 + 2×2^n. The first term is indeed 3; the next is 5. The sequence diverges to +∞. A solution of L = 2L − 1 is not automatically the limit of every sequence defined by this rule.

D - An insufficient bound. “A sequence bounded above converges.” **Solution:** (−1)^n is bounded above but does not converge. Its subsequences with even and odd indices have different limits, which makes a common limit impossible. A sufficient assumption would be that the sequence is increasing and bounded above: the monotone convergence theorem then ensures a finite limit. A sequence that is only decreasing and bounded above could tend to −∞, such as u_n = −n; monotonicity without its direction and the appropriate bound is therefore insufficient.

10. Revision and references

Be able to obtain an explicit formula from an affine relation, prove a bound by induction, and justify the first index at which a threshold is reached. For every result, distinguish observation, a proof of the direction of change, a proof of convergence and identification of the limit. These stages answer different questions and may require different arguments.

The progression follows the published official programme for the French Terminale mathematics specialism: https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm, annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. All examples, developed proofs and exercises are original LFDP writing. No real dataset, forecast of returns or reproduction of an examination paper is presented.

Chapter 3 - Limits: establish behaviour without confusing calculation with proof

Historical perspective - Cauchy: check the conditions of analysis

Cauchy's Cours d'analyse was published in 1821 for teaching at the École Polytechnique. It forms part of the historical effort to state results in analysis rigorously. The useful connection with this chapter is attention to conditions: recognising an expression or observing a trend is not enough to justify an operation on limits. The modern tools presented here must be read together with their domains and assumptions. This perspective identifies a historical stage; it does not equate every formulation used by Cauchy with today's definitions and theorems.

Historical source: Bibliothèque nationale de France - notice bibliographique : https://catalogue.bnf.fr/ark:/12148/cb302073247 ; University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Cauchy/

1. What approaching a value means

A sequence converges to L when its terms eventually remain as close to L as we wish. More precisely, for every accuracy ε > 0, there is an index N such that, for every n ≥ N, |u_n − L| < ε. The index may depend on the accuracy required. It is not enough for one term to be very close to L: all subsequent terms must satisfy the condition from some index onwards.

This formulation explains why a limit need not be a value that is reached. For u_n = 1 − 1/(n + 1), no term equals 1, but the distance is 1/(n + 1) and becomes arbitrarily small. Conversely, a sequence can take exactly the same value at infinitely many indices without converging to it: (−1)^n equals 1 at every even

index and −1 at odd indices. The frequency of a value does not replace a property concerning all sufficiently large indices.

A sequence cannot have two different finite limits. To understand the idea, if L and M are distinct, choose neighbourhoods with radii smaller than one third of |L − M|. They are disjoint. The terms cannot all be in both neighbourhoods from the same sufficiently large index onwards. This uniqueness allows us to speak of “the limit”, but only after establishing its existence.

2. A first proof from the definition

For u_n = 1/(n + 1), prove convergence to zero. The distance from zero is exactly 1/(n + 1). For an accuracy ε > 0, choosing an integer N such that N + 1 > 1/ε ensures that, for every n ≥ N, 1/(n + 1) < ε. The proof directly meets the definition and shows how to construct a suitable index. It does not rely on a diagram that appears to approach the axis.

For u_n = 3 + 2/(n + 1), the limit is 3 because |u_n − 3| = 2/(n + 1). An index satisfying N + 1 > 2/ε is suitable. The coefficient changes the rate at which an accuracy is reached, not the limit. Two sequences can have the same limit and very different distances from it at each stage; do not confuse the destination with the rate of convergence.

Terminale exercises do not always require this formal presentation. Understanding the quantifiers nevertheless helps us assess an argument. “For ε = 0.01, the terms are close” establishes at most one particular accuracy. Convergence concerns every positive accuracy, including much more demanding ones. “There is an index at which the distance is small” also forgets that the property must continue to hold beyond that index.

3. Divergence to infinity and absence of a limit

A sequence tends to +∞ when, for every real bound A, all its terms exceed A from some index onwards. Infinity is not a number equal to the last term; a sequence generally has no last term. The symbol describes behaviour. The definition of −∞ is analogous, with terms eventually below every bound.

The sequence 2n + 1 tends to +∞: simply choose N large enough that 2N + 1 > A. The sequence $-n^2$ tends to −∞. In contrast, (−1)^n n tends neither to +∞ nor to −∞: its terms alternate between large positive and negative values. It has no finite limit either. The word “divergent” can therefore describe several situations that a precise conclusion must distinguish.

An unbounded sequence does not necessarily tend to +∞. We must examine all sufficiently large indices, not merely a few very large values. Similarly, a bounded sequence need not converge: alternating between −1 and 1 is enough to show this. These counterexamples make a useful revision notebook because they prevent a theorem from being used with its assumptions removed.

4. Operations on limits and indeterminate forms

If u_n and v_n converge to finite limits L and M, their sum converges to L + M and their product to LM. If M ≠ 0 and the quotient is defined from some index onwards, u_n/v_n converges to L/M. The condition on the denominator is essential. The usual rules simplify a study when they apply; they do not constitute unrestricted arithmetic on infinite symbols.

The forms ∞ − ∞, 0×∞, ∞/∞ and 0/0 are indeterminate: the separate limits alone do not allow a conclusion. For example, n^2/n tends to +∞, n/n tends to 1, and n/n^2 tends to zero; all three quotients have the same form ∞/∞. Similarly, (n + 1) − n equals 1,

whereas $n^2 - n$ tends to $+\infty$. The lack of an automatic conclusion does not mean there is no method.

We need to transform the expression, reveal a dominant term, factorise or rationalise. Transformations must respect the domains. Calculation rules can then be applied to a more suitable expression. Immediately seeking a very large numerical value can suggest an answer, but does not replace reasoning and may even hide a rounding error.

5. Dominant terms in a quotient

For u_n = $(3n^2 - 2n + 5)/(2n^2 + n + 1)$, divide the numerator and denominator by n^2 for $n \geq 1$. This gives $[3 - 2/n + 5/n^2]/[2 + 1/n + 1/n^2]$. Every term involving $1/n$ or $1/n^2$ tends to zero; the denominator tends to 2, which is nonzero. The limit is therefore 3/2. The argument does not discard terms because they appear small: it establishes precisely that they tend to zero in the normalised expression.

If the numerator's degree is greater than the denominator's, the quotient may diverge, with a sign determined by the dominant terms. If it is smaller, the quotient may tend to zero. For $(-2n^3 + n)/(n^2 + 4)$, dividing by n^2 gives $(-2n + 1/n)/(1 + 4/n^2)$, hence divergence to $-\infty$. Reading the sign is essential; merely saying “higher degree, therefore infinity” is not enough.

The same transformations apply to functions when x tends to $+\infty$ or $-\infty$. For $x \to -\infty$, the sign of odd powers changes. A formula studied only at natural-number indices should not be extended to negative real numbers without examination, particularly if it contains a root, logarithm or noninteger power.

6. Rationalise a difference of square roots

Consider u_n = sqrt(n² + n) − n for n ≥ 1. Both terms tend to +∞ and their difference is indeterminate. Multiplying by the conjugate expression gives u_n = n/[sqrt(n² + n) + n]. Dividing by n, which is positive, gives 1/[sqrt(1 + 1/n) + 1]. The limit is 1/2. Rationalisation transforms a difficult subtraction into a quotient whose limits can be used immediately.

The sign of n was used in sqrt(n²) = |n| = n. For a variable x tending to −∞, writing sqrt(x²) = x would be wrong. For example, sqrt(x² + x) − x then behaves like −2x and tends to +∞, in the domain where the root is defined. The same algebraic appearance does not impose the same limit at both ends of the axis.

On a calculator, a large value of n can make sqrt(n² + n) and n so close in relative terms that rounding loses part of their difference. The rationalised form is often more numerically stable. This connection between symbolic reasoning and numerical reliability shows that simplification is not merely a convenience on paper: it can improve the quality of the actual calculation.

7. Comparison and bounding

The squeeze theorem states that, if a_n ≤ u_n ≤ b_n from some index onwards and a_n and b_n have the same finite limit L, then u_n converges to L. The condition “same limit” is essential. Bounding a sequence between 0 and 1 does not prove its convergence; it establishes only a bound.

For u_n = (−1)^n/(n + 1), we have −1/(n + 1) ≤ u_n ≤ 1/(n + 1). Both bounds tend to zero, so the sequence converges to zero. It is not monotonic, but its oscillation decreases. This example clearly distinguishes convergence from monotonicity. A bound on the absolute value, |u_n| ≤ 1/(n + 1), gives the same conclusion in more compact notation.

For infinite limits, a comparison is sometimes sufficient. If u_n ≥ v_n from some index onwards and v_n tends to +∞, then u_n also tends to +∞. Thus 5n + (−1)^n ≥ 5n − 1 shows divergence to +∞ despite a small oscillation. An upper bound tending to +∞ would not allow the same conclusion: a constant or decreasing sequence would still be possible.

8. Bounded monotonic sequences and fixed points

An increasing sequence bounded above converges to a finite limit. A decreasing sequence bounded below also converges. The bounds do not necessarily give the value of that limit. If u_n ≤ 10, there is no reason to conclude that its limit is 10. The theorem ensures convergence; identifying the limit requires an additional argument.

For a relation u_(n+1) = f(u_n), if the sequence converges to L and f is continuous at L, then L = f(L). This equation supplies candidates. It proves neither convergence nor the absence of other candidates. The sequence u_(n+1) = 2u_n with u_0 = 1 diverges, although the equation L = 2L has the solution L = 0. The existence of a fixed point does not mean every trajectory reaches or approaches it.

Another trap is choosing a bound merely to simplify the calculation. The bound must be proved at every index. For an increasing sequence, checking u_0 < 5 does not ensure that every term remains below 5. Induction or an explicit expression must establish that the interval is preserved before the monotone convergence theorem is invoked.

9. Limits of functions and one-sided limits

For a function, x can approach a value a without equalling it. The limit of f(x) as x tends to a concerns nearby values in the domain. For $f(x) = (x^2 - 1)/(x - 1)$, defined for $x \neq 1$, factorisation gives $f(x) = x + 1$ on that domain. The limit at 1 is 2, although f(1) is undefined. A continuous extension can be constructed by assigning the value 2 at that point; this is a new definition, not a value already present.

The left-hand and right-hand limits can differ. For $f(x) = 1/(x - 2)$, the denominator is negative when x approaches 2 from below and positive when it approaches from above. The left-hand limit is $-\infty$, and the right-hand limit is $+\infty$. The line $x = 2$ is a vertical asymptote, but there is no single infinite limit without specifying the side.

For $f(x) = 2 + 1/x$, the limit at $+\infty$ is 2: $y = 2$ is a horizontal asymptote in that direction. A curve is not required to remain on just one side of a horizontal asymptote for every function. An asymptote expresses a distance tending to zero, not a physical barrier that the curve can never cross.

An oblique asymptote can be identified by showing $f(x) - (ax + b) \to 0$. For $f(x) = (x^2 + 1)/(x - 1)$, division gives $f(x) = x + 1 + 2/(x - 1)$. The final term tends to zero at infinity; $y = x + 1$ is therefore an oblique asymptote. The function itself does not converge to a finite number. The distance from a line and the value of the function are two different objects.

10. Exponentials and comparative growth

For $0 < q < 1$, q^n tends to zero. An accessible justification uses $1/q = 1 + h$ with $h > 0$. Bernoulli's inequality, $(1 + h)^n \geq 1 + nh$, which can be proved by induction, gives $0 < q^n \leq 1/(1 + nh)$. The upper bound tends to zero. If $-1 < q < 0$, the absolute value is $|q|^n$ and the same bounding argument still ensures convergence to zero, despite the alternating signs.

When $q > 1$, q^n tends to $+\infty$. For functions, the usual comparative-growth results say, in particular, that $\ln(x)/x \to 0$ as $x \to +\infty$ and that the exponential dominates powers at $+\infty$. Here these results are course properties to know; they are not “proved” by a table of values. They can be used to establish $e^x/x^2 \to +\infty$ or $x^2e^{-x} \to 0$, after examining the domain and rewriting the product as a quotient.

There is no need to introduce an advanced differentiation rule for ratios when calculating limits to handle all these cases. Using a result not yet studied without justification can conceal the assumptions and the expected method. Factorisation, comparisons and course properties already offer a wide range of reasoning suited to Terminale.

11. Worked exercises

A - Quotient. Find the limit of $(4n^2 + 3)/(n^2 - n + 2)$. **Solution:** divide by n^2: $[4 + 3/n^2]/[1 - 1/n + 2/n^2] \to 4$. The denominator tends to 1, which permits the quotient rule for limits.

B - Damped alternation. $u_n = 7 + (-1)^n/n$ for $n \geq 1$. **Solution:** $|u_n - 7| = 1/n \to 0$, so $u_n \to 7$. The sequence is not monotonic; this does not prevent convergence.

C - Conjugate. v_n = sqrt(n² + 4n) − n. **Solution:** v_n = 4/[sqrt(1 + 4/n) + 1], so the limit is 2. The factor 4 is not discarded: it remains in the numerator after normalisation.

D - Limit and value. f(x) = x + 2 for x ≠ 0 and f(0) = 10. **Solution:** the limit at zero is 2, but the value is 10. The function is not continuous at zero. Nearby values determine the limit; the value chosen at the point also determines continuity.

E - Fixed point. u_0 = 2 and u_(n+1) = −u_n. **Solution:** the terms alternate between 2 and −2. The equation L = −L gives 0, but the sequence does not converge. We cannot take limits in a relation before establishing, or explicitly assuming, that this limit exists.

12. Revision and references

For every study, state the direction or side of the limit, check the domain and look for a possible indeterminate form. Then choose a transformation, followed by a theorem whose assumptions have been verified. A conclusion distinguishes a finite limit, divergence to an infinity and absence of a limit. Rounding and graphs are checks, not replacements for these stages.

Progression framework: the published official programme for the French Terminale mathematics specialism, https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and its annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. Examples, developed proofs and exercises are original LFDP writing; no third-party textbook text, diagram or examination paper is reproduced.

Chapter 4 - Continuity: prove existence, uniqueness and bounds

Historical perspective - Bolzano: existence requires a proof

In a paper published in 1817, Bernard Bolzano gave a proof of the intermediate value theorem while seeking an analytical justification. This fact illustrates the difference between seeing a curve cross a level and establishing that a solution exists under precise assumptions. Our chapter applies that distinction to continuity, then separates existence, uniqueness and numerical approximation. The historical perspective does not replace a proof or allow an assumption to be forgotten; it invites us to ask exactly why the result holds.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Bolzano/

1. Three questions that must not be confused

When an equation f(x) = k appears, three questions arise: does it have at least one solution, does it have only one, and how can that solution be approximated? A calculator displaying a number sometimes answers the third without having answered the first two. An existence proof does not necessarily provide an exact expression. A uniqueness proof does not automatically prove that a solution exists. Separating these questions makes the solution process safer and clearer.

Continuity plays a central role in existence. Strict monotonicity can ensure uniqueness on an interval. A numerical method then constructs bounds or an approximation. Reasoning therefore complements calculation: it identifies what the calculation is seeking and why it can succeed. The last number displayed must

not be presented as an exact solution when the algorithm has provided only an approximate value.

The domain is part of the problem. A function containing a root or logarithm imposes restrictions; a denominator can exclude a point. The intermediate value theorem requires continuity throughout the interval under consideration. Checking only the values at the two endpoints does not replace this condition. A discontinuity or a missing point can prevent the argument from working.

2. Continuity at a point and on an interval

A function f is continuous at a point a in its domain when the limit of f(x) as x tends to a equals f(a), with the approach adapted to the domain. This definition connects nearby values with the value at the point. If f(a) is undefined, continuity at a cannot be asserted; a continuous extension may perhaps be constructed after studying the limit.

Polynomial functions are continuous on the real numbers. A rational function is continuous wherever its denominator is nonzero. The exponential is continuous on the real numbers; the logarithm, on the strictly positive real numbers. Sums, products and compositions of continuous functions are continuous in their appropriate domains. A quotient requires a nonzero denominator. This toolkit allows continuity to be justified without returning to its definition in every exercise.

The intuitive image of drawing “without lifting the pencil” is useful but limited. It depends on the diagram's scale and does not constitute a general definition. A curve calculated from a few points can hide a hole or very rapid change. Conversely, a piecewise-defined function can be continuous if its values and limits join correctly. The name of the formula alone does not decide the property.

3. Joining a piecewise-defined function

Take $f(x) = x^2$ for $x \leq 1$ and $f(x) = ax + b$ for $x > 1$. To the left of 1, the limit is 1; the value $f(1)$ is also 1. To the right, the limit is $a + b$. Continuity at 1 therefore requires $a + b = 1$. Infinitely many pairs satisfy this condition. Requiring continuity alone does not necessarily determine all the parameters.

If differentiability at the join is also required, the derivative from the left is 2 and that from the right is a; we need $a = 2$, then $b = -1$. This additional condition answers a distinct question. A function can be continuous without being differentiable, such as $|x|$ at zero. Its graph has no jump, but the slopes on the left and right differ.

In modelling, continuity can express that a quantity does not change its value instantaneously when crossing a threshold. Differentiability can express a transition in slope. These properties must nevertheless be justified by the real problem; they are not universal aesthetic requirements. A price structure defined in steps, for example, can intentionally have jumps. A model should describe what is relevant, not merely a curve that is pleasant to look at.

4. The intermediate value theorem

If f is continuous on a closed interval $[a; b]$, every value k between $f(a)$ and $f(b)$ is attained at least once on that interval. To seek a root, take $k = 0$. If $f(a)$ and $f(b)$ have opposite signs, continuity guarantees at least one c such that $f(c) = 0$. If one endpoint value is already zero, that endpoint is itself a solution.

The theorem asserts existence, not necessarily uniqueness. For $f(x) = x^2 - 1$ on $[-2; 2]$, both endpoint values are positive; nevertheless, the function has two roots. A change of sign at the endpoints is a sufficient condition for a root in the usual case, not a

necessary condition for every instance in which a root exists. Do not infer “no sign change, therefore no solution”.

Continuity is essential. A function equal to -1 for $x < 0$ and to 1 for $x \geq 0$ changes sign between -1 and 1 without taking the value zero. The function $1/x$ also has opposite signs on either side of zero, but is not defined throughout that interval. These examples refute an incorrect version of the theorem that retains only the endpoint signs.

5. Add uniqueness through monotonicity

If f is strictly monotonic on an interval, it cannot take the same value at two different points of that interval. Thus continuity and bounds around a value give existence, whereas strict monotonicity gives uniqueness. The justifications can be short, but their roles must remain distinct in the written argument.

Study $f(x) = e^x - 3x$ on $[0; 1]$. The function is continuous, $f(0) = 1$ and $f(1) = e - 3 < 0$. There is therefore a root in $(0; 1)$. Its derivative $f'(x) = e^x - 3$ is strictly negative on $[0; 1]$, because $e^x \leq e < 3$. The function is strictly decreasing; the root is unique there. No exact elementary expression is required to establish these two results.

On a larger interval, the same equation may have another solution. Here $f(1) < 0$ and $f(2) = e^2 - 6 > 0$ guarantee at least one root between 1 and 2. Uniqueness proved on $[0; 1]$ therefore does not mean uniqueness over all the real numbers. The domain of the conclusion must be stated as carefully as the domain of the function.

A derivative that is strictly positive everywhere is a sufficient condition for strict increase, not a necessary one. The function x^3 is strictly increasing although its derivative is zero at zero. For $a < b$, the difference $b^3 - a^3 = (b - a)(a^2 + ab + b^2)$ is positive: the first

factor is positive and the second cannot be zero for two distinct real numbers. This algebraic argument avoids an unjustified generalisation from a single point where the derivative vanishes.

6. The image of an interval and extrema

Continuity ensures that the image of an interval contains no missing value between two attained values. It does not mean that the extreme values are necessarily those at the two endpoints. For $f(x) = x^2$ on $[-2; 3]$, the endpoint values are 4 and 9, but zero is attained at $x = 0$. The image is $[0; 9]$, not $[4; 9]$. Changes or extrema must be studied to determine the entire image.

A continuous function on a closed, bounded interval attains a minimum and a maximum. This result cannot be transferred to an open interval without examination. The function $f(x) = x$ on $(0; 1)$ is bounded but attains neither zero nor one; it therefore has neither a minimum nor a maximum on that domain. The lower and upper bounds exist as numbers without necessarily being values of the function.

This distinction is useful in optimisation. A calculation may identify a value outside the domain or at an excluded endpoint. The answer must then state whether an attained maximum exists or merely an upper bound that can be approached. Replacing an open interval with a closed interval to simplify the study changes the problem; that modification must be justified rather than made silently.

7. Bisection: preserve an invariant

Bisection seeks a root of a continuous function in an interval whose endpoint values have opposite signs. Calculate the midpoint m = (a + b)/2. If f(m) = 0 exactly, the problem is solved. Otherwise, retain the half whose endpoint values still have opposite signs. Each stage halves the interval's length and preserves the existence of at least one root in the retained interval.

The sign change is the method's logical invariant. If the root is unique, the successive intervals always enclose it. After N halvings of an initial interval of length L, the length is L/2^N. Taking the final midpoint as an approximation gives an absolute error of at most L/(2^(N+1)), under the exact assumptions of the mathematical model. The criteria “interval width” and “midpoint error” must be distinguished.

To obtain a width less than or equal to 0.001 from an interval of length 1, we need 2^N ≥ 1000, so N = 10 halvings suffice: 1/1024 ≈ 0.0009766. Nine halvings give approximately 0.001953 and are insufficient. The count depends on when the width is measured; clearly stating an initialisation and a stage avoids an index shift.

8. Explain a first execution

Return to f(x) = e^x − 3x on [0; 1], where the root is unique. At the midpoint 0.5, f is approximately 0.1487, hence positive. Since f(1) is negative, retain [0.5; 1]. The new midpoint 0.75 gives approximately −0.1330, so retain [0.5; 0.75]. At the midpoint 0.625, f is approximately −0.00675; retain [0.5; 0.625].

The interval should not be chosen because the midpoint “seems close” to zero. The sign and continuity justify which half is retained. A small value of f(m) alone does not ensure that m is very close to the root: if the function is nearly flat, a large horizontal distance

can correspond to a small vertical distance. The bounds provide geometric information different from the residual f(m).

A program can test for “same signs” between f(a) and f(m). In exact arithmetic, a positive product indicates this case. In floating-point arithmetic, extremely large or small values can overflow or be rounded to zero; comparing signs separately can be more robust. This observation does not require mastery of all numerical analysis, but encourages us not to confuse a computer operation with an exact proof.

9. Other methods and reasons to compare them

Newton's method uses the tangent: $x_{n+1} = x_n - f(x_n)/f'(x_n)$, when the denominator is nonzero. It can converge very quickly in favourable conditions, but a poor starting point can lead to divergence, a cycle or a value outside the domain. It is presented here as additional insight, not an automatic replacement for bisection or for the teacher's expectations.

For $f(x) = x^3 - 2x + 2$ and $x_0 = 0$, Newton gives $x_1 = 0 - 2/(-2) = 1$. Then $f(1) = 1$ and $f'(1) = 1$, so $x_2 = 0$. The values alternate between 0 and 1 without approaching a root. Software that stops after a few iterations can display a number without having obtained the solution. A method that is faster in favourable circumstances is therefore not a universally more reliable method.

Bisection is slow compared with some local methods, but its bounds are easy to guarantee when the assumptions are respected. Methods can be combined, for example by using reliable bounds and a controlled faster step. This comparison illustrates a general idea: numerical efficiency, mathematical guarantees and ease of verification are distinct criteria.

10. From an equation to an inequality

A root can organise the sign of a function. If f is strictly decreasing on [a; b] and has a unique root α in that interval, f is positive before α and negative after it. Solving f(x) ≥ 0 on [a; b] then gives [a; α]. The root can be described exactly as “the unique solution of...” even if only decimal bounds are known.

For f(x) = e^x − 3x on [0; 1], the previous result provides this sign structure. Do not replace α with a rounded approximation in an answer described as exact. The exact set can first be given using α, followed by bounds for interpretation. This separation avoids including or excluding values near the threshold solely on the basis of rounding.

A nonmonotonic function requires a more complete sign table. Several roots can divide the domain into several intervals; signs do not always alternate when a root is repeated. For example, $(x - 1)^2$ vanishes at 1 without becoming negative. The existence theorem and the roots therefore do not replace a study of the expression's structure.

11. Worked exercises

A - Continuity at a join. f(x) = 2x + 1 for x < 2 and f(x) = x² + c for x ≥ 2. **Solution:** the left-hand limit at 2 is 5; the value and right-hand limit are 4 + c. Continuity requires c = 1. This does not ensure differentiability: the derivatives at the join would be 2 and 4.

B - Existence. $P(x) = x^3 - 4x + 1$. Show that a root lies between 0 and 1. **Solution:** the polynomial is continuous, P(0) = 1 and P(1) = −2. The intermediate value theorem ensures at least one root. On [0; 1], $P'(x) = 3x^2 - 4 < 0$, which ensures uniqueness in this interval. This conclusion does not concern all real numbers.

C - Number of halvings. The initial interval has width 8. What width is obtained after six halvings? **Solution:** 8/2^6 = 0.125. The midpoint has a maximum error of 0.0625 relative to any root still enclosed. The function value at the midpoint was not used for this error bound.

D - An incorrectly applied IVT. $f(x) = 1/x$ and $f(-1)f(1) < 0$. **Solution:** the function is not continuous on $[-1; 1]$, since it is undefined at zero. The theorem does not apply. In fact, $1/x$ has no root anywhere in its domain.

E - Incorrectly inferred uniqueness. A continuous function has $f(a) < 0$ and $f(b) > 0$. Does it have exactly one root? **Solution:** not necessarily; it may cross the axis several times. An additional property is needed, such as strict monotonicity on the interval, or another proof of uniqueness.

12. Revision and references

For practice, write separate paragraphs for existence, uniqueness and approximation. In the first, check continuity and endpoint values; in the second, justify the property that prevents two solutions; in the third, state the invariant and accuracy. Then reconstruct an example in which each missing assumption causes a conclusion to fail.

The published official programme for the French Terminale mathematics specialism provides the progression framework: https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. The situations, formulations and exercises are original LFDP writing. The additional material on Newton illustrates the limits of a numerical method; it is neither a stated examination requirement nor a guarantee that software will work.

Chapter 5 - Differentiation and optimisation: move from local change to a conclusion

Historical perspective - Leibniz: calculate changes, then justify them

In 1684, Leibniz published an article on his differential calculus in Acta Eruditorum, including differentiation rules and the notation d. The institutional biography also notes that the article did not provide proofs of those rules. When revising today, we should therefore distinguish a computational rule from the reasons and conditions that make it valid. This chapter connects the calculation of derivatives with domains and the interpretation of signs. The date marks a publication, without attributing the invention of infinitesimal calculus exclusively to one person.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Leibniz/

1. Understand the derivative before using it

The quotient $[f(a + h) - f(a)]/h$ measures an average change between two points for $h \neq 0$. If this quotient has a finite limit as h tends to zero, f is differentiable at a and that limit is $f'(a)$. The derivative therefore expresses local information: it describes change near a point, not necessarily the overall behaviour throughout the domain.

Geometrically, $f'(a)$ is the slope of the tangent to the graph at the point with coordinates $(a; f(a))$. Its equation is $y = f(a) + f'(a)(x - a)$. The equation contains both a slope and a point; merely writing $y = f'(a)x$ generally forgets the vertical intercept. A horizontal tangent means $f'(a) = 0$; it does not automatically mean $f(a) = 0$.

If f represents a distance in metres and x a duration in seconds, the derivative has unit m/s. If f represents an area in m^2 as a function of a length in m, its derivative has unit m. Differentiation therefore changes the units according to the variable. This consistency helps detect confusion between a function and its derivative in an applied problem.

2. Differentiability and local approximation

A function differentiable at a point is continuous there, but the converse is false. The function |x| is continuous at zero and has no derivative there, because the rates of change from the left and right tend to -1 and 1. Continuity joins the values; differentiability additionally requires local change compatible with a single slope.

When f is differentiable at a, we can write $f(a + h) = f(a) + f'(a)h +$ a remainder that becomes negligible compared with h. The affine approximation $f(a + h) \approx f(a) + f'(a)h$ is therefore local. It is not a general equality, and its accuracy depends on the function and the size of h. For $f(x) = x^2$ near $a = 2$, we obtain $f(2 + h) = 4 + 4h + h^2$: the remainder is explicitly h^2.

For $h = 0.10$, the approximation gives 4.40 and the exact value is 4.41. For $h = 1$, it gives 8 rather than 9. The same linear formula becomes less accurate as we move away from the point. Without knowing or bounding the remainder, we cannot turn a tangent into a guaranteed numerical error. This caution connects a local concept with its use in approximate calculations.

3. Differentiation rules and domains

For two differentiable functions u and v, the derivative of a sum is u' + v'. The derivative of a product is u'v + uv', not u'v'. For a quotient u/v defined with v ≠ 0, the derivative is (u'v − uv')/v². A constant multiplier is retained: (ku)' = ku'. Domains must be determined before or alongside the calculation.

For a composite function g(u(x)), the derivative is g'(u(x))u'(x). The factor u' accounts for the change in the inner expression. Thus (e^(3x²))' = 6x e^(3x²). For ln(u(x)), with u(x) > 0, the derivative is u'/u. For sqrt(u(x)), the formula u'/(2sqrt(u)) can be used at points where u > 0 and the other differentiability assumptions hold. An expression defined at a point is not necessarily differentiable there.

Consider f(x) = ln(x² + 1). Its domain is all real numbers, since x² + 1 > 0, and f'(x) = 2x/(x² + 1). For g(x) = ln(x² − 1), the domain is (−∞; −1) ∪ (1; +∞). The formula g'(x) = 2x/(x² − 1) resembles the previous one, but the domains differ. A correct formal calculation on an expression does not remove the need to define where it is allowed.

4. From the derivative's sign to the function's behaviour

On an interval where f is differentiable, a nonnegative derivative ensures a nondecreasing function; a nonpositive derivative ensures a nonincreasing function. A strictly positive derivative ensures strict increase. Exact formulations and any isolated points with zero derivative must be examined before concluding strict monotonicity. A complete study can use a sign table and additional arguments.

The interval matters. A rational function may have a derivative of the same sign on two parts of its domain separated by a forbidden value; this is not enough to declare it monotonic on their union. For f(x) = 1/x, the derivative is negative on each interval (−∞; 0) and

(0; +∞). Yet comparing x = −1 and x = 1 across the disconnected set gives f(−1) < f(1), which does not correspond to global decrease.

A zero derivative at a point does not guarantee an extremum. For x^3, the derivative vanishes at zero but the function continues to increase. A change in the sign of f' from positive to negative produces a local maximum; the reverse change produces a local minimum. We must then compare other candidate values before identifying a global maximum or minimum on the required domain.

5. A complete study of a function

Take f(x) = (x + 1)e^(−x). It is defined and differentiable on the real numbers. By the product rule, f'(x) = e^(−x) − (x + 1)e^(−x) = −xe^(−x). Since the exponential is always positive, the derivative has the sign of −x. The function is strictly increasing on (−∞; 0] and strictly decreasing on [0; +∞). Its global maximum is f(0) = 1.

At +∞, the product (x + 1)e^(−x) tends to zero by comparative growth, remaining positive for x > −1. At −∞, x + 1 tends to −∞ and e^(−x) to +∞; the product tends to −∞. The function vanishes at x = −1. These facts organise a consistent graph: passing through (−1; 0), rising to (0; 1), then decreasing towards the horizontal asymptote y = 0 on the right.

If the problem restricts x to [0; 4], the maximum remains 1 and the minimum is f(4) = 5e^−4, approximately 0.0916. A result on a restricted domain must be studied together with its endpoints. A globally relevant critical point may be excluded by a constraint; conversely, an endpoint can become the location of an optimum even when the derivative there is nonzero.

6. Construct an objective function under a constraint

Optimisation begins before differentiation. Define the quantity to be optimised, the variables, their units and the constraints. For an open box with a square base of side x and height h, a prescribed volume V gives $x^2h = V$, hence $h = V/x^2$ for $x > 0$. The idealised material surface area is $A = x^2 + 4xh$, or $A(x) = x^2 + 4V/x$. The base and four sides are counted; this model has no lid.

For $V = 32$ dm³, $A(x) = x^2 + 128/x$ in dm², with x in dm. Its derivative is $A'(x) = 2x - 128/x^2 = 2(x^3 - 64)/x^2$. The denominator is positive, so the derivative is negative for $0 < x < 4$ and positive for $x > 4$. The minimum surface area is obtained for $x = 4$ dm; the height is then $32/16 = 2$ dm and the area is $16 + 32 = 48$ dm².

Next, interpret the real constraints given in the exercise. If x had to belong to [2; 3], the value 4 would be inadmissible and A would be decreasing throughout this interval: its minimum there would be at $x = 3$. Calculating an optimum without a domain answers an unfinished problem. Thickness, cutting, strength and material losses are absent from the model; the fictional box is not a certified manufacturing plan.

7. Local and global optimisation

A local minimum means that nearby values are greater than or equal to it; a global minimum compares with every value in the domain. A function can have several local extrema. Solving $f' = 0$ supplies interior candidates, not a universal list of extrema: nondifferentiable points and endpoints must also be examined according to the domain.

For $f(x) = x^2$ on $[-1; 2]$, the minimum is at zero and the maximum at 2. The derivative at 2 is 4, not zero. The maximum lies at an endpoint of the domain; there is no contradiction with the usual condition for a differentiable interior extremum. For $|x|$ on $[-1; 1]$,

the minimum is at zero although the function is not differentiable there. These two examples clearly delimit the assumptions.

The word “candidate” is therefore useful. Establish the critical points, endpoints and any admissible singularities, then compare their values or the direction of change. For an unbounded domain, limits must also be examined. A function can have an upper bound without attaining it: f(x) = 1 − e^(−x) on [0; +∞) approaches 1 but never equals it.

8. Derivatives and motion: distance or displacement?

A fictional position is $x(t) = 12t - t^3$ for t in [0; 3], with suitable coefficients and units. The velocity is $v(t) = 12 - 3t^2$ and the acceleration $a(t) = -6t$. Velocity is positive before t = 2 s, zero at that time, then negative. Position increases to x(2) = 16 m, then decreases to x(3) = 9 m.

The displacement between zero and three seconds is 9 m. The distance travelled is 16 + (16 − 9) = 23 m, because the direction changes. Differentiating position and studying the sign of velocity therefore allows a geometric question about the journey to be solved. Using only x(3) − x(0) as the distance would forget the return phase.

The sign of acceleration does not directly tell us whether speed increases or decreases. Before the reversal, $v > 0$ and $a < 0$: the magnitude of velocity decreases. Afterwards, $v < 0$ and $a < 0$: it increases. This connection with physics shows why signs and vectors must be interpreted, not merely calculated. The derivative supplies a relation; the meaning of the variables supplies the answer to the problem.

9. Check a derivative in several ways

A check can use the structure, the expected sign and a numerical difference. For f(x) = e^(−2x), a positive derivative would be suspicious because the function decreases. The correct formula is −2e^(−2x). For a polynomial, the derivative's degree falls by one if the leading term is not constant. For a composite function, explicitly look for the often-forgotten inner factor.

A difference quotient with a small h can check a derivative's value: [f(a + h) − f(a)]/h should approach f'(a). It does not prove the formula throughout the domain and can suffer rounding errors when h is too small. A central difference [f(a + h) − f(a − h)]/(2h) can be useful, but requires both points to lie in the domain.

Dimensional consistency provides another check. In a problem concerning area as a function of length, a derivative stated in m² rather than m reveals an error of representation or units. An independent check does not mean copying the same calculation with the same steps: look for a property or method capable of detecting another kind of error.

10. Worked exercises

A - Product and composition. Differentiate f(x) = x²e^(3x) on the real numbers. **Solution:** f' = 2xe^(3x) + 3x²e^(3x) = xe^(3x)(2 + 3x). The factor 3 comes from the inner expression. The derivative's zeros are 0 and −2/3; their nature still requires a sign table.

B - Quotient. Differentiate g(x) = (x + 1)/(x − 2). **Solution:** the domain excludes 2 and g' = [(x − 2) − (x + 1)]/(x − 2)² = −3/(x − 2)². The function is decreasing on each interval of its domain, without allowing global decrease to be inferred on their union.

C - Tangent. For f(x) = x² + 1, give the tangent at a = −1. **Solution:** f(−1) = 2 and f'(−1) = −2, so y = 2 − 2(x + 1) = −2x. The line

does indeed pass through $(-1; 2)$. A negative slope does not mean that the point's vertical coordinate is negative.

D - Extrema. Study $h(x) = x^3 - 3x$ on $[-2; 2]$. **Solution:** $h' = 3(x^2 - 1)$, positive outside $[-1; 1]$ and negative inside. The candidate values are $h(-2) = -2$, $h(-1) = 2$, $h(1) = -2$ and $h(2) = 2$. The global maximum 2 is attained at two points, and so is the minimum -2. A global extremum is not necessarily unique.

E - Domain. Differentiate $k(x) = \ln(2 - x)$. **Solution:** the domain is $x < 2$ and $k' = -1/(2 - x)$. The function is strictly decreasing on that domain. Writing $1/(2 - x)$ would forget the inner derivative; studying $x > 2$ would forget the logarithm's real domain.

F - Approximation. Estimate $(3.02)^2$ using a tangent to x^2 at 3, then give the exact difference. **Solution:** the approximation is $9 + 6 \times 0.02 = 9.12$. The exact value is 9.1204; the difference is $(0.02)^2 = 0.0004$. The remainder can be calculated here because the function is explicitly known.

11. Revision and references

Useful revision involves four actions: determine the domain, differentiate using the right structure, study signs on intervals, then interpret extrema and constraints. Repeat a differentiation without a formula sheet and explain every factor. For optimisation, check that the objective function properly represents all the quantities counted, then check that the solution belongs to the admissible domain.

Progression reference: the published programme for the French Terminale mathematics specialism, https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm ; annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. All explanations and situations are original LFDP writing, without reproduced examination papers or textbook text. The applications support understanding of methods and do not certify any

real installation, manufacturing process or forecast.

Chapter 6 - Convexity: compare a curve, its chords and its tangents

Historical perspective - Jensen: convexity makes comparisons possible

In 1906, Johan Ludwig Jensen published an inequality concerning convex functions in Acta Mathematica. This milestone gives context to the central idea of this chapter: the shape of a function imposes comparisons between values, chords and tangents. Here we work with proved properties and criteria accessible at the intended level, without making the general form of Jensen's inequality an additional examination requirement. The history reminds us why convexity is useful; an illustration or a famous name never removes the need to check the interval and assumptions.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Jensen/

1. A property different from increase and decrease

The direction of change tells us whether a function's values increase or decrease. Convexity describes another structure: the position of its curve relative to the chords joining two of its points. A convex function can decrease and then increase, or decrease throughout its domain. It is therefore incorrect to use “convex” as a synonym for “increasing”.

For $f(x) = (x - 1)^2$, the function decreases up to 1 and then increases, but is convex over all real numbers. For f(x) = e^(−x), it is decreasing and convex. These examples separate two questions that a diagram can sometimes blur. In a complete study, the sign

of f' gives information about increase and decrease; changes in f' and, when it exists, the sign of f'' give information about convexity.

The usefulness is concrete. Convexity allows values to be compared without calculating each one exactly, bounds a function using straight lines, and justifies certain inequalities. It also illuminates optimisation: for a differentiable convex function, an interior point where the derivative vanishes is a global minimum on the interval under consideration. Checking that this point belongs to the problem's domain remains essential.

2. Definition using chords

A function f is convex on an interval I if, for every a and b in I and every λ in [0; 1], $f((1 - \lambda)a + \lambda b) \leq (1 - \lambda)f(a) + \lambda f(b)$. The point $(1 - \lambda)a + \lambda b$ lies between a and b. The right-hand side is the vertical coordinate of the corresponding point on the graph's chord. The curve therefore lies below this chord.

For a concave function, the inequality is reversed. An affine function is both convex and concave, since the curve and all its chords coincide. Strict convexity requires a strict inequality for $a \neq b$ and $0 < \lambda < 1$. This distinction separates a general structure from the absence of affine segments.

At the midpoint, $\lambda = 1/2$, the definition gives $f((a + b)/2) \leq [f(a) + f(b)]/2$. For $f(x) = x^2$, this becomes $((a + b)/2)^2 \leq (a^2 + b^2)/2$. Expanding shows that the difference between the right-hand and left-hand sides is $(a - b)^2/4$, which is nonnegative. This direct proof confirms a particular inequality and reveals the role of the difference between a and b.

The average of the images is therefore not generally the image of the average. Automatic equality holds only in particular situations, notably for an affine function. This difference is useful in applications, but a practical conclusion also requires us to check

that the chosen function really represents the phenomenon under study.

3. The derivative criterion and second derivatives

On an interval where f is differentiable, f is convex if and only if f' is nondecreasing. If f is twice differentiable, a nonnegative second derivative on the interval ensures convexity; a nonpositive one ensures concavity. The second derivative expresses change in the slope. A negative slope that becomes less negative can therefore correspond to convexity.

For f(x) = e^(−x), f' = −e^(−x) is negative but increasing, because f'' = e^(−x) > 0. The function falls less and less quickly. For f(x) = ln(x), on x > 0, f' = 1/x is positive but decreasing and f'' = $-1/x^2$ < 0: the logarithm increases, with a slope that decreases. Increase and decrease and convexity therefore supply two complementary kinds of information.

The sign must be studied on an interval, not at just one point. The fact that f''(a) > 0 gives local information if appropriate conditions allow it to extend to a neighbourhood, but does not prove convexity throughout the domain. Similarly, the second derivative can fail to exist at some points of a function that is nevertheless convex: |x| is an example. The chord definition is more general than the practical second-derivative criterion.

4. Tangents and a proof of an inequality

For a differentiable convex function, the curve lies above each of its tangents on the interval of convexity. At the point a, this is written f(x) ≥ f(a) + f'(a)(x − a). For a concave function, the curve lies below. The descriptions “below the chords” and “above the tangents” do not contradict each other: a chord joins two distinct points, whereas a tangent describes a local slope.

One justification for the convex-function result defines h(x) = f(x) − f(a) − f'(a)(x − a). Its derivative is f'(x) − f'(a). Since f' is nondecreasing, h' is nonpositive before a and nonnegative after a. Thus h attains its minimum at a, where h(a) = 0. We obtain h(x) ≥ 0. The proof uses the behaviour of an auxiliary function and gives a reusable tool.

For f(x) = e^x, the tangent at zero is y = 1 + x. Convexity therefore gives e^x ≥ 1 + x for every real x. This inequality is a bound, not an approximation that is always accurate. For very negative x, the straight line can be negative while the exponential remains positive. Near zero, it gives a local approximation whose error requires further study.

5. Bound the logarithm

The logarithm is concave on the positive real numbers. Its tangent at 1 is y = x − 1. We obtain ln(x) ≤ x − 1 for x > 0, with equality at 1. Replacing x with 1 + h gives ln(1 + h) ≤ h for h > −1. The domain is essential; ln(1 + h) is not real if h ≤ −1.

Another tangent supplies a lower bound. Take the tangent to ln at the point x > 0 and evaluate its upper bound at 1: ln(1) ≤ ln(x) + (1/x)(1 − x). Since ln(1) = 0, we obtain ln(x) ≥ 1 − 1/x. Together, the two bounds give 1 − 1/x ≤ ln(x) ≤ x − 1 for x > 0.

For x = 1.10, these bounds become 0.090909... ≤ ln(1.10) ≤ 0.10. The numerical value is approximately 0.09531. The bounds do not depend on this last value: they are proved before the approximate calculation. For x far from 1, the bounds may not be tight. A valid bound is not necessarily the best bound for every question.

6. Inflection points: a change in convexity

An inflection point is associated with a change in convexity near a point on the curve. For a sufficiently regular function, a change in the sign of f'' on either side often identifies one. The equality f''(a) = 0 alone is insufficient. The sign or convexity around the point must be examined, and the point must belong to the curve.

For f(x) = x³, f''(x) = 6x changes sign at zero: the function is concave on the left and convex on the right. The origin is an inflection point. For g(x) = x^4, g''(x) = 12x² vanishes at zero but remains positive on either side; there is no change in convexity. Zero is a minimum point, not an inflection point.

A value excluded from the domain cannot be an inflection point on the curve, even if the signs of f'' differ on two separated intervals. For 1/x, the second derivative 2/x³ changes sign on either side of zero, but the function is undefined at zero. The correct conclusion concerns convexity on each interval, not a nonexistent point.

7. A detailed study of a polynomial

Consider f(x) = x^4 − 4x³. The derivatives are f'(x) = 4x²(x − 3) and f''(x) = 12x(x − 2). The sign of f'' is positive for x < 0, negative for 0 < x < 2, then positive for x > 2. The function is therefore convex, then concave, then convex on those intervals. The points with coordinates (0; 0) and (2; −16) are inflection points.

The first derivative is negative for x < 3, apart from vanishing at the isolated point zero, and positive for x > 3. The function decreases up to 3, then increases; it has a minimum at x = 3, with value −27. The inflection point at zero is therefore not an extremum. The same graph illustrates why a zero derivative, an extremum and an inflection point are three distinct concepts.

The tangent at 2 has slope f'(2) = −16 and passes through (2; −16). Its equation is y = −16 − 16(x − 2), or y = −16x + 16. Checking the slope and that the line passes through the point is essential. Near an inflection point, the curve changes its position relative to this tangent in this regular case; it can be studied directly by factorising the difference if the question requires it.

For the tangent at zero, both slope and vertical coordinate are zero: it is the horizontal axis. Near zero, $f(x) = x^3(x - 4)$ is positive on the left and negative on the right. The curve therefore crosses its tangent in this example. This local observation should not be turned into a universal criterion without the necessary regularity assumptions.

8. Convexity and a global minimum

If f is convex and differentiable on an interval, and f'(a) = 0 at an interior point a, the tangent inequality gives $f(x) \geq f(a)$ for every x in the interval. The point a is therefore a global minimum. If f is strictly convex, this minimum is unique. The proof explains why a local slope property can produce a global conclusion here: convexity provides the extra information.

For $f(x) = x^2 + 8/x$ on $x > 0$, $f''(x) = 2 + 16/x^3 > 0$. The function is strictly convex. Its derivative $2x - 8/x^2$ vanishes when $x^3 = 4$, hence at a = the cube root of 4. This point gives the unique global minimum. A long table of increase and decrease is not needed again to establish uniqueness, although limits or a table can provide a check.

If the admissible domain is a subinterval that does not include a, the minimum must be sought on that domain. Convexity does not make a forbidden value admissible. An application therefore requires two levels: understand the mathematical function, then respect the problem's constraints. A theoretical optimum does not

validate a real decision whose costs or risks have not been modelled.

9. Averages and spread: an original illustration

Take two measurements a and b. Convexity of the square function gives that the square of their average is less than or equal to the average of their squares. The difference is $(a - b)^2/4$. For $a = 2$ and $b = 6$, the square of the average is 16 and the average of the squares is 20; the difference is 4. This difference is connected with spread and disappears only when the two values coincide.

For several numbers with positive weights adding to 1, an extension of the convexity inequality compares the function of the weighted average with the weighted average of the function values. This extension, often called Jensen's inequality, is offered as additional understanding; it is not a stated particular examination requirement. It can be proved by repeatedly applying the two-point definition.

The practical scope depends on the model. If a fictional consumption is proportional to the square of a speed, varying the speed around the same average can increase the average consumption. This does not establish a real vehicle's balance without accounting for duration, distance, acceleration and the actual mechanisms involved. The mathematical inequality remains exact; its empirical interpretation requires additional data.

10. Read and construct a convexity table

An efficient study begins with the domain and the intervals where the derivatives exist. Calculate f'', factorise if possible, and study its sign. Deduce the intervals of convexity and concavity, then check candidate inflection points. Their coordinates still require f to be evaluated at the retained horizontal coordinates.

Do not mix up the sign table for f with those for f' and f''. The first describes the function's positive and negative values; the second describes increase and decrease; the third, under the stated conditions, describes convexity. A negative function can be convex; a positive function can be concave; an inflection point can have a nonzero vertical coordinate. The three tables answer different questions.

On a graph, looking at whether the slope increases or decreases can help. This remains a qualitative check dependent on the scale and accuracy of the drawing. A table based on an exact expression provides a stronger justification. If only a graphical document is supplied, state conclusions at the level of precision it permits and identify the assumptions made in reading it.

11. Worked exercises

A - Increase, decrease and convexity. Study both properties for e^(−2x). **Solution:** f' = −2e^(−2x) < 0 and f'' = 4e^(−2x) > 0. The function is decreasing and strictly convex. These properties are not opposites.

B - Tangent to the logarithm. Write the tangent to ln(x) at 2 and compare it with the function. **Solution:** y = ln(2) + (x − 2)/2. Concavity gives ln(x) ≤ ln(2) + (x − 2)/2 for x > 0. The result holds throughout the concavity domain, not merely in a small neighbourhood of 2.

C - Inflection. f(x) = x^5. **Solution:** f''(x) = 20x³ changes sign at zero. The origin is an inflection point. The first derivative 5x^4 is nonnegative; the function remains increasing, confirming that inflection does not necessarily change the direction of change.

D - Exponential bound. Show that e^(2x) ≥ 1 + 2x. **Solution:** apply the inequality e^t ≥ 1 + t to t = 2x, or use the tangent at zero to

the function e^(2x). Both approaches produce the same bound with assumptions valid over all real numbers.

E - Average. For 1 and 5, compare the square of the average with the average of the squares. **Solution:** the average is 3, its square 9; the average of the squares is (1 + 25)/2 = 13. The difference 4 equals $(5 - 1)^2/4$. Replacing an average of squares with a square of the average would therefore remove information about spread.

F - A zero of f''. A function has f''(a) = 0. Can an inflection point be concluded? **Solution:** no. A change in convexity must be established. The counterexample x^4 at zero shows that a zero of the second derivative alone is insufficient.

12. Revision and references

Revise convexity through three representations: the curve below chords, the curve above tangents, and a nondecreasing derivative. Be able to turn one representation into an inequality, then explain its domain. For inflection points, check the change in convexity and the coordinates. A notebook containing the counterexamples x^4, |x| and 1/x helps preserve the exact assumptions.

Progression framework: the published official programme for the French Terminale mathematics specialism, https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. Proofs, examples and exercises are original LFDP writing. Additional material is explicitly distinguished from an announcement of examination content; no third-party textbook or paper is reproduced.

Chapter 7 - Exponentials, logarithms and trigonometry: transform without losing the domain

Historical perspective - Napier: transform a difficult calculation

John Napier presented his logarithms in a book published in 1614, aiming to simplify lengthy calculations. His logarithms are not identical to the natural logarithm taught today. The useful comparison is the transformation of a problem: replacing a product by a sum can make a relation easier to study, but only when the logarithm's conditions are met. The chapter therefore distinguishes identities, their domains and their use in equations or models. A computational tool becomes more powerful when we understand what it preserves and what it requires.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Napier/

1. Three families and three structures

The exponential describes, in particular, change proportional to a current value. The logarithm transforms multiplicative ratios into additive differences and allows the exponential to be inverted on positive values. Trigonometry represents angles, projections and periodic changes. These functions can appear in the same problem, but do not obey the same identities or have the same domains.

A reliable method defines the domain before making a transformation, then identifies the structure: product, power, composition or periodicity. A graphical resemblance is insufficient for choosing a law. Exponential growth is not faster linear growth; a

sinusoid does not represent every real oscillation; a logarithm does not allow the rules it has for products to be applied to sums.

The arguments of functions also require interpretation. In a model e^(kt), the product kt must be dimensionless. If t is expressed in seconds, k has unit s^−1. In ln(q/q_0), the ratio of two quantities with the same unit is dimensionless. Writing the logarithm of a numerical value assumes a unit convention or a normalisation; otherwise, arbitrarily changing the unit can change the calculated value.

2. The exponential: properties and growth

The exponential is defined over the real numbers, strictly positive and strictly increasing. It satisfies e^0 = 1, e^(a+b) = e^a e^b and e^(−a) = 1/e^a. Its derivative is itself; for e^(u(x)), the derivative becomes u'(x)e^(u(x)). The algebraic properties allow a product or quotient to be transformed, not the exponents to be added in a sum of exponentials.

Thus e^a + e^b is not e^(a+b). With a = b = 0, the first expression is 2 and the second 1. Similarly, e^(2x) = (e^x)², but e^(x²) is not (e^x)². The position of the brackets describes different functions. A simple counterexample can detect a doubtful identity before it spreads throughout a calculation.

The limits are e^x → +∞ as x → +∞ and e^x → 0 as x → −∞. The exponential never equals zero for a finite real number. A calculator displaying zero for a very negative value may have reached its representation limit; that display is not a new mathematical property.

3. Solve an exponential equation by substitution

Consider e^(2x) − 5e^x + 6 = 0. Set Y = e^x, with Y > 0. The equation becomes Y² − 5Y + 6 = 0, or (Y − 2)(Y − 3) = 0. Both candidates, 2 and 3, respect Y's domain. Since the exponential is a bijection from the real numbers to the positive real numbers, the solutions are x = ln(2) and x = ln(3).

Substitution does not remove the domain requirement. If another polynomial supplied a negative root Y, it would correspond to no real e^x. First solve the algebraic equation, then filter the candidates and return to the original variable. These stages should appear in the written solution to show that all the solutions, and only those, have been retained.

For the inequality e^(2x) − 5e^x + 6 ≤ 0, the polynomial's sign table gives 2 ≤ Y ≤ 3. The increasing logarithm allows us to conclude ln(2) ≤ x ≤ ln(3). An equation supplies points; an inequality here supplies an interval. Reusing only the two roots as the inequality's solution set would be insufficient.

4. The logarithm: inverse and exact rules

The natural logarithm is defined over the strictly positive real numbers. It is the inverse of the exponential: ln(e^x) = x for every real x and e^(ln y) = y for y > 0. It is strictly increasing and its derivative is 1/x. Its value at 1 is zero; it is negative between zero and one and positive above one.

For a > 0 and b > 0, ln(ab) = ln(a) + ln(b). This can be justified by exponentiating the sum: e^(ln a + ln b) = ab, then using uniqueness of the preimage under the exponential. We also obtain ln(a/b) = ln(a) − ln(b). For an integer power of a positive base, the rule follows by repetition; for a real power defined by a^r = e^(r ln a), it gives ln(a^r) = r ln(a).

Domains do not automatically simplify with the formula. For $x \neq 0$, $\ln(x^2) = 2\ln(|x|)$, because $|x|$ is positive. Writing $2\ln(x)$ would exclude negative x although x^2 is positive. For $\ln(x) + \ln(y)$, both arguments must be positive; the fact that $xy > 0$ alone is insufficient if x and y are both negative.

5. Logarithmic equations and checking

Solve $\ln(x - 1) + \ln(x + 1) = \ln(8)$. The domain requires $x - 1 > 0$ and $x + 1 > 0$, hence $x > 1$. On that domain, the sum becomes $\ln(x^2 - 1)$. The logarithm's injectivity gives $x^2 - 1 = 8$, or $x^2 = 9$. The candidates are 3 and -3, but only 3 satisfies the original domain.

For $\ln(x - 1) + \ln(x + 1) \leq \ln(8)$, the increasing logarithm gives $x^2 - 1 \leq 8$ on $x > 1$. The solution set is $(1; 3]$. It is not $[-3; 3]$, because the original arguments must remain positive. The domain is a constraint maintained through every transformation, not an optional check at the end of a calculation.

An equation $\ln(u(x)) = 0$ is equivalent to $u(x) = 1$ when the logarithm is defined. An equation $\ln(u(x)) = \ln(v(x))$ requires $u > 0$ and $v > 0$ before $u = v$ can be concluded. Solving only $u = v$ can introduce points where both expressions are negative. The sequence of steps must therefore retain the positivity conditions.

6. Multiplicative growth and an integer threshold

A fictional change increasing by 5% at each stage follows $q_n = q_0(1.05)^n$ if the model and percentage remain constant. To find the first index at which $q_n \geq 2q_0$ with $q_0 > 0$, we obtain $(1.05)^n \geq 2$, then $n \geq \ln(2)/\ln(1.05)$, approximately 14.21. The first admissible integer is 15.

Checking gives $(1.05)^{14} \approx 1.980$ and $(1.05)^{15} \approx 2.079$. Index 14 is insufficient; 15 is sufficient. The two comparisons establish

minimality within the model. The reasoning remains purely mathematical: it forecasts no actual return, population or capacity without validation of the change rule.

For a decrease with a factor between zero and one, the logarithm of the factor is negative and division reverses the inequality. This is the critical point in many threshold problems. One way to check an answer is to test the two neighbouring indices in the original expression rather than trusting a sequence of signs copied too quickly.

7. Logarithms with different bases

For a base $a > 0$, $a \neq 1$, the logarithm to base a is $\log_a(x) = \ln(x)/\ln(a)$, for $x > 0$. It is increasing if $a > 1$ and decreasing if $0 < a < 1$. The common logarithm corresponds to $a = 10$. Changing base allows a calculator with only ln or common log to be used.

Multiplying a ratio by 100 increases its common logarithm by 2. This structure explains the usefulness of a logarithmic scale for comparing orders of magnitude. If a level is defined by $10\log_{10}(I/I_0)$, multiplying I by 100 adds 20 to the level, with the same reference I_0. It does not mean that the physical quantity has been multiplied by 20.

A difference of logarithms corresponds to the logarithm of a ratio. Adding logarithmic levels therefore does not automatically represent the addition of physical quantities. To add intensities, return to the intensities and then apply the definition of the level. Interpreting the scale is as important as using the calculator keys.

8. Angles and the trigonometric circle

The radian measures an angle by the ratio between an arc's length and the circle's radius. A full turn corresponds to 2π radians, or 360°. A half-turn is π radians and a right angle π/2. The usual derivatives of sin and cos use arguments expressed in radians. A calculator set to degrees can therefore produce values inconsistent with a formula when this convention is forgotten.

On the unit circle, a point associated with an angle x has coordinates (cos x; sin x). The identity $\cos^2 x + \sin^2 x = 1$ follows from its membership of the circle. Cosine is even, sine is odd, and both are periodic with period 2π. A trigonometric value does not identify a unique angle without an interval or other information.

The addition formulae connect angles: $\cos(a + b) = \cos a \cos b - \sin a \sin b$ and $\sin(a + b) = \sin a \cos b + \cos a \sin b$. Taking $a = b = x$ recovers $\sin(2x) = 2\sin x \cos x$ and $\cos(2x) = \cos^2 x - \sin^2 x$. These identities can transform an expression; they must not be confused with $\cos(a + b) = \cos a + \cos b$, which is false.

9. Solve a trigonometric equation

The equation $\cos x = 1/2$ has solutions $x = \pi/3 + 2k\pi$ or $x = -\pi/3 + 2k\pi$, with k an integer. On [0; 2π], this gives π/3 and 5π/3. The circle explains the two families through symmetry. Writing only π/3 would omit some solutions; writing an infinite family without filtering the interval would overlook the question asked.

For $\sin x = -\text{sqrt}(2)/2$ on [0; 2π], the solutions are 5π/4 and 7π/4. Sine corresponds to the vertical coordinate, so the points lie in the circle's lower half. Knowing the reference angles and signs by quadrant allows the values to be reconstructed rather than memorising an unstructured list.

Periodicity requires care at the endpoints. If the equation admits x = 0 and the domain is [0; 2π], the value 2π can also be a distinct numerical solution, even though it represents the same point on the circle. On [0; 2π), it is excluded. Interval notation must therefore be maintained through to the final answer.

10. Derivatives and the periodicity of a composite function

The usual derivatives are (sin x)' = cos x and (cos x)' = −sin x. For f(x) = cos(3x), f' = −3sin(3x). Its period is 2π/3, because increasing x by that amount increases the argument by 2π. The frequency of a function cos(ωt) corresponds to |ω|/(2π) when t is a duration and ω an angular frequency in radians per unit of time.

The function $\sin^2 x$ has period π, although sin x has period 2π: changing sine's sign does not change its square. A transformation can therefore change the smallest period. Not every period of a function is necessarily its smallest positive period. Distinguish a value that works from the fundamental characteristic being sought.

For sin x + cos x, an identity gives sqrt(2)sin(x + π/4). This expression reveals an amplitude sqrt(2), a period 2π and a phase shift. The sum of two sinusoids with the same frequency can therefore remain a sinusoid with a different amplitude and phase. With different frequencies, such a simple reduction is not generally possible.

11. Local approximations and checks

Near zero, differentiability gives e^h ≈ 1 + h, ln(1 + h) ≈ h and sin h ≈ h, with suitable domains and conventions. These approximations are local. For sin h, h must be in radians. They do not become equalities because an exercise states “small”; the required accuracy must be discussed if a numerical value is used.

Convexity supplies information about the sign of some differences. We know e^h ≥ 1 + h, and concavity of the logarithm gives ln(1 + h) ≤ h for h > −1. The approximations are thus respectively a lower and an upper bound in those domains. This is insufficient to determine the size of the error; tighter bounds or a study of the remainder may be needed.

To check a calculation, examine reference values: e^0 = 1, ln 1 = 0, sin 0 = 0, cos 0 = 1. Check signs, domain and periodicity. A transformation that changes one of these known properties should be re-examined. This strategy detects errors before a longer calculation without claiming to replace a proof.

12. Worked exercises and revision

A - Exponential. Solve e^(2x) = 7. **Solution:** 2x = ln 7, so x = (ln 7)/2. Solving e^x = 7 and then forgetting the factor 2 would give another value. The positivity of 7 makes the transformation possible.

B - Logarithmic domain. Solve ln(3 − x) ≥ 0. **Solution:** the domain is x < 3. The logarithm is nonnegative for an argument at least equal to 1, so 3 − x ≥ 1, or x ≤ 2. The solution set is (−∞; 2], already contained in the domain.

C - Identity. Simplify ln(5x) − ln(x) for x > 0. **Solution:** ln(5x/x) = ln 5. If x were negative, the original expressions would not be defined over the real numbers. The constant result does not automatically extend the domain.

D - Trigonometry. Solve sin(2x) = 0 on [0; π]. **Solution:** 2x = kπ, so x = kπ/2. Filtering gives 0, π/2 and π. The argument's interval is [0; 2π], with both endpoints included.

E - Differentiation. Differentiate ln(1 + e^x). **Solution:** the argument is positive over all real numbers and the derivative is e^x/(1 +

e^x). It lies between zero and one, providing a check of sign and size.

F - Period. Give a fundamental period of cos(4t). **Solution:** 2π/4 = π/2. The value 2π is also a period, but not the smallest positive one. Do not confuse argument, angular frequency and frequency.

Revise by reconstructing domains, identities and transformations rather than memorising isolated equations. For each solution, indicate whether it is exact, approximate, unique or drawn from a periodic family. Progression framework: the published Terminale programme, https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. Text and exercises are original LFDP writing; no actual forecast or reproduction of a third-party paper is presented.

Chapter 8 - Antiderivatives and integrals: measure accumulation and check an approximation

Historical perspective - Cavalieri: think about area before modern integral calculus

Bonaventura Cavalieri's Geometria was published in 1635 and presented his method of indivisibles. It is a historical milestone in the study of areas and volumes. In our chapter, dividing and summing helps us understand the integral, but modern definitions and justifications do not reduce to that earlier method. The teaching connection is the move from a whole figure to comparable contributions, then to a controlled result. This perspective does not attribute the invention of integration to Cavalieri alone or confuse geometric intuition with a modern proof.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Cavalieri/

1. An accumulated quantity, not an instantaneous value

Power gives energy per unit of time; the energy received over a duration depends on the power's entire evolution. Velocity gives displacement per unit of time; displacement depends on how velocity changes during the journey. An integral calculates this accumulation. It therefore does not replace the initial function with an “average” without specifying how that average is defined.

For a continuous nonnegative function on [a; b], the integral can be introduced through the area under its curve between the two endpoints. Increasingly narrow rectangles approximate that area. For a function of any sign, the integral becomes signed:

contributions below the axis are negative. Total geometric area and the integral can then differ.

The integral's unit is the function's unit multiplied by the variable's unit. Power in watts integrated over seconds gives joules; velocity in m/s integrated over seconds gives metres. If the axes are unitless numbers, we speak of area units for a geometric interpretation. This check connects the calculation with the question actually asked.

2. Antiderivatives: a family of functions

An antiderivative F of f on an interval I is a differentiable function such that F' = f on I. If F is an antiderivative, so is F + C for any constant C. Conversely, two antiderivatives of the same function on an interval differ by a constant. The word “interval” matters: on a domain consisting of separate pieces, the constants can differ on each piece.

For f(x) = 2x, F(x) = x² is an antiderivative. For f(x) = e^(3x), F(x) = e^(3x)/3 works. For f(x) = cos(4x), F(x) = sin(4x)/4. Check every proposal by differentiation. The inner factor must be compensated for; writing e^(3x) as an antiderivative of e^(3x) would forget the factor 3 in its derivative.

An additional condition fixes the constant. If F' = 2x and F(1) = 5, then F(x) = x² + C and 1 + C = 5, so C = 4. The initial condition turns a family into a determined function. This idea will be reused in differential equations: integration gives a family, and the data select a member.

3. The fundamental theorem and calculation with an antiderivative

For a continuous function f on an interval containing a and b, any antiderivative F allows ∫_a^b f(x) dx = F(b) − F(a) to be calculated. Adding a constant to F does not change this difference. The lower-endpoint value is subtracted after evaluating the antiderivative; incorrect brackets can reverse the result.

The function G(x) = ∫_a^x f(t) dt is an antiderivative of f when the continuity assumptions hold. The variable t under the integral sign is an integration variable; here x represents the variable upper limit. This distinction avoids writing the same letter in several roles without knowing what varies. The integral thus connects accumulated area with differentiation.

For f(x) = 3x² − 2 on [1; 2], an antiderivative is F(x) = x³ − 2x. The integral is (8 − 4) − (1 − 2) = 5. The function is positive on this interval, which makes a positive result consistent. Checking the sign is independent of calculating the brackets and can reveal an error before the final answer.

4. Signed area and geometric area

Take f(x) = x − 1 on [0; 3]. An antiderivative is x²/2 − x. The integral is 3/2. Yet the region between the curve and axis has a negative part between 0 and 1 and a positive part between 1 and 3. The total geometric area is the sum of both positive areas: 1/2 + 2 = 5/2.

To calculate this area, integrate |f| or split the domain at sign changes. The formula ∫f gives the difference between positive and negative contributions. A zero integral therefore does not imply a zero function: the contributions can cancel. For example, ∫_(−1)^1 x dx = 0 although the graph has two nonzero regions.

For the area between two curves f and g, integrate |f − g|. If f ≥ g is known throughout the interval, the area is ∫(f − g). Intersections may require several pieces. Choosing the “upper” curve from a single point without studying the sign can produce a negative area or unintended cancellation.

5. Properties of the integral

Linearity gives ∫(αf + βg) = α∫f + β∫g, with the same integration limits and appropriate assumptions. The interval-addition relation allows splitting: ∫_a^b f + ∫_b^c f = ∫_a^c f. Reversing the limits changes the sign, and the integral between equal limits is zero. These rules organise a calculation without requiring a new antiderivative for every piece.

If f ≥ 0 on [a; b] with a ≤ b, its integral is nonnegative. If f ≤ g, then ∫f ≤ ∫g. A pointwise inequality can therefore bound an accumulation. The order of the integration limits matters: signs are reversed when the orientation changes. The properties must be used with their conventions.

If m ≤ f(x) ≤ M on [a; b], then m(b − a) ≤ ∫_a^b f ≤ M(b − a). This bound is not always very precise, but provides an immediate check. A calculated value outside these bounds is incompatible with the data or calculation. For a function between 2 and 5 on an interval of length 3, the integral must be between 6 and 15.

6. Recognise a composite derivative in reverse

Finding antiderivatives benefits from recognising structure. If u is differentiable, an antiderivative of u'e^u is e^u. For u > 0, an antiderivative of u'/u is ln(u). An antiderivative of u'cos(u) is sin(u). The factor u' is decisive; the mere presence of an exponential, quotient or cosine is insufficient.

For f(x) = 2x/(x² + 1), recognise u = x² + 1, always positive, and u' = 2x. An antiderivative is ln(x² + 1). For g(x) = 6x e^(3x²), an antiderivative is e^(3x²). If the numerator in the first example were x, a factor 1/2 would be needed. Checking by differentiation recovers the exact coefficient.

Do not invent an antiderivative by “integrating separately” the numerator and denominator of a quotient. Similarly, an antiderivative of a product is not the product of two antiderivatives. The differentiation rules explain why these shortcuts are false. When no immediate structure fits, another method or an approximation may be necessary.

7. Integration by parts: use the derivative of a product

The formula (uv)' = u'v + uv' gives, after integration, ∫_a^b u'v = [uv]_a^b − ∫_a^b uv'. This integration-by-parts relation transforms one integral into another. Its usefulness depends on the choice: seek a derivative that simplifies one factor and an accessible antiderivative for the other. No choice is automatically effective for every product.

For ∫_0^1 xe^x dx, take v = x and u' = e^x, hence u = e^x and v' = 1. The integral is [xe^x]_0^1 − ∫_0^1 e^x dx = e − (e − 1) = 1. An equivalent antiderivative is (x − 1)e^x, which can be checked by differentiation. The positive result is consistent with the positive integrand on the interval.

For ∫_1^e ln(x) dx, take v = ln(x) and u' = 1, hence u = x and v' = 1/x. This gives [xln(x)]_1^e − ∫_1^e 1 dx = e − (e − 1) = 1. Writing an antiderivative xln(x) without the term −x would forget part of its derivative. The method is understood through the product rule, not through an unjustified memorised selection recipe.

8. The average value of a function

The average value of f on [a; b], with a < b, is μ = [1/(b − a)]∫_a^b f(x) dx. It has the same unit as f. It corresponds to the height of a rectangle with the same width and signed area as the integral. The average of a few sampled values is not automatically this value; it depends on the sampling and approximation used.

For f(x) = x² on [0; 2], the integral is 8/3 and the average 4/3. The average of just the endpoint values, (0 + 4)/2 = 2, is different. If f is continuous, its average value is attained at at least one point in the interval; here x = 2/sqrt(3) works. This property does not always give a unique point.

In a power model P(t) = 3t + 2 on [0; 4] seconds, with appropriate units for the coefficients, the energy is ∫P dt = 32 J and the average power is 8 W. For this affine function, the average of the initial and final values also gives 8 W. The equality follows from the affine structure; it must not be generalised to every curve.

9. Rectangles and numerical bounds

Divide [0; 1] into four intervals of width 1/4 for the increasing function f(x) = x². Rectangles using left-endpoint values underestimate the area; those using right-endpoint values overestimate it. The left sum is (1/4)[0² + (1/4)² + (1/2)² + (3/4)²] = 0.21875. The right sum is 0.46875. The exact integral 1/3 lies between them.

Increasing the number of subdivisions tightens these bounds for a continuous function in this setting. A numerical method must state the number and width of intervals and the choice of points. A list of results without a method does not allow accuracy to be assessed. If the function is not monotonic, the choice of left or right points alone does not guarantee a bound; extrema on each piece must be studied or other arguments used.

The trapezoidal method uses the segments joining endpoint values. For this convex example, it overestimates the area and gives 0.34375 with four pieces. The midpoint method gives 0.328125, an underestimate in this case. The position of the chords and convexity explain the direction of these differences. Without this information, these methods provide approximations but not necessarily certified bounds.

10. A sequence of integrals and a recurrence relation

Define I_n = ∫_0^1 x^n e^x dx for n ≥ 0. The integrands are positive in the interior, hence I_n > 0. Since 0 ≤ x ≤ 1, x^(n+1) ≤ x^n, so I_(n+1) ≤ I_n. The sequence is decreasing and bounded below by zero; it converges. To identify the limit, use e^x ≤ e: I_n ≤ e∫_0^1 x^n dx = e/(n + 1), which tends to zero. The bounds therefore give I_n → 0.

Integration by parts for n ≥ 1 gives I_n = [x^n e^x]_0^1 − n∫_0^1 x^(n−1)e^x dx = e − nI_(n−1). The first term is I_0 = e − 1; we recover I_1 = 1. The case n = 0 must be handled separately because the boundary expression at zero and the power n − 1 change their status.

This relation is exact but can be numerically difficult for large n, since it subtracts close quantities and amplifies some rounding errors. A correct recurrence formula is not automatically the best floating-point algorithm. The theoretical bounds provide a valuable check here: a negative result would be impossible for these positive integrals.

11. Worked exercises

A - Antiderivative and endpoint. Calculate $\int_0^2 (4x + 1)\,dx$. **Solution:** an antiderivative is $2x^2 + x$, so the integral is 10. The function is positive on this interval; a negative result would indicate an error in signs or integration limits.

B - Area between curves. Compare x and x^2 on $[0; 1]$, then calculate the area between their graphs. **Solution:** $x - x^2 = x(1 - x) \geq 0$. The area is $\int_0^1 (x - x^2)\,dx = 1/2 - 1/3 = 1/6$. The ordering of the functions is established before integration.

C - Product. Calculate $\int_0^1 xe^{2x}\,dx$. **Solution:** integration by parts gives $[xe^{2x}/2]_0^1 - (1/2)\int_0^1 e^{2x}\,dx = e^2/2 - (e^2 - 1)/4 = (e^2 + 1)/4$. The factor 1/2 appears twice and must be retained.

D - Displacement. A velocity is $v(t) = 2 - t$ on $[0; 4]$. **Solution:** its integral is $[2t - t^2/2]_0^4 = 0$. Distance is not zero: it is $2 + 2 = 4$ in the length unit, separating the phases before and after $t = 2$. The integral of signed velocity gives displacement.

E - Bounds. A function satisfies $1 \leq f \leq 3$ on $[2; 5]$. **Solution:** $3 \leq \int_2^5 f \leq 9$ and its average value lies between 1 and 3. These bounds do not determine a unique value; they establish what is compatible with the given bounds alone.

12. Revision and references

Revise an integral through its meaning, units, sign and calculation method. Check an antiderivative by differentiation. For a geometric area, examine sign changes; for an approximation, state the assumptions that make it a bound or only an estimate. Retain constants and composite-function factors until the final check.

The published official Terminale framework includes antiderivatives, integrals, average values and integration by parts:

https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. The proofs, situations and exercises in this chapter are original LFDP writing. No text from a curriculum document or textbook, or third-party examination paper, is reproduced.

Chapter 9 - Differential equations: connect change with its rate

Historical perspective - Euler: find functions from relations

Euler published his three-volume treatise on integral calculus between 1768 and 1770; it also studies differential equations. This milestone reminds us that the unknown can be an entire function rather than an isolated number. In this chapter, a relation between a function and its derivative is accompanied by an initial condition, then examined through an interpretation and a numerical approximation. The institutional archive dates the first volume to 1768; the file it provides is a third edition published in 1824, a distinction to keep when consulting the historical document.

Historical source: University of St Andrews - MacTutor : https://mathshistory.st-andrews.ac.uk/Biographies/Euler/ ; University of the Pacific - Euler Archive : https://scholarlycommons.pacific.edu/euler-works/342/

1. An equation whose unknown is a function

A differential equation connects an unknown function with one or more of its derivatives. Solving y' = 2y does not mean finding a single number y: we seek all differentiable functions satisfying this relation on an interval. An initial condition, such as y(0) = 3, then selects a solution from a family. The distinction between the change relation and the initial state is as important as it is for sequences.

The variables must be named. If t represents time, y(t) may represent a temperature, a fictional concentration or an abstract quantity. The coefficient multiplying y in y' = ay has an inverse-time unit if y keeps the same unit. A term b in y' = ay + b has the unit of y per unit of time. An equation can be algebraically correct and physically inconsistent when the units do not match.

The chapter mainly addresses first-order linear equations with constant coefficients. Additional material shows why this model does not cover every kind of change. The situations are constructed for learning; they forecast no return, health development or real phenomenon without validation of the assumptions and parameters.

2. Solve y' = ay and justify the entire family

The function y(t) = Ce^(at) satisfies y' = aCe^(at) = ay for every constant C. To show that all solutions have this form, take a solution y and define z(t) = y(t)e^(−at). By the product rule, z' = y'e^(−at) − aye^(−at) = 0. On an interval, z is therefore constant and y = Ce^(at). This proof establishes the entire family, not merely some examples of solutions.

If y(t_0) = y_0, then C = y_0e^(−at_0), hence y(t) = y_0e^[a(t − t_0)]. The form with t − t_0 clearly preserves the origin of the initial condition. For a = 0, we recover constant solutions. For y_0 = 0, the solution is identically zero whatever a is. The sign of y_0 is preserved because the exponential is positive.

If a > 0 and y_0 > 0, the function increases and diverges in the long term within the model. If a < 0 and y_0 > 0, it decreases to zero. A negative a means a relative decrease proportional to the current value, not a fixed decrease at each unit of time. The latter would be described by y' = b with constant b.

3. A constant term and an equilibrium value

For y' = ay + b with a ≠ 0, seek a constant solution L. It must satisfy 0 = aL + b, so L = −b/a. Setting z = y − L gives z' = az. The solutions are therefore y(t) = L + Ce^(at). The change of variable measures the distance from equilibrium and transforms the equation into an already-solved case.

With y(t_0) = y_0, the solution becomes y(t) = L + (y_0 − L)e^[a(t − t_0)]. Check the initial condition, then the equation by differentiation. If a < 0, the distance from L tends to zero. A solution can rise towards L or fall towards L according to the sign of y_0 − L. The same model does not fix a direction of change independently of the initial state.

If a > 0, a nonzero initial distance grows in absolute value; the constant solution y = L nevertheless still exists. It would therefore be wrong to assert that all solutions diverge for a > 0. For a = 0, the previous transformation is not allowed; the equation is y' = b and its solutions are y(t) = bt + C. Separating this case avoids division by zero.

4. An equation with a condition at another time

Solve y' + 3y = 9 with y(1) = 4. The equation is equivalent to y' = −3y + 9, so L = 3 and y(t) = 3 + Ce^(−3t). The condition gives 4 = 3 + Ce^−3, hence C = e³. A clearer expression is y(t) = 3 + e^[−3(t − 1)]. At t = 1, the exponential factor is one and the value is indeed four.

The equation and condition do not necessarily require an origin at t = 0. A common mistake is to take C = 4 because the case y(0) = y_0 for y' = ay has been memorised. Here the equilibrium term and initial time modify the constant. Checking at the specified point is short and immediately detects this mistake.

The solution remains above 3 and decreases for t ≥ 1. Its limit is 3. The limit follows from the sign of the coefficient in the exponent, not merely from the fact that the equation has a constant solution. For an equation with a > 0, the constant value would generally be unstable in response to an initial departure. The existence of equilibrium and attraction towards equilibrium are different properties.

5. Cooling: construct the model from a balance

A fictional object of heat capacity C exchanges energy with surroundings maintained at T_ext. A received-power law −H(T − T_ext), with constant H > 0, leads to C T' = −H(T − T_ext). The equation is written T' = −(1/τ)(T − T_ext), where τ = C/H is a characteristic time. The solution is T = T_ext + (T_0 − T_ext)e^(−t/τ) when the initial state is specified at zero.

For T_0 = 70 °C, T_ext = 10 °C and τ = 4 min, we obtain T(t) = 10 + 60e^(−t/4), with t in minutes. The model approaches 10 °C, not zero. To reach at most 25 °C, the difference must become at most 15: e^(−t/4) ≤ 1/4, so t ≥ 4ln(4) ≈ 5.55 min. The logarithm applies to a dimensionless ratio.

This calculation provides no protocol for contact with a hot object. The actual surroundings may vary, the object's temperature may not be uniform, and the coefficient H may depend on conditions. The model is useful when these approximations are acceptable for the question. The equation does not become a physical validation merely because its solution is mathematically exact.

6. Inflow, outflow and conservation in a fictional tank

An ideal tank of constant volume V receives and discharges the same volumetric flow rate Q. Assume an instantly homogeneous mixture and a constant incoming concentration c_in. If M = Vc is the amount of substance present, the balance is M' = Qc_in − Qc. With constant V, this gives c' = (Q/V)(c_in − c). The relation follows from the two flows, not a preference for an exponential.

The equilibrium concentration is c_in and the characteristic time V/Q. For V = 100 litres and Q = 5 litres per minute, this time is 20 min. If the initial concentration is 3 g/L and the incoming one 1 g/L, the solution is c(t) = 1 + 2e^(−t/20). After 20 min, it is

approximately 1.736 g/L. The total volume remains constant in this model, but the amount of substance changes.

If the inflow contains none of the substance, c_in = 0 and c decreases exponentially. For c_0 = 5 g/L, the initial amount in 100 litres is 500 g; after one time constant, it is approximately 184 g. These are exercise values, not doses, treatments or recommendations for using a fluid. A nonhomogeneous mixture or variable volume would require another equation.

7. Linear, exponential and saturating models

Constant growth y' = b gives an affine evolution. Proportional growth y' = ry gives an exponential. The two can be close over a short range and very different in the long term. Choosing between them requires assumptions and observations. Fitting two points does not always determine which rule will remain relevant beyond those points.

As additional material, a nonlinear equation y' = ry(1 − y/K), with r > 0 and K > 0, models reduced growth as a capacity K is approached. For 0 < y_0 < K, one solution is y(t) = K/[1 + A e^(−rt)], with A = K/y_0 − 1. It can be checked by differentiation. The constant solutions zero and K must be examined separately; the formula for A is undefined for y_0 = 0.

This additional example is not a universal population model. The parameter K may be artificial or variable in a real situation, and other effects may dominate. It shows that an unlimited exponential is not the only possible mathematical structure. Competing models are judged by their assumptions and predictions, not by the more reassuring appearance of a curve.

8. Euler's method: a discrete approximation

For y' = f(t,y), the explicit Euler method uses y_(n+1) = y_n + h f(t_n,y_n), with t_(n+1) = t_n + h. It replaces evolution over a small interval with the local change specified at the start. The step h must have the unit of time. The calculated values approximate the solution; in general, they are not the exact values at the same times.

For y' = −ky with k > 0, Euler gives y_(n+1) = (1 − kh)y_n. Over one step, the exact solution multiplies y by e^(−kh), which is always positive. If 0 < kh ≤ 1, the Euler factor remains nonnegative and produces a decrease from a positive initial state. If 1 < kh < 2, the values alternate in sign while decreasing in absolute value. If kh > 2, they can grow in absolute value although the continuous solution decreases.

With k = 1 and h = 3 in compatible units, the numerical factor is −2: the method diverges for y_0 ≠ 0. This failure does not refute the exact solution e^−t; it demonstrates an unsuitable step and a limitation of the algorithm. Reducing the step can improve an approximation, but regularity, duration and accumulated errors must still be examined. A program running without a computer error is not necessarily a mathematically reliable simulation.

9. Identify a parameter and check independently

If equilibrium L is known and observations are assumed to follow y(t) = L + Ce^(at), two values whose differences from L are nonzero and have the same sign allow a to be estimated by ln[(y(t_2) − L)/(y(t_1) − L)]/(t_2 − t_1). A positive ratio and two distinct times are necessary. The estimated coefficient depends on the stated model.

For fictional data L = 2, y(0) = 10 and y(4) = 6, the ratio of the differences is 4/8 = 1/2, so a = −ln(2)/4. The model then predicts

y(8) = 4. An independent measurement at that time could test the prediction. Recovering the two values used to fit the model is not an independent test: it checks the construction.

If L is also unknown, two observations do not necessarily determine the three parameters L, C and a. Uncertainties and the observation range can also make an estimate fragile. Solving a differential equation, fitting its parameters and validating the model are three separate stages. An attractive calculated curve replaces none of them.

10. A glimpse of second order

An equation y'' + ω²y = 0, with ω > 0, appears in an ideal oscillator. The functions A cos(ωt) + B sin(ωt) satisfy it, as two differentiations show. With y(0) = y_0 and y'(0) = v_0, we obtain A = y_0 and B = v_0/ω. Two conditions are needed to determine these two constants. This additional example illuminates the connection between trigonometry and dynamic change.

The period is 2π/ω. The model includes neither dissipation nor a sustaining force; a real oscillation may therefore require other terms. Do not solve this equation as though it were y' = ay: the order and type of relation differ. Identifying the equation precedes choosing the method.

The main chapter remains focused on first-order linear equations with constant coefficients. The oscillator and saturating growth are explicitly distinguished extensions, useful for understanding the limits of a family of models. They are not presented as announcements of topics or requirements for an examination session.

11. Worked exercises

A - Complete family. Solve y' = −2y and impose y(0) = 5. **Solution:** the solutions are Ce^(−2t); the condition gives C = 5. The solution is positive and decreases to zero. Checking the derivative and initial value checks the obtained function, whereas the family proof guarantees that no other solution has been forgotten within this framework.

B - Equilibrium. y' = 2y − 6, y(0) = 4. **Solution:** L = 3 and y(t) = 3 + e^(2t). Equilibrium exists but this solution moves away from it. With y(0) = 3, it would be constant; with y(0) = 2, it would be 3 − e^(2t), which can become negative. If y had to represent a nonnegative quantity, that later phase would indicate a limit of the physical model.

C - Constant term. y' = −4 and y(0) = 7. **Solution:** y(t) = 7 − 4t. This is not an exponential. The value becomes zero at t = 7/4. A physical domain requiring y ≥ 0 would need to stop or change beyond this threshold.

D - Proposed solution. Does the function y = e^(2t) satisfy y' = 2y + 1? **Solution:** its derivative is 2e^(2t), whereas the right-hand side is 2e^(2t) + 1. The constant difference shows that the proposal is unsuitable, independently of the initial condition.

E - Numerical step. For y' = −y, y_0 = 1 and h = 0.5, calculate the first Euler step and compare it with the exact value. **Solution:** Euler gives 0.5; the exact solution gives e^−0.5 ≈ 0.6065. The difference is approximately 0.1065. The numerical step describes an approximation, not another exact solution of the same equation.

12. Revision and references

For revision, establish the family, use the initial condition, then check by differentiation. Study equilibrium, sign and limit without mechanically inferring them merely from the equation's name. For a model, state the exchanges and units. For a simulation, distinguish exact solution from approximation and check an independent invariant or bound.

Published official Terminale framework: https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm and annex https://cache.media.education.gouv.fr/file/SPE8_MENJ_25_7_2019/90/7/spe246_annexe_1158907.pdf, consulted on 11 October 2026. Developments, examples and exercises are original LFDP writing. The cooling and tank models are educational and provide no handling protocol, diagnosis, dosage or prediction of a real situation.

Chapter 10 - Counting, probability and the binomial distribution

Historical perspective - Pascal: an arrangement reveals a structure

The BnF records a 1665 publication of Blaise Pascal's Traité du triangle arithmétique, after his death. This perspective accompanies our reading of binomial coefficients: an arrangement of numbers can help us recognise a relation, then a proof explains why it holds. In this chapter, we separate construction of the triangle, counting and probabilistic applications. Associating the triangle with Pascal's name does not assert that he was the first to know this arrangement; it identifies a specific historical work rather than a single origin for all the chapter's ideas.

Historical source: Bibliothèque nationale de France - Catalogue collectif de France : https://ccfr.bnf.fr/portailccfr/ark:/16871/00114268519

1. A probability model must begin with its outcomes

A probability calculation describes a model of an experiment, not uncertainty in general. Define the sample space of possible outcomes, events and a rule assigning probabilities. In a finite sample space, elementary probabilities are nonnegative and add up to one. An event is a set of outcomes; its probability is the sum of the probabilities assigned to them.

The word “random” does not guarantee equal probabilities. A coin can be biased; a draw depends on its mechanism; a sample can be selected with bias. The formula “favourable cases divided by possible cases” is justified when the counted outcomes are equally probable. It becomes misleading when counting categories that group different numbers of outcomes.

For two rolls of a fair die, the thirty-six ordered pairs are equally probable. Possible sums range from two to twelve, but these eleven values are not equally probable: just one pair gives two, whereas six give seven. Changing how an experiment is represented therefore does not automatically preserve a uniform distribution over the new categories.

Theoretical probability must be distinguished from observed frequency. A frequency is the number of occurrences divided by the number of trials performed. It is an observed value that can vary from one series to another. A probability is a model parameter. Their relationship will be justified under repetition assumptions, not through immediate identification.

2. Count without repetition or double counting

If a first decision offers a possibilities and a second offers b possibilities for each, there are a b pairs of choices. This principle assumes that the second stage really offers b possibilities for every first decision. When that number varies, the corresponding contributions must be added. A tree shows this distinction instead of mechanically applying a product.

To form a four-character code with ten possible choices at each position and repetition allowed, we obtain 10^4 codes. If the characters must be distinct, the successive choices are ten, nine, eight and seven: we obtain 5,040 codes. Changing the condition changes the count even though the code's length remains the same.

Order matters in a code or a podium. Choosing three people from five to occupy three successive positions gives $5 \times 4 \times 3 = 60$ possibilities. Choosing a group of three from the same five, without distinct roles, gives only ten groups. Each group has been counted $3! = 6$ times among the ordered lists, hence $60/6 = 10$.

For n distinct objects, the number of permutations is n!. Define 0! = 1, a convention consistent with several identities. The number of k-element subsets selected from n is the binomial coefficient:

$C(n,k) = n! / [k! (n - k)!]$, for $0 \leq k \leq n$.

The formula precisely expresses the removal of ordering. We also have $C(n,k) = C(n,n - k)$: choosing retained elements is equivalent to choosing their complement. This symmetry provides a check and simplifies some calculations.

For $n \geq 2$ and $1 \leq k \leq n - 1$, Pascal's relation, $C(n,k) = C(n - 1,k - 1) + C(n - 1,k)$, is proved by separating groups that contain a fixed object from those that do not. The proof gives meaning to both terms; it is more useful than a factorial identity learned without interpretation.

3. Union, intersection and complement

The complement of A, written not A or A^c, contains every outcome not belonging to A. Its probability is $1 - P(A)$. This relation is particularly effective for “at least one”, by first calculating “none” when the latter event has a simpler description.

The intersection $A \cap B$ means that both events occur. The union $A \cup B$ means that at least one occurs, possibly both. The general formula is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Subtraction removes the intersection's double counting. If the events are mutually exclusive, their intersection is empty and this contribution is zero.

Mutual exclusivity and independence are different concepts. Two mutually exclusive events with strictly positive probabilities are not independent: one occurring excludes the other. Conversely, independent events can occur together. An even die result and a die result of at least four can share outcomes without being

automatically independent; their relationship is calculated within the model.

A diagram of two sets can help explain the operations, but does not provide a reliable measure of their probabilities unless the areas have been constructed for that purpose. Diagrams organise reasoning; data and the distribution provide the calculation.

4. Conditional probability: change the reference population

When P(A) is strictly positive, the probability of B given A is defined by P(B|A) = P(A ∩ B)/P(A). The denominator corresponds to the event that becomes the reference. We can rewrite P(A ∩ B) = P(A) P(B|A). This relation calculates a branch of a tree without assuming independence.

In a fictional population of batches, 40% come from workshop A and 60% from workshop B. The probability that a batch meets a stated criterion is 90% within A and 60% within B. The event R denotes meeting the criterion. The joint probabilities are P(A ∩ R) = 0.40 × 0.90 = 0.36 and P(B ∩ R) = 0.60 × 0.60 = 0.36.

Since the two workshops form a partition of the population, P(R) = 0.72. Among batches meeting the criterion, the probability of originating from A is 0.36/0.72 = 0.50. It is neither 0.40 nor 0.90. The calculation has changed the reference population: we no longer consider every batch, only those satisfying the criterion.

This example shows why P(R|A) and P(A|R) are not interchangeable. The first describes results within the workshop; the second describes origins among the retained results. The same events appear in both expressions, but their denominators differ.

A two-way table provides another representation. Among a thousand fictional batches matching the model's proportions, 400 come from A and 600 from B; 360 in each group meet the criterion.

Of the 720 retained batches, half come from A. The table can make conditioning more tangible than symbols, without replacing the definition.

5. Total probability and the effect of base rates

A partition is a family of mutually exclusive events covering the sample space. For a partition A_1, …, A_m whose probabilities are positive, P(B) is the sum of P(A_i) P(B|A_i). Contributions belonging to distinct groups are added. The formula does not require B to be independent of the groups.

A fictional checking device flags 95% of defective parts and also flags 5% of nondefective parts. The model assumes that 1% of parts are defective. Among ten thousand theoretical parts, there are one hundred defective ones, ninety-five of which are flagged, and nine thousand nine hundred nondefective ones, four hundred and ninety-five of which are flagged. The total number flagged is therefore five hundred and ninety.

The probability that a flagged part is actually defective is 95/590, approximately 16.1%. A high proportion of defective parts detected therefore does not mean that every flag is reliable. The base rate of defects and false flags contribute to the result. The example concerns no medical decision: it illustrates conditional reasoning in imaginary manufacturing only.

The same device applied to a population where defects are much more frequent would produce a different conditional probability after a flag. It would therefore be wrong to present the result as an isolated property of the device. This caution becomes essential when a percentage is communicated without describing the starting population.

6. Independence: an assumption allowing a product

Two events A and B are independent when $P(A \cap B) = P(A)\,P(B)$. If $P(A)$ is positive, this condition is equivalent to $P(B|A) = P(B)$. Learning that A occurred then does not change the probability of B within the model. Independence expresses a probability structure, not the absence of spatial proximity or a general absence of common causes.

Two successive tosses of a coin are often modelled as independent with a constant success probability. Draws with replacement can also be independent if the mixing and mechanism are suitable. Draws without replacement from a small urn generally are not: the remaining composition changes after the first draw.

An urn contains three red balls and two blue ones. Without replacement, P(red on the first draw) = 3/5, and the probability of red on the second given red on the first is 2/4. The probability of two reds is therefore 3/10. With replacement and mixing consistent with the model, it would be $(3/5)^2 = 9/25$. These are different experiments, not alternative calculations for the same procedure.

For several trials, use joint independence appropriate to the sequence considered. Checking pairwise relations is not generally sufficient to establish a family's full independence structure. In Bernoulli-trial exercises, independence of all the trials is a stated model assumption; it must be mentioned before multiplying.

7. A random variable counts what is studied

A random variable assigns a number to each outcome. Different outcomes can produce the same value. Its distribution specifies the possible values and their probabilities. Its expectation is the weighted average of these values by their probabilities, not a value the experiment must necessarily produce.

A Bernoulli variable is one on success and zero otherwise, with success probability p. Its expectation is p. Since its square equals itself, its variance is $p - p^2 = p(1 - p)$. The word success is a counting convention: it can denote a defect, failure or undesirable event without a value judgement.

For a variable X, variance is written $V(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$. It measures quadratic spread around the expectation. Standard deviation is its square root and has the same unit as X. Variance in squared seconds and standard deviation in seconds are not interchangeable.

Linearity $E(aX + bY) = aE(X) + bE(Y)$ does not require independence. In contrast, adding the variances of two variables requires conditions, notably independence in the framework studied here. If $Y = X$, the variance of $X + Y$ is $4\,V(X)$, not $2\,V(X)$. This counterexample shows why the condition cannot be removed.

8. Construct the binomial distribution rather than guessing it

A Bernoulli scheme repeats n independent trials, each with two outcomes and the same success probability p. The variable X counts successes. It then follows a binomial distribution with parameters n and p. Check the four elements-repetition, two categories, constant probability and independence-before choosing this distribution.

For a specified sequence containing k successes and $n - k$ failures, the probability is $p^k (1 - p)^{(n - k)}$. There are $C(n,k)$ sequences giving exactly k successes, because the success positions must be chosen. These sequences are mutually exclusive, so their identical probabilities are added:

$P(X = k) = C(n,k)\, p^k (1 - p)^{(n - k)}$.

The binomial coefficient is not a mysterious factor added to a power. It accounts for the fact that the same number of successes can appear in several orders. For n = 3 and k = 2, success-success-failure, success-failure-success and failure-success-success are distinct sequences, hence the factor three.

The formula naturally extends to p = 0 or p = 1 through the corresponding certain distributions, rather than ambiguous manipulations of 0^0. When p = 0, X is always zero; when p = 1, X is always n. These limiting cases provide a conceptual check of the distribution.

9. Expectation, variance and the scope of averages

Write X as the sum of n independent Bernoulli variables X_1, …, X_n. By linearity, E(X) = n p. By adding independent variances, V(X) = n p(1 − p). The standard deviation is √[n p(1 − p)]. These results concern the number of successes, not a frequency; the frequency X/n has a different variance.

In fifteen independent trials with p = 0.80, the expected number of successes is twelve and the variance 2.4. The standard deviation is approximately 1.55. This does not mean that every series produces twelve successes or that every result lies between twelve minus 1.55 and twelve plus 1.55. Expectation and standard deviation summarise the distribution without replacing probabilities of specific events.

An expectation can be noninteger even when X is always integer-valued. With three trials of probability 0.50, the expectation is 1.5 successes. No experiment observes one and a half successes; the number describes a theoretical average. Distinguishing a possible value from an average avoids rounding an expectation and treating it as a guaranteed result.

If probability p depends on the trial, the total number of successes generally no longer follows the classic binomial distribution, even under independence. Its expectation can still be found by adding the individual probabilities. Another method is then needed for its distribution or variance. A slightly more complicated model can be more faithful than a binomial distribution chosen out of habit.

10. Calculate events and solve a threshold problem

For ten independent trials with p = 0.20, the probability of two successes is $C(10,2) \times 0.20^2 \times 0.80^8$, approximately 0.302. The probability of at least one success is $1 - 0.80^{10}$, approximately 0.893. The second question is more easily handled through the complement; adding ten terms of the distribution would be possible but unnecessarily lengthy.

Find the smallest number n of independent trials with probability 0.20 such that the probability of at least one success reaches 95%. We need $1 - 0.80^n \geq 0.95$, hence $0.80^n \leq 0.05$. The logarithm gives $n \ln 0.80 \leq \ln 0.05$. Since ln 0.80 is negative, division reverses the inequality: $n \geq \ln 0.05 / \ln 0.80$, approximately 13.43. The smallest integer is fourteen.

Check the neighbouring values: thirteen trials give approximately 94.5%, fourteen approximately 95.6%. This check confirms both the integer rounding and minimality. A calculator displaying 13.43 does not directly answer a question about a number of trials.

For a probability of the form $P(X \leq k)$, add the terms from zero to k inclusive. For $P(X > k)$, use $1 - P(X \leq k)$ if convenient. Read inclusion of the threshold precisely: “more than five” does not mean “at least five”. Many binomial-distribution errors arise from this translation, even when the formula and parameters are correct.

11. Worked exercises and interpretation

Exercise A. Form a committee of two from six people, then appoint a chair from the two. There are C(6,2) = 15 committees and two choices of chair for each: thirty possibilities. Alternatively, choose the chair first, then the other member, giving 6 × 5. Both methods recover thirty because they count the same configurations.

Exercise B. An urn contains four red balls and one blue. Two draws without replacement are performed. The probability of obtaining exactly one red is P(red then blue) + P(blue then red), or (4/5)(1/4) + (1/5)(4/4) = 2/5. Using a binomial distribution with p = 4/5 would be unsuitable because the second conditional probability changes.

Exercise C. Six independent trials have success probability 1/2. The probability of at least five successes is [C(6,5) + C(6,6)]/2^6 = 7/64, approximately 10.9%. The calculation includes five and six. Symmetry of the distribution gives the same probability for at most one success, providing a check.

Exercise D. Define a score Y = 3X − 2, where X follows a binomial distribution with parameters ten and 0.20. Its expectation is 3 × 2 − 2 = 4. Its variance is 9 × 1.6 = 14.4. Subtracting two does not change the spread; multiplying by three multiplies standard deviation by three and variance by nine.

Exercise E. “An event has probability 1%, so it must occur once in one hundred trials.” This assertion is false. Under independence and constant probability, the probability of never observing it in one hundred trials is 0.99^{100}, approximately 36.6%. The expected number of occurrences is indeed one, but does not impose the result of every series.

12. What a probability solution must make visible

A solution must describe events before assigning notation, state the procedure, justify conditioning and recall the assumptions allowing products. It must present a probability between zero and one, or the corresponding percentage, then interpret it in the right population. An expectation, frequency and conditional probability cannot be substituted for one another.

Numerical accuracy must suit the question. To compare two probabilities close to a threshold, retain enough figures until the conclusion. For a general illustration, three significant figures may suffice. In every case, final rounding must not hide the direction of an inequality or turn an estimate into certainty.

Framework reference

The French Terminale mathematics specialism programme provides the framework for counting, Bernoulli schemes and sums of variables. The proofs, fictional situations and exercises above are original: https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm

Chapter 11 - Geometry in space: vectors, planes and distances

Historical perspective - Descartes: connect a figure with equations

La Géométrie accompanied Descartes's Discours de la méthode in 1637. The BnF's institutional presentation highlights his approach of converting geometric problems into equations. The connection with our chapter is the movement between a geometric object and an algebraic expression: a vector, line or plane must retain its meaning when we calculate its coordinates. This historical perspective does not mean that Descartes already used our complete modern vector formalism in three dimensions. It helps us understand a method of representation without replacing a geometric check of the result.

Historical source: Bibliothèque nationale de France - institutional presentation: https://www.bnf.fr/fr/mediatheque/les-equations-la-methode-de-descartes-pour-resoudre-les-problemes-geometriques

1. See a figure without asking it to prove anything

A representation of space on a sheet of paper preserves some relationships and distorts others. Two segments drawn parallel may represent different directions; a right angle may not look right in perspective. The figure helps organise points, faces and assumptions. A proof rests on established geometric or algebraic relationships, rather than on a visual impression alone.

A cube drawn in perspective, for example, prompts us to distinguish a face diagonal from a body diagonal. The first lies in the plane of a face; the second connects opposite vertices through space. In a cube with edge length a, their lengths are $a\sqrt{2}$ and $a\sqrt{3}$ respectively. The drawing does not allow us to compare them with

a ruler if the projection shortens different directions by different amounts.

Coordinate geometry translates configurations into coordinates and equations. It does not remove geometry: it supplies another language for expressing collinearity, parallelism, orthogonality or intersection. A complete solution must return to the spatial meaning after the calculation. A system with no solution tells us something about the objects being studied; it is more than a technical result.

2. Points, vectors and coordinate systems: three different objects

A point describes a position. A vector describes a directed displacement, or a direction with a magnitude. In a coordinate system, the coordinates of the vector AB are those of B minus those of A. If A = (1,2,−1) and B = (4,0,3), then AB = (3,−2,4). Calling the coordinates of B alone the “coordinates of AB” would be wrong, unless A is the origin of the coordinate system.

A vector can be represented by directed segments in different locations. Equality of two vectors means the same direction, the same sense and the same length, without necessarily having the same endpoints. This distinction lets us use the sides of a parallelogram or a cube without confusing displacement with absolute position.

A basis of space consists of three vectors such that every vector can be expressed uniquely as a linear combination of them. In an orthonormal coordinate system, the basis vectors are unit vectors and are pairwise orthogonal. The usual coordinate formulae for the norm and the scalar product depend on this property. They cannot be applied unchanged to an arbitrary basis.

For example, if two unit basis vectors make an angle of sixty degrees, the norm of their sum is $\sqrt{3}$, rather than $\sqrt{2}$. Coordinates alone are therefore insufficient to calculate a length without knowing the geometry of the basis. In this chapter's exercises, the coordinate system is orthonormal unless explicitly stated otherwise.

3. Linear combinations, collinearity and coplanarity

Two non-zero vectors are collinear if one is a real multiple of the other. The multiplier may be positive or negative; its sign indicates the same sense or the opposite sense. To check collinearity using coordinates, we look for one coefficient common to all components. We do not blindly divide by a component that might be zero.

The vectors $(2,-1,0)$ and $(-6,3,0)$ are collinear with coefficient -3. The vectors $(2,-1,0)$ and $(-6,3,1)$ are not, because the third component prevents equality. The zero vector is a linear combination of any vector, but it cannot serve as the direction of a line. This particular feature must be kept in mind whenever a proof uses a direction vector.

Three points A, B and C are collinear when AB and AC are collinear. To show that four points are coplanar, we can ask whether AD is a combination of AB and AC, provided these last two vectors are not collinear. We then solve $AD = \alpha\,AB + \beta\,AC$. The existence of a solution means that D belongs to the plane through A spanned by these two directions.

Let A be at the origin, $AB = (1,0,1)$ and $AC = (0,1,1)$. Every point in the plane has coordinates $(\alpha,\beta,\alpha+\beta)$. The equation of the plane is therefore $z = x + y$. The point $(1,1,2)$ belongs to it; the point $(1,1,3)$ does not. This interpretation directly connects a linear combination with a Cartesian equation.

When the two chosen vectors are collinear, they span only a line. We must then choose other directions to describe the plane, or handle a degenerate case. A formula written with two parameters is not automatically a representation of a plane if its directions are not independent.

4. Represent a line with one parameter

A line through A with non-zero direction vector u is the set of points M such that AM = t u, where t ranges over the real numbers. If A = (a,b,c) and u = (u_1,u_2,u_3), we write x = a + t u_1, y = b + t u_2, z = c + t u_3. All three coordinates depend on the same parameter t. Using three independent parameters would describe a different set.

The starting point and the length of the direction vector are not uniquely determined. Replacing u with 2u preserves the line; the parameter is simply rescaled. Replacing A with another point on the line also preserves the set. Two different expressions can therefore represent exactly the same object.

To test whether a point belongs to the line, we seek a single value of t that satisfies all the coordinates simultaneously. For the line (1 + 2t, t, 2 − t), the point (3,1,1) corresponds to t = 1 in all three equations. The point (3,1,2) does not work: the first two coordinates require t = 1, but the third would then be one.

A whole line, a ray and a segment differ in the set of allowed parameter values. For the segment from A to B, the expression A + t AB with t between zero and one is appropriate. If t ranges over all real numbers, the expression describes the line, including the portions beyond the endpoints. Forgetting this restriction could produce an intersection outside the segment actually being studied.

5. Scalar product, norm and angle

In an orthonormal coordinate system, $u \cdot v = u_1v_1 + u_2v_2 + u_3v_3$. The norm of u is $\sqrt{(u \cdot u)}$. For non-zero vectors, the scalar product also equals $||u||\ ||v|| \cos \theta$, where θ is the angle between the vectors, between zero and π. These two expressions connect calculation using coordinates with geometric interpretation.

Two vectors are orthogonal when their scalar product is zero. For $u = (1,2,-1)$ and $v = (2,-1,0)$, the product is two minus two, hence zero. The check requires all components; the apparent perpendicularity of their projections in a figure is insufficient.

In a cube with edge length one, the vector along an edge can be $(1,0,0)$ and the vector along a body diagonal $(1,1,1)$. The cosine of the angle between them is $1/\sqrt{3}$. The angle is therefore approximately 54.7 degrees. It is not forty-five degrees, which is the angle between a face diagonal and an edge of that face. These are different configurations.

Expanding $||u + v||^2 = ||u||^2 + 2u \cdot v + ||v||^2$ recovers Pythagoras's theorem when the scalar product is zero. It also allows us to express the scalar product in terms of lengths. The sign of the cross term indicates whether the sum is longer or shorter than it would be in the orthogonal case.

Two lines with orthogonal directions do not necessarily intersect. In space, the line $(t,0,0)$ and the line $(0,1,s)$ have orthogonal directions, but no common point, since their second coordinates remain zero and one. We speak of orthogonality of directions; describing lines as perpendicular usually also requires them to meet. This distinction prevents us from importing an intuition from the plane too quickly.

6. Define a plane using a normal vector

A non-zero normal vector n is orthogonal to every direction in a plane. If the plane passes through A, a point M belongs to it when $n \cdot AM = 0$. For n = (a,b,c) and A = (x_A,y_A,z_A), we obtain $a(x - x_A) + b(y - y_A) + c(z - z_A) = 0$, giving an equation $a\,x + b\,y + c\,z + d = 0$.

At least one of the coefficients a, b and c must be non-zero. Otherwise, the equation would be an identity or a contradiction, rather than the equation of a plane. Multiplying all coefficients by the same non-zero constant preserves the plane. Its normal vector is therefore not unique, and the sense of that vector does not impose a compulsory orientation on the plane.

The plane $x + 2y - z - 3 = 0$ has normal vector $(1,2,-1)$ and contains A = (3,0,0). The point (1,1,0) also belongs to it. A direction with vector u lies in this plane if $n \cdot u = 0$; a point lies in the plane if its coordinates satisfy the equation. Confusing these two tests means overlooking the constant term that fixes the position.

To construct a normal from two non-collinear directions, we can seek n whose scalar product with each of them is zero, by solving a system. There is no need to introduce an unfamiliar technique to solve a simple case. The vector product provides another method, but here it is a possible extension, rather than a hidden prerequisite.

7. Relative positions: a system gives a geometric conclusion

To intersect the line A + t u with the plane $n \cdot M + d = 0$, we substitute the coordinates of M. We obtain $n \cdot A + d + t(n \cdot u) = 0$. If $n \cdot u$ is non-zero, there is a unique value of t, and therefore a unique intersection point. If $n \cdot u$ is zero, the line is parallel to the plane; it is contained in the plane if $n \cdot A + d$ is also zero, and otherwise does not meet it.

The line $M(t) = (1 + 2t, t, 2 - t)$ and the plane $x + 2y - z - 3 = 0$ give $-4 + 5t = 0$. The intersection corresponds to $t = 4/5$ and the point $(13/5,4/5,6/5)$. We check this point in the plane's equation. If the same parametrisation described a segment with t between zero and one, the intersection would remain admissible; with t between two and three, it would no longer be admissible.

For two lines, we use distinct parameters t and s, then equate their coordinates. A unique solution gives an intersection. No solution may correspond to distinct parallel lines or to skew lines; their directions allow us to distinguish these situations. Infinitely many solutions can express coincidence, when the lines describe the same set.

Two planes whose normal vectors are collinear are parallel or coincident. With non-collinear normals, their intersection is a line. For example, the planes $x + y + z = 1$ and $x - y = 0$ give $x = y = t$ and $z = 1 - 2t$. The direction of their intersection line is $(1,1,-2)$. Every proposed point must satisfy both equations, rather than only one of them.

8. Orthogonal projection onto a line

The orthogonal projection H of a point M onto a line through A with direction u belongs to the line and has MH orthogonal to u. Writing $H = A + t\,u$ turns the condition into $(A + t\,u - M)\cdot u = 0$. Since u is non-zero, we obtain $t = [(M - A)\cdot u]/||u||^2$.

For A at the origin, $u = (1,1,0)$ and $M = (2,2,1)$, we find $t = 4/2 = 2$. The projected point is $H = (2,2,0)$, and the distance from M to the line is one. Had the direction vector been replaced by $(2,2,0)$, the parameter would have been one; the point H and the distance would have been identical.

Why is this distance minimal? For every other point P on the line, MH is orthogonal to HP. Pythagoras's theorem gives $MP^2 = MH^2 +$

HP^2, so $MP \geq MH$. Equality requires $P = H$. This proof connects the construction with the optimisation problem, rather than presenting a distance formula without a reason.

Finding the distance to a segment may require an additional step. If the projection onto the line lies outside the segment, the nearest point on the segment is an endpoint. The problem is no longer exactly that of an infinite line. Checking the value of the parameter avoids calculating the correct distance for the wrong object.

9. Orthogonal projection onto a plane

For the plane $a\,x + b\,y + c\,z + d = 0$, with normal $n = (a,b,c)$, the projection H of M lies on the line through M with direction n. We can write $H = M - \alpha\, n$. Requiring H to belong to the plane gives $\alpha = (a\,x_M + b\,y_M + c\,z_M + d)/(a^2 + b^2 + c^2)$.

For $M = (1,1,4)$ and the plane $x + 2y - z - 3 = 0$, the numerator is -4 and the denominator six. We obtain $H = M + (2/3)(1,2,-1)$, giving $H = (5/3,7/3,10/3)$. Substitution into the plane's equation gives zero. The vector MH is collinear with the normal, so both conditions for the projection are satisfied.

The distance to the plane is the norm of MH, namely $|a\,x_M + b\,y_M + c\,z_M + d|/\sqrt{(a^2 + b^2 + c^2)}$. In the example, this is $4/\sqrt{6}$, approximately 1.63 units of length. The absolute value removes the sign associated with the side of the plane; a distance cannot be negative.

Minimality is proved as for the line. For any point P in the plane, HP is a direction in the plane and is therefore orthogonal to MH. Thus $MP^2 = MH^2 + HP^2$. The point H is the only point attaining the minimum distance. The formula is therefore more than a convenient calculation: it answers a proved geometric property.

10. Geometric loci and volumes

The perpendicular bisector plane of segment AB is the set of points equidistant from A and B. Expanding $MA^2 = MB^2$ cancels the quadratic terms in the coordinates of M and yields the equation of a plane. This plane passes through the midpoint of the segment and has AB as a normal vector. For A = (0,0,0) and B = (2,0,0), its equation is $x = 1$.

A sphere with centre O and radius r is the set of points M such that OM = r. In an orthonormal coordinate system centred at O, its equation is $x^2 + y^2 + z^2 = r^2$. The corresponding solid ball uses a less-than-or-equal-to inequality. The surface and the solid are different sets.

The sphere $x^2 + y^2 + z^2 = 4$ meets the plane $z = 1$ in the circle $x^2 + y^2 = 3$, lying in that plane and with radius $\sqrt{3}$. A plane at a distance greater than two from the centre does not meet the sphere; at a distance of exactly two, it is tangent to it. These cases can be understood through the distance from the centre to the plane, before any detailed calculation.

For a tetrahedron with vertices O = (0,0,0), A = (2,0,0), B = (0,3,0), C = (0,0,4), the base OAB is a right-angled triangle with area three. The height from C is four. The volume is therefore one third of three times four, giving four units of volume. The height used in this definition must be perpendicular to the base; the length of an oblique edge is not always the height.

11. Worked exercises and checks

Exercise A. Are the points A = (1,0,2), B = (3,1,1) and C = (5,2,0) collinear? AB = (2,1,−1) and AC = (4,2,−2) = 2 AB. They are collinear. The coefficient two means that C lies in the same direction from A, twice as far away as B; it is not between A and B.

Exercise B. Find the plane through A = (1,0,2) with normal n = (2,−1,3). Its equation is 2(x − 1) − y + 3(z − 2) = 0, giving 2x − y + 3z − 8 = 0. First check A; the direction (1,2,0) is indeed a direction in the plane, since its scalar product with n is zero.

Exercise C. Does the line (t,2 − t,1) lie in the plane x + y + z = 3? The sum of the coordinates is always three. The whole line is therefore contained in the plane, rather than just the point obtained for a particular value of t. A symbolic substitution here allows a stronger conclusion than an isolated numerical test.

Exercise D. Calculate the distance from M = (0,0,0) to the plane 2x − y + 2z − 6 = 0. The absolute value of the numerator is six, and the norm of the normal is three. The distance is two. Multiplying the plane's equation by five multiplies both numerator and denominator by five; the result must remain unchanged.

Exercise E. Do a line with direction (1,0,0) and a line with direction (0,1,0) necessarily intersect? No. Their directions are orthogonal, but they can lie in distinct horizontal planes. To conclude that they meet, we must also study their points and solve the intersection conditions.

12. Organise a proof in space

The method depends on the question. For membership, we substitute coordinates or seek a parameter. For parallelism, we compare directions or normals. For orthogonality, we use a scalar product while specifying the objects involved. For a minimum distance, we construct a projection and justify the minimum using Pythagoras's theorem. This classification turns a complex figure into a sequence of targeted tests.

Calculations must retain their interpretation. A parameter value is not a distance; a normal vector is not a point; an equation satisfied by one point does not automatically describe a whole line. After

each step, restate what the result establishes and what remains to be shown. This discipline allows us to use algebra without losing sight of the geometric object.

Curriculum reference

French Terminale mathematics specialism programme, sections on vectors, orthogonality and representations of space. The configurations and educational proofs are original:
https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm

Chapter 12 - Statistics, concentration and algorithmic methods

Historical perspective - Bienaymé and Chebyshev: bound deviations

Bienaymé published the inequality now associated with the names Bienaymé and Chebyshev in 1853; Chebyshev published it in turn in 1867. This history clarifies the role of a probability bound: it can control the risk of a deviation without describing the entire distribution exactly. The chapter therefore distinguishes an observed frequency, a theoretical probability and an upper bound that may be imprecise but remains valid under its assumptions. This perspective does not guarantee that a particular sample will be close to the mean or attribute the discovery to just one name.

Historical source: University of St Andrews - MacTutor: https://mathshistory.st-andrews.ac.uk/Biographies/Bienayme/

1. Describe data before interpreting them

A series of numbers has no meaning independent of its origin. We ask what quantity is measured, in which unit, for which population, on which dates and according to which procedure. Durations, masses and categories encoded as numbers are not handled in the same way. A postcode is not a quantity whose mean would have a direct geographical meaning.

Descriptive statistics summarise the available data. Inference seeks to extract information about a population or mechanism beyond those observations. Moving from one to the other requires assumptions about sampling. A graph based on one hundred voluntary responses can describe those responses; it does not automatically represent every inhabitant of a territory.

Before calculating, check missing values, duplicates, changes of unit and categories. An absent value should not be replaced with zero without justification. Zero is a value, and introducing it can change a mean or a proportion. Similarly, a duplicate is not always an independent repetition; it may be a data-entry error.

An honest summary keeps the sample size and collection method close to the result. Saying “the mean is six” does not allow us to assess a series without its unit, dispersion or size. Display precision does not create information precision: writing six to ten decimal places corrects neither a coarse measurement nor a biased selection.

2. Mean, median and dispersion

For n observations x_1, …, x_n, the arithmetic mean is their sum divided by n. It uses every value and can be sensitive to extreme observations. For groups of different sizes, the overall mean must weight the group means by their sizes; simply taking the mean of the means is correct only in particular cases.

The median divides an ordered series into two parts. With an odd number of observations, it is the central value. With an even number, the usual calculation convention is to take the mean of the two central values. This convention must be understood when comparing software or presentations. Quartiles also require an explicit convention: some methods choose observed values, while others interpolate.

For the series 2, 3, 3, 4, 18, the mean is six and the median three. The value eighteen pulls the mean upwards. This does not make the mean wrong; it shows that the mean alone does not describe a “typical” value. Depending on the question, we can present mean, median and dispersion together, rather than choosing the indicator that best supports a desired conclusion.

The descriptive variance with divisor n is the mean of the squared deviations from the mean. It can also be calculated as the mean of the squares minus the square of the mean. In the preceding series, the mean of the squares is 72.4 and the square of the mean thirty-six; the variance is therefore 36.4 and the standard deviation approximately 6.03. The standard deviation retains the unit of the observations; variance uses its square.

A sample variance with divisor n − 1 follows a different convention and serves a particular estimation objective. It is not a correction to apply automatically to every descriptive series. Before using a calculation command, we must know which divisor and which purpose it adopts. Two different results can arise from this convention without an arithmetic error.

3. Weighting and composition effects

Two groups have respective means of ten and twenty. If the first contains ninety observations and the second ten, the overall mean is eleven, rather than fifteen. Weighting expresses the amount of information in each group. It cannot be replaced with an intuitive choice when the group sizes are known.

Comparisons between populations may change with their composition. Suppose there are two fictional activities, one easy and one difficult. An organisation can have a higher success rate in each activity yet a lower overall rate if it handles many more difficult cases. Before announcing an overall advantage, examine the categories and their distributions.

This phenomenon does not justify rejecting every aggregate indicator. It invites us to make the question explicit: do we want to measure performance within a comparable category, or an organisation's total outcome in its actual activity? Both questions are worthwhile, but do not necessarily require the same weighting.

Rigour means avoiding an unannounced move from one to the other.

A graph must also retain a readable scale. A truncated vertical axis can make a small change look enormous; a logarithmic scale can be useful for comparing ratios, but it must be identified. The representation is not additional proof independent of the numbers it depicts.

4. Sums and means of random variables

Consider n independent, identically distributed variables with expectation μ and finite variance σ². Their sum S_n has expectation nμ and variance nσ². Their mean M_n = S_n/n has expectation μ and variance σ²/n. Its standard deviation is σ/√n. This relationship expresses the reduction in the mean's dispersion within this model of repetition.

Halving the standard deviation therefore requires multiplying the sample size by four, rather than two. Multiplying n by one hundred divides the mean's standard deviation by ten. Improvement does not progress linearly with the number of observations. This dependence explains the cost of increased sampling precision.

The assumptions are essential. If all observations repeat the same shared random error, their mean does not benefit from the same reduction. If the values do not follow the same distribution, their interpretation requires adaptation. If a collection method always selects the same category, a large sample can make a mean very precise without making it representative of the intended population.

The relationship σ/√n is a theoretical property of the distribution of the mean, rather than a promise that every observed mean will lie at exactly that distance from μ. Nor does it mean that the standard deviation of individual observations decreases as more data are

collected. The observations retain their distribution; it is their mean that becomes less dispersed in the model.

5. Prove the Bienaymé-Chebyshev inequality

For a variable X with expectation μ and finite variance V, and for a strictly positive δ, the inequality states:

$P(|X - \mu| \geq \delta) \leq V/\delta^2$.

It relates the probability of a substantial deviation to a measure of dispersion. If the right-hand side exceeds one, the bound remains mathematically true but provides less information than the elementary bound of one for a probability. We can then retain the minimum of the two values.

A short proof consists of setting $Y = (X - \mu)^2$, a non-negative variable. On the event $|X - \mu| \geq \delta$, we have $Y \geq \delta^2$. The expectation of Y is at least δ^2 multiplied by the probability of this event, since all remaining contributions are non-negative. Now $E(Y) = V$. Dividing by δ^2 yields the inequality. This proof shows why the square and non-negativity are involved.

The inequality does not assume a normal distribution. Its generality is an advantage, but the bound may be loose. A result $P \leq 0.20$ does not mean that the probability is 0.20. The true probability may be much smaller. Confusing an upper bound with an equality would undermine the very meaning of this tool.

The event includes deviations greater than or equal to the stated threshold. When X is discrete, inclusion of the boundary can matter. Translate the condition before calculating an exact probability or comparing it with a bound.

6. Concentration of a mean and the law of large numbers

Applying the inequality to M_n, whose variance is σ²/n, gives P(|M_n − μ| ≥ δ) ≤ σ²/(nδ²). For a fixed positive δ, the right-hand side tends to zero as n increases. Thus, the probability of a deviation of at least δ becomes arbitrarily small in the model. This is a form of the law of large numbers.

The conclusion concerns a probability, rather than a monotonic evolution of every sequence of observed means. A new observation can temporarily move the mean further from μ. Nor does the law state that a future series must compensate for a past one. After several independent failures, the probability of success on the next trial remains p; it does not increase to “restore the balance”.

For a success frequency F_n in a Bernoulli scheme, μ = p and σ² = p(1 − p). The bound becomes P(|F_n − p| ≥ δ) ≤ p(1 − p)/(nδ²). Since p(1 − p) ≤ 1/4 for every p between zero and one, we have a uniform bound of 1/(4nδ²), even when p is unknown.

To ensure an upper bound of 5% on the probability of a deviation of at least 0.05, the uniform bound requires n ≥ 1/(4 × 0.05 × 0.05²), giving two thousand. This is a sufficient size according to this inequality, rather than necessarily the smallest possible size for a particular distribution. Targeting 0.02 with the same upper bound on the risk gives a sufficient size of twelve thousand five hundred.

These numbers rely on independent trials and a constant probability. They guarantee neither the representativeness of a voluntary survey nor the validity of a questionnaire. Sample size alone does not repair selection bias. Probabilistic precision and the suitability of the procedure are two different requirements.

7. What can be said after an observation

Before the trials are conducted, F_n is a random variable. Afterwards, its value f is a datum. The concentration bound provides control over possible repetitions of the procedure. By itself, it does not assign a posterior probability to the parameter p after f has been observed.

We can construct an interval around the observed frequency with a coverage guarantee derived from a stated method. This means that, over repetitions of the procedure satisfying the assumptions, the method produces intervals containing p with a controlled frequency. Without a further framework, this should not be rephrased as “this fixed parameter now has such-and-such a probability of being in this interval”.

For revision purposes, it is often sufficient to distinguish estimation, uncertainty and a guarantee. A frequency of 0.42 does not establish p = 0.42 exactly. A risk bound does not establish that the procedure will succeed in every case. A high level of confidence in a statistical framework does not remove the possibility of a poorly designed procedure.

When a decision concerns people, a technical system or a sensitive phenomenon, statistical analysis is only one element of the reasoning. The cost of errors, the constraints and the validity of the measurement must also be examined. This guide does not turn a concentration method into personalised medical or financial advice.

8. Design an algorithm with an invariant and termination

An algorithm is a sequence of instructions whose result can be explained. It is insufficient for a programme to display a plausible number. We specify its inputs, their domain, its output, its steps and its stopping conditions. A loop should be accompanied by a preserved property, called an invariant, and an argument explaining why it terminates when this is required.

In bisection, the invariant is that the current interval contains a solution, under the assumptions of continuity and a change of sign being used. Its length is halved at each step. Termination for a positive tolerance follows from this geometric decrease. The proof does not depend on a particular programming language.

To find the smallest integer n such that q^n ≤ a, with 0 < q < 1 and 0 < a < 1, we can initialise n to zero and a variable v to one, then multiply v by q and increment n while v > a. The invariant is v = q^n. Convergence of q^n to zero ensures termination in exact arithmetic. A numerical implementation must also consider rounding and representation limits.

The condition test must correspond to the intended threshold. Replacing v > a with v ≥ a changes the answer when equality is possible. This is a logical error, rather than an aesthetic detail. Testing a case in which the threshold is reached exactly can reveal a defect invisible with ordinary values.

9. Simulation: explore a distribution without proving it

A Bernoulli simulation can generate a value U uniformly between zero and one, then count a success if U < p. Repeating this step n times simulates a series within the chosen numerical model. Repeating whole series lets us observe the distribution of success counts or frequencies.

The numbers used are generally pseudorandom: they come from a deterministic mechanism initialised by a seed. Fixing this seed can make a study easier to reproduce. This does not make the generator suitable for every application, particularly those requiring security properties. Here it serves only an educational exploration.

A simulation allows us to conjecture a relationship, check a formula or observe dispersion. It does not prove an identity valid for every parameter. If a rare event occurs in none of the thousand simulated series, we cannot conclude that its probability is zero. The number of series and their variability must accompany the conclusion.

We can also estimate an integral by simulation. For U uniform on [0,1], the expectation of U^2 is the integral of x^2 over this interval, namely one third. The mean of many values of U^2 approaches this quantity under the model's conditions. The individual variance is $1/5 - 1/9 = 4/45$; for ten thousand independent observations, the theoretical standard deviation of the mean is approximately 0.00298. A result close to 0.33 is expected, but does not prove the exact result.

10. Numerical stability and independent checks

An exact identity can lead to a fragile numerical calculation. Subtracting two very close numbers can lose some significant digits. Calculating enormous factorials and then dividing them can exceed a representation's capacity even when the final binomial coefficient remains manageable. The order and organisation of operations therefore also matter.

For a binomial distribution with $0 < p < 1$, we can start from $q_0 = (1 - p)^n$ and successively calculate $q_{k+1} = q_k \times (n - k)/(k + 1) \times p/(1 - p)$. This recurrence avoids repeatedly calculating

factorials. It does not solve every problem: for very large n, q_0 can become numerically too small, and a different treatment would be required. The cases p = 0 and p = 1 must be handled separately.

A useful check is to verify that all probabilities are positive and that their sum is close to one within the available precision. We can check the expectation calculated from the distribution and compare it with n p. These checks do not replace a proof of the algorithm, but detect parameter or indexing errors.

For the square root of two, a Newton iteration gives u_(n+1) = (u_n + 2/u_n)/2 with u_0 = 1. The first values are 1.5 and then approximately 1.41667. From u_1 onwards, the values remain above √2 and decrease. Setting e_n = u_n − √2 allows us to establish e_(n+1) = e_n²/(2u_n), a relationship explaining the rapid improvement in this case. It does not justify a guarantee of convergence for every Newton method applied to every function.

11. Worked exercises and discussion

Exercise A. Two groups of ten and thirty people have respective means of twelve and sixteen. The overall mean is (10 × 12 + 30 × 16)/40 = fifteen. Fourteen, the unweighted mean of the two means, would assign the groups equal weight despite their different sizes.

Exercise B. A variable has expectation ten and variance four. What upper bound does Bienaymé-Chebyshev give for a deviation of at least five? The bound is 4/25 = 0.16. It is correct to say that the probability is at most 16%. It would be incorrect to say that it is exactly 16%, or that 84% of the values must be uniformly distributed within the central interval.

Exercise C. Independent, identically distributed variables have variance nine. For their mean over one hundred observations, the variance is 0.09 and the standard deviation 0.30. Over four

hundred observations, the standard deviation is 0.15. Halving the standard deviation corresponds to multiplying the sample size by four, in accordance with the square-root dependence.

Exercise D. For a Bernoulli frequency with parameter p = 0.40 and n = one thousand, bound the probability of a deviation of at least 0.05. We find 0.40 × 0.60/(1 000 × 0.05²) = 0.096. The uniform bound would give 0.10. Knowing p slightly improves this upper bound, but does not give the exact probability.

Exercise E. A programme initialises v to one and multiplies v by 0.5 while v ≥ 0.125. It stops at v = 0.0625 after four multiplications. To find the first index at which v ≤ 0.125, the test should be v > 0.125, and it would stop after three multiplications. Analysing the equality case reveals the difference between the questions.

Exercise F. One hundred volunteers choose to respond to a survey after reading a message about a difficulty they experience. Can we apply a guarantee of independent, representative sampling to all users? Not without justification. Volunteers may have different characteristics from those who do not respond. Increasing the number of volunteers does not automatically remove this bias.

12. Build a synthesis for revision

For each concept, prepare a definition, a property with its assumptions, a short proof, an exercise and a counterexample. In probability, the counterexample can concern independence; in analysis, missing continuity; in geometry, a wrongly assumed intersection; in algorithms, a mistranslated loop condition.

A good synthesis does not recite every formula. It chooses the one that corresponds to the question and explains why the others would be insufficient. It preserves the distinction between proof, approximation and simulation. Finally, it checks that the answer

addresses the original problem, with appropriate precision and interpretation.

Curriculum and methodological references

French Terminale mathematics specialism programme, for concentration, the law of large numbers and algorithms; all situations and explanations are original: https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm

NIST, descriptive guidance on measures of location, consulted to distinguish the mean and median without reproducing its wording: https://www.itl.nist.gov/div898/handbook/eda/section3/eda351.htm

Appendices - Extended problems, worked solutions and revision strategies

A. Study a family of functions with a parameter

Consider f_a(x) = ln x − a x on the strictly positive real numbers, where a is a strictly positive parameter. We must distinguish the unknown x from the parameter a. While studying one function in the family, a remains fixed. The derivative is f_a'(x) = 1/x − a; it is positive before x = 1/a, zero at this point and negative afterwards.

The function therefore has a maximum at 1/a. Its value is ln(1/a) − 1, or −ln a − 1. The limit as x approaches zero through positive values is minus infinity. The limit as x tends to plus infinity is also minus infinity, since the negative linear term dominates the logarithm for a fixed positive a. These limits and the intervals of increase and decrease allow us to count the solutions of f_a(x) = 0.

If a < 1/e, the maximum is strictly positive. Continuity and strict monotonicity on each of the two intervals then give two distinct solutions, one before 1/a and the other after it. If a = 1/e, the maximum is zero: the unique solution is x = e. If a > 1/e, the maximum is negative and no solution exists. We have classified the entire family without having to calculate each root explicitly.

The reasoning separates existence from uniqueness. Continuity with values of opposite signs ensures a root; strict monotonicity on the interval under consideration ensures that there cannot be two. The maximum and the limits allow us to decide on which intervals these arguments apply. Simply inspecting a graph might suggest the result, but would not justify it for every parameter.

The second derivative is $-1/x^2$: all functions in this family are strictly concave. In the limiting case a = 1/e, the curve touches the axis at its maximum. A change of sign is therefore unnecessary for

a root to exist. This example complements the intermediate value method by showing a situation in which a search based only on opposite signs would miss the solution.

For revision, revisit the problem with ln x + a x, for positive a. The derivative becomes 1/x + a, which is always positive. The limits run from minus infinity to plus infinity, so the equation has a unique solution. A change of sign within the expression alters the structure of the problem; copying the previous variation table without recalculating would be an error.

B. A sequence defined by a fraction, with a justified limit argument

Define $u_0 = 1$ and u_(n+1) = (2u_n + 3)/(u_n + 2). Before seeking a limit, we must show that all terms exist and control their domain. The function $\varphi(x) = (2x + 3)/(x + 2)$ is defined on $[1,\sqrt{3}]$, because its denominator is positive there. Its derivative is $1/(x + 2)^2$, so it is increasing on this interval.

We have $\varphi(1) = 5/3$ and $\varphi(\sqrt{3}) = \sqrt{3}$. Since $1 \leq 5/3 \leq \sqrt{3}$, the function maps the interval into itself. Induction therefore proves that every term lies between one and $\sqrt{3}$. The calculation using this function is now guaranteed at every step.

The difference $\varphi(x) - x$ is $(3 - x^2)/(x + 2)$. It is non-negative on $[1,\sqrt{3}]$. Consequently, u_(n+1) ≥ u_n: the sequence is increasing and bounded above. It therefore converges to a limit ℓ belonging to this interval. Only at this stage does continuity of φ allow us to write $\ell = \varphi(\ell)$.

The fixed-point equation leads to $\ell^2 = 3$. The two algebraic roots are $\sqrt{3}$ and $-\sqrt{3}$, but the limit belongs to $[1,\sqrt{3}]$. It is therefore $\sqrt{3}$. Solving the equation at the outset would have supplied candidates without establishing the existence of a limit or selecting the correct candidate.

The first terms are $u_1 = 5/3$, $u_2 = 19/11$ and $u_3 = 71/41$, approximately 1.6667, 1.7273 and 1.7317. They are consistent with the proof, but do not replace it. A series of apparently stabilised values would be insufficient to justify convergence for every subsequent step.

An extension can bound the difference between images by a constant less than one on the interval. The derivative is at most 1/9 there, suggesting a strong contraction. This technique goes beyond the main requirements of the exercise and should be presented as an extension. The Terminale-level solution already rests on invariance, monotonicity, an upper bound and continuity.

C. Optimisation with a geometric constraint

A fictional rectangular plate must have an area of 48 square units. We seek the rectangle with minimum perimeter among those with strictly positive sides. If one side is x, the other is 48/x, and the perimeter is $P(x) = 2x + 96/x$ for positive x. The domain is part of the model; zero and negative values do not describe an admissible rectangle.

The derivative is $2 - 96/x^2$. It vanishes at $x = \sqrt{48} = 4\sqrt{3}$. It is negative before this value and positive afterwards. The global minimum is therefore attained by a square, whose sides are both $4\sqrt{3}$ and whose perimeter is $16\sqrt{3}$. The infinite limits at both ends of the domain are consistent with the result, but the variation table is already sufficient to establish the minimum.

A proof without differentiation uses the inequality $x + 48/x \geq 2\sqrt{48}$ for positive x, obtained by expanding a suitably chosen square. Equality requires $x = 48/x$. Both approaches establish the same result; the second can illuminate the problem's symmetry, while the first adapts easily to other cost functions.

Now suppose that one side must be at least ten. The domain becomes $x \geq 10$. Since ten exceeds $4\sqrt{3}$, the function is increasing there, and the constrained minimum is at $x = 10$, with the other side equal to 4.8. The optimal rectangle is no longer a square. Recalling a conclusion without accounting for the domain would produce an inadmissible answer.

Finally, if the sides represent real dimensions, thickness, margins, material costs and manufacturing constraints can change the objective. The mathematical problem is exact within its definition; applying it requires checking whether this definition addresses the need. An optimisation does not choose the correct criterion for itself.

D. A probability for a decision and its assumptions

A fictional system completes a task when two independent components work, each with probability 0.90. If both are essential, the probability that the system works is 0.81. If one is sufficient and the backup model is correctly represented by the event “at least one works”, the probability is $1 - 0.10^2 = 0.99$.

These results are not two competing estimates for the same arrangement. The operating events differ. The first configuration corresponds to an intersection; the second to the complement of both components failing. Identifying the components' roles therefore precedes calculating probabilities.

Independence is also an important assumption. If the components depend on the same power supply, which can fail, their states are no longer necessarily independent. A calculation retaining 0.99 without accounting for this common cause may overestimate reliability. The problem cannot be corrected by arbitrarily adjusting decimal places; the model must be revised.

For a procedure of ten independent trials with success probability 0.90, the probability of ten successes is 0.90^{10}, approximately 34.9%. A high probability on each trial therefore does not imply certainty of success throughout the series. The probability of at least one failure is approximately 65.1%. These results do not contradict the fact that the expected number of successes is nine.

We can explain this distinction by comparing three questions: success on one trial, the mean number of successes and success on every trial. They use the same parameters, but describe different events or indicators. Rigorous writing identifies which question it addresses and does not use one result to answer all three.

E. Geometry and optimisation of a distance

In an orthonormal coordinate system, the point M = (2,−1,3) is fixed and a line is described by H(t) = (t, t, 0). The squared distance is $D(t) = (t - 2)^2 + (t + 1)^2 + 9 = 2t^2 - 2t + 14$. Studying the derivative gives $D'(t) = 4t - 2$, hence a minimum at t = 1/2.

The nearest point is H = (1/2,1/2,0). The squared distance is 13.5 and the distance $\sqrt{13.5}$. The vector MH = (−3/2,3/2,−3) has zero scalar product with the direction vector (1,1,0). The algebraic optimisation method therefore recovers the orthogonal projection, and the geometric condition supplies an independent check.

If the problem concerns only the segment with t between one and two, the point t = 1/2 is inadmissible. The function D is increasing on this segment; its minimum there is attained at t = 1, at the point (1,1,0). The squared distance is fourteen. A minimum value on a line does not automatically answer a question about a segment.

Working with the squared distance is useful because the square root is increasing on non-negative numbers: minimising D is

equivalent to minimising √D. This simplifies the calculations without changing the objective. The justification should be given, rather than assuming that any outer function can be removed from an optimisation problem.

F. Prove an inequality and recognise equality cases

For every positive x, we want to establish $x \ln x - x + 1 \geq 0$. Define $g(x) = x \ln x - x + 1$. Its derivative is ln x. It is negative on (0,1), zero at one and positive on $(1,+\infty)$. The function therefore has a global minimum at one, and $g(1) = 0$. The inequality is proved; equality holds only for $x = 1$.

The second derivative is 1/x, which is positive. We can also interpret the result through the convexity of x ln x: its tangent at one has equation $y = x - 1$, and the curve remains above it. The two proofs express aspects of the same reasoning. The first uses increase and decrease; the second uses a geometric property of convex functions.

The condition $x > 0$ is essential for the logarithm to be defined. We could study a continuous extension of x ln x at zero, but that extension must be constructed and stated; it cannot be used silently within the initial proof. Domains are more than presentational precautions: they determine where objects exist.

Recognising the equality case is often essential in optimisation. A lower bound without a condition for attaining it does not always suffice to identify a solution. Here, the bound zero is attained, and uniqueness follows from strict variation on either side. A good worked solution distinguishes three facts: the bound, its attainment and uniqueness.

G. Prepare an extended solution in several stages

For a synthesis exercise, begin by setting out dependencies: which question uses which previous result? Calculating a limit may establish a sign, which establishes the existence of a root, which then allows bounds to be found. If an intermediate result is given or admitted, we can sometimes use it to continue, clearly identifying what remains unproved.

Retain exact values when subsequent steps use them. Rounding too early can shift an integer threshold or alter a very small difference. Present approximations at the end, with an appropriate symbol. When several algebraic solutions appear, return to the domain and the problem's conditions to decide which are admissible.

If using a calculation tool, record what it actually did: evaluation, plotting, approximate solution or simulation. A graph that seems to show a unique intersection is not a proof of uniqueness. A list of ten terms is not a proof of convergence. The tool can guide the choice of proof and check a result; it does not turn an observation into a general proposition.

H. Revise using explained errors

Build a short error notebook organised by category: forgotten domain, reversed implication, missing assumption, loop condition, mistranslation of an event, confusion between a point and a vector, premature rounding. For every error, write an incorrect example and its justified correction. The aim is to recognise the structure of the error in a different exercise, rather than memorising only an answer.

Alternate problems in which a familiar technique works with problems in which it would be inappropriate. For a binomial distribution, compare sampling with and without replacement. For

optimisation, compare an open domain with a constrained domain. For an integral, compare geometric area with signed accumulation. This alternation teaches us to choose, rather than apply automatically.

A useful personal assessment is to redo an exercise several days after reading it, then explain why another possible method is longer or insufficient. This checks the reconstruction of the reasoning. A feeling of fluency while reading may come from the solution already being present and should not be confused with independent mastery.

The guide offers examples and methods, rather than a promise of academic results. A correct answer must also be clear, relevant and coherent: state what you are proving, check that your assumptions match the question and conclude with the exact scope of the result. This requirement remains useful even when the calculations are already correct.