FSc/ICS Part-II (2nd Year) · PECTAA Board · English Medium

Mathematics 2nd Year

Complete Notes — All 11 Units

freebooks.pk

Table of Contents

Prepared by freebooks.pk — original student notes covering the PECTAA Board (Punjab Education, Curriculum, Training and Assessment Authority) Mathematics Class 12 (2nd Year, ICS) syllabus. Each unit includes key concepts, definitions, formulas, diagrams, solved examples, short and long questions, MCQs, revision and exam tips.

Unit 1: Graphical Representation of Functions

This unit builds a graphical toolkit for functions: sketching a quadratic from its factored form and predicting a quadratic equation from a given graph, classifying functions as algebraic or transcendental, and studying the key transcendental function families -- exponential, hyperbolic, and logarithmic -- through their domains, ranges, intercepts, monotonicity, and asymptotes.

It then connects these ideas visually: the graph of an inverse function is the reflection of the original graph about the line y=x (which is why y=ln x mirrors y=e^x), and any graph can be moved or resized using horizontal/vertical shifts and horizontal/vertical scaling. The unit closes with solving exponential and logarithmic equations and inequalities, and applying these functions to real-world growth and decay, sound intensity (decibels), and compound interest problems.

Learning Objectives

- Sketch the graph of a quadratic function given in factored form, and predict a quadratic function from its graph
- Classify functions as algebraic or transcendental, and as fundamental or non-fundamental transcendental functions
- State and apply the domain, range, intercepts, monotonicity, and asymptotic behaviour of exponential functions and sketch their graphs
- Define the hyperbolic and inverse hyperbolic functions and derive their logarithmic forms
- State and apply the laws of logarithms and sketch the graph of a logarithmic function as the inverse of an exponential function
- Construct the graph of an inverse function by reflecting the original graph about the line y=x
- Apply horizontal/vertical shifts and horizontal/vertical scaling to transform the graph of a function
- Solve exponential and logarithmic equations and inequalities, and apply these functions to growth/decay, sound intensity, and compound interest problems

Key Concepts

1.1 Sketching and Predicting Quadratic Graphs Using Factors

A quadratic y=a(x-h)(x-k) (h<=k) can be sketched directly from its factors: the graph cuts the x-axis at x=h and x=k, its vertex lies at x=(h+k)/2, its end behaviour (y -> +infinity or y -> -infinity as x -> +-infinity) is fixed by the sign of a, and its y-intercept is found by setting x=0 (giving y=ahk). When h=k, the factor (x-h)^2 makes the parabola touch rather than cross the x-axis, with vertex (h,0).

Conversely, if the x-intercepts x=h and x=k of a quadratic graph are known, the function can be written as f(x)=a(x-h)(x-k), and substituting the coordinates of any third point on the curve determines the remaining constant a -- letting a full equation be recovered directly from a picture of its graph.

1.2 Algebraic and Transcendental Functions

A function y=f(x) is algebraic if it satisfies a polynomial equation in y with polynomial coefficients in x -- equivalently, if it is built from x using finitely many algebraic operations (addition, subtraction, multiplication, division, and nth roots). A function that is not algebraic is called transcendental.

Transcendental functions split further into fundamental transcendental functions -- the standard building blocks such as y=a^x, y=ln x, y=sin x, y=sin^-1 x -- and non-fundamental transcendental functions, formed by combining these with algebraic functions or with each other through composition (e.g. y=sin x + ln x, y=e^(x^2)).

1.3 Exponential, Hyperbolic, and Inverse Hyperbolic Functions

The exponential function f(x)=a^x (a>0, a<>1) has domain]-infinity,infinity[, range]0,infinity[, is continuous and always positive, cuts the y-axis at (0,1), has no x-intercept, and has horizontal asymptote y=0. It is strictly increasing when a>1 and strictly decreasing when 0<a<1, and it is injective in both cases, so every horizontal line meets its graph at most once.

Hyperbolic functions $\sinh x=(e^x-e^{-x})/2$, $\cosh x=(e^x+e^{-x})/2$, and $\tanh x=\sinh x/\cosh x$ (plus their reciprocals coth, sech, csch) are built from e^x and behave like trigonometric functions but arise from the hyperbola $x^2-y^2=1$ rather than the circle. Their inverses -- such as $\sinh^{-1} x=\ln(x+\sqrt{x^2+1})$ and $\cosh^{-1} x=\ln(x+\sqrt{x^2-1})$ on $[1,\infty)$ -- are derived by solving the defining exponential equation for x algebraically using the quadratic formula.

1.4 Logarithmic Functions and Their Graphs

The logarithmic function $y=\log_a x$ is defined as the inverse of the exponential function $y=a^x$, with domain $]0,\infty[$ and range $]-\infty,\infty[$. It is injective, has vertical asymptote $x=0$, cuts the x-axis at $(1,0)$, has no y-intercept, and is strictly increasing when $a>1$ and strictly decreasing when $0<a<1$. The product, quotient, and power laws of logarithms, together with the change-of-base formula $\log_a x=\ln x/\ln a$, are the standard tools for simplifying logarithmic expressions and equations.

The modulus function $f(x)=|x|$ produces a V-shaped graph made of the two lines $y=x$ (for $x>=0$) and $y=-x$ (for $x<0$), meeting at the origin.

1.5 Graphs of Inverse Functions and Transformations

A function has an inverse exactly when it is bijective (one-to-one and onto); its inverse graph is the reflection of the original graph about the line $y=x$, since interchanging the roles of input and output corresponds exactly to reflecting across the line where input equals output. This is why the graph of $y=\ln x$ is the mirror image of $y=e^x$ about $y=x$.

A graph can be transformed by shifting -- $y=f(x-h)+k$ moves the graph h units horizontally (right if $h>0$, left if $h<0$) and k units vertically (up if $k>0$, down if $k<0$) -- or by scaling -- $y=a f(bx)$ stretches/compresses the graph vertically by a and horizontally by $1/b$ (for $a,b>0$), where $a>1$ stretches vertically and $0<a<1$ compresses vertically, while $b>1$ compresses horizontally and $0<b<1$ stretches horizontally.

1.6 Exponential/Logarithmic Equations, Inequalities, and Real-World Applications

Exponential and logarithmic equations are solved by matching bases (using injectivity of a^x or $\log_a x$) or by taking logarithms of both sides; the domain must always be checked first, since logarithmic expressions require positive arguments and positive, non-unit bases. The same injectivity/monotonicity facts convert exponential and logarithmic inequalities into ordinary algebraic inequalities, again subject to the domain restrictions of the original expressions.

These functions model real growth and decay via $N=N_0 e^{kt}$ ($k>0$ for growth, $k<0$ for decay), sound intensity in decibels via $L=10 \log(I/I_0)$ with threshold $I_0=10^{-12}$ W/m^2, and compound interest via the discrete formula $A=P(1+r/n)^{nt}$ or the continuous formula $A=Pe^{rt}$ -- each a direct application of solving an exponential or logarithmic equation for the unknown quantity.

Important Definitions

What is an algebraic function?

A function $y=f(x)$ obtained from x using a finite number of algebraic operations (addition, subtraction, multiplication, division, and nth roots), i.e. one satisfying a polynomial equation in y with polynomial coefficients in x.

What is a transcendental function?

A function that is not algebraic, such as $y=e^x$, $y=\ln x$, $y=\sin x$, or $y=\sin^{-1} x$.

What is the natural exponential function?

$f(x)=e^x$, where e is the number defined by the infinite series $\sum(1/n!)$ for $n=0$ to infinity, with e is approximately 2.71828.

What is a horizontal asymptote?

A horizontal line $y=c$ that a function's graph approaches (but need not touch) as $x \to +\infty$ or $x \to -\infty$.

What is a vertical asymptote?

A vertical line $x=c$ such that $f(x) \to \pm\infty$ as $x \to c$ from the left or right.

What are the hyperbolic functions sinh x and cosh x?

sinh x = (e^x - e^-x)/2 and cosh x = (e^x + e^-x)/2; they satisfy cosh^2 x - sinh^2 x = 1, the hyperbola analogue of sin^2+cos^2=1.

What is a logarithmic function?

y=log_a x (a>0, a<>1, x>0), defined as the inverse of the exponential function y=a^x, i.e. a^y=x.

How is the graph of an inverse function obtained from the original graph?

By reflecting the original graph about the line y=x.

What does the general shift formula y=f(x-h)+k represent?

A horizontal shift of h units (right if h>0, left if h<0) and a vertical shift of k units (up if k>0, down if k<0).

What is the continuous compound interest formula?

A = P e^(rt), where P is the principal, r is the annual interest rate (decimal), and t is time.

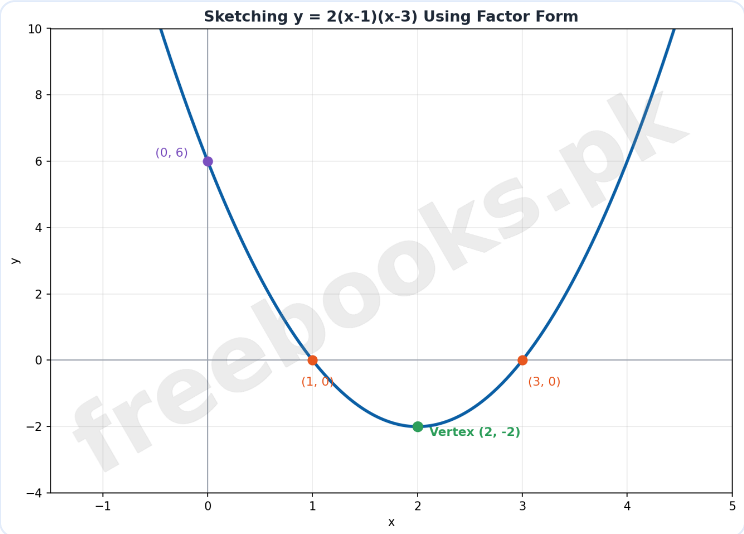
Key Facts and Relations

Topic	Key Fact / Relation
Quadratic graph via factors	y = a(x-h)(x-k); x-intercepts x=h,k; vertex at x=(h+k)/2; y-intercept y=ahk
Exponential function	y = a^x (a>0, a<>1); domain]-inf,inf[, range]0,inf[, y-intercept (0,1), asymptote y=0
Hyperbolic functions	sinh x=(e^x-e^-x)/2, cosh x=(e^x+e^-x)/2, tanh x=sinh x/cosh x; cosh^2 x - sinh^2 x = 1
Inverse hyperbolic sine/cosine	sinh^-1 x = ln(x+sqrt(x^2+1)); cosh^-1 x = ln(x+sqrt(x^2-1)), x>=1
Laws of logarithms	log_a(xy)=log_a x+log_a y; log_a(x/y)=log_a x-log_a y; log_a(x^t)=t log_a x; log_a x = ln x/ln a
Logarithmic function	y = log_a x (a>0, a<>1, x>0); domain]0,inf[, range]-inf,inf[, x-intercept (1,0), asymptote x=0
General shift of a graph	y = f(x-h) + k (h: horizontal shift, k: vertical shift)

Topic	Key Fact / Relation
General scaling of a graph	y = a f(bx), a>0, b>0 (a: vertical scale factor, 1/b: horizontal scale factor)
Growth/decay model	N = N0 e^(kt) (k>0 growth, k<0 decay)
Sound level (decibels) / Compound interest	L = 10 log(I/I0), I0=10^-12 W/m^2; A = P(1+r/n)^(nt) (discrete); A = Pe^(rt) (continuous)

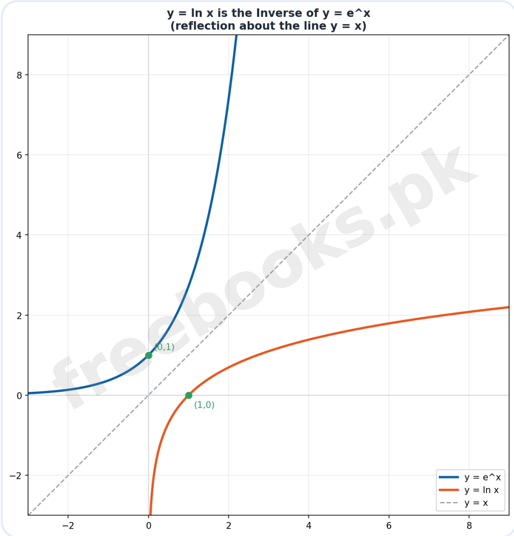
Diagrams

Sketching y = 2(x-1)(x-3) Using Factor Form: A parabola showing the x-intercepts at x=1 and x=3, the vertex at (2,-2), and the y-intercept at (0,6), illustrating the factor-form sketching method



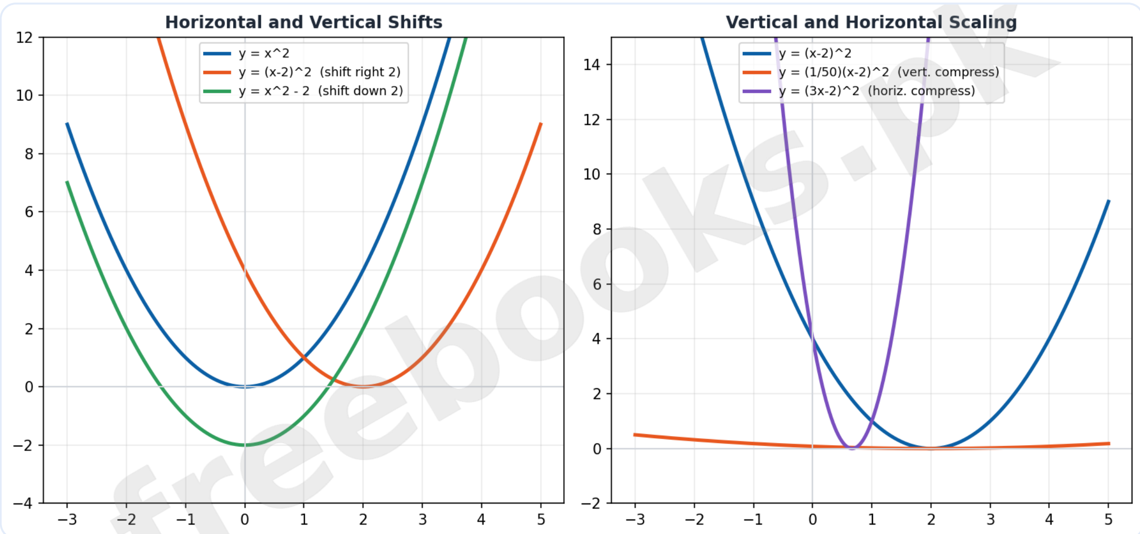
Sketching y = 2(x-1)(x-3) Using Factor Form

y = ln x is the Inverse of y = e^x: The graphs of y=e^x and y=ln x plotted together with the line y=x, showing that the two curves are reflections of each other about y=x



y = ln x is the Inverse of y = e^x

Horizontal/Vertical Shifts and Scaling of Graphs: Two panels showing y=x^2 transformed by horizontal and vertical shifts (left) and by vertical and horizontal scaling (right)



Horizontal/Vertical Shifts and Scaling of Graphs

Solved Examples

Example 1: Sketching a Quadratic Graph Using Factors

Problem: Sketch the graph of y = 2(x-1)(x-3).

1. Find the x-intercepts by setting each factor to zero: x-1=0 and x-3=0, giving x=1 and x=3, so the graph meets the x-axis at (1,0) and (3,0).

2. Find the vertex by evaluating y at $x=(1+3)/2=2$: $y=2(2-1)(2-3)=-2$, so the vertex is $(2,-2)$.
3. Determine end behaviour: since $a=2>0$, $y \to +\infty$ as $x \to \pm\infty$.
4. Find the y-intercept by setting $x=0$: $y=2(-1)(-3)=6$, so the y-intercept is $(0,6)$.
5. Using the intercepts $(1,0)$, $(3,0)$, $(0,6)$ and vertex $(2,-2)$, sketch the upward-opening parabola.

Example 2: Predicting a Quadratic Function from Its Graph

Problem: Find an equation of the graph whose x-intercepts are $x=1$ and $x=2$ and which passes through the point $(3,4)$.

1. Since the x-intercepts are $x=1$ and $x=2$, write the function in factored form: $y=a(x-1)(x-2)$.
2. Substitute the known point $(3,4)$: $4=a(3-1)(3-2)=2a$.
3. Solve for a: $a=4/2=2$.
4. Substitute $a=2$ back into the factored form: $y=2(x-1)(x-2)=2(x^2-3x+2)$.
5. Final answer: $y=2x^2-6x+4$.

Example 3: Proving a Hyperbolic Identity

Problem: Prove that $\cosh^2 x - \sinh^2 x = 1$.

1. Write both functions in terms of e^x: $\cosh x=(e^x+e^{-x})/2$ and $\sinh x=(e^x-e^{-x})/2$.
2. Square each: $\cosh^2 x = (e^{2x}+e^{-2x}+2)/4$ and $\sinh^2 x = (e^{2x}+e^{-2x}-2)/4$.
3. Subtract: $\cosh^2 x - \sinh^2 x = [(e^{2x}+e^{-2x}+2)-(e^{2x}+e^{-2x}-2)]/4$.
4. Simplify the numerator: $(e^{2x}+e^{-2x}+2-e^{2x}-e^{-2x}+2)/4 = 4/4 = 1$.
5. Final answer: $\cosh^2 x - \sinh^2 x = 1$.

Example 4: Deriving the Logarithmic Form of sinh^-1 x

Problem: Prove that sinh^-1 x = ln(x + sqrt(x^2+1)).

1. Let y=sinh^-1 x, so x=sinh y = (e^y-e^-y)/2, giving 2x=e^y-1/e^y.
2. Multiply through by e^y: 2x e^y = e^(2y) - 1, i.e. (e^y)^2 - 2x e^y - 1 = 0, a quadratic in e^y.
3. Apply the quadratic formula: e^y = [2x +- sqrt(4x^2+4)]/2 = x +- sqrt(x^2+1).
4. Since x - sqrt(x^2+1) < 0 for all real x but e^y must be positive, reject the negative sign: e^y = x + sqrt(x^2+1).
5. Take the natural log of both sides: y = ln(x+sqrt(x^2+1)). Final answer: sinh^-1 x = ln(x+sqrt(x^2+1)).

Example 5: Solving an Exponential Equation by Matching Bases

Problem: Solve the equation 4^(x-1) = 8.

1. Note the domain: 4^(x-1) is defined for all real x, so the domain is]-infinity,infinity[.
2. Write both sides with base 2: 4=2^2 and 8=2^3, so (2^2)^(x-1)=2^3, i.e. 2^(2x-2)=2^3.
3. Since the exponential function is one-to-one (injective), the exponents must be equal: 2x-2=3.
4. Solve for x: 2x=5, so x=5/2.
5. Final answer: solution set = {5/2}.

Example 6: Solving a Logarithmic Inequality

Problem: Solve the inequality log(2x-1) >= log(5-3x).

1. Find the domain: 2x-1>0 gives x>1/2, and 5-3x>0 gives x<5/3, so the domain is 1/2 < x < 5/3.

2. Since the logarithmic function with base 10 is strictly increasing, the inequality is equivalent to comparing the arguments directly: 2x-1 >= 5-3x.
3. Solve the resulting linear inequality: 2x+3x >= 5+1, so 5x >= 6, giving x >= 6/5.
4. Combine with the domain restriction 1/2 < x < 5/3: this gives 6/5 <= x < 5/3.
5. Final answer: solution set = [6/5, 5/3).

Example 7: Applying Multiple Graph Transformations in Order

Problem: Let f(x) = x/(x+1). Find the function after: (i) vertical stretch by 2, (ii) horizontal stretch by 3, (iii) shift 2 units right, (iv) shift 1 unit down.

1. (i) Vertical stretch by 2: y=2f(x)=2x/(x+1). Call this g(x)=2x/(x+1).
2. (ii) Horizontal stretch by 3 means y=g(x/3): substituting gives y=2(x/3)/((x/3)+1)=2x/(x+3). Call this s(x)=2x/(x+3).
3. (iii) Shift 2 units right means y=s(x-2)=2(x-2)/((x-2)+3)=(2x-4)/(x+1). Call this t(x)=(2x-4)/(x+1).
4. (iv) Shift 1 unit down means y=t(x)-1=(2x-4)/(x+1) - 1 = (2x-4-(x+1))/(x+1) = (x-5)/(x+1).
5. Final answer: y = (x-5)/(x+1).

Example 8: Real-World Application: Continuous Compound Interest

Problem: A sum of Rs. 1000 is invested at a rate of 5% per year compounded continuously for 3 years. Find the final amount.

1. Identify the given values: P=1000, r=5%=0.05, t=3.
2. Apply the continuous compounding formula: A = P e^(rt).
3. Substitute the values: A = 1000 e^(0.05x3) = 1000 e^0.15.
4. Evaluate: e^0.15 is approximately 1.1618, so A is approximately 1000 x 1.1618 = 1161.8.

> 5. Final answer: the final amount after 3 years is approximately Rs. 1161.8.

Short Questions & Answers

What determines whether an exponential function $y=a^x$ is increasing or decreasing?

The base a: strictly increasing if $a>1$, strictly decreasing if $0<a<1$.

What is the horizontal asymptote of every exponential function $y=a^x$?

$y=0$, the x-axis.

What is the vertical asymptote of every logarithmic function $y=\log_a x$?

$x=0$, the y-axis.

Through which point does every exponential function $y=a^x$ pass?

$(0,1)$, since $a^0=1$.

Through which point does every logarithmic function $y=\log_a x$ pass?

$(1,0)$, since $\log_a 1=0$.

How is the graph of an inverse function related to the original graph?

It is the reflection of the original graph about the line $y=x$.

What is the shape of the graph of the modulus function $f(x)=|x|$?

A V-shape, made of the lines $y=x$ for $x>=0$ and $y=-x$ for $x<0$, meeting at the origin.

Long Questions & Answers

Explain how to sketch and predict the graph of a quadratic function using its factored form, with the roles of intercepts and vertex.

How do the x-intercepts of $y=a(x-h)(x-k)$ relate to its factors?

Setting each factor to zero gives $x=h$ and $x=k$ directly, so the graph crosses the x-axis exactly at these two points (or touches it once at $x=h$ if $h=k$).

How is the vertex of the parabola located?

The vertex lies midway between the two x-intercepts, at x=(h+k)/2; substituting this x-value back into the function gives the vertex's y-coordinate.

How does the sign of a determine the graph's end behaviour?

If a>0, y -> +infinity as x -> +-infinity (the parabola opens upward); if a<0, y -> -infinity as x -> +-infinity (it opens downward).

How can a quadratic function be predicted from its graph?

If the x-intercepts h and k are known, write f(x)=a(x-h)(x-k), then substitute the coordinates of any third point on the curve to solve for the remaining constant a.

Compare exponential and logarithmic functions: their domains, ranges, monotonic behaviour, and how their graphs relate to each other.

What are the domain and range of y=a^x versus y=log_a x?

The exponential function y=a^x has domain]-infinity,infinity[and range]0,infinity[; the logarithmic function y=log_a x has domain]0,infinity[and range]-infinity,infinity[-- domain and range are exactly swapped, since the two are inverses.

How does the base a affect monotonic behaviour in each case?

For both y=a^x and y=log_a x, the function is strictly increasing when a>1 and strictly decreasing when 0<a<1 -- the same base condition governs monotonicity in both families.

What asymptotes do these two graphs have?

y=a^x has horizontal asymptote y=0 (the x-axis); y=log_a x has vertical asymptote x=0 (the y-axis) -- again mirror images of each other.

Why is the graph of y=log_a x the reflection of y=a^x about y=x?

Because y=log_a x is defined as the inverse of y=a^x, and the graph of any inverse function is obtained by reflecting the original graph about the line y=x, where input and output roles are interchanged.

Multiple Choice Questions (MCQs)

The graph of y=a(x-h)(x-k) crosses the x-axis at:

A. x=a only

B. x=h and x=k

C. x=(h+k)/2 only

D. It never crosses the x-axis

Answer: B. x=h and x=k

Explanation: Setting each factor to zero gives the x-intercepts x=h and x=k directly.

A function is called transcendental if it is:

A. Always positive

B. Not algebraic

C. A polynomial

D. Always increasing

Answer: B. Not algebraic

Explanation: A transcendental function is, by definition, any function that is not algebraic.

The exponential function y=a^x (a>1) is:

A. Strictly decreasing

B. Strictly increasing

C. Constant

D. Undefined for x<0

Answer: B. Strictly increasing

Explanation: When a>1, y=a^x is strictly increasing on its whole domain.

The horizontal asymptote of every exponential function y=a^x is:

A. x=0

B. y=1

C. y=0

D. y=a

Answer: C. y=0

Explanation: As x -> -infinity (for a>1) or x -> +infinity (for 0<a<1), a^x -> 0, so y=0 is the horizontal asymptote.

cosh^2 x - sinh^2 x equals:

A. 0

B. 1

C. 2x

D. e^x

Answer: B. 1

Explanation: This is the hyperbolic identity analogous to sin^2+cos^2=1, proved directly from the exponential definitions of sinh and cosh.

The logarithmic function y=log_a x has vertical asymptote:

A. x=0

B. y=0

C. x=1

D. y=1

Answer: A. x=0

Explanation: As x -> 0+, log_a x -> +-infinity depending on the base, making x=0 the vertical asymptote.

The graph of an inverse function is obtained from the original graph by:

A. Reflecting about the x-axis

B. Reflecting about the line y=x

C. Rotating 90 degrees

D. Shifting up by 1 unit

Answer: B. Reflecting about the line y=x

Explanation: Interchanging input and output corresponds geometrically to reflecting the graph about the line y=x.

In y=f(x-h)+k, a positive value of h shifts the graph:

A. Left

B. Right

C. Up

D. Down

Answer: B. Right

Explanation: A positive h shifts the graph h units to the right; the sign is opposite to what might be expected because it replaces x with x-h.

In y=af(bx) with a>1 and b>1, the graph is:

A. Compressed vertically and stretched horizontally

B. Stretched vertically and compressed horizontally

C. Unchanged

D. Reflected about the origin

Answer: B. Stretched vertically and compressed horizontally

Explanation: a>1 stretches the graph vertically by factor a, while b>1 compresses it horizontally by factor b.

The continuous compound interest formula is:

A. A=P(1+r)^t

B. A=Prt

C. A=Pe^(rt)

D. A=P+rt

Answer: C. A=Pe^(rt)

Explanation: Continuous compounding is modeled by A=Pe^(rt), the limiting case of discrete compounding as the number of periods grows without bound.

Quick Revision Summary

- y=a(x-h)(x-k): x-intercepts x=h,k; vertex at x=(h+k)/2; y-intercept y=ahk; end behaviour set by sign of a
- Algebraic functions use only +,-,x,/,and nth roots of x; transcendental functions are everything else
- Fundamental transcendental functions: a^x, ln x, sin x, sin^-1 x; combinations of these are non-fundamental
- Exponential y=a^x: domain]-inf,inf[, range]0,inf[, y-intercept (0,1), asymptote y=0, increasing if a>1, decreasing if 0<a<1
- sinh x=(e^x-e^-x)/2, cosh x=(e^x+e^-x)/2; cosh^2 x - sinh^2 x = 1; sinh^-1 x = ln(x+sqrt(x^2+1))
- Logarithmic y=log_a x: domain]0,inf[, range]-inf,inf[, x-intercept (1,0), asymptote x=0, is the inverse of y=a^x
- Laws of logarithms: log_a(xy)=log_a x+log_a y; log_a(x/y)=log_a x-log_a y; log_a(x^t)=t log_a x

- The graph of an inverse function is the reflection of the original graph about the line y=x
- Shift: y=f(x-h)+k moves h units horizontally, k units vertically; Scale: y=af(bx) scales vertically by a, horizontally by 1/b
- Exponential/logarithmic equations: match bases using injectivity, or take logs of both sides; always check the domain first
- Growth/decay: N=N0 e^(kt); Sound level: L=10 log(I/I0); Compound interest: A=P(1+r/n)^(nt) (discrete) or A=Pe^(rt) (continuous)

Exam Tips

- When sketching a quadratic from factors, always find the x-intercepts, vertex, and y-intercept in that order -- three points are enough to sketch an accurate parabola
- To classify a function as algebraic or transcendental, check whether it can be built using only +,-,x,/,and roots of x -- any exponential, logarithmic, or trigonometric piece makes it transcendental
- Memorize the four exponential/logarithmic facts as a pair: exponential passes through (0,1) with asymptote y=0; logarithmic passes through (1,0) with asymptote x=0
- For hyperbolic identity proofs, always rewrite sinh x and cosh x in terms of e^x and e^-x first, then simplify algebraically
- Before solving any exponential or logarithmic equation/inequality, find the domain first -- many 'solutions' get rejected at the end for violating the original domain
- When applying multiple graph transformations in sequence, apply them strictly in the given order, one function substitution at a time -- reordering changes the final answer

Unit 2: Further Differentiation

This unit extends differentiation beyond polynomials and rational functions to trigonometric, inverse trigonometric, exponential, and logarithmic functions, and introduces the chain rule for composite, parametric, and implicit relations. Together these rules make it possible to differentiate almost any function encountered in calculus.

The unit then applies differentiation to real problems: finding equations of tangent and normal lines, computing higher-order derivatives, determining where a function is increasing or decreasing, locating local and global extrema (using critical points and the second derivative test), approximating function values with linearization and differentials (including relative/percentage error), and solving real-world optimization problems by finding the extrema of a modeling function.

Learning Objectives

- Differentiate trigonometric and inverse trigonometric functions, and apply the chain rule to composite, parametric, and implicit relations
- Differentiate exponential and logarithmic functions, including composite forms
- Find equations of tangent and normal lines to a curve at a given point
- Find higher-order derivatives of algebraic, implicit, parametric, and transcendental functions
- Determine intervals where a function is increasing or decreasing using critical points and the sign of the first derivative
- Find local and global extrema of a function using the second derivative test and by checking endpoints
- Use linearization and differentials to approximate function values and compute relative/percentage error
- Apply extrema to solve real-world optimization problems

Key Concepts

2.1 Derivatives of Trigonometric and Inverse Trigonometric Functions, and the Chain Rule

From the limit definition, $d/dx(\sin x)=\cos x$ and $d/dx(\cos x)=-\sin x$; the remaining trig derivatives (tan, cot, sec, csc) follow using the quotient rule, e.g. $d/dx(\tan x)=\sec^2 x$. If $f(u)$ is differentiable at $u=g(x)$ and g is differentiable at x, the chain rule gives $(f \circ g)'(x)=f'(g(x)) \cdot g'(x)$, or in Leibniz notation $dy/dx=(dy/du)(du/dx)$ -- this same idea extends to parametric equations ($dy/dx=(dy/dt)/(dx/dt)$) and implicit differentiation (differentiate both sides of $F(x,y)=0$ with respect to x, applying the chain rule to every y-term, then solve for dy/dx).

The derivatives of the six inverse trigonometric functions are proved by writing $y=f^{-1}(x)$ as $x=f(y)$, differentiating implicitly, and simplifying using the appropriate Pythagorean identity: $d/dx(\sin^{-1} x)=1/\sqrt{1-x^2}$, $d/dx(\cos^{-1} x)=-1/\sqrt{1-x^2}$, $d/dx(\tan^{-1} x)=1/(1+x^2)$, $d/dx(\cot^{-1} x)=-1/(1+x^2)$, $d/dx(\sec^{-1} x)=1/(|x|\sqrt{x^2-1})$, and $d/dx(\csc^{-1} x)=-1/(|x|\sqrt{x^2-1})$.

2.2 Derivatives of Exponential and Logarithmic Functions

Using the limit definition and the standard limit $\lim_{h\to 0}(a^h-1)/h=\ln a$, the derivative of the exponential function is $d/dx(a^x)=a^x \ln a$, which reduces to $d/dx(e^x)=e^x$ since $\ln e=1$. Composite exponential functions are differentiated with the chain rule: $d/dx(a^u)=a^u \ln a \cdot du/dx$.

Similarly, using the limit definition and $\lim_{h\to 0}(1+h/a)^{a/h}=e$, the derivative of the logarithmic function is $d/dx(\log_a x)=1/(x \ln a)$, which reduces to $d/dx(\ln x)=1/x$. Composite logarithmic functions follow the same chain rule pattern: $d/dx(\log_a u)=(1/(u \ln a))\, du/dx$.

2.3 Tangent/Normal Lines and Higher-Order Derivatives

The tangent line to $y=f(x)$ at $(x_0,f(x_0))$ is $y-y_0=m(x-x_0)$ where $m=f'(x_0)$; the normal line is perpendicular to the tangent, given by $y-y_0=-(1/m)(x-x_0)$ when $m\neq 0$, or $x=x_0$ when $m=0$ (and $x=x_0$ is a vertical tangent when m is undefined/infinite).

Since the derivative f' is itself a function, it can be differentiated again to give the second derivative f'', and repeatedly to give third, fourth, and higher-order derivatives, written y', y'', y''', $y^{(n)}$ or dy/dx, d^2y/dx^2, ..., $d^n y/dx^n$ -- these are found by simply differentiating the previous derivative again, and can be computed for algebraic, implicit, parametric, and transcendental functions alike.

2.4 Increasing/Decreasing Functions and Extrema

A differentiable function f is increasing on an interval I if $f'(x) \geq 0$ throughout I, and decreasing if $f'(x) \leq 0$ throughout I. A critical point is a point $x=c$ in the domain where $f'(c)=0$ (a stationary point) or $f'(c)$ does not exist; critical points divide the domain into subintervals on which the sign of f' stays constant, and testing one point per subinterval determines whether f is increasing or decreasing there.

A function has a global (absolute) maximum/minimum at $x=c$ if $f(c)$ is the largest/smallest value on the whole domain, and a local (relative) maximum/minimum if $f(c)$ is the largest/smallest value only in some neighborhood of c. Global extrema on a closed interval $[a,b]$ are found by comparing f at all critical points and at both endpoints; local extrema are classified using the second derivative test: if $f'(c)=0$, then f has a local max at c if $f''(c)<0$, a local min if $f''(c)>0$, and the test is inconclusive if $f''(c)=0$.

2.5 Linear Approximation and Differentials

Since a differentiable curve behaves locally like its tangent line, the linearization of f at $x=a$ is $L(x)=f(a)+f'(a)(x-a)$, and $f(x)$ is approximated by $L(x)$ for x near a -- useful for quickly estimating values like square roots or trigonometric values near a known point.

For a differentiable function with independent differential dx, the differential dy is defined by $dy=f'(x)dx$; for small $\Delta x=dx$, the actual change $\Delta y=f(x+\Delta x)-f(x)$ is approximately equal to dy, so $f(x+\Delta x)$ is approximately $f(x)+dy$. The error in this approximation is $\Delta y-dy$, the relative error is $(\Delta y-dy)/\Delta y$, and the percentage error is $\left|\left(\frac{\Delta y-dy}{\Delta y}\right)\right| \times 100$ -- this same framework estimates the error propagated into a computed quantity (like volume or area) from a small measurement error in a linear dimension.

2.6 Applying Extrema to Real-World Optimization Problems

Optimization problems are solved by: understanding what is to be optimized, assigning variables to the unknowns, writing a function for the quantity to be optimized (using any given constraint to eliminate a variable), finding critical points by differentiating and setting the derivative to zero, confirming maximum or minimum using the second derivative test, and finally interpreting the mathematical answer back in the real-world context.

This method applies broadly -- maximizing the area of a fenced garden or an inscribed rectangle, minimizing the material (surface area) of a container of fixed volume, minimizing distance (e.g. signal strength problems), and minimizing algorithm runtime or other modeled quantities -- always by reducing the problem to finding the extremum of a single-variable function.

Important Definitions

What is the chain rule for a composite function y=f(g(x))?

dy/dx = f'(g(x)) . g'(x), or in Leibniz notation dy/dx=(dy/du)(du/dx) where y=f(u) and u=g(x).

What is implicit differentiation?

Differentiating both sides of an equation F(x,y)=0 with respect to x, applying the chain rule to every term containing y, then solving the resulting equation for dy/dx.

What is the derivative of a^x?

d/dx(a^x) = a^x ln a; in particular, d/dx(e^x)=e^x since ln e=1.

What is the derivative of log_a x?

d/dx(log_a x) = 1/(x ln a); in particular, d/dx(ln x)=1/x.

What is a critical point of a function f?

A point x=c in the domain of f where f'(c)=0 (a stationary point) or f'(c) does not exist.

What does the second derivative test state?

If f'(c)=0, then f has a local maximum at c if f''(c)<0, a local minimum if f''(c)>0, and the test is inconclusive if f''(c)=0.

What is the linearization of f at x=a?

L(x) = f(a) + f'(a)(x-a), the equation of the tangent line, used to approximate f(x) for x near a.

What is the differential dy?

dy = f'(x) dx, where dx is a small change in the independent variable x; for small dx, dy approximates the actual change Deltay = f(x+Deltax)-f(x).

What is percentage error in a differential approximation?

Percentage Error = |(Deltay - dy)/Deltay| x 100, the size of the approximation error relative to the actual change, expressed as a percentage.

What is a global (absolute) maximum of f on an interval I?

A value f(c), c in I, such that f(c) >= f(x) for every x in I.

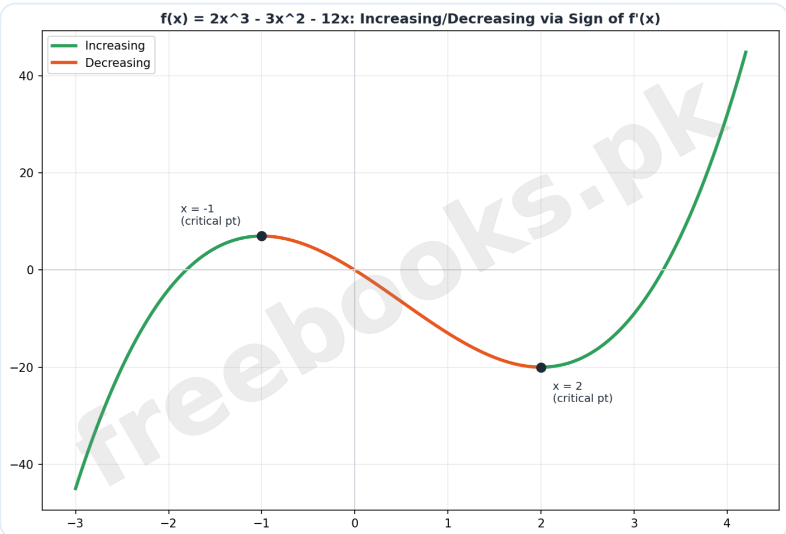
Key Facts and Relations

Topic	Key Fact / Relation
Derivatives of trig functions	d/dx(sin x)=cos x, d/dx(cos x)=-sin x, d/dx(tan x)=sec^2 x, d/dx(cot x)=-csc^2 x, d/dx(sec x)=sec x tan x, d/dx(csc x)=-csc x cot x
Derivatives of inverse trig functions	d/dx(sin^-1 x)=1/sqrt(1-x^2), d/dx(tan^-1 x)=1/(1+x^2), d/dx(sec^-1 x)=1/(|x| sqrt(x^2-1))
Chain rule	dy/dx = dy/du . du/dx (composite); dy/dx = (dy/dt)/(dx/dt) (parametric)
Exponential and logarithmic derivatives	d/dx(a^x) = a^x ln a; d/dx(log_a x) = 1/(x ln a); d/dx(e^x)=e^x; d/dx(ln x)=1/x
Tangent and normal lines	Tangent: y-y0 = m(x-x0), m=f'(x0); Normal: y-y0 = -(1/m)(x-x0), m<>0
Second derivative test	f'(c)=0: local max if f''(c)<0; local min if f''(c)>0; inconclusive if f''(c)=0
Linearization	L(x) = f(a) + f'(a)(x-a); f(x) is approximately L(x) for x near a
Differential and error	

Topic	Key Fact / Relation
	dy = f'(x)dx; Deltay is approximately dy; Percentage Error = |(Deltay-dy)/Deltay| x 100
Global extrema on [a,b]	Compare f at all critical points in [a,b] and at both endpoints x=a, x=b
Optimization method	Define variables -> write function to optimize -> find critical points -> confirm with 2nd derivative test -> interpret result

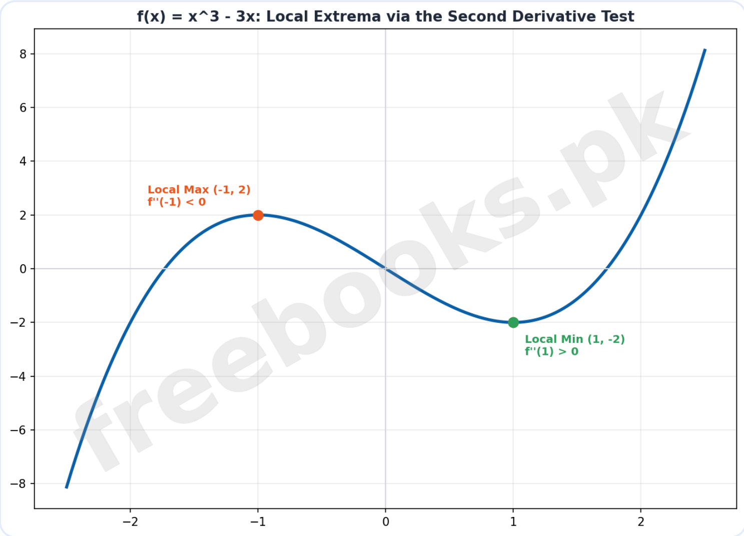
Diagrams

Increasing/Decreasing via Sign of f'(x): The graph of f(x)=2x^3-3x^2-12x, colored green where increasing and orange where decreasing, with critical points marked at x=-1 and x=2



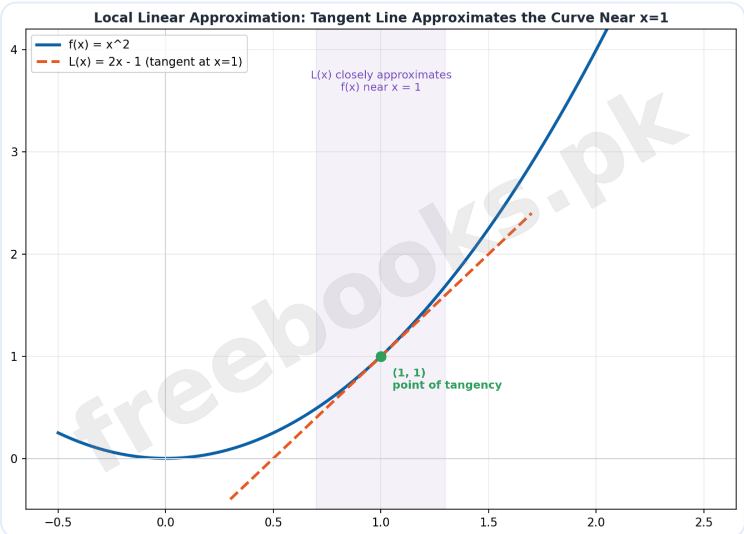
Increasing/Decreasing via Sign of f'(x)

Local Extrema via the Second Derivative Test: The graph of f(x)=x^3-3x showing a local maximum at (-1,2) where f''<0 and a local minimum at (1,-2) where f''>0



Local Extrema via the Second Derivative Test

Local Linear Approximation: Tangent Line Near a Point: The graph of f(x)=x^2 together with its tangent line L(x)=2x-1 at x=1, showing that the tangent closely approximates the curve near the point of tangency



Local Linear Approximation: Tangent Line Near a Point

Solved Examples

Example 1: Differentiating a Composite Function Using the Chain Rule

Problem: Differentiate y = sin(x^2 - x) with respect to x.

1. Let u = x^2 - x, so that y = sin u.

2. Differentiate y with respect to u: $dy/du = \cos u$.
3. Differentiate u with respect to x: $du/dx = 2x - 1$.
4. Apply the chain rule: $dy/dx = (dy/du)(du/dx) = \cos u \cdot (2x-1)$.
5. Substitute $u = x^2-x$ back: $dy/dx = \cos(x^2-x)(2x-1)$.

Example 2: Implicit Differentiation

Problem: Find dy/dx if $x + y^2 + \sin(xy^2) = 0$.

1. Differentiate every term with respect to x, treating y as a function of x: $d/dx(x) + d/dx(y^2) + d/dx(\sin(xy^2)) = 0$.
2. This gives: $1 + 2y(dy/dx) + \cos(xy^2) \cdot d/dx(xy^2) = 0$.
3. Expand $d/dx(xy^2)$ using the product rule: $d/dx(xy^2) = y^2 + x(2y\, dy/dx)$, so the equation becomes $1 + 2y(dy/dx) + \cos(xy^2)\{y^2 + 2xy(dy/dx)\} = 0$.
4. Collect all dy/dx terms on one side: $2y(dy/dx)\{1 + x \cos(xy^2)\} = -1 - y^2 \cos(xy^2)$.
5. Solve for dy/dx: $dy/dx = -[1 + y^2 \cos(xy^2)] / [2y\{1 + x \cos(xy^2)\}]$.

Example 3: Differentiating a Composite Exponential Function

Problem: Find dy/dx if $y = e^{(\sin^2 x)}$.

1. Apply the chain rule for exponentials: $dy/dx = e^{(\sin^2 x)} \cdot d/dx(\sin^2 x)$, since $d/dx(e^u)=e^u \cdot du/dx$.
2. Differentiate $\sin^2 x$ using the chain rule again: $d/dx(\sin^2 x) = 2 \sin x \cdot d/dx(\sin x) = 2 \sin x \cos x$.
3. Substitute back: $dy/dx = e^{(\sin^2 x)} \cdot 2 \sin x \cos x$.
4. Final answer: $dy/dx = 2 \sin x \cos x \cdot e^{(\sin^2 x)}$.

Example 4: Finding Higher-Order Derivatives

Problem: Find all higher-order derivatives of $y = e^{(ax)}$.

1. Differentiate once using the chain rule: $dy/dx = e^{(ax)} \cdot a = a e^{(ax)}$.
2. Differentiate again: $d^2y/dx^2 = a(a e^{(ax)}) = a^2 e^{(ax)}$.
3. Differentiate a third time: $d^3y/dx^3 = a^2(a e^{(ax)}) = a^3 e^{(ax)}$.
4. The pattern continues: $d^4y/dx^4 = a^4 e^{(ax)}$, and in general $d^n y/dx^n = a^n e^{(ax)}$ for $n>=1$.
5. Final answer: $d^n y/dx^n = a^n e^{(ax)} = a^n y$.

Example 5: Equations of the Tangent and Normal Lines

Problem: Find the equations of the tangent and normal to the curve $y = \sec x$ at the point where $x = \pi/4$.

1. Differentiate $y = \sec x$: $dy/dx = \sec x \tan x$.
2. Evaluate the slope at $x=\pi/4$: $m = \sec(\pi/4) \tan(\pi/4) = \sqrt{2} \cdot 1 = \sqrt{2}$.
3. Find the y-coordinate at $x=\pi/4$: $y = \sec(\pi/4) = \sqrt{2}$, so the point is $(\pi/4, \sqrt{2})$.
4. Tangent line: $y - \sqrt{2} = \sqrt{2}(x - \pi/4)$, which simplifies to $4\sqrt{2}x - 4y + 4\sqrt{2} - \pi \cdot \sqrt{2} = 0$.
5. Normal line (slope $-1/\sqrt{2}$): $y - \sqrt{2} = -(1/\sqrt{2})(x - \pi/4)$, which simplifies to $x + 4\sqrt{2}y - 8 - \pi = 0$.

Example 6: Finding Intervals of Increase and Decrease

Problem: Find the intervals on which $f(x) = 2x^3 - 3x^2 - 12x$ is increasing or decreasing.

1. Differentiate: $f'(x) = 6x^2 - 6x - 12$, which is defined for all real x.

2. Find critical points by solving $f'(x)=0$: $6(x^2-x-2)=0$, i.e. $6(x+1)(x-2)=0$, giving $x=-1$ and $x=2$.
3. These critical points divide the domain into (-infinity,-1), (-1,2), and (2,infinity). Test a point in each: $f'(-2)=24>0$, $f'(0)=-12<0$, $f'(3)=24>0$.
4. Interpret the signs: positive f' means increasing, negative f' means decreasing.
5. Final answer: f is increasing on (-infinity,-1) and (2,infinity), and decreasing on (-1,2).

Example 7: Local Linear Approximation

Problem: (a) Find the local linear approximation of f(x)=sqrt(2x-1) at x=5. (b) Use it to approximate sqrt(9.2).

1. (a) Differentiate: f'(x) = 1/sqrt(2x-1). Evaluate at x=5: f(5)=sqrt(9)=3 and f'(5)=1/sqrt(9)=1/3.
2. Form the linearization: L(x) = f(5) + f'(5)(x-5) = 3 + (1/3)(x-5) = (1/3)(x+4).
3. So sqrt(2x-1) is approximately (1/3)(x+4) near x=5.
4. (b) To approximate sqrt(9.2), note 2x-1=9.2 gives x=5.1. Substitute: L(5.1) = (1/3)(5.1+4) = (1/3)(9.1) is approximately 3.0333.
5. Final answer: sqrt(9.2) is approximately 3.0333, very close to the calculator value 3.0331.

Example 8: Real-World Optimization: Maximizing a Fenced Garden's Area

Problem: A rectangular garden is to be fenced on three sides with 100 metres of fencing; the fourth side is along a wall. Find the dimensions that maximize the area, and the maximum area.

1. Let x be the side opposite the wall and y each side perpendicular to the wall. The fencing constraint is $x + 2y = 100$, so $y = (100-x)/2$.

2. Write the area function: A = xy = x(100-x)/2 = (1/2)(100x - x^2).
3. Differentiate and set to zero: dA/dx = (1/2)(100-2x) = 50-x; setting 50-x=0 gives the critical point x=50.
4. Confirm it's a maximum: d^2A/dx^2 = -1 < 0, so A is maximum at x=50.
5. Substitute back: y = (100-50)/2 = 25, and A = (1/2)(100x50 - 50^2) = 1250. Final answer: dimensions 50m by 25m, giving a maximum area of 1250 square metres.

Short Questions & Answers

What is the derivative of tan x?

sec^2 x, found using the quotient rule on sin x/cos x.

What is the derivative of e^x?

e^x itself, since ln e = 1.

What condition makes x=c a critical point of f?

f'(c)=0 (a stationary point) or f'(c) does not exist.

If f''(c) > 0 and f'(c) = 0, what does this indicate?

f has a local (relative) minimum at x=c, by the second derivative test.

What is the formula for the differential dy?

dy = f'(x) dx.

How is the normal line to a curve related to the tangent line at the same point?

It is perpendicular to the tangent line at that point.

What is the first step in solving a real-world optimization problem?

Understand the problem and identify exactly what quantity needs to be optimized.

Long Questions & Answers

Explain the chain rule and how it extends to differentiate composite, parametric, and implicit functions.

What does the chain rule state for a composite function y=f(g(x))?

If f is differentiable at u=g(x) and g is differentiable at x, then (f o g)'(x) = f'(g(x)) . g'(x), written in Leibniz notation as dy/dx = (dy/du)(du/dx).

How does the chain rule apply to a function given parametrically?

If x=f(t) and y=g(t), then dy/dx = (dy/dt)/(dx/dt), obtained by applying the chain rule to y=g(t) and t=f^-1(x).

What is implicit differentiation, and when is it needed?

It is used when a relation F(x,y)=0 cannot easily be solved explicitly for y; both sides are differentiated with respect to x, with the chain rule applied to every y-term, and the equation is then solved for dy/dx.

What are the three steps of implicit differentiation?

Differentiate both sides of F(x,y)=0 with respect to x treating y as a function of x, apply the chain rule to every term containing y, then algebraically solve the resulting equation for dy/dx.

Explain how critical points are used to find both the intervals where a function is increasing/decreasing and its local/global extrema.

What is a critical point, and how is it found?

A point x=c in the domain of f where f'(c)=0 or f'(c) does not exist; it is found by differentiating f and solving f'(x)=0, then checking for any points where f' is undefined.

How do critical points determine intervals of increase and decrease?

They split the domain into subintervals in which the sign of f' does not change; testing one point in each subinterval and checking whether f' is positive (increasing) or negative (decreasing) there classifies the whole subinterval.

How does the second derivative test classify a critical point?

If f'(c)=0, the function has a local maximum at c when f''(c)<0, a local minimum when f''(c)>0, and the test is inconclusive when f''(c)=0.

How are global extrema found on a closed interval [a,b]?

By evaluating f at every critical point inside [a,b] and at both endpoints a and b, then taking the largest value as the global maximum and the smallest as the global minimum.

Multiple Choice Questions (MCQs)

The derivative of sin x is:

A. -cos x

B. cos x

C. -sin x

D. sec^2 x

Answer: B. cos x

Explanation: d/dx(sin x) = cos x, proved directly from the limit definition of the derivative.

The chain rule for y=f(g(x)) states dy/dx equals:

A. f'(x)

B. f'(g(x))

C. f'(g(x)) . g'(x)

D. g'(x)

Answer: C. f'(g(x)) . g'(x)

Explanation: The chain rule multiplies the outer derivative (evaluated at the inner function) by the derivative of the inner function.

The derivative of a^x is:

A. a^x

B. a^x ln a

C. x a^(x-1)

D. ln a

Answer: B. a^x ln a

Explanation: d/dx(a^x) = a^x ln a, derived from the limit lim(h->0)(a^h-1)/h = ln a.

The derivative of ln x is:

A. 1/x

B. x

C. ln x

D. 1/(x ln x)

Answer: A. 1/x

Explanation: d/dx(ln x) = 1/x, the special case of d/dx(log_a x)=1/(x ln a) when a=e.

A critical point of f occurs where:

A. f(x)=0

B. f'(x)=0 or f'(x) does not exist

C. f''(x)=0 always

D. f is undefined

Answer: B. f'(x)=0 or f'(x) does not exist

Explanation: Critical points are exactly where the first derivative is zero (stationary points) or fails to exist.

By the second derivative test, if f'(c)=0 and f''(c)>0, then f has:

A. A local maximum at c

B. A local minimum at c

C. No extremum at c

D. An inflection point at c

Answer: B. A local minimum at c

Explanation: A positive second derivative at a stationary point indicates the graph is concave up there, giving a local minimum.

The linearization of f at x=a is given by:

A. f(a)

B. f'(a)

C. f(a) + f'(a)(x-a)

D. f(x) - f(a)

Answer: C. f(a) + f'(a)(x-a)

Explanation: The linearization is the equation of the tangent line: L(x) = f(a) + f'(a)(x-a).

The differential dy is defined as:

A. f(x+dx) - f(x)

B. f'(x) dx

C. dx/f'(x)

D. f''(x) dx

Answer: B. f'(x) dx

Explanation: dy = f'(x) dx is the definition of the differential, used to approximate the actual change Deltay for small dx.

The normal line to a curve at a point is:

A. Parallel to the tangent line

B. Perpendicular to the tangent line

C. Always horizontal

D. Always vertical

Answer: B. Perpendicular to the tangent line

Explanation: By definition, the normal line is perpendicular to the tangent line at the same point on the curve.

In a real-world optimization problem, after finding a critical point, the next step is to:

A. Stop, since the answer is found

B. Confirm it is a maximum or minimum using the second derivative test

C. Ignore endpoints

D. Assume it is always a maximum

Answer: B. Confirm it is a maximum or minimum using the second derivative test

Explanation: A critical point alone does not confirm whether it is a maximum or minimum; the second derivative test (or another method) is needed to verify this before interpreting the result.

Quick Revision Summary

- $d/dx(\sin x)=\cos x$, $d/dx(\cos x)=-\sin x$, $d/dx(\tan x)=\sec^2 x$ -- proved from the limit definition and quotient rule
- Chain rule: $dy/dx = (dy/du)(du/dx)$; extends to parametric equations ($dy/dx=(dy/dt)/(dx/dt)$) and implicit differentiation
- Inverse trig derivatives: $d/dx(\sin^{-1} x)=1/\sqrt{1-x^2}$, $d/dx(\tan^{-1} x)=1/(1+x^2)$, $d/dx(\sec^{-1} x)=1/(|x|\sqrt{x^2-1})$
- $d/dx(a^x)=a^x \ln a$ and $d/dx(\log_a x)=1/(x \ln a)$; reduce to e^x and $1/x$ respectively when $a=e$
- Tangent: $y-y_0=m(x-x_0)$, $m=f'(x_0)$; Normal: $y-y_0=-(1/m)(x-x_0)$
- Higher-order derivatives: differentiate f' repeatedly to get f'', f''', ..., $f^{(n)}$
- Critical points: $f'(c)=0$ or $f'(c)$ undefined; sign of f' on each subinterval determines increasing/decreasing behavior
- Second derivative test: $f'(c)=0$ and $f''(c)<0$ gives local max; $f''(c)>0$ gives local min; $f''(c)=0$ is inconclusive
- Linearization $L(x)=f(a)+f'(a)(x-a)$; differential $dy=f'(x)dx$ approximates Δy for small dx; percentage error$=|(\Delta y-dy)/\Delta y|\times 100$
- Optimization: define variables, write the function to optimize using any constraint, find critical points, confirm with 2nd derivative test, interpret the result

Exam Tips

- Memorize the six trig derivatives as a pair pattern: sin/cos and tan/sec share structure with cot/csc, each getting a negative sign for the 'co-' function
- When applying the chain rule, always identify the 'outer' and 'inner' functions first, and write out u=(inner function) explicitly before differentiating
- For implicit differentiation, remember every derivative of a y-term needs an extra factor of dy/dx from the chain rule -- this is the most common mistake
- To find tangent/normal line equations, always compute the point (x_0,y_0) and slope $m=f'(x_0)$ separately before writing the line equation
- When finding intervals of increase/decrease, always check for points where f' fails to exist (not just where $f'=0$) -- these can also be critical points
- For optimization problems, always express the objective function in a single variable using the given constraint before differentiating

Unit 3: Integration

Integration is the reverse process of differentiation: given a function's derivative, integration recovers the family of functions it could have come from. This unit builds a complete toolkit for evaluating integrals -- the basic power/sum rules, the method of substitution (reversing the chain rule), trigonometric substitution for radicals, integration by parts (reversing the product rule), and partial fractions for rational functions -- then introduces the definite integral, the Fundamental Theorem of Calculus, and a rich set of properties that make definite integrals easy to evaluate using symmetry.

The unit closes with real applications: finding the area under a curve and between two curves (including consumer and producer surplus in economics), finding volumes of solids of revolution using the disk and cross-sectional methods, and computing the moment of inertia of physical bodies -- showing integration as the natural tool for accumulation, area, and volume problems across mathematics, physics, and economics.

Learning Objectives

- Define the antiderivative and indefinite integral, and apply the basic rules and standard integration formulas
- Evaluate integrals using the method of substitution (change of variable)
- Evaluate integrals involving radicals using trigonometric substitution
- Evaluate integrals using integration by parts, including reduction and repeated applications
- Evaluate integrals of rational functions using partial fractions (all four standard cases)
- Define the definite integral, state the Fundamental Theorem of Calculus, and apply the properties of definite integrals
- Find the area under a curve, the area between two curves, and consumer/producer surplus using definite integrals
- Find the volume of a solid of revolution using the disk and cross-sectional methods, and compute the moment of inertia of simple bodies

Key Concepts

3.1 Antiderivatives, Basic Rules, and Substitution

A function F is an antiderivative of f if F'(x)=f(x); the indefinite integral integral f(x)dx = F(x)+c collects every antiderivative of f, differing only by a constant c. The basic rules are: integral kf(x)dx = k integral f(x)dx, integral[f(x)+-g(x)]dx = integral f(x)dx +- integral g(x)dx, the power rule integral x^n dx = x^(n+1)/(n+1)+c (n<>-1), and integral f'(x)/f(x) dx = ln|f(x)|+c.

The method of substitution reverses the chain rule: setting u=g(x) so du=g'(x)dx converts integral f(g(x))g'(x)dx into the simpler integral f(u)du = F(u)+c, then substituting back u=g(x). Recognizing a composite structure -- an 'inner function' whose derivative also appears in the integrand -- is the key skill for choosing a good substitution.

3.2 Trigonometric Substitution and Integration by Parts

Integrals containing sqrt(a^2-x^2), sqrt(a^2+x^2), or sqrt(x^2-a^2) can be simplified by substituting x=a sin(theta), x=a tan(theta), or x=a sec(theta) respectively, using the Pythagorean identities to eliminate the radical; this yields standard results such as integral dx/sqrt(a^2-x^2) = sin^-1(x/a)+c and integral dx/sqrt(x^2+a^2) = ln|x+sqrt(x^2+a^2)|+c.

Integration by parts reverses the product rule: integral f(x)g(x)dx = f(x)(integral g(x)dx) - integral[(integral g(x)dx)f'(x)]dx, often summarized as 'first function times integral of second, minus integral of derivative of first times integral of second.' Choosing which factor to differentiate and which to integrate (commonly guided by the order algebraic-logarithmic-trigonometric-exponential) is the central skill, and some integrals (like integral e^(ax) sin(bx) dx) require applying the rule twice and solving algebraically for the original integral.

3.3 Integration Using Partial Fractions

When P(x)/Q(x) is a proper rational function and Q(x) factors into linear and irreducible quadratic pieces, the fraction can be split into simpler pieces that are each easy to integrate. Four cases arise: non-repeated linear factors (A/(x-a) terms), repeated linear factors (A/(x-a) + B/(x-a)^2 + ... terms), non-repeated irreducible quadratic factors ((Ax+B)/(x^2+px+q) terms), and repeated irreducible quadratic factors.

In every case, the unknown constants are found by clearing denominators to get a polynomial identity, then either substituting convenient values of x (roots of the linear factors) or comparing coefficients of matching powers of x; once split, each piece integrates using the power rule, the ln|f(x)| rule, or a standard arctan/arcsin form.

3.4 The Definite Integral and Its Properties

The definite integral integral from a to b of f(x)dx = F(b)-F(a), where F'=f, is a real number (not a family of functions) -- this is the Second Fundamental Theorem of Calculus. The First Fundamental Theorem states that if F(x)=integral from a to x of f(t)dt, then F'(x)=f(x): differentiating an integral with variable upper limit just recovers the integrand.

A rich set of properties makes definite integrals easier to evaluate: reversing limits flips the sign (property II), an integral splits over subintervals (property III), integral from a to b of f(x)dx = integral from a to b of f(a+b-x)dx (property IV, very useful for symmetry tricks), even/odd functions on [-a,a] simplify to 2*integral from 0 to a or 0 respectively (property V), and linearity holds for sums of functions (property IX).

3.5 Applications: Area Under and Between Curves

If f(x)>=0 on [a,b], the area under y=f(x) from a to b is simply integral from a to b of f(x)dx; if f(x)<=0, the area is the absolute value of that same integral, since the definite integral there is negative (a 'signed area'). Where f changes sign, the total area is the sum of the absolute values of the integral over each sub-region between consecutive roots. For two curves with f(x)>=g(x) on [a,b], the area between them is integral from a to b of [f(x)-g(x)]dx, found after locating their intersection points.

In economics, the same idea gives consumer surplus CS = integral from 0 to x_e of [D(x)-p_e]dx (the area between the demand curve and the equilibrium price line) and producer surplus PS = integral from 0 to x_e of [p_e-S(x)]dx (the area between the equilibrium price line and the supply curve), where (x_e,p_e) is the market equilibrium point where demand meets supply.

3.6 Applications: Volumes of Revolution and Moment of Inertia

Revolving the region under y=f(x) (f>=0) on [a,b] about the x-axis generates a solid whose volume is V = integral from a to b of pi[f(x)]^2 dx (the disk method); revolving a region bounded by x=g(y) about the y-axis similarly gives V = integral from c to d of pi[g(y)]^2 dy. More generally, if a solid has cross-sectional area A(x) at each point x between a and b, its volume is V = integral from a to b of A(x)dx (the cross-sectional method), which reduces to the disk method when every cross-section is a circle.

The moment of inertia of a particle of mass m at distance r from an axis is mr^2; for a continuous body it becomes the integral I = integral r^2 dm taken over the whole body, found by expressing an element of mass dm in terms of its distance x from the axis (using the body's mass-per-unit-length or mass-per-unit-area) and integrating -- this same accumulation idea also appears in problems involving population growth, drug dosage, and other real-world total-change quantities.

Important Definitions

What is an antiderivative of f(x)?

A function F(x) such that F'(x)=f(x) for all x in the interval considered.

What is the indefinite integral of f(x)?

integral f(x)dx = F(x)+c, the family of all antiderivatives of f, where F is any one antiderivative and c is an arbitrary constant.

What is the method of substitution used for?

To simplify an integral by replacing x with a new variable u=g(x) (and dx with du), reversing the chain rule -- useful when the integrand contains a composite function.

When is trigonometric substitution useful?

When the integrand contains sqrt(a^2-x^2), sqrt(a^2+x^2), or sqrt(x^2-a^2); substituting x=a sin(theta), a tan(theta), or a sec(theta) respectively eliminates the radical.

What is the formula for integration by parts?

integral f(x)g(x)dx = f(x)(integral g(x)dx) - integral[(integral g(x)dx) f'(x)]dx.

What is a definite integral?

integral from a to b of f(x)dx = F(b)-F(a), where F'=f; it is a real number, the net signed area between the curve and the x-axis from a to b.

What does the First Fundamental Theorem of Calculus state?

If F(x)=integral from a to x of f(t)dt, then F'(x)=f(x) -- differentiating an integral with respect to its upper limit recovers the original function.

What is consumer surplus?

The total benefit consumers gain from paying the equilibrium price rather than the higher price they were willing to pay; geometrically, the area between the demand curve and the equilibrium price line.

What is the disk method for volume of revolution?

V = integral from a to b of pi[f(x)]^2 dx, the volume of the solid formed by revolving the region under y=f(x) about the x-axis.

What is the moment of inertia of a continuous body about an axis?

I = integral r^2 dm, taken over the whole body, where r is the distance from each mass element dm to the axis.

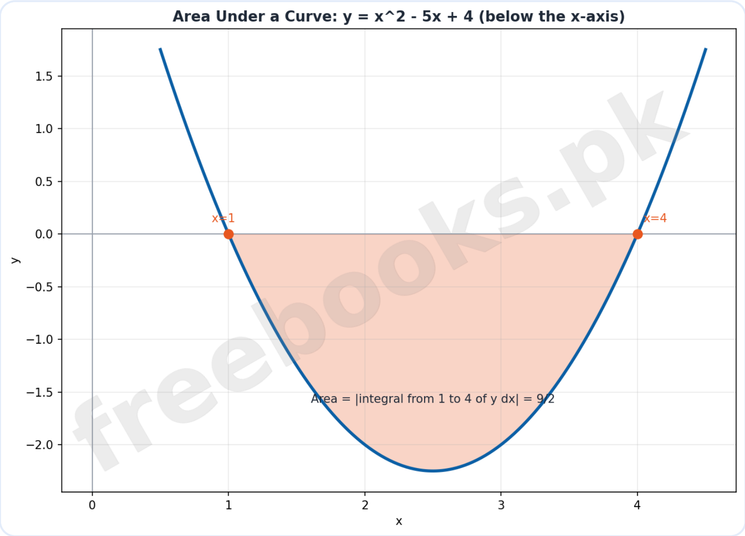
Key Facts and Relations

Topic	Key Fact / Relation
Basic integration rules	integral kf(x)dx=k integral f(x)dx; integral x^n dx = x^(n+1)/(n+1)+c (n<>-1); integral f'(x)/f(x) dx = ln|f(x)|+c
Standard integrals	integral e^x dx=e^x+c; integral a^x dx=a^x/ln a+c; integral sec^2 x dx=tan x+c; integral dx/sqrt(1-x^2)=sin^-1 x+c
Substitution rule	u=g(x), du=g'(x)dx => integral f(g(x))g'(x)dx = integral f(u)du = F(u)+c
Trigonometric substitution cases	sqrt(a^2-x^2): x=a sin(theta); sqrt(a^2+x^2): x=a tan(theta); sqrt(x^2-a^2): x=a sec(theta)
	integral f g dx = f(integral g dx) - integral[(integral g dx) f'] dx

Topic	Key Fact / Relation
Integration by parts	
Definite integral / FTC	integral from a to b of f(x)dx = F(b)-F(a); d/dx[integral from a to x of f(t)dt] = f(x)
Key properties of definite integrals	integral from a to b = -integral from b to a; integral from a to b of f(x)dx = integral from a to b of f(a+b-x)dx; even f: integral from -a to a = 2 integral from 0 to a; odd f: integral from -a to a = 0
Area formulas	Under curve: integral from a to b of f(x)dx (or its absolute value if f<=0); Between curves: integral from a to b of [f(x)-g(x)]dx, f>=g
Consumer / Producer surplus	CS = integral from 0 to x_e of [D(x)-p_e]dx; PS = integral from 0 to x_e of [p_e-S(x)]dx
Volume of revolution (disk method)	V = integral from a to b of pi[f(x)]^2 dx (about x-axis); V = integral from c to d of pi[g(y)]^2 dy (about y-axis)

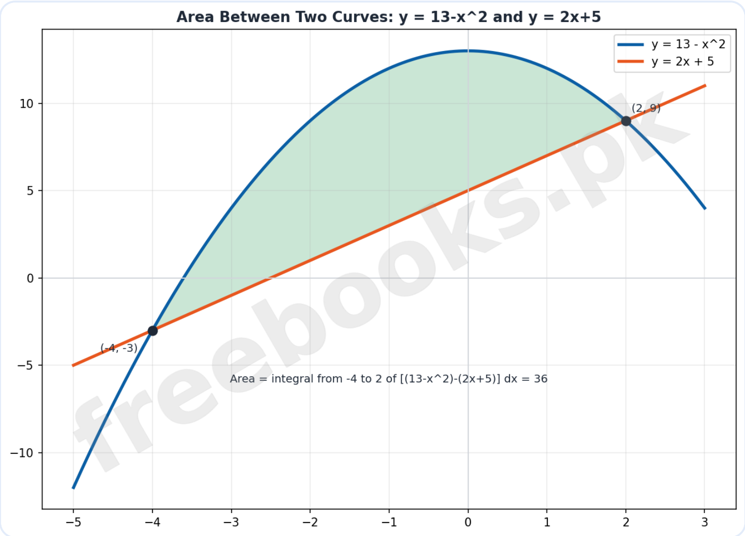
Diagrams

Area Under a Curve (Signed Area): The graph of y=x^2-5x+4, shaded between x=1 and x=4 where the curve lies below the x-axis, illustrating that the area equals the absolute value of the definite integral there



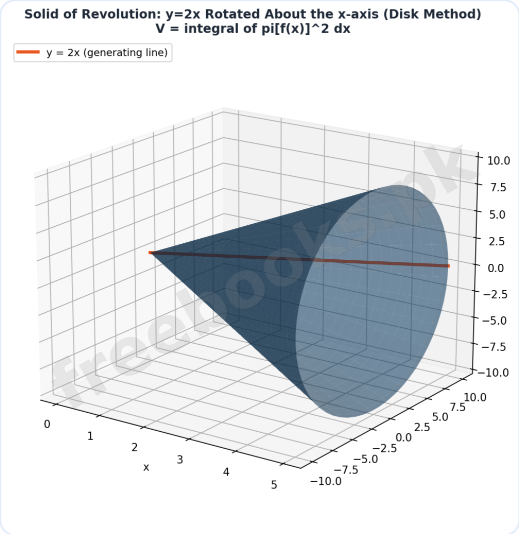
Area Under a Curve (Signed Area)

Area Between Two Curves: The graphs of y=13-x^2 and y=2x+5 with the shaded region between their intersection points at x=-4 and x=2, illustrating area = integral of (upper - lower) dx



Area Between Two Curves

Solid of Revolution (Disk Method): The line y=2x revolved about the x-axis from x=0 to x=5, generating a cone whose volume is found via V = integral of pi[f(x)]^2 dx



Solid of Revolution (Disk Method)

Solved Examples

Example 1: Basic Integration Using the Power Rule

Problem: Evaluate integral of (8x^7 - 6x^5 + 4x^3 - 2x - 3) dx.

1. Split the integral term by term using the sum/difference rule: integral 8x^7 dx - integral 6x^5 dx + integral 4x^3 dx - integral 2x dx - integral 3 dx.
2. Pull out constants using the constant multiple rule: 8 integral x^7 dx - 6 integral x^5 dx + 4 integral x^3 dx - 2 integral x dx - 3 integral 1 dx.
3. Apply the power rule to each term: 8(x^8/8) - 6(x^6/6) + 4(x^4/4) - 2(x^2/2) - 3x + c.
4. Simplify each coefficient: x^8 - x^6 + x^4 - x^2 - 3x + c.
5. Final answer: integral (8x^7-6x^5+4x^3-2x-3)dx = x^8 - x^6 + x^4 - x^2 - 3x + c.

Example 2: Integration by Substitution

Problem: Evaluate integral of (5x^5+5)^3 x^4 dx.

1. Notice the composite structure: let u = 5x^5+5, so du = 25x^4 dx.

2. Rewrite the integral to introduce the factor 25: integral (5x^5+5)^3 x^4 dx = (1/25) integral (5x^5+5)^3 (25x^4) dx.
3. Substitute u and du: (1/25) integral u^3 du.
4. Integrate using the power rule: (1/25)(u^4/4) + c = u^4/100 + c.
5. Substitute back u=5x^5+5: final answer = (1/100)(5x^5+5)^4 + c.

Example 3: Integration by Parts

Problem: Evaluate integral of x e^(ax) dx.

1. Apply integration by parts with 'first function' x (differentiate) and 'second function' e^(ax) (integrate): integral x e^(ax) dx = x(e^(ax)/a) - integral (e^(ax)/a)(1)dx.
2. Simplify the remaining integral: = x e^(ax)/a - (1/a) integral e^(ax) dx.
3. Integrate e^(ax): (1/a) integral e^(ax)dx = (1/a)(e^(ax)/a) = e^(ax)/a^2.
4. Combine terms: x e^(ax)/a - e^(ax)/a^2 + c = e^(ax)(x/a - 1/a^2) + c.
5. Final answer: integral x e^(ax) dx = e^(ax)[(ax-1)/a^2] + c.

Example 4: Integration Using Partial Fractions (Non-Repeated Linear Factors)

Problem: Evaluate integral of (3x-2)/[x(x-1)(x-2)] dx.

1. Write the partial fraction decomposition: (3x-2)/[x(x-1)(x-2)] = A/x + B/(x-1) + C/(x-2).
2. Clear denominators: 3x-2 = A(x-1)(x-2) + Bx(x-2) + Cx(x-1).
3. Substitute x=0: -2 = A(-1)(-2) = 2A, so A=-1. Substitute x=1: 1 = B(1)(-1) = -B, so B=-1. Substitute x=2: 4 = C(2)(1) = 2C, so C=2.
4. Rewrite and integrate each simple term: integral[-1/x - 1/(x-1) + 2/(x-2)]dx = -ln|x| - ln|x-1| + 2ln|x-2| + c.
5. Final answer: integral (3x-2)/[x(x-1)(x-2)] dx = -ln|x| - ln|x-1| + 2ln|x-2| + c.

Example 5: Evaluating a Definite Integral

Problem: Evaluate integral from 0 to sqrt(7) of x/sqrt(x^2+9) dx.

1. Recognize the substitution: let f(x)=x^2+9, so f'(x)=2x appears (up to a constant) in the numerator.
2. Rewrite with the needed factor of 2: (1/2) integral from 0 to sqrt(7) of (x^2+9)^(-1/2) (2x) dx.
3. Integrate using the power rule (or recognizing the antiderivative sqrt(x^2+9)): this equals [sqrt(x^2+9)] evaluated from 0 to sqrt(7).
4. Substitute the limits: sqrt(7+9) - sqrt(0+9) = sqrt(16) - sqrt(9) = 4 - 3.
5. Final answer: integral from 0 to sqrt(7) of x/sqrt(x^2+9) dx = 1.

Example 6: Area Bounded by a Curve and the x-axis

Problem: Find the area bounded by the parabola y = x^2 - 5x + 4 and the x-axis.

1. Find the x-intercepts: x^2-5x+4=0 factors as (x-1)(x-4)=0, giving x=1 and x=4.
2. Since the parabola opens upward and its roots are 1 and 4, the curve lies below the x-axis between x=1 and x=4.
3. Because y<=0 on [1,4], the area is the absolute value: Area = integral from 1 to 4 of -(x^2-5x+4)dx = integral from 1 to 4 of (-x^2+5x-4)dx.
4. Integrate: [-x^3/3 + 5x^2/2 - 4x] evaluated from 1 to 4.
5. Evaluate: at x=4 this is -64/3+40-16=8/3; at x=1 this is -1/3+5/2-4=-11/6. Subtracting: 8/3-(-11/6)=27/6=9/2. Final answer: Area = 9/2 square units.

Example 7: Area Between Two Curves

Problem: Find the area bounded by the curves y = 13 - x^2 and y = 2x + 5.

1. Find the intersection points by setting the curves equal: 13-x^2 = 2x+5, which rearranges to x^2+2x-8=0, i.e. (x-2)(x+4)=0, giving x=-4 and x=2.

2. Determine which curve is on top by testing x=0: y=13-0=13 for the parabola, y=2(0)+5=5 for the line -- so the parabola is above the line on [-4,2].
3. Set up the area integral: Area = integral from -4 to 2 of [(13-x^2)-(2x+5)]dx = integral from -4 to 2 of (8-x^2-2x)dx.
4. Integrate: [8x - x^3/3 - x^2] evaluated from -4 to 2.
5. Evaluate: at x=2 this is 16-8/3-4=28/3; at x=-4 this is -32+64/3-16=-80/3. Subtracting: 28/3-(-80/3)=108/3=36. Final answer: Area = 36 square units.

Example 8: Volume of a Solid of Revolution

Problem: A cone is formed by rotating the line y=2x about the x-axis from x=0 to x=5 cm. Find the volume of the cone.

1. Apply the disk method formula: V = integral from a to b of pi[f(x)]^2 dx, with f(x)=2x, a=0, b=5.
2. Substitute: V = pi integral from 0 to 5 of (2x)^2 dx = pi integral from 0 to 5 of 4x^2 dx = 4pi integral from 0 to 5 of x^2 dx.
3. Integrate: 4pi[x^3/3] evaluated from 0 to 5.
4. Evaluate: 4pi(125/3 - 0) = 500pi/3.
5. Final answer: V = 500pi/3 is approximately 523.6 cubic centimetres.

Short Questions & Answers

What does the power rule for integration state?

integral x^n dx = x^(n+1)/(n+1) + c, for n not equal to -1.

What is integral 1/x dx?

ln|x| + c.

What idea does the substitution method reverse?

The chain rule for differentiation.

What idea does integration by parts reverse?

The product rule for differentiation.

What is the value of integral from a to a of f(x)dx?

0, since the interval has zero width.

If f is an odd function, what is integral from -a to a of f(x)dx?

0, because the positive and negative contributions exactly cancel.

What does the area between two curves f(x)>=g(x) on [a,b] equal?

integral from a to b of [f(x)-g(x)]dx.

Long Questions & Answers

Explain the four cases of partial fraction decomposition and how the unknown constants are found in each case.

What is Case I, non-repeated linear factors, and how are constants found?

Each linear factor (x-a) gets a term A/(x-a); after clearing denominators, substituting x=a directly isolates A (since every other term vanishes at x=a).

What is Case II, repeated linear factors?

A factor (x-a)^n contributes n terms: A1/(x-a) + A2/(x-a)^2 + ... + An/(x-a)^n; the highest-power constant can be found by substitution, while the rest typically require comparing coefficients.

What is Case III, non-repeated irreducible quadratic factors?

An irreducible quadratic factor (x^2+px+q) contributes a term (Bx+C)/(x^2+px+q); after clearing denominators, comparing coefficients of matching powers of x (and using any linear-factor substitutions first) solves for B and C.

What is Case IV, repeated irreducible quadratic factors, and why is it hardest?

A repeated quadratic factor (x^2+px+q)^n contributes multiple terms with rising powers of the quadratic in the denominator; solving typically requires a mix of substitution, coefficient comparison, and further trigonometric substitution (x=a tan(theta)) to integrate the highest-power term.

Explain how definite integrals are used to find the area under a curve, the area between two curves, and consumer/producer surplus.

How is the area under a single curve found when f(x)>=0?

Directly as integral from a to b of f(x)dx; if f(x)<=0 instead, the definite integral is negative, so the actual area is its absolute value.

How is the area between two curves found?

By identifying which curve is on top on the interval of interest (often by testing a point), then integrating (upper function - lower function) over that interval, after first finding the intersection points to determine the limits.

How is consumer surplus defined and computed?

As the benefit to consumers from paying the equilibrium price rather than a higher price; geometrically the area between the demand curve and the equilibrium price line, computed as CS = integral from 0 to x_e of [D(x)-p_e]dx.

How is producer surplus defined and computed?

As the benefit to producers from receiving the equilibrium price rather than a lower price; geometrically the area between the equilibrium price line and the supply curve, computed as PS = integral from 0 to x_e of [p_e-S(x)]dx.

Multiple Choice Questions (MCQs)

integral x^n dx equals (for n not equal to -1):

A. nx^(n-1)+c
B. x^(n+1)/(n+1)+c
C. x^n/n+c
D. x^(n-1)+c
Answer: B. x^(n+1)/(n+1)+c
Explanation: The power rule for integration is integral x^n dx = x^(n+1)/(n+1)+c, valid whenever n is not -1.

The method of substitution reverses which differentiation rule?

A. Product rule

B. Quotient rule

C. Chain rule

D. Power rule

Answer: C. Chain rule

Explanation: Substitution undoes the chain rule, replacing a composite integrand with a simpler one in terms of u=g(x).

Integration by parts reverses which differentiation rule?

A. Chain rule

B. Product rule

C. Quotient rule

D. Power rule

Answer: B. Product rule

Explanation: Integration by parts is derived directly from the product rule for derivatives.

For sqrt(a^2-x^2) in an integrand, the standard trigonometric substitution is:

A. x=a tan(theta)

B. x=a sec(theta)

C. x=a sin(theta)

D. x=a cos(theta)

Answer: C. x=a sin(theta)

Explanation: x=a sin(theta) makes a^2-x^2 = a^2 cos^2(theta), eliminating the radical using the Pythagorean identity.

A rational function P(x)/Q(x) can be split using partial fractions when:

A. P(x) is always zero

B. Q(x) factors into linear and/or irreducible quadratic factors

C. P(x) has a higher degree than Q(x) always

D. Q(x) is never zero

Answer: B. Q(x) factors into linear and/or irreducible quadratic factors

Explanation: Partial fractions requires Q(x) to factor into linear and/or irreducible quadratic pieces, over which the original fraction is decomposed.

The Second Fundamental Theorem of Calculus states that integral from a to b of f(x)dx equals:

A. F(a)-F(b)

B. F(b)-F(a)

C. F(a)+F(b)

D. f(b)-f(a)

Answer: B. F(b)-F(a)

Explanation: Where F is any antiderivative of f, the definite integral equals F(b)-F(a).

If f is an even function, integral from -a to a of f(x)dx equals:

A. 0

B. 2 integral from 0 to a of f(x)dx

C. integral from 0 to a of f(x)dx

D. -2 integral from 0 to a of f(x)dx

Answer: B. 2 integral from 0 to a of f(x)dx

Explanation: Even symmetry lets the integral over [-a,a] be replaced by twice the integral over [0,a].

The area between two curves f(x)>=g(x) on [a,b] is given by:

A. integral from a to b of f(x)dx

B. integral from a to b of g(x)dx

C. integral from a to b of [f(x)-g(x)]dx

D. integral from a to b of [f(x)+g(x)]dx

Answer: C. integral from a to b of [f(x)-g(x)]dx

Explanation: The area between two curves is the integral of the difference (upper minus lower function) over the interval.

The disk method for a solid revolved about the x-axis gives volume:

A. integral pi f(x) dx

B. integral pi [f(x)]^2 dx

C. integral 2 pi f(x) dx

D. integral [f(x)]^2 dx

Answer: B. integral pi [f(x)]^2 dx

Explanation: Each disk has radius f(x) and area pi[f(x)]^2, so the volume is the integral of pi[f(x)]^2 over the interval.

The moment of inertia of a continuous body about an axis is given by:

A. integral r dm

B. integral m dr

C. integral r^2 dm

D. integral dm

Answer: C. integral r^2 dm

Explanation: The moment of inertia sums (distance)^2 x (mass element) over the whole body, i.e. I = integral r^2 dm.

Quick Revision Summary

- Antiderivative: F'(x)=f(x); indefinite integral integral f(x)dx = F(x)+c
- Basic rules: constant multiple, sum/difference, power rule x^(n+1)/(n+1)+c, and integral f'(x)/f(x)dx=ln|f(x)|+c
- Substitution: u=g(x), du=g'(x)dx, reduces integral f(g(x))g'(x)dx to integral f(u)du
- Trig substitution: sqrt(a^2-x^2) uses x=a sin(theta); sqrt(a^2+x^2) uses x=a tan(theta); sqrt(x^2-a^2) uses x=a sec(theta)
- Integration by parts: integral fg dx = f(integral g dx) - integral[(integral g dx)f']dx
- Partial fractions: four cases based on linear/quadratic, repeated/non-repeated factors of the denominator
- Definite integral: integral from a to b of f(x)dx=F(b)-F(a) (FTC); key properties include reversal, splitting, and even/odd symmetry
- Area under a curve: integral f(x)dx (or its absolute value if f<=0); area between curves: integral of (upper-lower)dx
- Consumer/producer surplus: areas between demand/supply curves and the equilibrium price line
- Volume of revolution (disk method): V=integral pi[f(x)]^2 dx; moment of inertia: I=integral r^2 dm

Exam Tips

- Always try the simplest technique first: check if the integral matches a basic formula before reaching for substitution or parts
- For substitution, look for a function and its derivative (up to a constant) both appearing in the integrand -- that's the signal to set u equal to that function

- For integration by parts, choose the 'first function' (to differentiate) using the rough priority: inverse trig/log, then algebraic, then trig, then exponential -- whichever simplifies fastest when differentiated
- For partial fractions, always check the degree of the numerator is less than the denominator first; if not, divide before decomposing
- When finding area between curves, always test a point between the intersection limits to confirm which curve is really on top before setting up the integral
- For solids of revolution, sketch the region and the axis of revolution first -- confusing revolution about the x-axis with the y-axis is the most common setup mistake

Unit 4: Differential Equations

A differential equation is an equation that involves at least one derivative of a dependent variable with respect to an independent variable. First proposed by Newton and Leibniz during the development of calculus, differential equations describe how a quantity changes -- velocity is the derivative of position, population growth depends on current population, and cooling depends on the temperature gap with the surroundings. This unit begins by classifying differential equations by order and degree, then shows how real situations (population growth, Newton's law of cooling, inflation, and electrical circuits) are translated into first-order differential equations.

The unit then develops two core solving techniques for first-order, first-degree equations: separable differential equations, which split into an x-only side and a y-only side that can be integrated directly, and homogeneous differential equations, which reduce to a separable equation after the substitution $y=vx$. The unit closes with a wide range of real-life applications -- population growth, radioactive decay, inflation and deflation, Newton's law of cooling, and RL circuit currents -- each solved by setting up and solving a first-order differential equation with a given initial condition.

Learning Objectives

- Define a differential equation and identify its order and degree
- Construct first-order differential equations from practical situations such as population growth, Newton's law of cooling, inflation/deflation, and RC/RL circuits
- Recognize and solve separable differential equations by integrating both sides after separating variables
- Recognize and solve homogeneous differential equations using the substitution $y=vx$
- Solve initial value problems (IVPs) by applying a given condition to find the arbitrary constant

- Apply differential equations to real-life problems: population growth/decay, radioactive decay, inflation/deflation, Newton's law of cooling, and RL circuits

Key Concepts

4.1 Definition, Order, and Degree

A differential equation involves at least one derivative of a dependent variable (commonly y) with respect to an independent variable (commonly x); when only ordinary derivatives are involved, it is called an ordinary differential equation. The order of a differential equation is the order of its highest derivative, while the degree is the greatest exponent (power) of that highest-order derivative once the equation is written as a polynomial in its derivatives.

A differential equation of order 1 is called a first-order differential equation. Constructing such equations from real situations -- population growth ($dP/dt=kP$), Newton's law of cooling ($dT/dt=k(T_s-T)$), inflation/deflation ($dP/dt=kP$), and RC/RL circuits ($RC\, dq/dt=CV_S-q$ and $L\, dI/dt=V_S-RI$, from Kirchhoff's voltage law) -- is the essential first skill before any solving technique is introduced.

4.2 Separable Differential Equations

A first-order, first-degree differential equation $dy/dx=f(x,y)$ is separable if it can be written as $dy/dx=g(x)h(y)$, i.e. the right-hand side factors into a function of x only times a function of y only. It is solved by rewriting it as $(1/h(y))dy=g(x)dx$ and integrating both sides independently: integral $(1/h(y))dy$ = integral $g(x)dx$, giving the general solution.

A condition of the form $y(x_0)=y_0$ is called an initial value condition, and a differential equation together with such a condition is an Initial Value Problem (IVP); substituting the condition into the general solution pins down the arbitrary constant, giving a unique particular solution.

4.3 Homogeneous Differential Equations

A first-order, first-degree differential equation is homogeneous if it can be written as $dy/dx=F(y/x)$, where the right-hand side depends only on the ratio y/x -- every term of the numerator and denominator (if written as a fraction) has the same total degree. Equations like $dy/dx=x+y$ are not homogeneous, since $x+y$ cannot be expressed purely as a function of y/x.

The substitution $y=vx$ converts a homogeneous equation into a separable one: since $dy/dx=v+x(dv/dx)$, the equation $v+x(dv/dx)=F(v)$ rearranges to $x(dv/dx)=F(v)-v$, which is separable in v and x. After integrating and substituting back $v=y/x$, the general solution of the original homogeneous equation is obtained.

4.4 Real-Life Applications

Population growth and decay follows $dP/dt=kP$, whose solution is $P(t)=P_0 e^{kt}$; $k>0$ gives growth (e.g. bacteria cultures), $k<0$ gives decay (e.g. radioactive substances), and two given data points pin down both P_0 and k. The same model describes inflation ($k>0$) and deflation ($k<0$) of prices over time.

Newton's law of cooling/heating gives $dT/dt=-k(T-T_s)$, whose solution $T(t)=T_s+ce^{-kt}$ approaches the surrounding temperature T_s as t grows; and RL circuits governed by $L(dI/dt)=V_S-RI$ (from Kirchhoff's voltage law) give a current $I(t)$ that approaches the steady-state value V_S/R as t grows -- both are separable equations solved the same way, using given initial and later-time data to find the constants.

Important Definitions

What is a differential equation?

An equation that involves at least one derivative of a dependent variable with respect to an independent variable.

What is the order of a differential equation?

The order of the highest derivative that appears in the differential equation.

What is the degree of a differential equation?

The greatest exponent (power) of the highest-order derivative, once the equation is written as a polynomial in its derivatives.

What is an initial value condition?

A condition of the form $y(x_0)=y_0$, specifying the value of the dependent variable at a particular value of the independent variable.

What is an Initial Value Problem (IVP)?

A differential equation together with an initial value condition, whose solution is a unique particular solution rather than a general family.

What is a separable differential equation?

A first-order, first-degree equation $dy/dx=g(x)h(y)$, where the right-hand side factors into a function of x only times a function of y only.

What is a homogeneous differential equation?

A first-order, first-degree equation $dy/dx=F(y/x)$, where the right-hand side depends only on the ratio y/x.

What substitution solves a homogeneous differential equation?

$y=vx$, which converts the equation into a separable equation in the variables v and x.

What does Newton's law of cooling state?

The rate of change of a body's temperature is proportional to the difference between the body's temperature and the surrounding temperature: $dT/dt=k(T_s-T)$.

What is half-life?

The time required for a quantity of a radioactive substance to reduce to half of its initial amount due to radioactive decay.

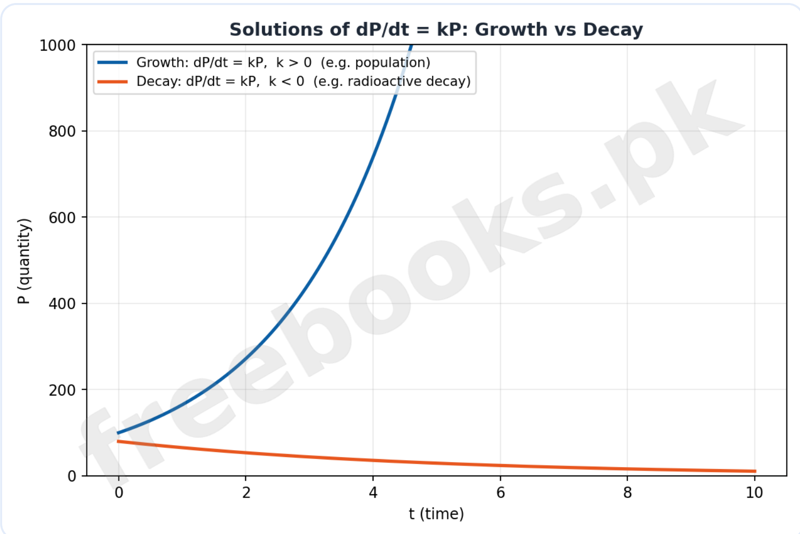
Key Facts and Relations

Topic	Key Fact / Relation
Population/price growth-decay model	$dP/dt = kP \Rightarrow P(t) = P_0 e^{kt}$; $k>0$ growth, $k<0$ decay
Newton's law of cooling/heating	$dT/dt = k(T_s - T) \Rightarrow T(t) = T_s + c e^{-kt}$
RC circuit (charging)	$RC\, dq/dt = CV_S - q$, from $V_S = q/C + IR$ and $I = dq/dt$
RL circuit (charging)	$L\, dI/dt = V_S - RI$, from $V_S = V_L + V_R = L(dI/dt) + RI$

Topic	Key Fact / Relation
Separable equation form and solution	dy/dx = g(x)h(y) => integral (1/h(y))dy = integral g(x)dx
Homogeneous equation form	dy/dx = F(y/x), where F depends only on the ratio y/x
Homogeneous substitution	y = vx => dy/dx = v + x(dv/dx)
Homogeneous reduces to separable	x(dv/dx) = F(v) - v => integral 1/(F(v)-v) dv = integral (1/x) dx
Radioactive decay model	dM/dt = -kM => M(t) = c e^(-kt), k>0 the decay constant
Initial value condition	y(x_0) = y_0, used to solve for the arbitrary constant in the general solution

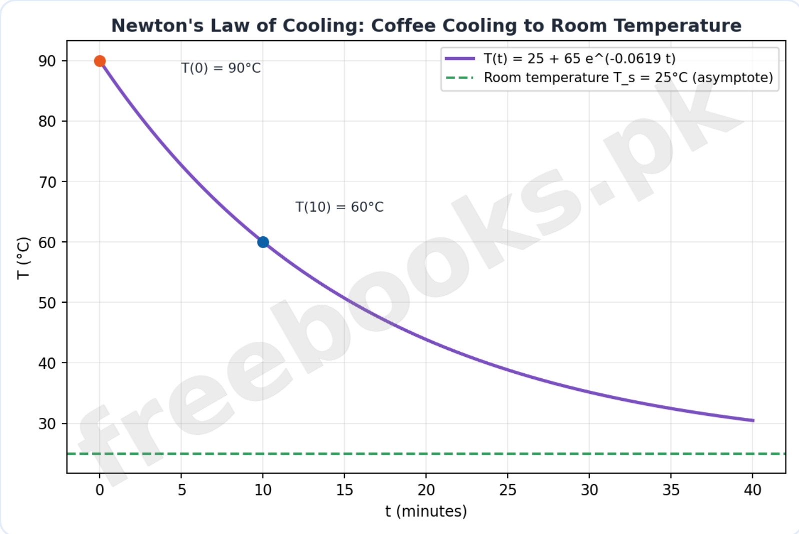
Diagrams

Growth vs Decay Solutions of dP/dt = kP: Two exponential curves illustrating the general solution P(t)=P_0 e^(kt): growth when k>0 (e.g. population, inflation) and decay when k<0 (e.g. radioactive decay, deflation)



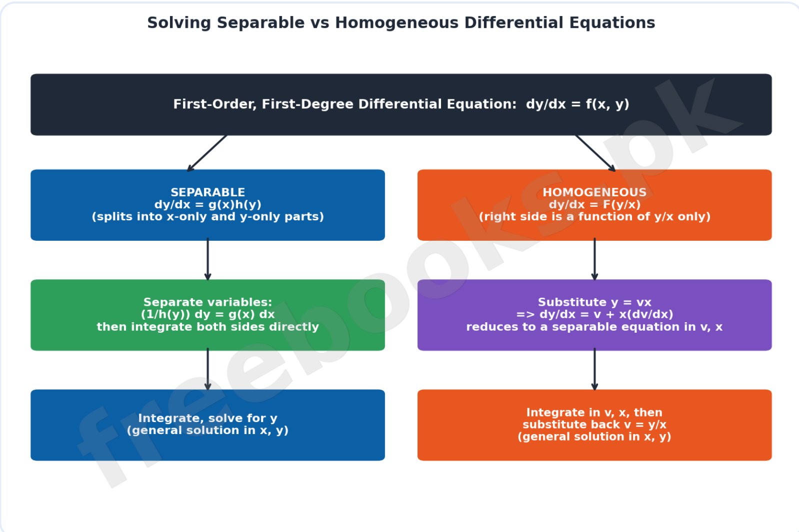
Growth vs Decay Solutions of dP/dt = kP

Newton's Law of Cooling Curve: The curve T(t)=25+65e^(-0.0619t) for a cooling cup of coffee, approaching the room temperature asymptote T_s=25°C, matching the given data points T(0)=90 and T(10)=60



Newton's Law of Cooling Curve

Solving Separable vs Homogeneous Differential Equations: A flowchart comparing the two solution methods: separating variables directly for a separable equation, versus substituting y=vx to reduce a homogeneous equation to a separable one before integrating



Solving Separable vs Homogeneous Differential Equations

Solved Examples

Example 1: Finding Order and Degree

Problem: Find the order and degree of the differential equation sqrt(xy)(dy/dx) + (dy/dx)^2 = sin y.

1. Identify the highest-order derivative present: only dy/dx (a first derivative) appears in the equation, so the order is 1.
2. Check that the equation is a polynomial in dy/dx: sqrt(xy)(dy/dx) + (dy/dx)^2 - sin y = 0 is indeed a polynomial in dy/dx (no square roots or fractional powers of the derivative itself).
3. Find the greatest exponent of the highest-order derivative dy/dx: the term (dy/dx)^2 has exponent 2, which is the greatest power present.
4. Conclude the order and degree.
5. Final answer: the differential equation has order 1 and degree 2.

Example 2: Solving a Separable Differential Equation

Problem: Find the general solution of x^2(2y+1)(dy/dx) - 1 = 0.

1. Rewrite in the form dy/dx = f(x,y): dy/dx = 1/[x^2(2y+1)].
2. Separate the variables by moving all y-terms to one side and x-terms to the other: (2y+1)dy = (1/x^2)dx.
3. Integrate both sides: integral (2y+1)dy = integral x^(-2) dx, giving y^2+y = -1/x + C.
4. Combine the constant and the 1/x term onto one side for a clean implicit form: y^2+y+1/x = C.
5. Final answer: y^2 + y + 1/x = C is the general solution, where C is an arbitrary constant.

Example 3: Solving a Homogeneous Differential Equation

Problem: Find the general solution of dy/dx = (x+y)/(y-x).

1. Confirm the equation is homogeneous: every term in the numerator and denominator has degree 1, so the right side is a function of y/x only.
2. Substitute y=vx, so dy/dx = v + x(dv/dx): v + x(dv/dx) = (x+vx)/(vx-x) = (1+v)/(v-1).
3. Isolate x(dv/dx): x(dv/dx) = (1+v)/(v-1) - v = (-v^2+2v+1)/(v-1), which separates as (v-1)/(-v^2+2v+1) dv = (1/x) dx.
4. Integrate both sides: -(1/2) integral (-2v+2)/(-v^2+2v+1) dv = ln|x| + ln|c|, giving -(1/2)ln|-v^2+2v+1| = ln|cx|, i.e. -v^2+2v+1 = 1/(c^2 x^2) (up to sign).
5. Substitute back v=y/x and simplify: -y^2/x^2 + 2y/x + 1 = 1/(c^2x^2), which simplifies to x^2+2xy-y^2 = c. Final answer: x^2 + 2xy - y^2 = c.

Example 4: Population Growth (IVP)

Problem: In a culture, there are 100 bacteria initially and they double in 2 hours. Find the number of bacteria 7 hours later.

1. Set up the model: dN/dt = kN with N(0)=100 and N(2)=200, since the population is growing (k>0).
2. Separate and integrate: integral (1/N)dN = integral k dt gives ln|N| = kt + C.
3. Apply N(0)=100: C = ln(100), so N = 100 e^(kt). Apply N(2)=200: 200=100e^(2k), so e^(2k)=2, giving k = (1/2)ln2 = ln(sqrt2).
4. Write the particular solution: N(t) = 100 e^((ln sqrt2) t) = 100(sqrt2)^t.
5. Substitute t=7: N(7) = 100(2)^(7/2) is approximately 1131.4. Final answer: about 1131 bacteria after 7 hours.

Example 5: Radioactive Decay (Three-Part IVP)

Problem: A radioactive substance has mass 80 grams initially; after 10 hours only 50 grams remain. (a) Find the decay constant. (b) Find the mass after 24 hours. (c) After how many hours will only 20 grams remain?

1. Set up the model: dM/dt = -kM, so M(t)=ce^(-kt). Apply M(0)=80 to get c=80, so M(t)=80e^(-kt).
2. (a) Apply M(10)=50: 50=80e^(-10k), so k = -(1/10)ln(5/8) is approximately 0.047. This is the decay constant.
3. (b) Substitute t=24 into M(t)=80e^(-0.047t): M(24) = 80e^(-0.047(24)) is approximately 25.89, so about 26 grams remain after 24 hours.
4. (c) Set M(t)=20 and solve for t: 20=80e^(-0.047t) gives e^(-0.047t)=1/4, so t = ln(4)/0.047 is approximately 29.5.
5. Final answer: (a) k is approximately 0.047; (b) about 26 grams remain after 24 hours; (c) 20 grams remain after approximately 29.5 hours.

Example 6: Deflation Problem in Economics

Problem: An economy experiences deflation at 4% per year. If the current price of an item is Rs. 500, find its expected price after 3 years.

1. Set up the model: since prices are falling at a rate proportional to the current price, dP/dt = -0.04P.
2. Separate and integrate: this is the same growth/decay form as before, so P(t) = Ce^(-0.04t).
3. Apply the initial condition P(0)=500: C=500, so P(t) = 500e^(-0.04t).
4. Substitute t=3: P(3) = 500e^(-0.04(3)) = 500e^(-0.12).
5. Final answer: P(3) is approximately Rs. 443.46, so the expected price after 3 years is about Rs. 443.5.

Example 7: Newton's Law of Cooling

Problem: A cup of coffee is served at 90°C in a room at 25°C. After 10 minutes, it cools to 60°C. Find the temperature after 20 minutes.

1. Set up the model: dT/dt = -k(T-25), which is separable: (1/(T-25))dT = -k dt.
2. Integrate: ln|T-25| = -kt + ln|c|, which rearranges to T(t) = 25 + ce^(-kt).
3. Apply T(0)=90: c=65, so T(t)=25+65e^(-kt). Apply T(10)=60: 60=25+65e^(-10k), giving e^(10k)=13/7, so k is approximately 0.0619.
4. Write the particular solution: T(t) = 25 + 65e^(-0.0619t).
5. Substitute t=20: T(20) = 25 + 65e^(-0.0619(20)) is approximately 43.85. Final answer: the coffee is about 43.85°C after 20 minutes.

Example 8: RL Circuit Current

Problem: An RL circuit has an emf of 5 volts, resistance 50 ohms, an inductor of 1 henry, and no initial current. Find the current I(t) at any time t.

1. Set up the model using L(dI/dt)=V_S-RI with L=1, V_S=5, R=50: dI/dt = 5-50I.
2. Separate variables: (1/5) integral 1/(1-10I) dI = integral dt, which simplifies to -(1/50)ln|1-10I| = t + C.
3. Apply the initial condition I(0)=0: -(1/50)ln|1|=C, so C=0, giving -(1/50)ln|1-10I| = t.
4. Solve for I: ln|1-10I| = -50t, so |1-10I| = e^(-50t); checking the initial condition confirms the positive sign, so 1-10I = e^(-50t).
5. Final answer: I(t) = (1/10)(1 - e^(-50t)), which approaches the steady-state current 1/10 A as t grows large.

Short Questions & Answers

What is the order of d^2y/dx^2 + xy(dy/dx)^2 = 0?

2, since the highest derivative present is the second derivative d^2y/dx^2.

Is $dy/dx = x+y$ a homogeneous differential equation?

No, because $x+y$ cannot be written purely as a function of the ratio y/x.

What equation results from Kirchhoff's voltage law for an RL circuit?

$L(dI/dt) + RI = V_S$, i.e. $L(dI/dt) = V_S - RI$.

How is a homogeneous differential equation converted into a separable one?

By substituting $y=vx$, which turns dy/dx into $v+x(dv/dx)$.

What is the general solution form of $dP/dt=kP$?

$P(t) = P_0 e^{(kt)}$, where P_0 is the value of P at $t=0$.

In Newton's law of cooling, what does the constant T_s represent?

The constant temperature of the surrounding medium, which the body's temperature approaches over time.

What condition turns a general solution into a particular (unique) solution?

An initial value condition $y(x_0)=y_0$, which pins down the value of the arbitrary constant.

Long Questions & Answers

Explain how to recognize and solve a separable differential equation, including the role of the initial value condition.

How is a separable differential equation recognized?

It is a first-order, first-degree equation that can be written as $dy/dx=g(x)h(y)$, where the right-hand side factors into a function of x only multiplied by a function of y only; if it cannot be factored this way, it is not separable.

What are the steps to solve it?

Rewrite the equation as $(1/h(y))dy = g(x)dx$, separating all y-terms on one side and all x-terms on the other, then integrate both sides independently to obtain the general solution.

What is an initial value condition and why is it needed?

A condition of the form $y(x_0)=y_0$ specifies the value of y at a particular x; without it, the general solution contains an arbitrary constant C representing an entire family of curves.

How does applying the initial condition give a unique solution?

Substituting x_0 and y_0 into the general solution and solving for C fixes its exact numerical value, turning the family of solutions into one particular solution -- this combination of equation plus condition is called an Initial Value Problem (IVP).

Explain how first-order differential equations model real-life growth, decay, and cooling problems, with reference to specific examples.

How is population growth or decay modeled?

The rate of change of population is assumed proportional to the population itself, giving $dP/dt=kP$; solving this separable equation gives $P(t)=P_0e^{kt}$, where $k>0$ models growth (e.g. bacteria) and $k<0$ models decay (e.g. a shrinking population).

How is radioactive decay modeled and why is it always decay?

The same proportional model applies, $dM/dt=-kM$ with $k>0$ the decay constant, so the mass $M(t)=ce^{-kt}$ always decreases over time -- the negative sign is built into the model because radioactive substances only lose mass, never gain it.

How is inflation or deflation modeled in economics?

The price of a good is assumed to change at a rate proportional to its current price, $dP/dt=kP$, exactly like population growth; $k>0$ models inflation (rising prices) and $k<0$ models deflation (falling prices), with the solution $P(t)=P_0e^{kt}$.

How is Newton's law of cooling different from the growth/decay model?

Instead of the rate of change being proportional to the quantity itself, it is proportional to the difference between the body's temperature and the surrounding temperature: $dT/dt=k(T_s-T)$; solving this separable equation gives $T(t)=T_s+ce^{-kt}$, which approaches T_s as t grows rather than approaching zero or infinity.

Multiple Choice Questions (MCQs)

The order of a differential equation is:

A. The number of terms in the equation

B. The order of the highest derivative present

C. The degree of the independent variable

D. The number of arbitrary constants

Answer: B. The order of the highest derivative present

Explanation: The order is defined as the order of the highest-order derivative appearing in the differential equation.

A differential equation $dy/dx=g(x)h(y)$ is called:

A. Homogeneous

B. Linear

C. Separable

D. Exact

Answer: C. Separable

Explanation: This form, where the right side factors into a function of x times a function of y, defines a separable differential equation.

A homogeneous differential equation has the form:

A. $dy/dx = x+y$

B. $dy/dx = F(y/x)$

C. $dy/dx = x^2+y^2$

D. $dy/dx = k$

Answer: B. $dy/dx = F(y/x)$

Explanation: A homogeneous equation's right-hand side must be expressible purely as a function of the ratio y/x.

The substitution used to solve a homogeneous differential equation is:

A. $y = x + v$

B. $y = vx$

C. $x = vy^2$

D. $v = dy/dx$

Answer: B. $y = vx$

Explanation: Setting $y=vx$ converts a homogeneous equation into a separable equation in v and x.

The general solution of $dP/dt = kP$ is:

A. $P(t) = kt + C$

B. $P(t) = P_0 e^{(kt)}$

C. $P(t) = P_0 + kt$

D. $P(t) = k/P_0$

Answer: B. $P(t) = P_0 e^{(kt)}$

Explanation: Separating and integrating $dP/dt=kP$ gives the exponential solution $P(t)=P_0e^{(kt)}$.

In Newton's law of cooling $dT/dt = k(T_s - T)$, as t increases, T(t) approaches:

A. Zero

B. Infinity

C. The surrounding temperature T_s

D. The initial temperature T_0

Answer: C. The surrounding temperature T_s

Explanation: The solution $T(t)=T_s+ce^{(-kt)}$ has the exponential term decaying to zero, so T(t) approaches T_s.

A condition of the form $y(x_0)=y_0$ is called:

A. A general solution

B. A differential equation

C. An initial value condition

D. A homogeneous condition

Answer: C. An initial value condition

Explanation: This is the definition of an initial value (one-point boundary) condition.

For an RL circuit, Kirchhoff's voltage law gives the differential equation:

A. RC dq/dt = CV_S - q

B. L dI/dt = V_S - RI

C. dT/dt = k(T_s-T)

D. dP/dt = kP

Answer: B. L dI/dt = V_S - RI

Explanation: Summing the voltage across the inductor and resistor and equating to the source voltage gives L dI/dt + RI = V_S, i.e. L dI/dt = V_S - RI.

If k<0 in the model dP/dt = kP, the quantity P is:

A. Growing exponentially

B. Constant

C. Decaying exponentially

D. Oscillating

Answer: C. Decaying exponentially

Explanation: A negative proportionality constant makes the exponential solution P_0e^(kt) decrease over time, representing decay.

The degree of a differential equation is defined using:

A. The independent variable's highest power

B. The greatest exponent of the highest-order derivative

C. The number of derivatives present

D. The lowest-order derivative's exponent

Answer: B. The greatest exponent of the highest-order derivative

Explanation: Degree is the greatest exponent (power) to which the highest-order derivative is raised, once the equation is a polynomial in its derivatives.

Quick Revision Summary

- Differential equation: involves at least one derivative of a dependent variable w.r.t. an independent variable
- Order: order of the highest derivative present; Degree: greatest exponent of that highest-order derivative
- First-order models: population/price growth-decay dP/dt=kP; Newton's cooling dT/dt=k(T_s-T); RC/RL circuits from Kirchhoff's voltage law
- Separable: dy/dx=g(x)h(y); solve by (1/h(y))dy=g(x)dx then integrate both sides

- Homogeneous: $dy/dx=F(y/x)$; solve via substitution $y=vx$, reducing to a separable equation in v, x
- Initial value condition $y(x_0)=y_0$ plus a differential equation forms an Initial Value Problem (IVP), giving a unique particular solution
- Growth/decay solution: $P(t)=P_0e^{kt}$, $k>0$ growth, $k<0$ decay (population, inflation, radioactive decay, deflation)
- Cooling/heating solution: $T(t)=T_s+ce^{-kt}$, approaches surrounding temperature T_s as t grows
- RL circuit current $I(t)$ approaches the steady-state value V_S/R as t grows large
- Always identify which model (growth-decay vs cooling) a word problem matches before setting up the differential equation

Exam Tips

- Before solving, always check whether an equation is separable or homogeneous first -- trying to separate a homogeneous equation directly usually fails and wastes time
- For homogeneous equations, replacing x with 1 and y with v in $f(x,y)$ is a fast way to find $F(v)$ without algebraic manipulation of the full fraction
- For growth/decay/cooling word problems, write down which known values represent $t=0$ and which represent a later time before setting up equations -- this avoids mixing up which condition finds P_0/c and which finds k
- Always double check the sign of k: growth and inflation use $k>0$, decay and deflation use $k>0$ with a negative sign already built into the model ($dM/dt=-kM$) -- read the model carefully rather than assuming
- When solving for a constant like k from an exponential equation, take the natural log of both sides only after isolating the exponential term completely
- In circuit problems, remember the steady-state current V_S/R is the value $I(t)$ approaches as t becomes large -- useful as a quick sanity check on your final answer

Unit 5: Analytical Geometry

Analytical geometry, also called coordinate geometry, combines algebra and geometry to study shapes and lines using a coordinate system. First developed by Rene Descartes in the 17th century, it lets us describe geometric figures with equations and calculate distances, midpoints, slopes and intersections algebraically instead of purely by drawing. This unit reviews the standard forms of a straight line's equation, then develops the special lines associated with a triangle -- medians, altitudes, and right bisectors -- and proves that each of these three families is always concurrent.

The unit continues with the condition for three lines to be concurrent, the family of lines through the intersection of two given lines, the angle between two intersecting lines, and the area of a triangular region from its vertices. It closes with homogeneous second-degree equations, which represent a pair of straight lines through the origin, the angle between such a pair, and real-life applications of analytical geometry in aviation and astronomy.

Learning Objectives

- Write the equation of a straight line in general, slope-intercept, point-slope, two-point, intercept, and normal form
- Transform a general linear equation $ax+by+c=0$ into each of the standard forms
- Find the equations of medians, altitudes, and right bisectors of a triangle, and prove each family is concurrent
- Apply the determinant condition for three lines to be concurrent
- Find the family of lines through the intersection of two given lines, and select a member satisfying an extra condition
- Find the angle between two intersecting lines using their slopes, including the parallel and perpendicular special cases
- Find the area of a triangular region given its vertices, and use it to test collinearity

- Recognize a homogeneous second-degree equation as a pair of lines through the origin, and find the angle between them
- Apply analytical geometry to real-life problems such as distances between locations and flight paths

Key Concepts

5.1 Equation of a Straight Line: Standard Forms

A straight line can be written in several equivalent forms depending on the information given: general form ax+by+c=0; slope-intercept form y=mx+c (slope m, y-intercept c); point-slope form y-y_1=m(x-x_1) (through a known point with known slope); two-point form through two known points; intercept form x/a+y/b=1 (x-intercept a, y-intercept b); and normal form x cos(theta)+y sin(theta)=p (p the perpendicular distance from the origin, theta the angle the perpendicular makes with the positive x-axis).

Any general equation ax+by+c=0 can be transformed into every standard form: dividing by b gives slope-intercept form with m=-a/b; the point (-c/a, 0) lying on the line combined with the slope gives point-slope form; two convenient points (-c/a,0) and (0,-c/c) give two-point and intercept form; and dividing by +-sqrt(a^2+b^2), choosing the sign so the right side is positive, gives normal form.

5.2 Medians, Altitudes, and Right Bisectors of a Triangle

A median joins a vertex of a triangle to the midpoint of the opposite side; an altitude is drawn from a vertex perpendicular to the opposite side; a right (perpendicular) bisector is perpendicular to a side and passes through that side's midpoint. Each is found using the midpoint formula together with either the two-point form (medians) or the point-slope form with the negative-reciprocal slope rule m_1m_2=-1 (altitudes and right bisectors).

Three key theorems establish that these three special families of lines are always concurrent: the medians of any triangle meet at a single point (the centroid), the altitudes meet at a single point (the orthocenter), and the right bisectors meet at a single point (the circumcenter) -- each proved using the determinant condition for concurrency and a row-reduction argument that forces the determinant to zero.

5.3 Concurrency and Families of Lines

Three non-parallel lines $a_1x+b_1y+c_1=0$, $a_2x+b_2y+c_2=0$, $a_3x+b_3y+c_3=0$ are concurrent (pass through one common point) if and only if the 3x3 determinant of their coefficients equals zero. If concurrent, the common point can be found as the intersection of any two of the three lines.

For a non-zero real k, the equation $a_1x+b_1y+c_1+k(a_2x+b_2y+c_2)=0$ represents a family of lines all passing through the intersection point of lines l_1 and l_2, without needing to compute that intersection point directly. Choosing k to satisfy an extra condition (parallel to a given line, perpendicular to a given line, etc.) picks out one specific member of the family.

5.4 Angle Between Two Lines and Area of a Triangle

If m_1 and m_2 are the slopes of two non-vertical, non-perpendicular lines, the angle theta measured from the first line to the second satisfies $\tan(\theta) = (m_2-m_1)/(1+m_1m_2)$. Two important special cases follow directly: the lines are parallel iff $m_1=m_2$ ($\theta=0$), and perpendicular iff $1+m_1m_2=0$ ($\theta=90$ degrees, undefined tangent).

The area of a triangular region with vertices (x_1,y_1), (x_2,y_2), (x_3,y_3) is $\Delta = (1/2)|x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)|$, equivalently the absolute value of half a 3x3 determinant with a column of 1s; if the three points are collinear, this determinant (and hence the area) equals zero, giving a quick collinearity test.

5.5 Homogeneous Second-Degree Equations and Applications

Multiplying two linear equations $a_1x+b_1y+c_1=0$ and $a_2x+b_2y+c_2=0$ gives a joint second-degree equation representing both lines together. A homogeneous second-degree equation $ax^2+2hxy+by^2=0$ (with a, h, b not all zero) always represents a pair of straight lines through the origin: real and distinct if $h^2>ab$, real and coincident if $h^2=ab$, and imaginary (but still meeting at the real point (0,0)) if $h^2<ab$.

The angle between the pair of lines represented by $ax^2+2hxy+by^2=0$ is given by $\tan(\theta) = 2\sqrt{h^2-ab}/(a+b)$; the lines are coincident when $h^2=ab$, and perpendicular exactly when $a+b=0$ (the coefficients of x^2 and y^2 sum to zero). Analytical geometry techniques like these have direct real-life applications -- computing distances between objects (e.g. planets, using coordinates in astronomical units) and finding the slope and length of a flight path between two locations.

Important Definitions

What is the general form of a straight line's equation?

$ax+by+c=0$, where a, b, c are real numbers and a and b are not both zero.

What is a median of a triangle?

A line segment joining a vertex of a triangle to the midpoint of the opposite side.

What is an altitude of a triangle?

A line drawn from a vertex of a triangle perpendicular to the opposite side.

What is a right (perpendicular) bisector of a triangle's side?

A line perpendicular to that side which passes through its midpoint.

What does it mean for three lines to be concurrent?

All three lines pass through one single common point.

What is the condition for three lines $a_ix+b_iy+c_i=0$ ($i=1,2,3$) to be concurrent?

The determinant of their coefficients, $|a_i \; b_i \; c_i|$ (3x3), must equal zero.

What is a family of lines through the intersection of two lines?

The set of lines $a_1x+b_1y+c_1+k(a_2x+b_2y+c_2)=0$ for varying real k, all passing through the common intersection point of the two original lines.

What is the formula for the angle from a line with slope m_1 to a line with slope m_2?

$\tan(\theta) = (m_2-m_1)/(1+m_1m_2)$.

What is a homogeneous equation of degree n?

An equation $f(x,y)=0$ such that $f(kx,ky) = k^n f(x,y)$ for any real number k.

What does a homogeneous second-degree equation $ax^2+2hxy+by^2=0$ represent geometrically?

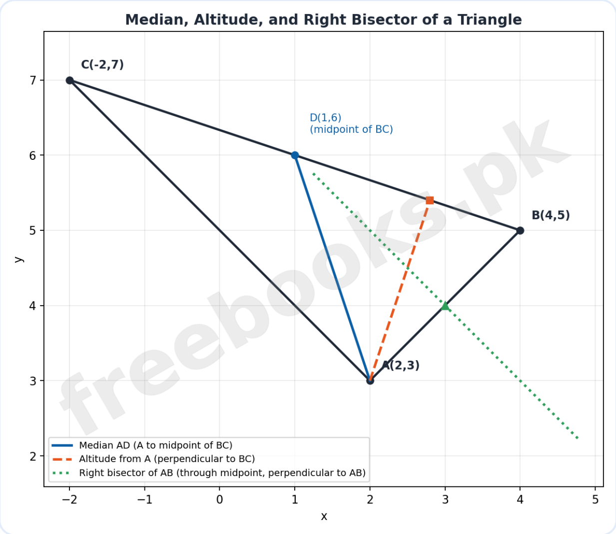
A pair of straight lines passing through the origin.

Key Facts and Relations

Topic	Key Fact / Relation
Standard line forms	General: $ax+by+c=0$; Slope-intercept: $y=mx+c$; Point-slope: $y-y_1=m(x-x_1)$; Intercept: $x/a+y/b=1$; Normal: $x\cos(\theta)+y\sin(\theta)=p$
Two-point form	$(y-y_1)/(y_1-y_2) = (x-x_1)/(x_1-x_2)$
Perpendicular slope condition	If m_1, m_2 are slopes of perpendicular lines, then $m_1 m_2 = -1$
Concurrency of three lines	Lines $a_ix+b_iy+c_i=0$ ($i=1,2,3$) are concurrent iff the 3x3 determinant of coefficients $[a_i, b_i, c_i]$ equals 0
Family of lines through an intersection	$a_1x+b_1y+c_1 + k(a_2x+b_2y+c_2) = 0$, for real k, non-zero
Angle between two lines (slopes)	$\tan(\theta) = (m_2-m_1)/(1+m_1m_2)$
Parallel / perpendicular conditions	Parallel: $m_1=m_2$ ($\theta=0$); Perpendicular: $1+m_1m_2=0$ ($\theta=90°$)
Area of a triangle from vertices	$\Delta = (1/2)\lvert x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)\rvert$; collinear iff $\Delta=0$
Homogeneous 2nd-degree equation (pair of lines through origin)	$ax^2+2hxy+by^2=0$; real & distinct if $h^2>ab$; coincident if $h^2=ab$; imaginary if $h^2<ab$
Angle between the pair of lines from $ax^2+2hxy+by^2=0$	$\tan(\theta) = 2\sqrt{h^2-ab} / (a+b)$; perpendicular iff $a+b=0$

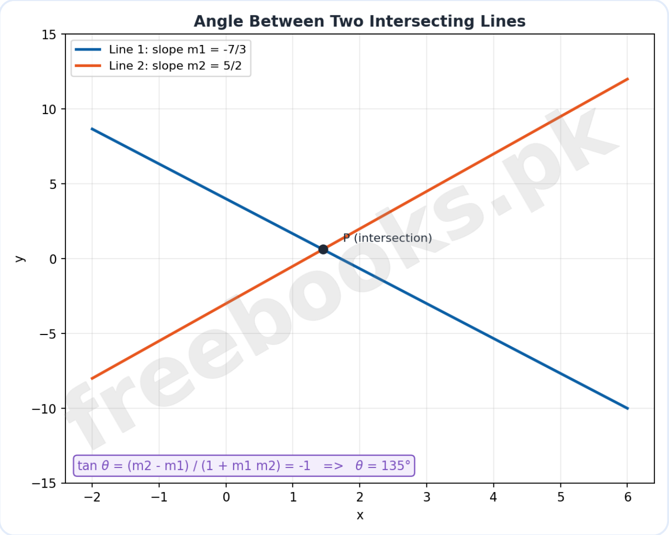
Diagrams

Median, Altitude, and Right Bisector of a Triangle: Triangle A(2,3), B(4,5), C(-2,7) with its median AD, altitude from A perpendicular to BC, and right bisector of side AB all drawn together to show how the three special lines differ



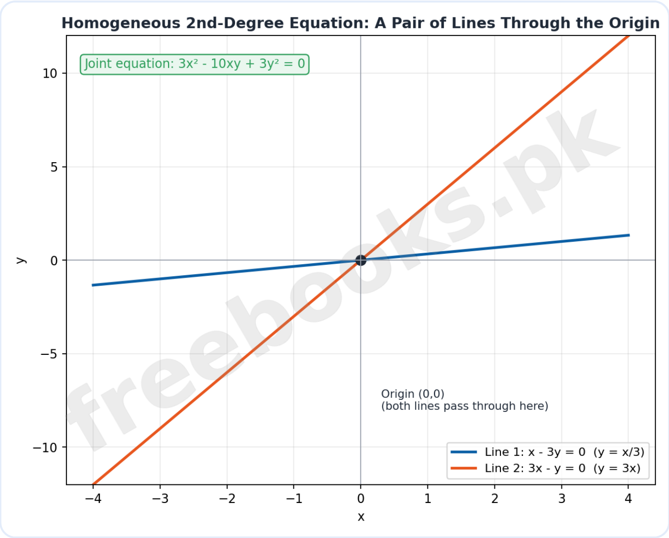
Median, Altitude, and Right Bisector of a Triangle

Angle Between Two Intersecting Lines: Two lines with slopes m1=-7/3 and m2=5/2 crossing at a point P, illustrating the angle theta from line 1 to line 2 given by tan(theta)=(m2-m1)/(1+m1m2), here giving theta=135°



Angle Between Two Intersecting Lines

Homogeneous 2nd-Degree Equation: A Pair of Lines Through the Origin: The joint equation $3x^2-10xy+3y^2=0$ factoring into two lines x-3y=0 and 3x-y=0, both passing through the origin, illustrating how a homogeneous second-degree equation represents a pair of straight lines



Homogeneous 2nd-Degree Equation: A Pair of Lines Through the Origin

Solved Examples

Example 1: Equation of a Line in Two-Point Form

Problem: Find an equation of the straight line passing through the points (-3,-2) and (5,4).

1. Apply the two-point form: (y-(-2))/(4-(-2)) = (x-(-3))/(5-(-3)).
2. Simplify the fractions: (y+2)/6 = (x+3)/8.
3. Cross-multiply: 8(y+2) = 6(x+3), giving 8y+16 = 6x+18.
4. Rearrange into general form: 6x-8y+2=0, which simplifies by dividing by 2 to 3x-4y+1=0.
5. Final answer: 3x - 4y + 1 = 0.

Example 2: Equations of the Medians of a Triangle

Problem: Triangle ABC has vertices A(2,3), B(4,5), C(-2,7). Find the equation of the median AD (from A to the midpoint of BC).

1. Find D, the midpoint of BC: D = ((4+(-2))/2, (5+7)/2) = (1, 6).

2. Use the two-point form through A(2,3) and D(1,6): (y-3)/(3-6) = (x-2)/(2-1).
3. Simplify: (y-3)/(-3) = (x-2)/1, so y-3 = -3(x-2).
4. Expand: y-3 = -3x+6.
5. Final answer: 3x + y - 9 = 0, the equation of median AD.

Example 3: Checking Concurrency of Three Lines

Problem: Check whether the lines 3x-2y-2=0, x+y-4=0, 2x-y-2=0 are concurrent, and if so find the point of concurrency.

1. Set up the 3x3 determinant of coefficients: |3 -2 -2; 1 1 -4; 2 -1 -2|.
2. Expand along the first row: 3(1(-2)-(-4)(-1)) - (-2)(1(-2)-(-4)(2)) + (-2)(1(-1)-1(2)) = 3(-6)+2(6)-2(-3).
3. Compute: 3(-6) = -18, 2(6) = 12, -2(-3) = 6, so the determinant is -18+12+6 = 0.
4. Since the determinant is zero, the three lines are concurrent. Find the point using any two lines: from x+y-4=0 and 2x-y-2=0, adding gives 3x-6=0, so x=2, then y=4-2=2.
5. Final answer: the lines are concurrent at the point (2, 2).

Example 4: Family of Lines Through an Intersection

Problem: Find the family of lines through the intersection of 2x-3y-14=0 and 2x+y-10=0, then find the member parallel to a line with slope -1/2.

1. Write the family: 2x-3y-14 + k(2x+y-10) = 0, which simplifies to (2+2k)x + (k-3)y + (-14-10k) = 0.
2. Find the slope of a general member: m = -(2+2k)/(k-3).
3. Set this equal to the required slope -1/2 and solve for k: -(2+2k)/(k-3) = -1/2 gives 4+4k = k-3, so 3k=-7, k=-7/3.
4. Substitute k=-7/3 back into the family equation and simplify: (2-14/3)x + (-7/3-3)y + (-14+70/3) = 0 becomes -8x-16y+28=0 after clearing fractions.

5. Final answer: dividing by -4, the required line is $2x + 4y - 7 = 0$.

Example 5: Angle Between Two Lines

Problem: Find the angle from the line with slope $-7/3$ to the line with slope $5/2$.

1. Identify $m_1=-7/3$ (first line) and $m_2=5/2$ (second line).
2. Apply the angle formula: $\tan(\theta) = (m_2-m_1)/(1+m_1m_2) = (5/2-(-7/3))/(1+(5/2)(-7/3))$.
3. Compute the numerator: $5/2+7/3 = 15/6+14/6 = 29/6$. Compute the denominator: $1-35/6 = 6/6-35/6 = -29/6$.
4. Divide: $\tan(\theta) = (29/6)/(-29/6) = -1$.
5. Final answer: $\tan(\theta) = -1$, so $\theta = 135°$.

Example 6: Area of a Triangle and Collinearity Test

Problem: Find the area of the region bounded by the triangle with vertices $A(1,2)$, $B(3,6)$, $C(5,10)$.

1. Set up the determinant formula: $\Delta = (1/2)|1\ 2\ 1;\ 3\ 6\ 1;\ 5\ 10\ 1|$.
2. Simplify using row operations $R2-R1$ and $R3-R1$: this gives $|1\ 2\ 1;\ 2\ 4\ 0;\ 4\ 8\ 0|$.
3. Expand along the third column (only the first row has a non-zero entry there): $\Delta = (1/2)[1 \times (2\times8-4\times4)] = (1/2)[16-16]$.
4. Compute: $(1/2)(0) = 0$.
5. Final answer: the area is 0, so the three points A, B, C are collinear (they lie on the same straight line).

Example 7: Finding Lines from a Homogeneous Second-Degree Equation

Problem: Find the equations of the lines represented by $3x^2-10xy+3y^2=0$, and the angle between them.

1. Factor the homogeneous equation by splitting the middle term: $3x^2-9xy-xy+3y^2=0$.
2. Group and factor: $3x(x-3y) - y(x-3y) = 0$, giving $(x-3y)(3x-y)=0$.
3. So the two lines are $x-3y=0$ (i.e. $y=x/3$) and $3x-y=0$ (i.e. $y=3x$), both passing through the origin.
4. Find the angle using $a=3$, $h=-5$, $b=3$: $\tan(\theta) = 2\sqrt{h^2-ab}/(a+b) = 2\sqrt{25-9}/6 = 2\sqrt{16}/6 = 8/6 = 4/3$.
5. Final answer: the lines are $x-3y=0$ and $3x-y=0$, and the angle between them satisfies $\tan(\theta)=4/3$, so theta is approximately 53.13°.

Example 8: Real-Life Application: Distance Between Two Locations

Problem: An aircraft flies from A(100,200) km to B(500,800) km. Find the slope of the track and the straight-line distance.

1. Find the slope using the two given points: $m = (800-200)/(500-100) = 600/400 = 1.5$.
2. Set up the distance formula between A and B: Distance = $\sqrt{(500-100)^2 + (800-200)^2}$.
3. Compute each term: $(400)^2 = 160000$ and $(600)^2 = 360000$, so the sum under the root is 520000.
4. Take the square root: $\sqrt{520000}$ is approximately 721.11.
5. Final answer: the slope of the flight track is 1.5, and the straight-line distance is approximately 721.11 km.

Short Questions & Answers

What is the slope-intercept form of a line's equation?

$y = mx + c$, where m is the slope and c is the y-intercept.

If two lines are parallel, what is true of their slopes?

Their slopes are equal: $m_1 = m_2$.

If two lines are perpendicular, what condition holds between their slopes?

The product of their slopes equals -1: $m_1 m_2 = -1$.

What point always lies on the family of lines $a_1x+b_1y+c_1+k(a_2x+b_2y+c_2)=0$?

The point of intersection of the two original lines l_1 and l_2, for every value of k.

How do you test whether three given points are collinear?

Compute the area of the triangle they form using the determinant formula; if the area is zero, the points are collinear.

What condition on a and b makes $ax^2+2hxy+by^2=0$ represent perpendicular lines?

$a + b = 0$, i.e. the coefficients of x^2 and y^2 sum to zero.

When does $ax^2+2hxy+by^2=0$ represent coincident (identical) lines?

When $h^2 = ab$.

Long Questions & Answers

Explain how the medians, altitudes, and right bisectors of a triangle are found, and why each family is always concurrent.

How is the equation of a median found?

The midpoint of the side opposite a vertex is found using the midpoint formula, and then the two-point form of a line's equation is applied through that vertex and the midpoint.

How is the equation of an altitude found?

The slope of the side opposite the vertex is computed, its negative reciprocal gives the altitude's slope (since altitude and side are perpendicular, $m_1m_2=-1$), and the point-slope form is applied through the vertex with that slope.

How is the equation of a right bisector found?

The midpoint of a side is found, the negative reciprocal of that side's slope gives the right bisector's slope, and the point-slope form is applied through the midpoint with that slope.

Why is each family (medians, altitudes, right bisectors) always concurrent?

In each case, writing the three line equations symbolically in terms of the triangle's vertices and forming the 3x3 determinant of their coefficients, adding the second and third rows to the first row always produces a row of zeros -- which forces the determinant to equal zero, proving concurrency regardless of the specific triangle.

Explain how a homogeneous second-degree equation represents a pair of lines through the origin, including how to find the lines and the angle between them.

What makes an equation 'homogeneous of degree 2'?

An equation $f(x,y)=0$ is homogeneous of degree n if $f(kx,ky)=k^n f(x,y)$ for any real k; a general second-degree homogeneous equation has the form $ax^2+2hxy+by^2=0$.

Why does $ax^2+2hxy+by^2=0$ always represent two lines through the origin?

Multiplying by b and completing the square in y gives $(by+hx)^2-x^2(h^2-ab)=0$, which factors as a difference of squares into two linear factors, each representing a line; since $x=0,y=0$ satisfies both factors, both lines pass through the origin.

How do you determine whether the two lines are real, coincident, or imaginary?

The discriminant-like quantity h^2-ab determines this: the lines are real and distinct if $h^2>ab$, real and coincident if $h^2=ab$, and imaginary (though they still meet at the real point (0,0)) if $h^2<ab$.

How is the angle between the pair of lines found?

Using the slopes derived from the two linear factors, the angle satisfies $\tan(\theta) = 2\sqrt{h^2-ab}/(a+b)$; the lines are perpendicular exactly when $a+b=0$, and coincident when $h^2=ab$ (giving $\theta=0$).

Multiple Choice Questions (MCQs)

The point-slope form of a line's equation is:

A. $y=mx+c$

B. $y-y_1=m(x-x_1)$

C. $x/a+y/b=1$

D. $ax+by+c=0$

Answer: B. $y-y_1=m(x-x_1)$

Explanation: The point-slope form uses a known point (x_1,y_1) and slope m: $y-y_1=m(x-x_1)$.

A median of a triangle joins a vertex to:

A. The opposite vertex

B. The midpoint of the opposite side

C. The foot of the altitude

D. The circumcenter

Answer: B. The midpoint of the opposite side

Explanation: By definition, a median connects a vertex to the midpoint of the side opposite that vertex.

Three lines are concurrent if and only if:

A. Their slopes are all equal

B. The determinant of their coefficients is zero

C. They are all parallel

D. Their y-intercepts are equal

Answer: B. The determinant of their coefficients is zero

Explanation: Concurrency of three lines is characterized by their coefficient determinant being equal to zero.

The family of lines $a_1x+b_1y+c_1+k(a_2x+b_2y+c_2)=0$ always passes through:

A. The origin

B. The intersection point of the two given lines

C. The midpoint of the two lines

D. A fixed point on the x-axis

Answer: B. The intersection point of the two given lines

Explanation: Every member of this family passes through the common intersection point of the two original lines, regardless of k.

If m_1 and m_2 are slopes of two perpendicular lines, then:

A. $m_1=m_2$

B. $m_1+m_2=0$

C. $m_1 m_2=-1$

D. $m_1 m_2=1$

Answer: C. $m_1 m_2=-1$

Explanation: Perpendicular lines satisfy the condition that the product of their slopes equals -1.

The area of a triangle with vertices (x1,y1), (x2,y2), (x3,y3) equals zero when:

A. The triangle is equilateral

B. The three points are collinear

C. The triangle is right-angled

D. The vertices are all positive

Answer: B. The three points are collinear

Explanation: A zero area indicates the three points do not form a proper triangle -- they lie on a single straight line.

A homogeneous second-degree equation $ax^2+2hxy+by^2=0$ represents:

A. A parabola

B. A pair of straight lines through the origin

C. A circle

D. A single line not through the origin

Answer: B. A pair of straight lines through the origin

Explanation: This is the defining geometric interpretation of a homogeneous second-degree equation in two variables.

The lines represented by $ax^2+2hxy+by^2=0$ are real and coincident when:

A. $h^2>ab$

B. $h^2=ab$

C. $h^2<ab$

D. $a+b=0$

Answer: B. $h^2=ab$

Explanation: Coincidence of the pair of lines occurs exactly when h^2 equals ab.

The lines represented by $ax^2+2hxy+by^2=0$ are perpendicular when:

A. $h=0$

B. $a=b$

C. $a+b=0$

D. $h^2=ab$

Answer: C. $a+b=0$

Explanation: Perpendicularity of the pair of lines requires the sum of the coefficients of x^2 and y^2 to be zero.

The angle theta from a line with slope m1 to a line with slope m2 satisfies:

A. tan(theta)=m1+m2

B. tan(theta)=(m2-m1)/(1+m1m2)

C. tan(theta)=m1m2

D. tan(theta)=(m1-m2)/(1-m1m2)

Answer: B. tan(theta)=(m2-m1)/(1+m1m2)

Explanation: This is the standard formula for the angle measured from the first line to the second, derived from the tangent subtraction identity.

Quick Revision Summary

- Standard line forms: general ax+by+c=0; slope-intercept y=mx+c; point-slope; two-point; intercept x/a+y/b=1; normal x cos(theta)+y sin(theta)=p

- Median: vertex to midpoint of opposite side; Altitude: vertex perpendicular to opposite side; Right bisector: perpendicular through a side's midpoint
- Medians, altitudes, and right bisectors of any triangle are each always concurrent (centroid, orthocenter, circumcenter)
- Three lines are concurrent iff the 3x3 determinant of their coefficients equals zero
- Family of lines through an intersection: $a_1x+b_1y+c_1+k(a_2x+b_2y+c_2)=0$, passes through the common point for every k
- Angle between two lines: $\tan(\theta)=(m_2-m_1)/(1+m_1m_2)$; parallel iff $m_1=m_2$; perpendicular iff $m_1m_2=-1$
- Area of a triangle from vertices: $\Delta=(1/2)|x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)|$; zero area means collinear points
- Homogeneous 2nd-degree equation $ax^2+2hxy+by^2=0$ represents a pair of lines through the origin
- Pair of lines: real & distinct if $h^2>ab$; coincident if $h^2=ab$; imaginary if $h^2<ab$; perpendicular iff $a+b=0$
- Angle between the pair of lines from a homogeneous equation: $\tan(\theta)=2\sqrt{h^2-ab}/(a+b)$

Exam Tips

- When a problem gives two points, default to the two-point form rather than trying to compute slope and intercept separately -- it is faster and less error-prone
- For altitude and right-bisector problems, always find the relevant side's slope first, then take its negative reciprocal -- writing this step out explicitly avoids sign errors
- To check concurrency, compute the 3x3 determinant carefully by expanding along the row or column with the most convenient (smallest/zero) entries
- For family-of-lines problems, always simplify the family equation into $ax+by+c=0$ form (collecting x, y, and constant terms) before applying the parallel/perpendicular slope condition

- Remember the angle formula is direction-sensitive: tan(theta)=(m2-m1)/(1+m1m2) measures the angle FROM line 1 TO line 2, so keep track of which slope is m1 and which is m2
- For homogeneous second-degree equations, always try factoring by splitting the middle term first before reaching for the general h, a, b formulas -- it's often faster and confirms the factorization is correct

Unit 6: Conic Section

A conic section is the curve obtained when a plane slices through a right circular cone. Depending on the angle of the cutting plane relative to the cone's axis, four distinct curves arise: a circle (plane perpendicular to the axis), an ellipse (plane inclined, cutting one nappe), a parabola (plane parallel to a generator, cutting one nappe), and a hyperbola (plane parallel to the axis, cutting both nappes). First studied geometrically by the Greek mathematicians Apollonius and Pappus, conics are studied here using analytic (coordinate) geometry, giving each curve a clean algebraic equation.

This unit develops each conic in turn: the circle (equation from centre and radius, general form, tangents, normals, and the position of a point relative to a circle), the parabola (standard equation from its focus-directrix definition, vertex, axis, latus rectum, tangents), the ellipse (standard equation from the constant-sum-of-distances definition, foci, eccentricity, tangents), and the hyperbola (standard equation from the constant-difference-of-distances definition, asymptotes, tangents). The unit closes with real-life applications -- satellite orbits, suspension bridge cables, parabolic reflectors, and planetary motion -- showing why conics matter far beyond the mathematics classroom.

Learning Objectives

- Derive and apply the standard and general equations of a circle, and find its centre and radius
- Find the equation of a circle satisfying given conditions (three points, tangency, diameter endpoints)
- Find equations of tangents and normals to a circle, and determine a point's position relative to a circle
- Derive the standard equation of a parabola from its focus-directrix definition and identify its elements
- Find equations of tangents and normals to a parabola, and the condition for a line to be tangent

- Derive the standard equation of an ellipse and identify its foci, vertices, eccentricity, and directrices
- Derive the standard equation of a hyperbola and identify its foci, vertices, asymptotes, and eccentricity
- Apply conic section concepts to real-life problems in astronomy, architecture, and optics

Key Concepts

6.1 The Circle: Equation, Tangents, and Normals

A circle is the set of points at a fixed distance (the radius r) from a fixed point (the centre). Its standard equation is $(x-h)^2+(y-k)^2=r^2$ for centre (h,k); expanding gives the general form $x^2+y^2+2gx+2fy+c=0$, which represents a real circle with centre $(-g,-f)$ and radius $\sqrt{g^2+f^2-c}$ whenever $g^2+f^2-c>0$. Three independent constants (g, f, c) mean a circle is fully determined by three conditions -- three non-collinear points, two points plus a line the centre lies on, or points plus a tangency condition.

The tangent to $x^2+y^2+2gx+2fy+c=0$ at a point (x_1,y_1) on the circle is $xx_1+yy_1+g(x+x_1)+f(y+y_1)+c=0$, obtained by a simple substitution rule (x^2 to xx_1, x to $(x+x_1)/2$, and so on) from the circle's own equation. A point lies outside, on, or inside a circle according as substituting its coordinates into the circle's general-form expression gives a value greater than, equal to, or less than zero; the length of the tangent from an external point is the square root of that same substituted value.

6.2 The Parabola: Standard Equation and Elements

A parabola is the locus of points equidistant from a fixed point (the focus) and a fixed line (the directrix). Placing the vertex at the origin with the focus at $(a,0)$ and directrix $x=-a$ gives the standard equation $y^2=4ax$; taking the directrix on the other side or perpendicular gives three more standard forms ($y^2=-4ax$, $x^2=4ay$, $x^2=-4ay$), each with focus, vertex, directrix, axis, and latus rectum (length $4a$) determined by the same pattern.

The tangent to $y^2=4ax$ at (x_1,y_1) is $yy_1=2a(x+x_1)$, found by implicit differentiation; a line $y=mx+c$ is tangent to the parabola exactly when $c=a/m$. The same substitution technique used for circles (x^2 to xx_1, $2x$ to $x+x_1$, etc.) extends to writing tangents for any conic once the general pattern is understood.

6.3 The Ellipse: Standard Equation and Key Features

An ellipse is the locus of points whose distances to two fixed foci sum to a constant 2a. Placing the foci on the x-axis at $(\pm c,0)$ gives the standard equation $x^2/a^2+y^2/b^2=1$, where $b^2=a^2-c^2$ and $a>b$; the major axis (length 2a) lies along the foci, the minor axis (length 2b) is perpendicular to it, and eccentricity $e=c/a$ satisfies $0<e<1$ ($e=0$ gives a circle).

Key features include vertices $(\pm a,0)$, co-vertices $(0,\pm b)$, directrices $x=\pm a/e$, and latus rectum length $2b^2/a$. The tangent to $x^2/a^2+y^2/b^2=1$ at (x_1,y_1) is $xx_1/a^2+yy_1/b^2=1$, and a line $y=mx+c$ is tangent to the ellipse exactly when $c^2=a^2m^2+b^2$.

6.4 The Hyperbola: Standard Equation and Key Features

A hyperbola is the locus of points whose distances to two fixed foci differ (in absolute value) by a constant 2a. With foci at $(\pm c,0)$, the standard equation is $x^2/a^2-y^2/b^2=1$, where now $c^2=a^2+b^2$ (so a and b can be in either order); the transverse axis (length 2a) joins the vertices, the conjugate axis (length 2b) is perpendicular, and eccentricity $e=c/a>1$.

A distinguishing feature not shared by the ellipse is the pair of asymptotes $y=\pm(b/a)x$, straight lines the hyperbola's two branches approach but never touch as x to $\pm\infty$; when $a=b$ the hyperbola is called rectangular and its asymptotes are perpendicular. The tangent to $x^2/a^2-y^2/b^2=1$ at (x_1,y_1) is $xx_1/a^2-yy_1/b^2=1$, and a line $y=mx+c$ is tangent exactly when $c^2=a^2m^2-b^2$.

6.5 Real-Life Applications of Conics

Parabolas describe suspension bridge cables (which hang in a parabolic curve under uniform load) and parabolic reflectors (headlights, satellite dishes, flashlights), which use the reflective property that rays from the focus reflect off the parabola parallel to its axis, and vice versa -- this is why bulbs are placed exactly at the focus.

Ellipses describe planetary and satellite orbits (Kepler's first law: planets orbit the sun in ellipses with the sun at one focus), with the apogee (farthest point) and perigee (closest point) related to the semi-major axis a and focal distance c by $A=a+c$ and $P=a-c$, giving eccentricity $e=(A-P)/(A+P)$. Hyperbolas appear in navigation systems like LORAN, where the constant time-difference between synchronized radio signals from two stations traces a hyperbolic path with the stations as foci.

Important Definitions

What is a conic section?

The curve formed by the intersection of a plane with a right circular cone; depending on the angle of the plane, the result is a circle, ellipse, parabola, or hyperbola.

What is the standard equation of a circle with centre (h,k) and radius r?

$(x-h)^2 + (y-k)^2 = r^2$.

What is a parabola?

The locus of points equidistant from a fixed point (the focus) and a fixed line (the directrix).

What is the latus rectum of a parabola?

The focal chord perpendicular to the axis; for $y^2=4ax$ it has length 4a and endpoints (a,2a) and (a,-2a).

What is an ellipse?

The locus of points for which the sum of the distances to two fixed points (the foci) is a constant, 2a.

What is the eccentricity of an ellipse?

$e = c/a$, where c is the distance from the centre to a focus; since $a>c>0$, we have $0<e<1$.

What is a hyperbola?

The locus of points for which the absolute value of the difference of the distances to two fixed points (the foci) is a constant, 2a.

What is an asymptote of a hyperbola?

A straight line that the hyperbola's branches approach arbitrarily closely as x tends to +-infinity, without ever touching it; for $x^2/a^2-y^2/b^2=1$ the asymptotes are $y=\pm(b/a)x$.

What is a rectangular hyperbola?

A hyperbola in which $a=b$, so its asymptotes are perpendicular to each other.

What is the tangent to a curve at a point, and the normal?

The tangent is the line touching the curve at exactly that point; the normal is the line through that point perpendicular to the tangent.

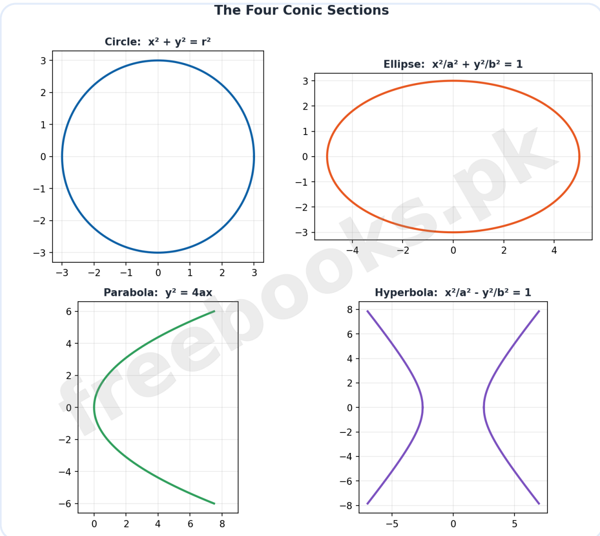
Key Facts and Relations

Topic	Key Fact / Relation
Circle: standard & general form	$(x-h)^2+(y-k)^2=r^2$; $x^2+y^2+2gx+2fy+c=0$, centre $(-g,-f)$, radius$=\sqrt{g^2+f^2-c}$
Tangent to a circle at (x1,y1)	$xx_1+yy_1+g(x+x_1)+f(y+y_1)+c=0$ (or $xx_1+yy_1=r^2$ for $x^2+y^2=r^2$)
Parabola standard forms	$y^2=4ax$ (focus $(a,0)$); $y^2=-4ax$; $x^2=4ay$; $x^2=-4ay$ -- all with vertex $(0,0)$, latus rectum length $4a$
Tangent to parabola $y^2=4ax$	$yy_1=2a(x+x_1)$ at (x_1,y_1); tangency condition for $y=mx+c$: $c=a/m$
Ellipse standard equation	$x^2/a^2+y^2/b^2=1$ $(a>b)$, $b^2=a^2-c^2$, foci $(\pm c,0)$, $e=c/a<1$
Ellipse tangent & tangency condition	$xx_1/a^2+yy_1/b^2=1$ at (x_1,y_1); $y=mx+c$ tangent iff $c^2=a^2m^2+b^2$
Hyperbola standard equation	$x^2/a^2-y^2/b^2=1$, $c^2=a^2+b^2$, foci $(\pm c,0)$, $e=c/a>1$
Hyperbola asymptotes & tangent	$y=\pm(b/a)x$; tangent at (x_1,y_1): $xx_1/a^2-yy_1/b^2=1$; $y=mx+c$ tangent iff $c^2=a^2m^2-b^2$
Latus rectum lengths	Parabola: $4a$; Ellipse and hyperbola: $2b^2/a$

Topic	Key Fact / Relation
Directrices	Ellipse & hyperbola (foci on x-axis): x = +-a/e

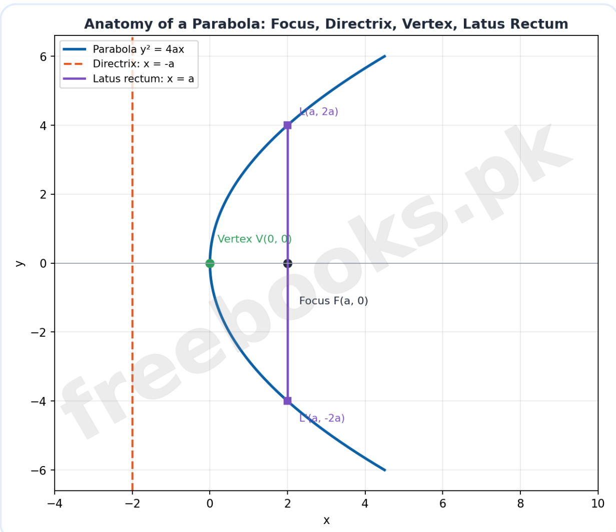
Diagrams

The Four Conic Sections: Circle, ellipse, parabola, and hyperbola shown together with their standard equations, illustrating how all four curves arise from slicing a cone at different angles



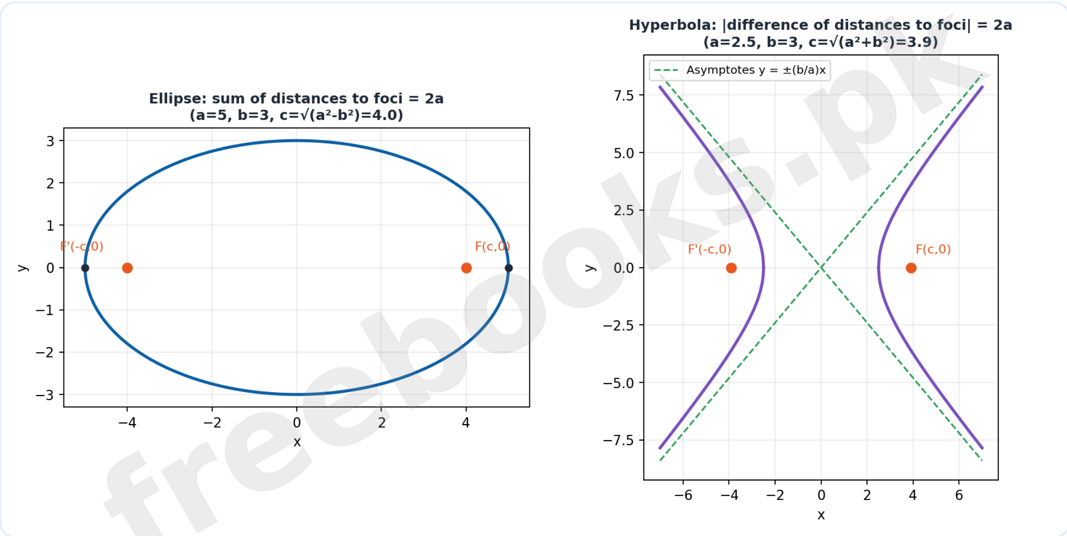
The Four Conic Sections

Anatomy of a Parabola: The parabola $y^2=4ax$ with its focus F(a,0), directrix x=-a, vertex V(0,0), and latus rectum endpoints L(a,2a) and L'(a,-2a) all labeled



Anatomy of a Parabola

Ellipse and Hyperbola: Foci and Asymptotes: Side-by-side comparison of an ellipse (sum of focal distances constant) and a hyperbola with its asymptotes (difference of focal distances constant), showing the parallel structure of their definitions



Ellipse and Hyperbola: Foci and Asymptotes

Solved Examples

Example 1: Equation of a Circle from Centre and Radius

Problem: Write an equation of the circle with centre (-2,1) and radius 5.

1. Use the standard form $(x-h)^2+(y-k)^2=r^2$ with $h=-2$, $k=1$, $r=5$.
2. Substitute: $(x-(-2))^2+(y-1)^2=5^2$, i.e. $(x+2)^2+(y-1)^2=25$.
3. Expand: $x^2+4x+4+y^2-2y+1=25$.
4. Simplify: $x^2+y^2+4x-2y-20=0$.
5. Final answer: $x^2 + y^2 + 4x - 2y - 20 = 0$.

Example 2: Centre and Radius from General Form

Problem: Equation $5x^2+5y^2+20x+25y+10=0$ represents a circle. Find its centre and radius.

1. Divide through by 5 to make the coefficients of x^2 and y^2 equal to 1: $x^2+y^2+4x+5y+2=0$.
2. Compare with $x^2+y^2+2gx+2fy+c=0$: $2g=4$ so $g=2$; $2f=5$ so $f=5/2$; $c=2$.
3. Find the centre: $(-g,-f) = (-2, -5/2)$.
4. Find the radius: $r = \sqrt{g^2+f^2-c} = \sqrt{4+25/4-2} = \sqrt{33/4}$.
5. Final answer: centre $(-2, -5/2)$, radius $= \sqrt{33}/2$.

Example 3: Tangent and Normal to a Circle

Problem: Find the equation of the tangent to the circle $x^2+y^2-4x+6y+12=0$ at the point on the circle whose ordinate (y-value) is -2.

1. Substitute $y=-2$ into the circle's equation to find x: $x^2+4-4x-12+12=0$, giving $x^2-4x+4=0$.
2. Factor: $(x-2)^2=0$, so $x=2$. The point of contact is $(2,-2)$.
3. Use the tangent rule $xx_1+yy_1+g(x+x_1)+f(y+y_1)+c=0$ with $g=-2$, $f=3$, $c=12$, $x_1=2$, $y_1=-2$: $2x-2y-2(x+2)+3(y-2)+12=0$.
4. Expand and simplify: $2x-2y-2x-4+3y-6+12=0$, which reduces to $y+2=0$.
5. Final answer: the tangent is $y = -2$ (a horizontal line), and by the perpendicularity of tangent and normal, the normal is $x = 2$.

Example 4: Elements of a Parabola from a General Equation

Problem: Analyze the parabola $x^2-4x-4y-2=0$: find its vertex, focus, and directrix.

1. Rewrite by completing the square in x: $x^2-4x=4y+2$, so $x^2-4x+4=4y+6$, giving $(x-2)^2=4(y+3/2)$.

2. Let $X=x-2$ and $Y=y+3/2$, so the equation becomes $X^2=4Y$, a standard parabola opening upward with $4a=4$, i.e. $a=1$.
3. The vertex is where $X=0$, $Y=0$, i.e. $x=2$, $y=-3/2$.
4. The focus is where $X=0$, $Y=a=1$, i.e. $x=2$, $y=-3/2+1=-1/2$.
5. Final answer: vertex $(2, -3/2)$, focus $(2, -1/2)$, and directrix $Y=-1$ i.e. $y = -3/2-1 = -5/2$.

Example 5: Tangent to a Parabola Using the Tangency Condition

Problem: Find the equation of the tangent to the parabola $y^2=12x$ which is parallel to the line $3x-y+4=0$.

1. Find the slope of the given line: $3x-y+4=0$ rearranges to $y=3x+4$, so its slope is $m=3$.
2. Since the required tangent is parallel, it has the same slope $m=3$. Compare $y^2=12x$ with $y^2=4ax$ to find a: $4a=12$, so $a=3$.
3. Apply the tangency condition $c=a/m$: $c = 3/3 = 1$.
4. Write the tangent line: $y = mx+c = 3x+1$.
5. Final answer: the tangent is $y = 3x + 1$, or equivalently $3x - y + 1 = 0$.

Example 6: Equation of an Ellipse from Foci and Major Axis

Problem: Find an equation of the ellipse with foci (2,1) and (2,7), and major axis of length 10.

1. Since both foci share the x-coordinate 2, the major axis is vertical. Major axis length $2a=10$, so $a=5$.
2. Find the centre as the midpoint of the foci: $((2+2)/2, (1+7)/2) = (2,4)$.
3. Find c as the distance from centre to a focus: $c=|4-1|=3$.
4. Use $b^2=a^2-c^2$: $b^2 = 25-9 = 16$.
5. Final answer: with vertical major axis, the equation is $(x-2)^2/16 + (y-4)^2/25 = 1$.

Example 7: Equation of a Hyperbola from Centre, Axis, and Eccentricity

Problem: A hyperbola has a horizontal transverse axis of length 12, centre at (-2,5), and eccentricity 5/3. Find its equation.

1. Since the transverse axis is horizontal, use the form $(x-h)^2/a^2 - (y-k)^2/b^2 = 1$, with centre $(h,k)=(-2,5)$.
2. Find a from the transverse axis length: $2a=12$, so $a=6$.
3. Find c using eccentricity $e=c/a$: $c/6 = 5/3$, so $c=10$.
4. Find b^2 using $c^2=a^2+b^2$: $100 = 36+b^2$, so $b^2=64$.
5. Final answer: $(x+2)^2/36 - (y-5)^2/64 = 1$.

Example 8: Real-Life Application: Parabolic Suspension Bridge Cable

Problem: A suspension bridge cable is parabolic. The roadway is 10 m below the lowest point of the cable, the span is 200 m, and the tops of the piers are 50 m above the roadway. Find the equation of the parabola, taking the vertex as the origin.

1. Take the vertex (lowest point of the cable) as the origin, with the parabola opening upward: $x^2 = 4ay$.
2. Half the span is 100 m, and the height of the pier top above the vertex is 50-10 = 40 m, so the point (100, 40) lies on the parabola.
3. Substitute into $x^2=4ay$: $(100)^2 = 4a(40)$, giving $10000 = 160a$.
4. Solve for 4a: $4a = 10000/40 = 250$.
5. Final answer: the equation of the parabola is $x^2 = 250y$.

Short Questions & Answers

What condition on g, f, c makes $x^2+y^2+2gx+2fy+c=0$ a real circle?

$g^2 + f^2 - c > 0$ (if it equals zero, the equation represents a single point; if negative, no real locus).

What are the coordinates of the vertex of the parabola $y^2=4ax$?

(0, 0), the origin, which is the midpoint of the perpendicular from the focus to the directrix.

For an ellipse $x^2/a^2+y^2/b^2=1$ with $a>b$, what is the relationship between a, b, and c?

$b^2 = a^2 - c^2$, i.e. $c = \sqrt{a^2-b^2}$.

Why can eccentricity of a hyperbola never be less than 1?

Because $c^2=a^2+b^2$ makes c always greater than a (since $b^2>0$), so $e=c/a$ is always greater than 1.

What happens to an ellipse's equation when its two foci coincide (c=0)?

It becomes a circle: $a=b$ and the equation reduces to $x^2+y^2=a^2$, with eccentricity $e=0$.

What is the reflective property of a parabola used for?

Rays from the focus reflect off the parabola parallel to its axis (or vice versa) -- used in flashlights, headlights, and satellite dishes to place the bulb/receiver exactly at the focus.

How is the length of the tangent from an external point P(x1,y1) to a circle found?

Substitute (x1,y1) into the circle's general-form left-hand side (with right side zero) and take the square root of the result.

Long Questions & Answers

Explain how the standard equations of the ellipse and hyperbola are derived from their focus-based definitions, and how their key features compare.

How is the ellipse's definition turned into an equation?

Starting from $|PF'|+|PF|=2a$ with foci at $(\pm c,0)$, the distance formula gives two square-root terms; isolating and squaring twice (to eliminate both radicals) leads to $(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2)$, which becomes $x^2/a^2+y^2/b^2=1$ after substituting $b^2=a^2-c^2$ and dividing through by a^2b^2.

How does the hyperbola's derivation differ?

The hyperbola uses the same distance-formula and double-squaring approach, but starts from $||PF'|-|PF||=2a$ instead of a sum; the key sign difference is that $a<c$ for a hyperbola (rather than $a>c$ for an ellipse), so the constant is defined as $-b^2=a^2-c^2$ instead of $b^2=a^2-c^2$, flipping a sign and producing $x^2/a^2-y^2/b^2=1$ with $c^2=a^2+b^2$ instead of $c^2=a^2-b^2$.

How do their eccentricities compare?

For an ellipse, $a>c$ always, so $e=c/a$ is between 0 and 1; for a hyperbola, $c>a$ always (since $c^2=a^2+b^2>a^2$), so $e=c/a$ is always greater than 1 -- this single number instantly tells you which conic you're looking at.

What feature does the hyperbola have that the ellipse does not?

Asymptotes: straight lines $y=\pm(b/a)x$ that the hyperbola's branches approach but never touch as $x\to\pm\infty$; an ellipse is a bounded, closed curve with no such lines, since both x and y stay within finite ranges ($|x|\leq a$, $|y|\leq b$).

Explain the general method for finding the tangent to a conic at a given point, and how it applies across the circle, parabola, ellipse, and hyperbola.

What is the general strategy for finding a tangent equation?

Differentiate the conic's equation implicitly with respect to x to find dy/dx as the slope of the tangent at any point, evaluate this slope at the given point (x_1,y_1), then use the point-slope form of a line and simplify using the fact that (x_1,y_1) satisfies the original conic's equation.

What is the substitution shortcut, and why does it work?

Once the tangent formula is derived once for a conic (e.g. $xx_1+yy_1=r^2$ for a circle $x^2+y^2=r^2$), the same result can be obtained directly from the conic's equation by substituting $x^2\to xx_1$, $y^2\to yy_1$, $x\to(x+x_1)/2$, $y\to(y+y_1)/2$, and leaving constants unchanged; this shortcut works because it exactly reproduces the algebra of the differentiation-based derivation.

How does the tangent equation differ between the ellipse and hyperbola?

Because the ellipse's equation has a plus sign between the x^2/a^2 and y^2/b^2 terms while the hyperbola's has a minus sign, their tangent equations mirror this: $xx_1/a^2+yy_1/b^2=1$ for the ellipse versus $xx_1/a^2-yy_1/b^2=1$ for the hyperbola -- the same substitution pattern, just preserving the sign structure of the original equation.

What is the tangency condition for a general line y=mx+c, and how does it differ across conics?

It comes from requiring the quadratic formed by substituting the line into the conic's equation to have equal roots (discriminant zero); this gives $c=a/m$ for the parabola $y^2=4ax$, $c^2=a^2m^2+b^2$ for the ellipse, and $c^2=a^2m^2-b^2$ for the hyperbola -- again the sign pattern of each conic's own equation carries through into its tangency condition.

Multiple Choice Questions (MCQs)

The general equation $x^2+y^2+2gx+2fy+c=0$ represents a real circle when:

A. $g^2+f^2-c<0$
B. $g^2+f^2-c=0$
C. $g^2+f^2-c>0$
D. $g=f=0$
Answer: C. $g^2+f^2-c>0$
Explanation: A real circle (with positive radius) requires g^2+f^2-c to be strictly positive, since the radius is $\sqrt{g^2+f^2-c}$.

The tangent to the circle $x^2+y^2=r^2$ at (x1,y1) is:

A. $xx_1+yy_1=r$
B. $xx_1+yy_1=r^2$
C. $x+y=r^2$
D. $xy_1+yx_1=r^2$
Answer: B. $xx_1+yy_1=r^2$
Explanation: The standard tangent formula for a circle centred at the origin is $xx_1+yy_1=r^2$.

The standard equation of a parabola with vertex at the origin and focus (a,0) is:

A. $x^2=4ay$

B. $y^2=4ax$

C. $x^2/a^2+y^2/b^2=1$

D. $x^2-y^2=a^2$

Answer: B. $y^2=4ax$

Explanation: This is the standard first form of the parabola, with axis along the x-axis and directrix $x=-a$.

The length of the latus rectum of the parabola $y^2=4ax$ is:

A. a

B. 2a

C. 4a

D. a^2

Answer: C. 4a

Explanation: The latus rectum, the focal chord perpendicular to the axis, has length 4a for this standard form.

For the ellipse $x^2/a^2+y^2/b^2=1$ with $a>b$, the relationship between a, b, and c is:

A. $c^2=a^2+b^2$

B. $c^2=a^2-b^2$

C. $c^2=b^2-a^2$

D. $c=ab$

Answer: B. $c^2=a^2-b^2$

Explanation: For an ellipse, $c^2=a^2-b^2$ (equivalently $b^2=a^2-c^2$), reflecting that $a>c$ always.

The eccentricity of an ellipse always satisfies:

A. $e=0$

B. $e=1$

C. $0<e<1$

D. $e>1$

Answer: C. $0<e<1$

Explanation: Since $a>c>0$ for an ellipse, eccentricity $e=c/a$ is strictly between 0 and 1.

For the hyperbola $x^2/a^2-y^2/b^2=1$, the relationship between a, b, and c is:

A. $c^2=a^2-b^2$

B. $c^2=a^2+b^2$

C. $c=a-b$

D. $c^2=ab$

Answer: B. $c^2=a^2+b^2$

Explanation: Unlike the ellipse, a hyperbola satisfies $c^2=a^2+b^2$, making c always greater than a.

The asymptotes of the hyperbola $x^2/a^2-y^2/b^2=1$ are:

A. $y=\pm(a/b)x$

B. $y=\pm(b/a)x$

C. $y=\pm ab\ x$

D. $x=\pm(b/a)y$

Answer: B. $y=\pm(b/a)x$

Explanation: The hyperbola's branches approach the lines $y=\pm(b/a)x$ as $x\to\pm\infty$.

A hyperbola with a=b is called:

A. A parabola

B. A degenerate conic

C. A rectangular hyperbola

D. An ellipse

Answer: C. A rectangular hyperbola

Explanation: When a=b, the hyperbola $x^2-y^2=a^2$ is called rectangular, and its asymptotes are perpendicular.

A point P(x1,y1) lies inside the circle $x^2+y^2+2gx+2fy+c=0$ when substituting its coordinates into the left side gives a value that is:

A. Positive

B. Zero

C. Negative

D. Undefined

Answer: C. Negative

Explanation: A negative value on substitution means the point is inside the circle (closer to the centre than the radius).

Quick Revision Summary

- Conic sections: circle, ellipse, parabola, hyperbola -- all arise from slicing a right circular cone at different angles
- Circle: $(x-h)^2+(y-k)^2=r^2$; general form $x^2+y^2+2gx+2fy+c=0$, centre $(-g,-f)$, radius $\sqrt{g^2+f^2-c}$
- Circle tangent at (x_1,y_1): $xx_1+yy_1+g(x+x_1)+f(y+y_1)+c=0$ (substitution rule from the circle's own equation)
- Parabola $y^2=4ax$: focus $(a,0)$, directrix $x=-a$, vertex $(0,0)$, latus rectum length $4a$; tangent condition $c=a/m$
- Ellipse $x^2/a^2+y^2/b^2=1$ $(a>b)$: foci $(\pm c,0)$, $b^2=a^2-c^2$, eccentricity $e=c/a$, $0<e<1$; tangent condition $c^2=a^2m^2+b^2$
- Hyperbola $x^2/a^2-y^2/b^2=1$: foci $(\pm c,0)$, $c^2=a^2+b^2$, eccentricity $e=c/a>1$; asymptotes $y=\pm(b/a)x$; tangent condition $c^2=a^2m^2-b^2$
- Rectangular hyperbola: $a=b$, asymptotes perpendicular
- Real-life applications: parabolic reflectors and bridge cables, elliptical planetary/satellite orbits, hyperbolic LORAN navigation
- Apogee/perigee of an elliptical orbit: $A=a+c$, $P=a-c$, giving eccentricity $e=(A-P)/(A+P)$
- General tangent-finding method: implicit differentiation for dy/dx, or the substitution shortcut ($x^2\to xx_1$, $x\to(x+x_1)/2$, etc.) applied directly to the conic's own equation

Exam Tips

- Before applying any formula, always identify which conic you're dealing with and whether its axis/foci are horizontal or vertical -- this determines whether a^2 or b^2 sits under the x^2 or y^2 term
- For circles, always divide through first so the coefficients of x^2 and y^2 are exactly 1 before reading off g, f, c -- skipping this step is the most common source of centre/radius errors
- Remember the sign difference between ellipse ($b^2=a^2-c^2$, $e<1$) and hyperbola ($c^2=a^2+b^2$, $e>1$) -- confusing these two relationships is the most common mistake between the two curves

- When completing the square to find a parabola/ellipse/hyperbola's centre or vertex from a general equation, work through the x-terms and y-terms as two completely separate completions before combining constants
- For tangency conditions ($c=a/m$ for parabola, $c^2=a^2m^2\pm b^2$ for ellipse/hyperbola), double check which conic's condition you're using -- the ellipse uses a plus sign, the hyperbola a minus sign
- In real-life application problems, always set up a clear coordinate system first (where is the origin, which way do the axes point) before translating the word problem into an equation -- most errors come from an unclear or inconsistent coordinate setup

Unit 7: Kinematics of Motion in a Straight Line

Kinematics is the branch of mechanics that describes how objects move without asking what causes that motion. This unit studies rectilinear motion -- motion of a particle along a straight line -- through the key quantities displacement, velocity, and acceleration, distinguishing scalar quantities (distance, speed) that have magnitude only from vector quantities (displacement, velocity, acceleration) that also have direction. Displacement-time and velocity-time graphs give a visual language for motion: the slope of a displacement-time graph is velocity, the slope of a velocity-time graph is acceleration, and the area under a velocity-time graph is displacement.

Using differentiation and integration with respect to time, this unit derives the three equations of motion for constant acceleration -- $v=u+at$, $x=ut+(1/2)at^2$, and $2ax=v^2-u^2$ -- which are the primary tools for solving motion problems. The unit closes by applying these ideas to vertical motion under gravity (free fall and upward projectiles, using $a=g$ or $a=-g$), and to problems where acceleration itself varies with time, solved by direct integration.

Learning Objectives

- Distinguish between scalar quantities (distance, speed) and vector quantities (displacement, velocity, acceleration)
- Calculate average velocity, average speed, instantaneous velocity, and instantaneous speed
- Find velocity and acceleration from a displacement function using differentiation
- Interpret displacement-time graphs: identify rest, uniform forward motion, and uniform backward motion from slope
- Interpret velocity-time graphs: find acceleration from slope and displacement from area under the graph
- Derive and apply the three equations of motion for constant acceleration

- Solve free-fall and vertical projectile motion problems using the equations of motion with a=g or a=-g
- Find velocity and displacement by integrating a variable (time-dependent) acceleration function

Key Concepts

7.1 Scalars, Vectors, and Rectilinear Motion

Rectilinear motion is motion of a particle along a straight line. Scalar quantities (distance, speed) are described by magnitude alone; vector quantities (displacement, velocity, acceleration) require both magnitude and direction, though in rectilinear motion direction is simplified to a plus or minus sign rather than a full vector arrow, with one direction along the line chosen as positive and the opposite as negative.

Average velocity is total displacement divided by total time (v_avg=Delta x/Delta t), while average speed is total distance divided by total time -- these can differ significantly, as when a body returns to its starting point (displacement zero, but distance and average speed nonzero). Instantaneous velocity is v(t)=dx/dt, the derivative of displacement with respect to time, and instantaneous speed is its absolute value |v(t)|.

7.2 Acceleration and Displacement-Time Graphs

Acceleration is the rate of change of velocity with respect to time: a(t)=dv/dt=d^2x/dt^2. When velocity is known as a function of displacement rather than time, the chain rule gives the useful alternate form a=v(dv/dx).

On a displacement-time graph (x on the vertical axis, t on the horizontal), the slope dx/dt gives velocity at each instant: a horizontal segment (zero slope) means the particle is at rest, a rising straight segment (constant positive slope) means uniform forward motion, and a falling straight segment (constant negative slope) means uniform backward motion.

7.3 Velocity-Time Graphs

On a velocity-time graph (v on the vertical axis, t on the horizontal), the slope dv/dt gives acceleration at each instant: a horizontal segment means uniform (constant) velocity with zero acceleration, a rising straight segment means uniform positive acceleration, and a falling straight segment means uniform negative acceleration (retardation/deceleration). A curved (non-straight) velocity-time graph indicates variable acceleration.

The area under a velocity-time graph between $t=t_1$ and $t=t_2$, above the time axis, equals the displacement over that interval, since integral from t_1 to t_2 of $v(t)dt$ = integral from t_1 to t_2 of $(dx/dt)dt = x(t_2)-x(t_1)$. If the graph dips below the time axis, the area above the axis (A_1) and the area below (A_2) combine differently for displacement (A_1-A_2) and total distance (A_1+A_2), since distance always adds magnitudes regardless of direction.

7.4 Equations of Motion for Constant Acceleration

Starting from a constant acceleration $a=dv/dt$, integrating once with the initial condition $v=u$ at $t=0$ gives $v=u+at$; integrating $v=dx/dt$ with $x=0$ at $t=0$ gives $x=ut+(1/2)at^2$. Eliminating t between these two (by substituting $t=(v-u)/a$) gives the third equation, $2ax=v^2-u^2$, useful whenever time is not given or not needed.

These three equations of motion -- $v=u+at$, $x=ut+(1/2)at^2$, $2ax=v^2-u^2$ -- are the standard toolkit for any constant-acceleration problem: given any three of the five quantities (u, v, a, t, x), the equations solve for the remaining two.

7.5 Vertical Motion Under Gravity and Variable Acceleration

Free fall and vertical projectile motion are constant-acceleration motion with a replaced by g (approximately 9.8 m/s^2) for downward motion, or $-g$ for upward motion, since gravity always decelerates an object thrown upward until its velocity becomes zero at the highest point, then accelerates it back downward. The same three equations of motion apply directly, simply substituting g or $-g$ for a.

When acceleration is not constant but given as a function of time, $a(t)$, the equations of motion no longer apply directly; instead, velocity and displacement are found by direct integration: $v(t)$ = integral $a(t)dt$ (using the initial velocity to find the constant of integration), then $x(t)$ = integral $v(t)dt$ (using the initial position to find the next constant).

Important Definitions

What is rectilinear motion?

The motion of a particle along a straight line.

What is the difference between a scalar and a vector quantity?

A scalar has magnitude only (e.g. distance, speed); a vector has both magnitude and direction (e.g. displacement, velocity, acceleration).

What is average velocity?

Total displacement divided by total time taken: v_{avg} = Delta x / Delta t.

What is instantaneous velocity?

The rate of change of displacement with respect to time at a specific instant: $v(t) = dx/dt$.

What is acceleration?

The rate of change of velocity with respect to time: $a(t) = dv/dt = d^2x/dt^2$.

What does the slope of a displacement-time graph represent?

The velocity of the particle at that instant.

What does the slope of a velocity-time graph represent?

The acceleration of the particle at that instant.

What does the area under a velocity-time graph represent?

The displacement of the particle over that time interval.

What are the three equations of motion for constant acceleration?

$v=u+at$; $x=ut+(1/2)at^2$; $2ax=v^2-u^2$, where u is initial velocity, v is final velocity, a is acceleration, t is time, and x is displacement.

What is free fall?

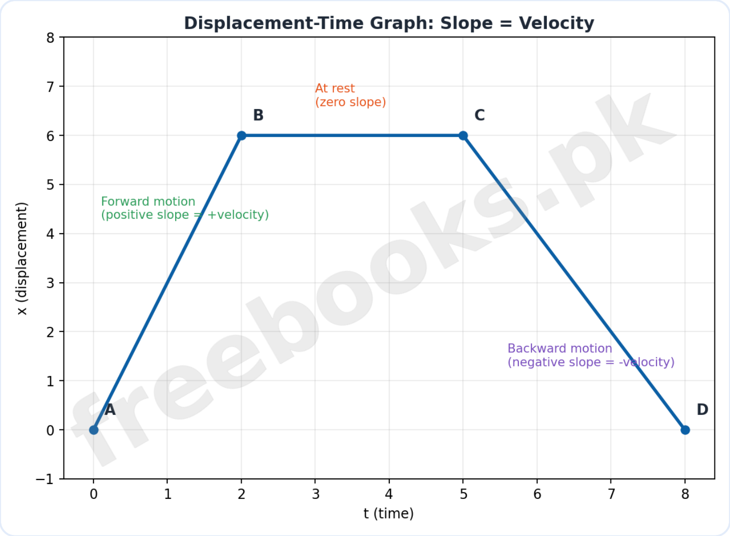
Motion of a particle under the influence of gravity alone (air resistance neglected), with acceleration $g \approx 9.8$ m/s² directed downward.

Key Facts and Relations

Topic	Key Fact / Relation
Average velocity & speed	v_avg = Total displacement / Total time; Average speed = Total distance / Total time
Instantaneous velocity & speed	$v(t) = dx/dt$; Speed = $\lvert v(t)\rvert$
Acceleration	$a(t) = dv/dt = d^2x/dt^2$; also $a = v(dv/dx)$ when v is a function of x
Displacement-time graph	Slope = dx/dt = velocity
Velocity-time graph	Slope = dv/dt = acceleration; Area under graph (t1 to t2) = displacement = x(t2)-x(t1)
Distance vs displacement from a v-t graph	If A1 = area above time axis, A2 = area below: Displacement = A1-A2; Distance = A1+A2
Equations of motion (constant acceleration)	$v = u + at$; $x = ut + (1/2)at^2$; $2ax = v^2 - u^2$
Vertical motion under gravity	Use a=g (downward motion) or a=-g (upward motion), $g \approx 9.8\ m/s^2$, in the same equations of motion
Variable acceleration a(t)	v(t) = integral a(t)dt + C (find C from initial velocity); x(t) = integral v(t)dt + C1 (find C1 from initial position)
Area of a trapezium (for v-t graph displacement)	Area = (1/2)(a+b)h, where a, b are the parallel sides and h is the distance between them

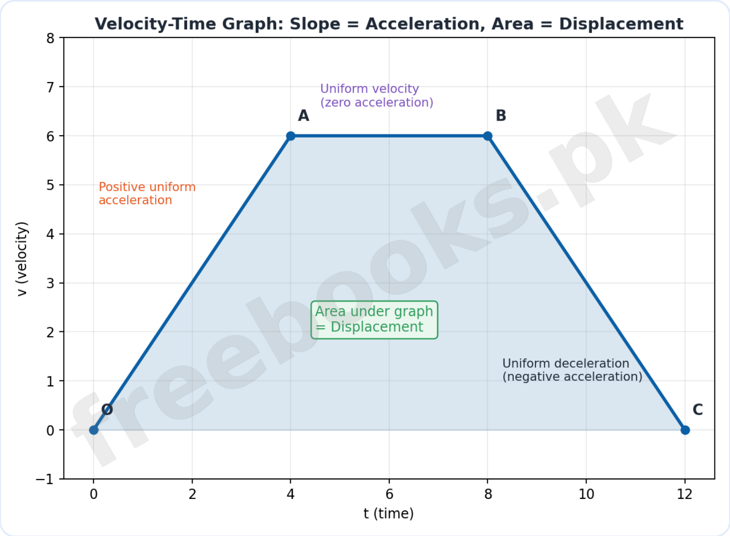
Diagrams

Displacement-Time Graph: Slope = Velocity: A displacement-time graph with four labeled points A, B, C, D showing forward motion (positive slope), rest (zero slope), and backward motion (negative slope) as three distinct straight-line segments



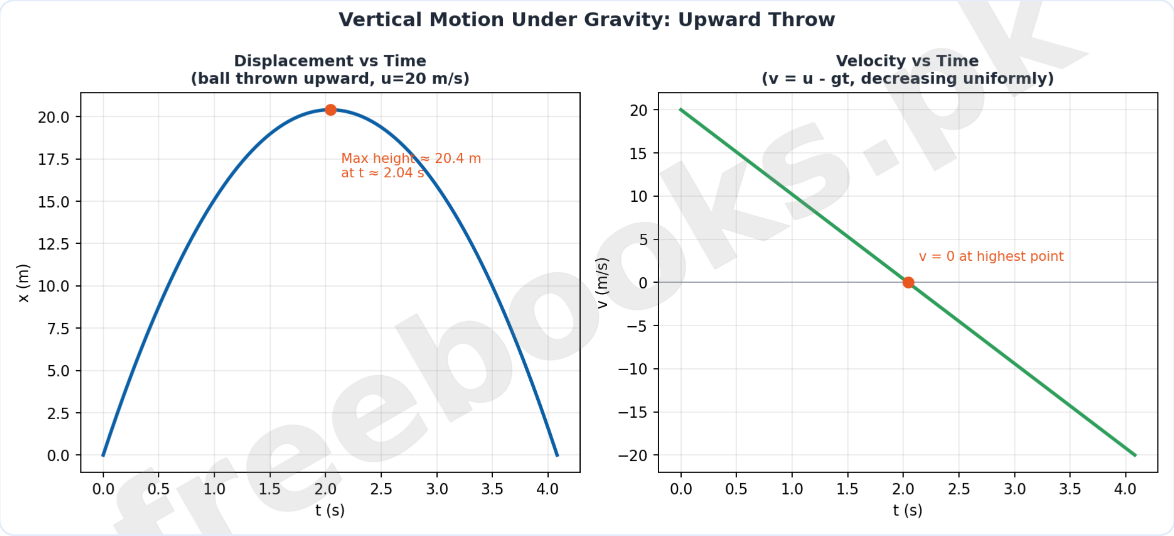
Displacement-Time Graph: Slope = Velocity

Velocity-Time Graph: Slope = Acceleration, Area = Displacement: A trapezoidal velocity-time graph showing uniform positive acceleration, uniform velocity (zero acceleration), and uniform deceleration, with the shaded area under the graph representing total displacement



Velocity-Time Graph: Slope = Acceleration, Area = Displacement

Vertical Motion Under Gravity: Upward Throw: Side-by-side displacement-vs-time and velocity-vs-time graphs for a ball thrown upward at 20 m/s, showing the parabolic displacement curve reaching maximum height where velocity crosses zero



Vertical Motion Under Gravity: Upward Throw

Solved Examples

Example 1: Average Speed and Velocity

Problem: A man walks 100 meters east in 2 minutes, then turns around and walks 60 meters west in 1 minute. Find his average speed and average velocity.

1. Find total distance travelled: 100+60 = 160 meters (distance always adds, regardless of direction).
2. Find total time: 2+1 = 3 minutes.
3. Compute average speed: Total distance/Total time = 160/3 ≈ 53.3 meters/min.
4. Taking east as positive and west as negative, find total displacement: 100+(-60) = 40 meters (eastward).
5. Compute average velocity: Total displacement/Total time = 40/3 ≈ 13.33 meters/min, eastward.

Example 2: Velocity and Speed from a Displacement Function

Problem: Let $x(t) = t^3-6t^2$ be the position of a particle (x in meters, t in seconds). Find the velocity and speed at any time t, and at t=1 second.

1. Differentiate x(t) with respect to t to find velocity: $v(t) = dx/dt = 3t^2-12t$.
2. Write the speed as the absolute value of velocity: Speed = $|v(t)| = |3t^2-12t|$.
3. Substitute t=1 into the velocity formula: $v(1) = 3(1)^2-12(1) = 3-12 = -9$ m/s.
4. Substitute t=1 into the speed formula: Speed = $|v(1)| = |-9| = 9$ m/s.
5. Final answer: at t=1 second, velocity is -9 m/s (moving in the negative direction) and speed is 9 m/s.

Example 3: Displacement from a Velocity-Time Graph (Trapezium Area)

Problem: A cyclist's velocity-time graph is a trapezium OABC where OC=12 minutes, AB=8 minutes (upper parallel side), and the vertical distance between OC and AB is 6 m/min. Find the total displacement over the 12 minutes.

1. Recognize that displacement equals the area under the velocity-time graph, and the graph forms a trapezium with parallel sides a=OC=12 and b=AB=8, height h=6.
2. Apply the trapezium area formula: Area = $(1/2)(a+b)h$.
3. Substitute the values: Area = $(1/2)(12+8)(6)$.
4. Compute: $(1/2)(20)(6) = 60$.
5. Final answer: the total displacement is 60 meters.

Example 4: Distance Using an Equation of Motion

Problem: A car accelerates from 10 m/s to 30 m/s in 4 seconds. Find the distance covered during this time.

1. First find the acceleration using $v=u+at$: $30 = 10 + a(4)$.
2. Solve for a: $4a = 20$, so $a = 5\ m/s^2$.
3. Use the equation $2ax = v^2-u^2$ to find distance x, since time is no longer needed: $2(5)x = 30^2-10^2$.
4. Compute the right side: $900-100 = 800$, so $10x = 800$.
5. Final answer: $x = 80$ meters.

Example 5: Velocity and Position from a Time-Dependent Acceleration

Problem: A particle starts from rest at $t=0$ and has acceleration $a(t)=t^2+\cos(t)$. Find its velocity and position at any time t.

1. Integrate acceleration to find velocity: $v(t)$ = integral $(t^2+\cos t)dt = (1/3)t^3+\sin t+C$.
2. Apply the initial condition $v=0$ at $t=0$: $0 = 0+0+C$, so $C=0$, giving $v(t) = (1/3)t^3+\sin t$.
3. Integrate velocity to find position: $x(t)$ = integral $[(1/3)t^3+\sin t]dt = (1/12)t^4-\cos t+C1$.
4. Apply the initial condition $x=0$ at $t=0$: $0 = 0-1+C1$, so $C1=1$.
5. Final answer: $v(t) = (1/3)t^3+\sin t$, and $x(t) = (1/12)t^4-\cos t+1$.

Example 6: Free Fall from a Height

Problem: A stone is dropped from the top of a tower 45 m high. Find the time taken to reach the ground and the final velocity on impact.

1. Identify the known values: $u=0$ (dropped, not thrown), $x=45$ m, $a=g=9.8\ m/s^2$.

2. Use $x=ut+(1/2)gt^2$ to find time: $45 = 0+(1/2)(9.8)t^2$.
3. Solve for t: $t^2 = 90/9.8$, so $t = \sqrt{90/9.8} \approx 3.03$ s.
4. Use $v=u+gt$ to find the final velocity: $v = 0+9.8(3.03)$.
5. Final answer: time to reach the ground is approximately 3.03 seconds, and the final velocity on impact is approximately 29.7 m/s.

Example 7: Maximum Height of a Vertical Throw

Problem: A ball is thrown vertically upward with a speed of 20 m/s. Find the maximum height attained and the time to reach it.

1. Identify the known values at the highest point: $u=20$ m/s, $v=0$ m/s (velocity is zero at the top), $a=g=-9.8$ m/s² (deceleration while rising).
2. Use $v=u+gt$ to find the time to reach maximum height: $0 = 20+(-9.8)t$.
3. Solve for t: $t = 20/9.8 \approx 2.04$ s.
4. Use $x=ut+(1/2)gt^2$ to find maximum height: $x = 20(2.04)+(1/2)(-9.8)(2.04)^2$.
5. Final answer: the time to reach maximum height is approximately 2.04 seconds, and the maximum height is approximately 20.4 meters.

Example 8: Constant Uniform Acceleration from Rest

Problem: A car starts from rest and accelerates uniformly at 3 m/s². Find the velocity of the car after 5 seconds.

1. Identify the known values: $u=0$ (starts from rest), $a=3$ m/s², $t=5$ s.
2. Apply the equation of motion $v=u+at$.
3. Substitute the values: $v = 0+3(5)$.
4. Compute: $v = 15$.
5. Final answer: the velocity of the car after 5 seconds is 15 m/s.

Short Questions & Answers

What is the difference between distance and displacement?

Distance is the total length of the path travelled (a scalar); displacement is the change in position from start to end, including direction (a vector).

What is the difference between speed and velocity?

Speed is the rate of covering distance regardless of direction (a scalar); velocity is the rate of change of displacement in a specified direction (a vector).

If a body goes and returns to its starting position, what are its displacement and distance?

Displacement is zero, but distance travelled is nonzero (equal to the total path length covered).

What does a horizontal segment on a displacement-time graph indicate?

The particle is at rest (zero velocity, since the slope is zero).

What does a horizontal segment on a velocity-time graph indicate?

The particle is moving with uniform (constant) velocity, since acceleration (the slope) is zero.

Write the equation of motion that does not involve time.

$2ax = v^2 - u^2$.

For an object thrown vertically upward, what is its velocity at the highest point?

Zero -- the velocity decreases uniformly under gravity until it momentarily reaches zero at the top before falling back down.

Long Questions & Answers

Derive the three equations of motion for constant acceleration, explaining the role of integration and initial conditions at each step.

How is the first equation, $v=u+at$, derived?

Starting from acceleration $a=dv/dt$ (constant), rewrite as $dv=a\,dt$ and integrate both sides: $v=at+c$; applying the initial condition $v=u$ at $t=0$ gives $c=u$, so $v=u+at$.

How is the second equation, $x=ut+(1/2)at^2$, derived?

Since $v=dx/dt$, substitute $v=u+at$ to get $dx/dt=u+at$, or $dx=(u+at)dt$; integrating gives $x=ut+(1/2)at^2+c1$, and applying the initial condition $x=0$ at $t=0$ gives $c1=0$, so $x=ut+(1/2)at^2$.

How is the third equation, $2ax=v^2-u^2$, derived?

From $v=u+at$, solve for t: $t=(v-u)/a$; substituting this into $x=ut+(1/2)at^2$ and simplifying algebraically eliminates t entirely, leaving the relation $2ax=v^2-u^2$.

Why are three equations needed instead of just one?

Each equation relates a different subset of the five motion quantities (u, v, a, t, x); having all three lets you solve any constant-acceleration problem regardless of which quantity is missing -- particularly useful when time is not given, since $2ax=v^2-u^2$ skips it entirely.

Explain how displacement-time and velocity-time graphs are used to analyze motion, including what their slopes and areas represent.

What does the slope of a displacement-time graph tell you?

The slope dx/dt equals velocity at that instant; a positive slope means forward motion, a negative slope means backward motion, and a zero slope (horizontal segment) means the particle is at rest.

What does the slope of a velocity-time graph tell you?

The slope dv/dt equals acceleration at that instant; a positive slope means increasing velocity (accelerating), a negative slope means decreasing velocity (decelerating/retarding), and a zero slope means constant velocity.

What does the area under a velocity-time graph represent, and why?

The area between the graph and the time axis over an interval equals the displacement during that interval, because integral of v(t)dt = integral of (dx/dt)dt = x(t2)-x(t1) by the fundamental theorem of calculus -- area under a rate-of-change graph gives total change.

How do you find total distance (rather than displacement) from a velocity-time graph that dips below the axis?

Split the area into the portion above the time axis (A1) and the portion below (A2); displacement is the signed difference A1-A2 (since velocity below the axis represents motion in the negative direction), while total distance is the sum A1+A2, since distance always accumulates regardless of direction.

Multiple Choice Questions (MCQs)

Which of the following is a scalar quantity?

A. Displacement
B. Velocity
C. Distance
D. Acceleration
Answer: C. Distance
Explanation: Distance is described by magnitude only, with no associated direction, making it a scalar quantity.

Instantaneous velocity is defined as:

A. Total displacement/Total time
B. dx/dt
C. d^2x/dt^2
D. |dx/dt|
Answer: B. dx/dt

Explanation: Instantaneous velocity is the derivative of displacement with respect to time, $v(t)=dx/dt$.

The slope of a displacement-time graph represents:

A. Acceleration

B. Velocity

C. Distance

D. Speed

Answer: B. Velocity

Explanation: Since slope = dx/dt, and velocity is defined as dx/dt, the slope of a displacement-time graph gives velocity.

The slope of a velocity-time graph represents:

A. Displacement

B. Distance

C. Acceleration

D. Speed

Answer: C. Acceleration

Explanation: Since slope = dv/dt, and acceleration is defined as dv/dt, the slope of a velocity-time graph gives acceleration.

The area under a velocity-time graph (above the time axis) represents:

A. Acceleration

B. Displacement

C. Speed

D. Time

Answer: B. Displacement

Explanation: By the fundamental theorem of calculus, integral of $v\,dt$ over an interval equals the displacement during that interval.

Which equation of motion does NOT involve time?

A. $v=u+at$

B. $x=ut+(1/2)at^2$

C. $2ax=v^2-u^2$

D. $a=dv/dt$

Answer: C. $2ax=v^2-u^2$

Explanation: $2ax=v^2-u^2$ relates displacement, velocities, and acceleration without any explicit time variable.

For a body thrown vertically upward, the acceleration used in the equations of motion is:

A. +g

B. -g

C. 0

D. 2g

Answer: B. -g

Explanation: Gravity decelerates an object moving upward, so acceleration is taken as -g in the equations of motion.

If a particle returns to its starting point after some motion, its total displacement is:

A. Equal to total distance

B. Zero

C. Negative of total distance

D. Undefined

Answer: B. Zero

Explanation: Displacement depends only on start and end position; if they coincide, displacement is zero regardless of the path taken.

For a particle with acceleration a(t) that varies with time, velocity is found by:

A. Differentiating a(t)

B. Integrating a(t) with respect to t

C. Multiplying a(t) by t

D. Taking the square root of a(t)

Answer: B. Integrating a(t) with respect to t

Explanation: Since a(t)=dv/dt, integrating a(t) with respect to time (and applying the initial velocity condition) gives v(t).

A horizontal segment on a velocity-time graph indicates:

A. The particle is at rest

B. The particle has zero displacement

C. The particle moves with constant velocity

D. The particle has infinite acceleration

Answer: C. The particle moves with constant velocity

Explanation: A horizontal (zero-slope) segment on a v-t graph means acceleration is zero, so the particle moves at constant velocity (not necessarily zero).

Quick Revision Summary

- Scalars (distance, speed) have magnitude only; vectors (displacement, velocity, acceleration) have magnitude and direction
- Average velocity = displacement/time; average speed = distance/time; these can differ, especially for round trips
- Instantaneous velocity $v(t)=dx/dt$; instantaneous speed = $|v(t)|$
- Acceleration $a(t)=dv/dt=d^2x/dt^2$; also $a=v(dv/dx)$ when v is given as a function of x
- Displacement-time graph: slope = velocity (zero slope = rest, positive = forward, negative = backward)
- Velocity-time graph: slope = acceleration; area under graph = displacement over that interval
- Equations of motion (constant a): $v=u+at$; $x=ut+(1/2)at^2$; $2ax=v^2-u^2$
- Vertical motion under gravity: use $a=g$ (downward) or $a=-g$ (upward) in the same equations of motion, $g\approx9.8$ m/s²
- Variable acceleration a(t): integrate once (with initial velocity) to find v(t), integrate again (with initial position) to find x(t)
- Distance = A1+A2 (sum of areas above and below time axis); Displacement = A1-A2 (signed sum)

Exam Tips

- Always write down which of the five motion quantities (u, v, a, t, x) are known and which is required before picking an equation of motion -- this instantly tells you which of the three equations to use
- Use $2ax=v^2-u^2$ whenever time isn't given or isn't needed -- it saves an extra step of finding time first
- For vertical motion, carefully choose your positive direction at the start (usually 'upward positive') and stay consistent -- this determines whether g is +9.8 or -9.8 throughout the problem

- On a displacement-time or velocity-time graph, always read off the slope as (change in y)/(change in x) between two clearly marked points rather than estimating visually
- When a velocity-time graph is a trapezium, split it into the standard formula (1/2)(a+b)h rather than manually breaking it into a rectangle and triangle -- it's faster and less error-prone
- For variable acceleration problems, always apply the initial condition immediately after each integration step (find C right after integrating for v, then C1 right after integrating for x) rather than saving both for the end

Unit 8: Numerical Solutions of Nonlinear Equations

A nonlinear equation $f(x)=0$ is one where f cannot be written as $ax+b$ -- the variable may appear with powers greater than one or inside trigonometric, exponential, or logarithmic functions. While quadratic, cubic, and quartic equations have algebraic formulas for their roots, no general formula exists for degree five or higher, and even when a formula exists it may be too complicated to use. Numerical methods solve this by iteratively producing closer and closer approximations to a root, starting from the Location of Root Theorem: if f is continuous on $[a,b]$ and $f(a)$ and $f(b)$ have opposite signs, at least one root exists in that interval.

This unit develops four methods that build on this idea: the graphical method (a quick visual estimate), the Bisection Method (repeatedly halving the sign-change interval), the Regula-Falsi Method (using a secant line instead of the midpoint for faster convergence), and the Newton-Raphson Method (using the tangent line at each approximation, converging fastest of all when it works). The unit also introduces the Trapezium Rule for numerically estimating definite integrals that resist exact evaluation, and closes with real-life applications in chemistry, medicine, electronics, and cryptography.

Learning Objectives

- Define a nonlinear equation and state the Location of Root Theorem (LRT)
- Identify intervals of sign change for a continuous function and use them to locate roots
- Approximate a root graphically to one or two decimal places
- Apply the Bisection Method to find a root to a specified accuracy using tolerance $|f(c)|<10^{(-n)}/2$
- Apply the Regula-Falsi (False Position) Method using the secant-line formula
- Apply the Newton-Raphson Method using the iterative tangent-line formula $x_{(n+1)}=x_n-f(x_n)/f'(x_n)$
- Apply the Trapezium Rule to numerically estimate a definite integral

- Apply numerical root-finding methods to real-life problems in chemistry, medicine, electronics, and cryptography

Key Concepts

8.1 The Location of Root Theorem and Graphical Method

A root (or zero) of $y=f(x)$ is a value of x where $f(x)=0$ -- geometrically, where the graph crosses the x-axis. The Location of Root Theorem (LRT) states: if f is continuous on $[a,b]$ and $f(a)f(b)<0$ (opposite signs), then at least one real root exists in the open interval (a,b), since a continuous graph cannot jump from below to above the x-axis without crossing it. An interval satisfying this condition is called an interval of sign change.

The graphical method approximates a root by plotting $y=f(x)$ over its interval of sign change and reading off where the curve crosses the x-axis; it gives one or two decimal places of accuracy and is often used to obtain a good starting guess for the more precise numerical methods that follow.

8.2 Accuracy, Tolerance, and the Bisection Method

Before applying a numerical method, a tolerance is set based on the number of decimal places n required: the root c is accepted once $|f(c)| < 10^{(-n)}/2$ (e.g. for 2 decimal places, $|f(c)|<0.005$, rounded to 0.01). All intermediate calculations should be carried out to at least $n+1$ decimal places, and calculators should be set to radian mode whenever trigonometric functions are involved.

The Bisection Method repeatedly halves the sign-change interval: compute the midpoint $c=(a+b)/2$, check $f(c)$; if $f(a)f(c)<0$ the root lies in $[a,c]$ (set $b=c$), otherwise it lies in $[c,b]$ (set $a=c$); repeat until $|f(c)|$ is within tolerance. It is simple and guaranteed to converge, but relatively slow since the interval only halves each step.

8.3 Regula-Falsi and Newton-Raphson Methods

The Regula-Falsi (False Position) Method improves on bisection by using the secant line through $(a,f(a))$ and $(b,f(b))$ instead of the plain midpoint: $c = [af(b)-bf(a)] / [f(b)-f(a)]$. Since this weights the estimate toward whichever endpoint is closer to zero, it usually converges faster than bisection, though it still relies only on function values (not derivatives) and the LRT sign-change condition.

The Newton-Raphson Method uses the tangent line to $y=f(x)$ at the current approximation x_n, and takes the tangent's x-intercept as the next approximation: $x_{(n+1)} = x_n - f(x_n)/f'(x_n)$. It requires f to be differentiable with $f'(x)\neq 0$ near the root, and typically converges much faster than bisection or regula-falsi, but does not directly use the LRT and can fail to converge if the initial guess is poor or $f'(x_n)$ is near zero.

8.4 The Trapezium Rule for Numerical Integration

Many definite integrals (like integral of $e^{(-x^2)}$) cannot be evaluated by elementary antiderivative methods. The Trapezium Rule approximates integral from a to b of $f(x)dx$ by dividing [a,b] into n equal subintervals of width $h=(b-a)/n$, treating the area under the curve on each subinterval as a trapezium with parallel sides $f(x_{(i-1)})$ and $f(x_i)$, and summing: integral $\approx (h/2)[f(x_0) + 2f(x_1) + 2f(x_2) + ... + 2f(x_{(n-1)}) + f(x_n)]$.

The approximation improves as n increases, since more (and narrower) trapeziums track the curve's shape more closely; the difference between the trapezium-rule estimate and the exact value (when known via the Fundamental Theorem of Calculus) shrinks toward zero as n grows. The rule works for any continuous function, including those that take negative values, where each trapezium contributes a signed area.

8.5 Real-Life Applications

Numerical root-finding methods solve practical nonlinear equations across many fields: in chemistry, equilibrium concentrations satisfy equations like $xe^x=1$, solved with Newton-Raphson; in medicine, steady-state heartbeat signal models like $x^3-x-1=0$ are solved with bisection; in electronics, diode/circuit voltage relationships like $\sin(V)-0.5V=0$ are solved with Regula-Falsi.

In cryptography, key parameters satisfying high-degree polynomial equations (like $x^5-10x+5=0$) are found using Regula-Falsi or Newton-Raphson to generate secure keys -- in every case, the underlying equation resists exact algebraic solution, making systematic numerical iteration the only practical path to an accurate answer.

Important Definitions

What is a root (zero) of a function f(x)?

A value of x for which $f(x)=0$; geometrically, the point where the graph of $y=f(x)$ crosses the x-axis.

What is a nonlinear equation?

An equation $f(x)=0$ where f cannot be written in the linear form $ax+b$; the variable may appear with powers greater than one, or inside trigonometric, exponential, or logarithmic functions.

What does the Location of Root Theorem (LRT) state?

If f is continuous on $[a,b]$ and $f(a)f(b)<0$, then f has at least one real root in the open interval (a,b).

What is an interval of sign change?

An interval $[a,b]$ on which a continuous function's endpoint values $f(a)$ and $f(b)$ have opposite signs.

What is the tolerance criterion for a root correct to n decimal places?

$|f(c)| < 10^{-n}/2$, meaning the function value at the approximate root is sufficiently close to zero.

What formula does the Bisection Method use at each step?

$c = (a+b)/2$, the midpoint of the current sign-change interval.

What formula does the Regula-Falsi Method use at each step?

$c = [af(b) - bf(a)] / [f(b) - f(a)]$, the x-intercept of the secant line through $(a,f(a))$ and $(b,f(b))$.

What formula does the Newton-Raphson Method use at each step?

$x_{n+1} = x_n - f(x_n)/f'(x_n)$, the x-intercept of the tangent line at the current approximation.

What is the Trapezium Rule used for?

Numerically estimating the value of a definite integral by approximating the area under the curve as a sum of trapezium areas.

What is the Trapezium Rule formula?

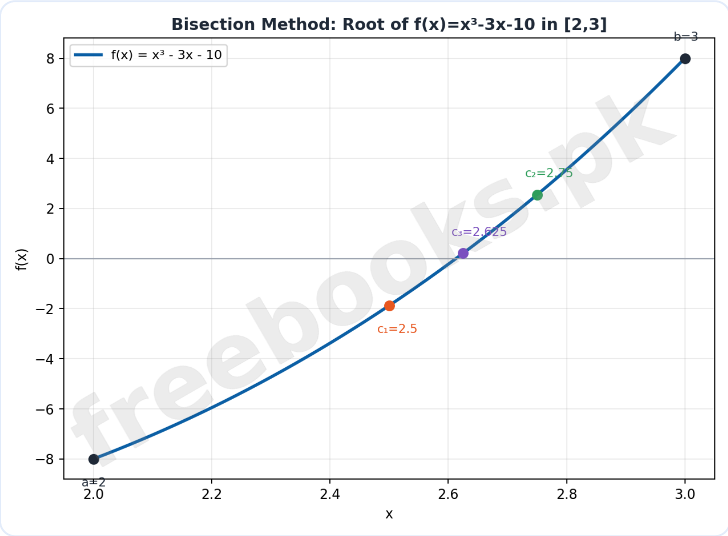
integral from a to b of $f(x)dx \approx (h/2)[f(x_0) + 2f(x_1) + 2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)]$, where $h=(b-a)/n$.

Key Facts and Relations

Topic	Key Fact / Relation
Location of Root Theorem condition	f continuous on [a,b] and $f(a)f(b) < 0 \Rightarrow$ at least one root exists in (a,b)
Tolerance for n decimal places	$\lvert f(c)\rvert < 10^{-n}/2$ (e.g. n=2: $\lvert f(c)\rvert < 0.005 \approx 0.01$)
Bisection Method	$c = (a+b)/2$; update a=c or b=c based on sign of $f(a)f(c)$
Regula-Falsi Method	$c = [af(b) - bf(a)] / [f(b) - f(a)]$
Newton-Raphson Method	$x_{n+1} = x_n - f(x_n)/f'(x_n)$, requires $f'(x_n) \neq 0$
Trapezium Rule	integral(a,b) $f(x)dx \approx (h/2)[f(x_0)+2f(x_1)+\ldots+2f(x_{n-1})+f(x_n)]$, $h=(b-a)/n$
Subinterval points for Trapezium Rule	$x_i = a + ih$, $i = 0,1,2,\ldots,n$
Single trapezium area	$A_i = (h/2)[f(x_{i-1}) + f(x_i)]$
Newton-Raphson tangent line	$y - f(x_n) = f'(x_n)(x - x_n)$
Regula-Falsi secant line	$y - f(a) = [f(b)-f(a)]/(b-a) * (x-a)$

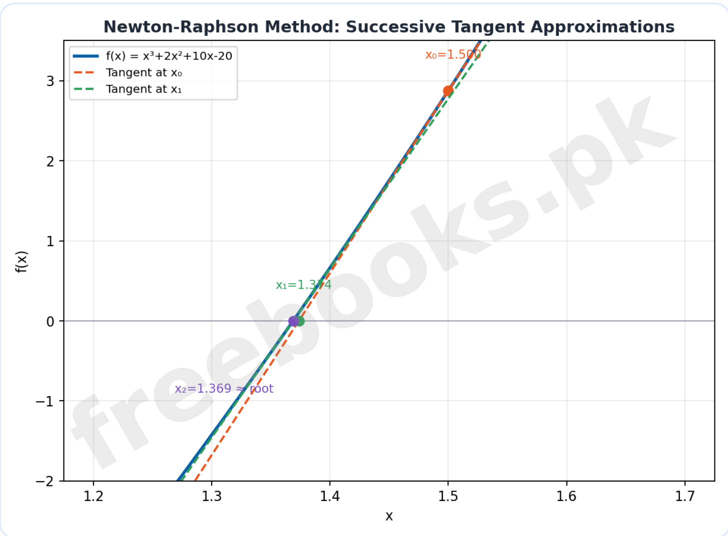
Diagrams

Bisection Method: Repeated Interval Halving: The cubic $f(x)=x^3-3x-10$ on [2,3] with the successive midpoints $c_1=2.5$, $c_2=2.75$, $c_3=2.625$ marked, showing how the Bisection Method narrows in on the root near $x\approx 2.61$



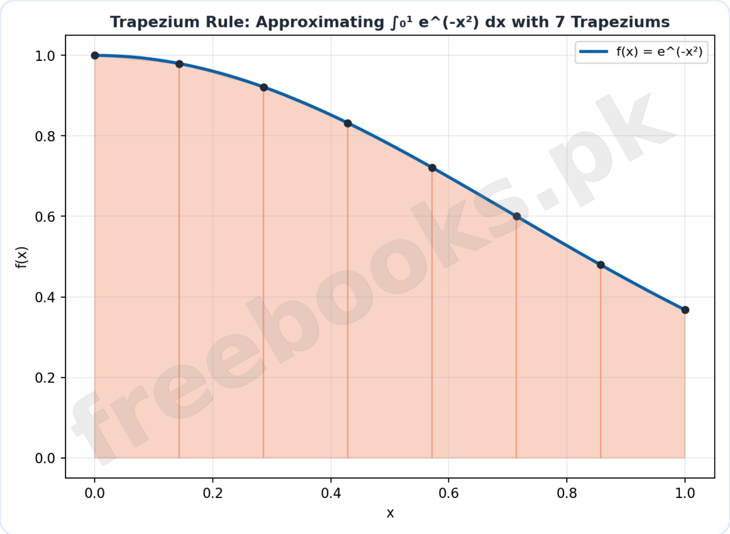
Bisection Method: Repeated Interval Halving

Newton-Raphson Method: Successive Tangent Approximations: The tangent lines drawn at $x_0=1.500$ and $x_1=1.374$ for $f(x)=x^3+2x^2+10x-20$, showing how each tangent's x-intercept gives a rapidly improving approximation converging to the root near $x\approx1.37$



Newton-Raphson Method: Successive Tangent Approximations

Trapezium Rule: Approximating a Definite Integral: The curve $f(x)=e^{(-x^2)}$ on [0,1] divided into 7 trapeziums, illustrating how the sum of trapezium areas approximates the exact area under the curve



Trapezium Rule: Approximating a Definite Integral

Solved Examples

Example 1: Finding an Interval of Sign Change

Problem: Show that $f(x)=x^4-200x-50$ has at least two real roots by finding its intervals of sign change.

1. Confirm f is continuous everywhere, since it is a polynomial: lim as $x \to c$ of $f(x) = f(c)$ for all real c.
2. Evaluate f at several integer points: $f(-1)=151>0$, $f(0)=-50<0$, $f(5)=-425<0$, $f(6)=46>0$.
3. Check the sign of $f(-1)\cdot f(0)$: $151\times(-50)<0$, so $[-1,0]$ is an interval of sign change.
4. Check the sign of $f(5)\cdot f(6)$: $(-425)(46)<0$, so $[5,6]$ is another interval of sign change.
5. Final answer: since f is continuous and changes sign on both $[-1,0]$ and $[5,6]$, by the LRT it has at least two real roots.

Example 2: Graphical Method for a Root

Problem: Find a root of $x-\cos(x)=0$ graphically, correct to two decimal places.

1. Define $f(x)=x-\cos(x)$ and confirm continuity everywhere (it's a sum of continuous functions).
2. Find an interval of sign change: $f(0)=0-\cos(0)=-1<0$ and $f(1)=1-\cos(1)\approx0.46>0$, so $[0,1]$ is an interval of sign change.
3. Compute several function values inside $[0,1]$ (in radian mode) and plot $y=x-\cos(x)$ over this interval.
4. Observe where the plotted curve crosses the x-axis.
5. Final answer: the curve crosses the x-axis at approximately $x\approx0.74$.

Example 3: Bisection Method

Problem: Use the Bisection Method to find a root of $f(x)=x^3-3x-10$ in $[2,3]$, accurate to two decimal places.

1. Confirm $f(2)=-8<0$ and $f(3)=8>0$, so $[2,3]$ is an interval of sign change, and f (a polynomial) is continuous there.
2. Iteration 1: $c=(2+3)/2=2.5$, $f(2.5)=-1.875<0$, so the root lies in $[2.5,3]$.
3. Iteration 2: $c=(2.5+3)/2=2.75$, $f(2.75)\approx2.547>0$, so the root lies in $[2.5,2.75]$. Continuing this halving process through further iterations narrows the interval further.
4. After 8 iterations, $c\approx2.6133$ with $f(2.6133)\approx0.0072$, satisfying the tolerance $|f(c)|<0.01$ for 2 decimal places.
5. Final answer: $x \approx 2.61$.

Example 4: Regula-Falsi Method

Problem: Use the Regula-Falsi Method to find a root of $f(x)=x^3-3x-10$ in [2,3], accurate to two decimal places.

1. Confirm $f(2)=-8<0$ and $f(3)=8>0$, so a root lies in [2,3].
2. Iteration 1: $c = [2(8)-3(-8)]/[8-(-8)] = 40/16 = 2.5$; $f(2.5)=-1.875<0$, so the root lies in [2.5,3].
3. Iteration 2: $c = [2.5(8)-3(-1.875)]/[8-(-1.875)] \approx 2.5949$; $f(2.5949)\approx-0.3119<0$, so the root lies in [2.5949,3].
4. Continuing this process, iteration 3 gives $c\approx2.6101$ with $f\approx-0.0487$, and iteration 4 gives $c\approx2.6125$ with $f\approx-0.0068$.
5. Final answer: since $|f(2.6125)|\approx0.0068<0.01$, the root is $x \approx 2.61$ (matching the Bisection Method result, but reached in fewer iterations).

Example 5: Newton-Raphson Method

Problem: Use the Newton-Raphson Method to find a root of $f(x)=x^3+2x^2+10x-20$, correct to two decimal places.

1. Confirm $f(1)=-7<0$ and $f(2)=16>0$, so a root lies in [1,2]; choose initial guess $x_0=1.500$.
2. Find the derivative: $f'(x)=3x^2+4x+10$, which is never zero (always positive), so the method applies.
3. Iteration 1: $f(1.5)=2.875$, $f'(1.5)=22.750$, so $x_1 = 1.500 - 2.875/22.750 \approx 1.374$.
4. Iteration 2: $f(1.374)\approx0.110$, $f'(1.374)\approx21.160$, so $x_2 = 1.374 - 0.110/21.160 \approx 1.369$.
5. Final answer: $f(1.369)\approx0.002<0.01$, satisfying the tolerance, so $x \approx 1.37$.

Example 6: Trapezium Rule for a Definite Integral

Problem: Use the Trapezium Rule to approximate integral from 0 to 1 of $e^{(-x^2)}dx$, dividing [0,1] into 7 equal subintervals.

1. Compute the step size: $h = (1-0)/7 = 1/7 \approx 0.1429$.
2. Compute function values at each x_i (rounded to 4 decimal places): $f(0)=1$, $f(1/7)=0.9798$, $f(2/7)=0.9216$, $f(3/7)=0.8322$, $f(4/7)=0.7214$, $f(5/7)=0.6004$, $f(6/7)=0.4797$, $f(1)=0.3679$.
3. Apply the Trapezium Rule: integral $\approx (h/2)[f(x_0) + 2(f(x_1)+...+f(x_6)) + f(x_7)]$.
4. Substitute: $(1/14)[1 + 2(0.9798+0.9216+0.8322+0.7214+0.6004+0.4797) + 0.3679] = (1/14)[1 + 2(4.5351) + 0.3679]$.
5. Final answer: integral from 0 to 1 of $e^{(-x^2)}dx \approx 0.7456$.

Example 7: Real-Life Application: Chemical Equilibrium Concentration

Problem: In a chemical reaction, the equilibrium concentration x (mol/L) satisfies $xe^x=1$. Find x correct to two decimal places using the Newton-Raphson Method.

1. Define $f(x)=xe^x-1$ and its derivative $f'(x)=e^x+xe^x$, which exist for all real x.
2. Check $f(0)=-1<0$ and $f(1)\approx1.718>0$, so a root lies in [0,1]; since $f(0.5)\approx-0.1756$ is close to zero, take $x_0=0.5$.
3. Iteration 1: $f(0.5)\approx-0.1757$, $f'(0.5)\approx2.4730$, so $x_1 = 0.5 - (-0.1757)/2.4730 \approx 0.5711$.
4. Iteration 2: $f(0.5711)\approx0.0107$, $f'(0.5711)\approx2.7813$, so $x_2 = 0.5711 - 0.0107/2.7813 \approx 0.5673$.
5. Final answer: $f(0.5673)\approx0.0004\approx0$, so the equilibrium concentration is $x \approx 0.57$ moles per liter.

Example 8: Real-Life Application: Electronic Circuit Voltage

Problem: In a circuit, the voltage V satisfies sin(V)-0.5V=0. Find a non-zero value of V correct to two decimal places using the Regula-Falsi Method.

1. Define f(V)=sin(V)-0.5V. Check f(1)=sin(1)-0.5≈0.3415>0 and f(2)=sin(2)-1≈-0.0907<0, so [1,2] is an interval of sign change.
2. Iteration 1: apply the Regula-Falsi formula to get $V_c \approx 1.7901$; f(1.7901)≈0.0810>0, so the root lies in [1.7901,2].
3. Iteration 2: $V_c \approx 1.8891$; f(1.8891)≈0.0052>0, so the root lies in [1.8891,2].
4. Iteration 3: $V_c \approx 1.8951$; f(1.8951)≈0.0003, very close to zero.
5. Final answer: since |f(1.8951)|<0.01, V ≈ 1.90 volts.

Short Questions & Answers

What condition must hold for the Location of Root Theorem to guarantee a root in [a,b]?

f must be continuous on [a,b], and f(a) and f(b) must have opposite signs, i.e. f(a)f(b)<0.

What is the tolerance condition for a root correct to 3 decimal places?

|f(c)| < 0.0005, which rounds to 0.001.

What update rule does the Bisection Method use if f(a)f(c)<0?

The root lies in [a,c], so set b=c and repeat.

What is the key difference between the Regula-Falsi and Bisection Methods?

Regula-Falsi uses the x-intercept of the secant line through the endpoints, while Bisection uses the plain midpoint of the interval.

What condition on f' is required for the Newton-Raphson Method to apply?

f must be differentiable near the current approximation, and f'(x) must not equal zero there.

In the Trapezium Rule, what happens to the approximation as n increases?

The approximation becomes more accurate, since more (narrower) trapeziums track the curve's shape more closely.

Why are numerical methods needed for equations like x^5-10x+5=0?

No general algebraic formula exists for solving degree-five (or higher) polynomial equations, so their roots must be approximated numerically.

Long Questions & Answers

Compare the Bisection Method, Regula-Falsi Method, and Newton-Raphson Method, explaining how each computes its next approximation and their relative strengths.

How does the Bisection Method compute its next approximation?

It always takes the exact midpoint c=(a+b)/2 of the current sign-change interval, checks which half still contains a sign change, and discards the other half -- guaranteed to converge, but relatively slow since it ignores how close each endpoint's function value is to zero.

How does the Regula-Falsi Method improve on this?

Instead of the plain midpoint, it uses the x-intercept of the secant line joining (a,f(a)) and (b,f(b)): c=[af(b)-bf(a)]/[f(b)-f(a)], which is pulled toward whichever endpoint has the smaller |f| value, typically converging faster than bisection while still only needing function values, not derivatives.

How does the Newton-Raphson Method differ from both?

It uses the tangent line to the curve at the current approximation x_n rather than any interval endpoints, taking the tangent's x-intercept x_(n+1)=x_n-f(x_n)/f'(x_n) as the next guess; this typically converges fastest of the three when it works, but requires f to be differentiable with f'≠0 near the root, and does not rely directly on the Location of Root Theorem.

When might Newton-Raphson be a poor choice compared to the other two?

If $f'(x_n)$ is very small or zero near the current approximation, the tangent line becomes nearly horizontal and the next approximation can shoot far away from the root, potentially failing to converge -- Bisection and Regula-Falsi, by staying within a guaranteed sign-change interval, avoid this risk entirely.

Explain how the Trapezium Rule is derived and applied, and how its accuracy can be assessed.

How is the interval [a,b] prepared for the Trapezium Rule?

It is divided into n equal subintervals of width $h=(b-a)/n$, with division points $x_i=a+ih$ for $i=0,1,\ldots,n$, so each subinterval $[x_{i-1},x_i]$ has the same width h.

How is the area under the curve approximated on each subinterval?

Each subinterval is treated as a trapezium with parallel sides of length $f(x_{i-1})$ and $f(x_i)$ and height (base) h, giving area $A_i=(h/2)[f(x_{i-1})+f(x_i)]$; summing all n trapezium areas and combining repeated terms gives the full formula $(h/2)[f(x_0)+2f(x_1)+\ldots+2f(x_{n-1})+f(x_n)]$.

How is the accuracy of a Trapezium Rule estimate checked?

When the exact value of the integral can be found using the Fundamental Theorem of Calculus, the two values are compared directly; the absolute difference gives the error of the approximation, and this error shrinks as n (the number of subintervals) increases.

Does the Trapezium Rule require f(x) to be positive?

No -- although it is motivated by approximating area under a positive curve, it is generally valid for any continuous function including one that takes negative values, where each trapezium then represents a signed area contributing to the total integral.

Multiple Choice Questions (MCQs)

The Location of Root Theorem requires f to be:

A. Differentiable on [a,b]

B. Continuous on [a,b] with f(a)f(b)<0

C. A polynomial

D. Always positive

Answer: B. Continuous on [a,b] with f(a)f(b)<0

Explanation: The LRT requires continuity on the closed interval and a sign change between the endpoint values.

The Bisection Method computes its next approximation as:

A. c=(a+b)/2

B. c=[af(b)-bf(a)]/[f(b)-f(a)]

C. c=a-f(a)/f'(a)

D. c=b-f(b)

Answer: A. c=(a+b)/2

Explanation: The Bisection Method always uses the exact midpoint of the current interval.

The Regula-Falsi formula for the next approximation is:

A. c=(a+b)/2

B. c=[af(b)-bf(a)]/[f(b)-f(a)]

C. c=x_n-f(x_n)/f'(x_n)

D. c=a+b

Answer: B. c=[af(b)-bf(a)]/[f(b)-f(a)]

Explanation: This is the x-intercept of the secant line through the two endpoint points (a,f(a)) and (b,f(b)).

The Newton-Raphson formula is:

A. x_(n+1)=(a+b)/2

B. x_(n+1)=x_n-f(x_n)/f'(x_n)

C. x_(n+1)=[af(b)-bf(a)]/[f(b)-f(a)]

D. x_(n+1)=f(x_n)/f'(x_n)

Answer: B. x_(n+1)=x_n-f(x_n)/f'(x_n)

Explanation: Newton-Raphson uses the tangent line's x-intercept, requiring the derivative f'(x_n).

For a root correct to 2 decimal places, the tolerance condition is:

A. |f(c)|<0.5

B. |f(c)|<0.05

C. |f(c)|<0.005

D. |f(c)|<5

Answer: C. |f(c)|<0.005

Explanation: The general tolerance rule is $|f(c)|<10^{-n}/2$, which for n=2 gives |f(c)|<0.005.

Which numerical method requires the function to be differentiable?

A. Bisection Method

B. Graphical Method

C. Newton-Raphson Method

D. Trapezium Rule

Answer: C. Newton-Raphson Method

Explanation: Newton-Raphson uses $f'(x_n)$ at every iteration, so f must be differentiable near each approximation.

In the Trapezium Rule with n subintervals, the step size h equals:

A. (b-a)

B. (b-a)/n

C. (b+a)/n

D. n/(b-a)

Answer: B. (b-a)/n

Explanation: The interval [a,b] is divided into n equal parts, each of width h=(b-a)/n.

In the Trapezium Rule formula, the coefficient of the interior function values $f(x_1),\dots,f(x_{n-1})$ is:

A. 1

B. 2

C. h

D. n

Answer: B. 2

Explanation: Each interior point is shared by two adjacent trapeziums, so its function value is counted twice, giving coefficient 2.

As the number of subintervals n increases in the Trapezium Rule, the approximation generally:

A. Becomes less accurate

B. Stays exactly the same

C. Becomes more accurate

D. Becomes undefined

Answer: C. Becomes more accurate

Explanation: More (narrower) trapeziums track the curve's shape more closely, reducing the gap between approximate and exact area.

A chemical equilibrium concentration satisfying $xe^x=1$ is best solved using which method, per the unit's worked example?

A. Graphical method only

B. Trapezium Rule

C. Newton-Raphson Method

D. None of these methods apply

Answer: C. Newton-Raphson Method

Explanation: The worked example in the unit solves $xe^x=1$ for equilibrium concentration using the Newton-Raphson Method.

Quick Revision Summary

- Nonlinear equation: $f(x)=0$ where f is not of the form $ax+b$; root/zero is where the graph crosses the x-axis
- Location of Root Theorem (LRT): f continuous on $[a,b]$, $f(a)f(b)<0 \Rightarrow$ at least one root in (a,b)
- Graphical method: plot f over the sign-change interval, read off where it crosses the x-axis (1-2 decimal places)
- Tolerance for n decimal places: $|f(c)| < 10^{(-n)}/2$
- Bisection Method: $c=(a+b)/2$, halve the interval each iteration based on sign of $f(a)f(c)$
- Regula-Falsi Method: $c=[af(b)-bf(a)]/[f(b)-f(a)]$, the secant-line x-intercept, converges faster than bisection
- Newton-Raphson Method: $x_{n+1}=x_n-f(x_n)/f'(x_n)$, tangent-line x-intercept, fastest convergence but needs $f'\neq 0$
- Trapezium Rule: $\int_a^b f(x)dx \approx (h/2)[f(x_0)+2f(x_1)+\ldots+2f(x_{n-1})+f(x_n)]$, $h=(b-a)/n$
- Trapezium Rule accuracy improves as n increases; works for functions taking negative values too

- Applications: chemical equilibrium (Newton-Raphson), heartbeat models (Bisection), circuit voltages (Regula-Falsi), cryptographic key parameters (Regula-Falsi/Newton-Raphson)

Exam Tips

- Always verify continuity and check $f(a)f(b)<0$ before applying any LRT-based method (Bisection, Regula-Falsi) -- skipping this check can lead to applying the method where no root is guaranteed
- Keep at least one extra decimal place in every intermediate calculation beyond the accuracy you need in the final answer, to avoid compounding rounding errors across iterations
- For Newton-Raphson, always compute $f'(x)$ symbolically once at the start, then substitute numerical values at each iteration -- recomputing the derivative formula each time wastes effort and invites errors
- Build an iteration table (columns: iteration, a, b, c, f(c), new interval) for Bisection and Regula-Falsi problems -- it keeps the halving/narrowing process organized and easy to check
- Remember trigonometric functions must be evaluated in radian mode throughout every numerical method calculation
- For the Trapezium Rule, always double-check the coefficient pattern (1, 2, 2, 2, ..., 2, 1) before summing -- forgetting to double the interior terms is the most common calculation error

Unit 9: Inverse Trigonometric Functions and Their Graphs

A function has an inverse only when it is one-to-one (bijective) -- every output must come from exactly one input. Trigonometric functions like sin x, cos x, and tan x fail this test on their natural domains because they are periodic: infinitely many values of x give the same sine, cosine, or tangent. To make each trigonometric function invertible, its domain is restricted to a principal interval on which it is one-to-one -- for sine, $[-\pi/2, \pi/2]$; for cosine, $[0, \pi]$; for tangent, $(-\pi/2, \pi/2)$ -- and the resulting inverse function, restricted to this principal range, is written $\sin^{-1}x$, $\cos^{-1}x$, and $\tan^{-1}x$ respectively (also called arcsine, arccosine, and arctangent).

This unit defines all six inverse trigonometric functions -- $\sin^{-1}$, $\cos^{-1}$, $\tan^{-1}$, $\csc^{-1}$, $\sec^{-1}$, and $\cot^{-1}$ -- together with their principal domains and ranges, and shows how each one's graph is obtained by reflecting the corresponding restricted trig graph about the line y=x. It covers evaluating principal values and composite expressions, finding the domain of composite inverse trig functions, proving reciprocal and complementary identities, and deriving the sum and difference formulas for inverse trigonometric functions, including the double-angle corollaries used to simplify identities and solve equations.

Learning Objectives

- Explain why trigonometric functions must have their domains restricted before they can be inverted
- State the principal domain and range of each of the six inverse trigonometric functions
- Sketch the graph of an inverse trigonometric function as the reflection of its restricted trig function about y=x
- Evaluate the principal value of expressions like $\sin^{-1}(1/2)$ and $\tan^{-1}(-1/\sqrt{3})$
- Evaluate composite expressions such as $\sec[\tan^{-1}(-\sqrt{3})]$ using reference triangles

- Determine the domain of composite inverse trigonometric functions like $\cos^{-1}(3x-1)$
- Prove and apply the reciprocal identities (e.g. $\csc^{-1}x=\sin^{-1}(1/x)$) and complementary identities (e.g. $\sin^{-1}x+\cos^{-1}x=\pi/2$)
- Apply the sum and difference formulas for inverse trigonometric functions, including the double-angle corollaries, to prove identities and solve equations

Key Concepts

9.1 Why Trigonometric Functions Need Domain Restriction

An inverse function f^{-1} exists only when f is one-to-one (bijective): each output value must correspond to exactly one input. Trigonometric functions are periodic -- sin x repeats every 2π, so infinitely many x-values share the same sine value -- meaning sin x, cos x, and tan x are NOT one-to-one over their full natural domains, and therefore have no inverse in the usual sense unless their domains are cut down.

To fix this, each trigonometric function's domain is restricted to a principal interval on which it is strictly one-to-one and still covers the full range of output values. On this restricted domain the function becomes invertible, and its inverse is called the corresponding inverse trigonometric function, defined only on this principal branch.

9.2 Inverse Sine, Cosine, and Tangent

$\sin^{-1}x$ is defined by restricting sin x to [-π/2, π/2], where sine increases monotonically from -1 to 1; thus $\sin^{-1}x$ has domain [-1,1] and range [-π/2, π/2]. $\cos^{-1}x$ is defined by restricting cos x to [0, π], where cosine decreases monotonically from 1 to -1; thus $\cos^{-1}x$ has domain [-1,1] and range [0, π].

$\tan^{-1}x$ is defined by restricting tan x to (-π/2, π/2), where tangent increases monotonically across all real values; thus $\tan^{-1}x$ has domain (-∞, ∞) and range (-π/2, π/2). Each graph is obtained by reflecting the corresponding restricted trig graph about the line y=x, swapping the roles of x and y.

9.3 Inverse Cosecant, Secant, and Cotangent

$\csc^{-1}x$ is defined by restricting csc x to $[-\pi/2,0)\cup(0,\pi/2]$, giving $\csc^{-1}x$ domain $(-\infty,-1]\cup[1,\infty)$ and range $[-\pi/2,0)\cup(0,\pi/2]$. $\sec^{-1}x$ is defined by restricting sec x to $[0,\pi/2)\cup(\pi/2,\pi]$, giving $\sec^{-1}x$ domain $(-\infty,-1]\cup[1,\infty)$ and range $[0,\pi/2)\cup(\pi/2,\pi]$.

$\cot^{-1}x$ is defined by restricting cot x to $(0,\pi)$, giving $\cot^{-1}x$ domain $(-\infty,\infty)$ and range $(0,\pi)$. These three reciprocal-function inverses are used less often directly but appear frequently through the reciprocal identities linking them back to $\sin^{-1}$, $\cos^{-1}$, and $\tan^{-1}$.

9.4 Evaluating Principal Values and Composite Expressions

To find the principal value of an expression like $\sin^{-1}(1/2)$, identify the angle θ within the principal range $[-\pi/2,\pi/2]$ for which $\sin\theta=1/2$, giving $\theta=\pi/6$. The same approach applies to $\cos^{-1}$ and $\tan^{-1}$, always selecting the answer that lies in the correct principal range -- this is why $\cos^{-1}(-\sqrt{3}/2)=5\pi/6$ (in $[0,\pi]$) rather than any other coterminal angle.

Composite expressions like $\sec[\tan^{-1}(-\sqrt{3})]$ are evaluated by first finding the angle $\alpha=\tan^{-1}(-\sqrt{3})$ in the principal range, then evaluating the outer function at α, often using a reference right triangle built from the known ratio to read off the remaining trigonometric values directly.

9.5 Domain of Composite Functions and Key Identities

The domain of a composite function like $\cos^{-1}(3x-1)$ is found by requiring the inner expression to lie within the domain of the outer inverse function: here, $-1\leq 3x-1\leq 1$, which solves to $0\leq x\leq 2/3$. The same technique applies to any composite inverse trig expression, including those with logarithms or other inner functions, by intersecting the inner function's own domain restriction with the inverse trig function's domain.

Reciprocal identities relate each inverse function to its reciprocal partner: $\csc^{-1}x=\sin^{-1}(1/x)$, $\sec^{-1}x=\cos^{-1}(1/x)$, $\cot^{-1}x=\tan^{-1}(1/x)$ for appropriate x. Complementary identities relate pairs summing to π/2: $\sin^{-1}x+\cos^{-1}x=\pi/2$, $\tan^{-1}x+\cot^{-1}x=\pi/2$, $\sec^{-1}x+\csc^{-1}x=\pi/2$, both provable directly from the definitions using right-triangle or algebraic arguments.

9.6 Sum and Difference Formulas for Inverse Trigonometric Functions

Six sum/difference formulas extend the inverse trig functions to combined expressions, including $\sin^{-1}x+\sin^{-1}y=\sin^{-1}[x\sqrt{(1-y^2)}+y\sqrt{(1-x^2)}]$ and $\tan^{-1}x+\tan^{-1}y=\tan^{-1}[(x+y)/(1-xy)]$ (valid when $xy<1$), each proved by setting $\alpha=\sin^{-1}x$ (or $\tan^{-1}x$) and $\beta=\sin^{-1}y$ (or $\tan^{-1}y$), applying the corresponding angle-sum trig identity, then converting back using the inverse function.

Setting $y=x$ in these formulas produces double-angle corollaries: $2\tan^{-1}x=\tan^{-1}(2x/(1-x^2))$, $2\sin^{-1}x=\sin^{-1}(2x\sqrt{(1-x^2)})$, and $2\cos^{-1}x=\cos^{-1}(2x^2-1)$. These are used throughout the unit's worked examples to prove identities and to solve equations involving sums of inverse trigonometric terms, always checking that the final angle found lies within the required principal range.

Important Definitions

What is required for a function to have an inverse?

The function must be one-to-one (bijective): every output value corresponds to exactly one input value.

Why can't sin x, cos x, and tan x be inverted directly?

Because they are periodic, so infinitely many input values give the same output value, making them not one-to-one over their full natural domains.

What are the domain and range of $\sin^{-1}x$?

Domain $[-1,1]$, range $[-\pi/2, \pi/2]$, obtained by restricting $\sin x$ to that principal interval.

What are the domain and range of $\cos^{-1}x$?

Domain $[-1,1]$, range $[0, \pi]$, obtained by restricting $\cos x$ to that principal interval.

What are the domain and range of $\tan^{-1}x$?

Domain $(-\infty,\infty)$, range $(-\pi/2, \pi/2)$, obtained by restricting $\tan x$ to that principal interval.

What are the domain and range of $\cot^{-1}x$?

Domain $(-\infty,\infty)$, range $(0, \pi)$, obtained by restricting $\cot x$ to that principal interval.

How is the graph of an inverse trigonometric function obtained from its trig function's graph?

By reflecting the restricted trig function's graph about the line $y=x$, which swaps the roles of the x- and y-coordinates.

What is the reciprocal identity linking $\csc^{-1}x$ and $\sin^{-1}x$?

$\csc^{-1}x = \sin^{-1}(1/x)$, for appropriate values of x.

What is the complementary identity linking $\sin^{-1}x$ and $\cos^{-1}x$?

$\sin^{-1}x + \cos^{-1}x = \pi/2$, for all x in $[-1,1]$.

What is the double-angle corollary for $\tan^{-1}x$?

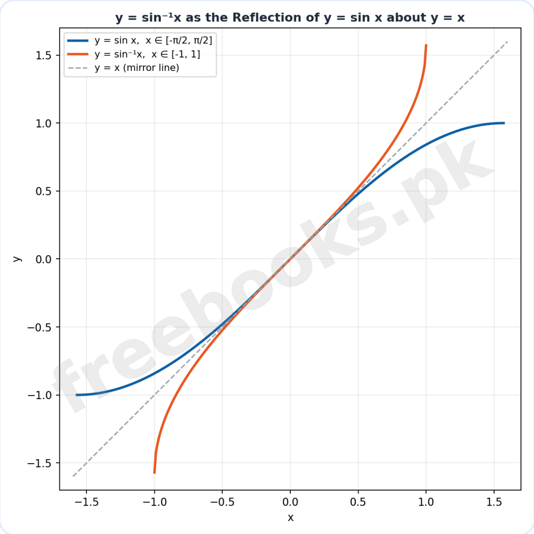
$2\tan^{-1}x = \tan^{-1}(2x/(1-x^2))$, obtained by setting $y=x$ in the sum formula for $\tan^{-1}x+\tan^{-1}y$.

Key Facts and Relations

Topic	Key Fact / Relation
$\sin^{-1}x$ principal range	$[-\pi/2, \pi/2]$, domain $[-1,1]$
$\cos^{-1}x$ principal range	$[0, \pi]$, domain $[-1,1]$
$\tan^{-1}x$ principal range	$(-\pi/2, \pi/2)$, domain $(-\infty,\infty)$
Reciprocal identities	$\csc^{-1}x=\sin^{-1}(1/x)$, $\sec^{-1}x=\cos^{-1}(1/x)$, $\cot^{-1}x=\tan^{-1}(1/x)$
Complementary identities	$\sin^{-1}x+\cos^{-1}x=\pi/2$, $\tan^{-1}x+\cot^{-1}x=\pi/2$, $\sec^{-1}x+\csc^{-1}x=\pi/2$
Sum formula for $\sin^{-1}$	$\sin^{-1}x+\sin^{-1}y = \sin^{-1}[x\sqrt{(1-y^2)}+y\sqrt{(1-x^2)}]$
Sum formula for $\tan^{-1}$	$\tan^{-1}x+\tan^{-1}y = \tan^{-1}[(x+y)/(1-xy)]$, valid for $xy<1$
Difference formula for $\tan^{-1}$	$\tan^{-1}x-\tan^{-1}y = \tan^{-1}[(x-y)/(1+xy)]$
Double-angle corollary for $\tan^{-1}$	$2\tan^{-1}x = \tan^{-1}(2x/(1-x^2))$
Double-angle corollaries for $\sin^{-1}$ and $\cos^{-1}$	$2\sin^{-1}x = \sin^{-1}(2x\sqrt{(1-x^2)})$; $2\cos^{-1}x = \cos^{-1}(2x^2-1)$

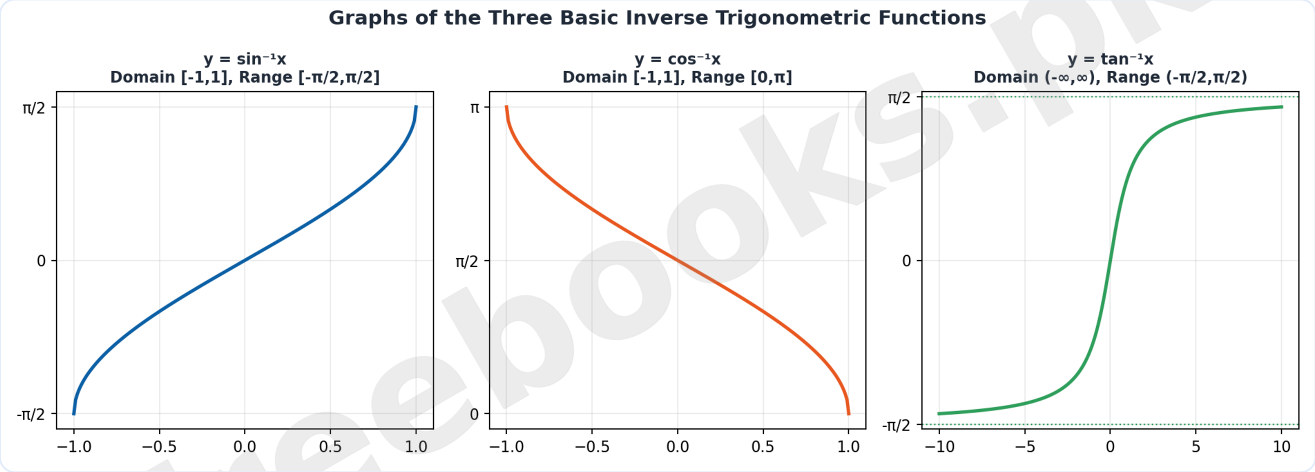
Diagrams

y = sin⁻¹x as the Reflection of y = sin x about y = x: The restricted graph of y=sin x on [-π/2,π/2] plotted alongside y=sin⁻¹x on [-1,1], with the line y=x shown as a mirror, illustrating how every inverse trig graph is obtained by reflection



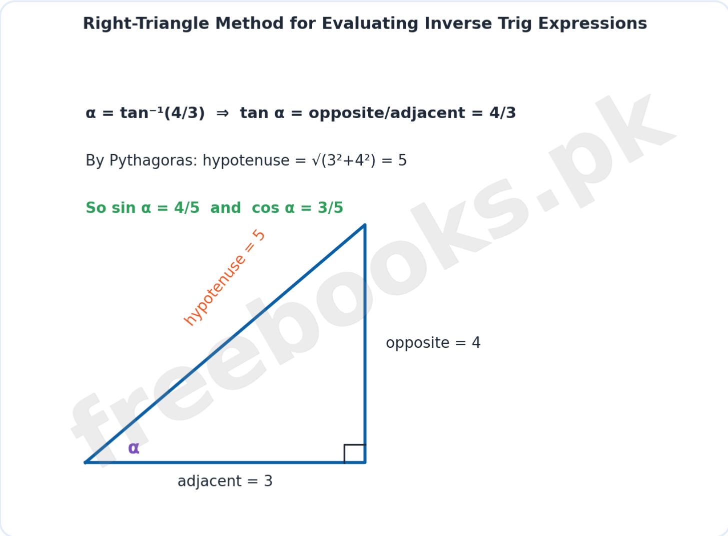
y = sin⁻¹x as the Reflection of y = sin x about y = x

Graphs of the Three Basic Inverse Trigonometric Functions: Side-by-side graphs of y=sin⁻¹x, y=cos⁻¹x, and y=tan⁻¹x with their domains and ranges labelled, showing the characteristic shape and asymptotic behaviour of each



Graphs of the Three Basic Inverse Trigonometric Functions

Right-Triangle Method for Evaluating Inverse Trig Expressions: A labelled 3-4-5 right triangle used to evaluate composite expressions like sin[tan⁻¹(4/3)] by reading off the remaining trigonometric ratios directly from the triangle's sides



Right-Triangle Method for Evaluating Inverse Trig Expressions

Solved Examples

Example 1: Principal Values of Basic Inverse Trig Expressions

Problem: Find the principal values of $\sin^{-1}(1/2)$, $\cos^{-1}(-\sqrt{3}/2)$, and $\tan^{-1}(-1/\sqrt{3})$.

1. For $\sin^{-1}(1/2)$: find θ in $[-\pi/2,\pi/2]$ with $\sin\theta=1/2$ -- this is $\theta=\pi/6$.
2. For $\cos^{-1}(-\sqrt{3}/2)$: find θ in $[0,\pi]$ with $\cos\theta=-\sqrt{3}/2$ -- since cosine is negative in the second quadrant, $\theta=5\pi/6$.
3. For $\tan^{-1}(-1/\sqrt{3})$: find θ in $(-\pi/2,\pi/2)$ with $\tan\theta=-1/\sqrt{3}$ -- since tangent is negative for negative θ here, $\theta=-\pi/6$.
4. Check each answer lies within the correct principal range: $\pi/6\in[-\pi/2,\pi/2]$ ✓, $5\pi/6\in[0,\pi]$ ✓, $-\pi/6\in(-\pi/2,\pi/2)$ ✓.
5. Final answer: $\sin^{-1}(1/2)=\pi/6$, $\cos^{-1}(-\sqrt{3}/2)=5\pi/6$, $\tan^{-1}(-1/\sqrt{3})=-\pi/6$.

Example 2: Evaluating a Composite Expression

Problem: Evaluate $\sec[\tan^{-1}(-\sqrt{3})]$.

1. Let $\alpha=\tan^{-1}(-\sqrt{3})$, so $\tan\alpha=-\sqrt{3}$ with α in the principal range $(-\pi/2,\pi/2)$; since tangent is negative, α is a negative angle: $\alpha=-\pi/3$.

2. Build a reference right triangle for the reference angle $\pi/3$: opposite=$\sqrt{3}$, adjacent=1, hypotenuse=$\sqrt{(1+3)}=2$.
3. Since $\alpha=-\pi/3$ lies in the fourth quadrant direction, $\cos\alpha=\cos(\pi/3)=1/2$ (cosine is even, so the sign is unaffected by the negative angle).
4. $\sec\alpha = 1/\cos\alpha = 1/(1/2) = 2$.
5. Final answer: $\sec[\tan^{-1}(-\sqrt{3})] = 2$.

Example 3: Finding the Domain of a Composite Inverse Trig Function

Problem: Find the domain of $y = \cos^{-1}(3x-1)$.

1. Recall $\cos^{-1}u$ is defined only when $u\in[-1,1]$, so require $-1 \le 3x-1 \le 1$.
2. Add 1 to all three parts: $0 \le 3x \le 2$.
3. Divide by 3: $0 \le x \le 2/3$.
4. Verify with an endpoint: at $x=0$, $3(0)-1=-1$, and $\cos^{-1}(-1)=\pi$, which is defined; at $x=2/3$, $3(2/3)-1=1$, and $\cos^{-1}(1)=0$, also defined.
5. Final answer: the domain of $\cos^{-1}(3x-1)$ is $[0, 2/3]$.

Example 4: Proving a Reciprocal Identity

Problem: Prove that $\csc^{-1}x = \sin^{-1}(1/x)$ for $x\ge1$.

1. Let $\theta = \csc^{-1}x$, so by definition $\csc\theta = x$ with θ in the principal range for $\csc^{-1}$.
2. Since $\csc\theta = 1/\sin\theta$, this gives $\sin\theta = 1/x$.
3. Because the principal range of $\csc^{-1}$ maps directly onto the same values as the principal range of $\sin^{-1}$ (both avoiding 0), $\theta = \sin^{-1}(1/x)$ follows directly from $\sin\theta=1/x$.
4. Therefore $\theta = \csc^{-1}x = \sin^{-1}(1/x)$.
5. Final answer: $\csc^{-1}x = \sin^{-1}(1/x)$ is proved for $x\ge1$ (and similarly for $x\le-1$).

Example 5: Proving a Complementary Identity

Problem: Prove that $\sin^{-1}x + \cos^{-1}x = \pi/2$ for all x in [-1,1].

1. Let $\theta = \sin^{-1}x$, so $\sin\theta = x$ with θ in $[-\pi/2,\pi/2]$.
2. Using the co-function identity, $\cos(\pi/2-\theta) = \sin\theta = x$.
3. Since $\theta\in[-\pi/2,\pi/2]$, the angle $(\pi/2-\theta)$ lies in $[0,\pi]$, which is exactly the principal range of $\cos^{-1}$.
4. Therefore $\cos^{-1}x = \pi/2 - \theta = \pi/2 - \sin^{-1}x$.
5. Final answer: rearranging gives $\sin^{-1}x + \cos^{-1}x = \pi/2$ for every x in [-1,1].

Example 6: Showing an Identity Between Two Inverse Functions

Problem: Show that $\cos^{-1}(7/25) = \csc^{-1}(25/24)$.

1. Let $\theta = \cos^{-1}(7/25)$, so $\cos\theta = 7/25$ with θ in $[0,\pi]$; build the reference right triangle with adjacent=7, hypotenuse=25.
2. Find the opposite side: $\sqrt{(25^2-7^2)} = \sqrt{(625-49)} = \sqrt{576} = 24$.
3. Since $\theta\in[0,\pi]$ and $\cos\theta>0$, θ is in the first quadrant, so $\sin\theta = 24/25$ (positive).
4. Then $\csc\theta = 1/\sin\theta = 25/24$, so $\theta = \csc^{-1}(25/24)$ as well (both principal ranges agree for this positive value).
5. Final answer: $\cos^{-1}(7/25) = \csc^{-1}(25/24)$ is confirmed.

Example 7: Using the Half-Angle Method with a Reference Triangle

Problem: Find $\sin[\frac{1}{2}\tan^{-1}(4/3)]$.

1. Let $\alpha = \tan^{-1}(4/3)$, so $\tan\alpha = 4/3$ with α in $(-\pi/2,\pi/2)$; since $4/3>0$, α is in the first quadrant.
2. Build the reference right triangle: opposite=4, adjacent=3, hypotenuse=$\sqrt{(16+9)}=5$, giving $\cos\alpha = 3/5$.

3. Apply the half-angle formula: $\sin(\alpha/2) = \sqrt{[(1-\cos \alpha)/2]} = \sqrt{[(1-3/5)/2]} = \sqrt{[(2/5)/2]} = \sqrt{(1/5)}$.
4. Simplify: $\sqrt{(1/5)} = 1/\sqrt{5} = \sqrt{5}/5$.
5. Final answer: $\sin[½ \tan^{-1}(4/3)] = \sqrt{5}/5$.

Example 8: Solving an Equation Involving a Sum of Inverse Tangents

Problem: Solve $\tan^{-1}[(x-1)/(x+1)] + \tan^{-1}[2x/(x+1)] = \pi/4$ for x.

1. Apply the sum formula $\tan^{-1}A+\tan^{-1}B = \tan^{-1}[(A+B)/(1-AB)]$ with $A=(x-1)/(x+1)$ and $B=2x/(x+1)$.
2. Compute $A+B = [(x-1)+2x]/(x+1) = (3x-1)/(x+1)$, and $AB = 2x(x-1)/(x+1)^2$.
3. Set the combined expression equal to $\tan(\pi/4)=1$, then cross-multiply and simplify the resulting equation in x.
4. After expanding and collecting terms, the equation reduces to a quadratic that factors to give $x=1$ as the solution lying within the valid domain.
5. Final answer: $x = 1$ (checked to satisfy the original equation and keep both inverse tangent arguments within valid ranges).

Short Questions & Answers

Why must the domain of sin x be restricted before defining $\sin^{-1}x$?

Because sin x is periodic and not one-to-one over its full domain; restricting it to $[-\pi/2,\pi/2]$ makes it one-to-one so an inverse can be defined.

What is the range of $\cos^{-1}x$?

$[0, \pi]$.

How is the graph of $\tan^{-1}x$ related to the graph of tan x?

It is the reflection of the restricted graph of tan x (on $(-\pi/2,\pi/2)$) about the line $y=x$.

What is the domain of $\sec^{-1}x$?

$(-\infty,-1] \cup [1,\infty)$.

State the complementary identity linking $\tan^{-1}x$ and $\cot^{-1}x$.

$\tan^{-1}x + \cot^{-1}x = \pi/2$.

What condition is required for the sum formula $\tan^{-1}x+\tan^{-1}y=\tan^{-1}[(x+y)/(1-xy)]$ to hold directly?

The product xy must be less than 1 ($xy<1$); otherwise an adjustment of $\pm\pi$ is needed.

What double-angle formula results from setting y=x in the $\sin^{-1}$ sum formula?

$2\sin^{-1}x = \sin^{-1}(2x\sqrt{1-x^2})$.

Long Questions & Answers

Explain how each of the six inverse trigonometric functions is defined, including why domain restriction is necessary and what principal domain and range each one is given.

Why is domain restriction necessary before defining any inverse trigonometric function?

Trigonometric functions are periodic, so the same output value is produced by infinitely many input angles, meaning none of them are one-to-one over their natural domain; a function must be one-to-one to have a genuine inverse, so each trig function's domain is first cut down to a principal interval on which it is strictly increasing or decreasing.

What are the principal domains and ranges of $\sin^{-1}x$, $\cos^{-1}x$, and $\tan^{-1}x$?

$\sin^{-1}x$ has domain $[-1,1]$ and range $[-\pi/2,\pi/2]$; $\cos^{-1}x$ has domain $[-1,1]$ and range $[0,\pi]$; $\tan^{-1}x$ has domain $(-\infty,\infty)$ and range $(-\pi/2,\pi/2)$ -- each obtained by restricting the corresponding trig function to the interval on which it is one-to-one.

What are the principal domains and ranges of $\csc^{-1}x$, $\sec^{-1}x$, and $\cot^{-1}x$?

$\csc^{-1}x$ and $\sec^{-1}x$ both have domain $(-\infty,-1]\cup[1,\infty)$, with $\csc^{-1}x$'s range $[-\pi/2,0)\cup(0,\pi/2]$ and $\sec^{-1}x$'s range $[0,\pi/2)\cup(\pi/2,\pi]$; $\cot^{-1}x$ has domain $(-\infty,\infty)$ and range $(0,\pi)$, obtained by restricting cot x to that interval.

How does the graph of each inverse trig function relate to its original trig function's graph?

Each inverse trig function's graph is the mirror reflection of its restricted trig function's graph about the line $y=x$, which swaps the x- and y-axes' roles -- so a steep, near-vertical section of the trig graph becomes a near-horizontal section of the inverse graph, and vice versa.

Explain how the sum formula for inverse tangent is derived and used, including its double-angle corollary and a worked application.

How is the sum formula $\tan^{-1}x+\tan^{-1}y=\tan^{-1}[(x+y)/(1-xy)]$ derived?

Set $\alpha=\tan^{-1}x$ and $\beta=\tan^{-1}y$, so $\tan\alpha=x$ and $\tan\beta=y$; apply the trigonometric angle-sum identity $\tan(\alpha+\beta)=(\tan\alpha+\tan\beta)/(1-\tan\alpha\tan\beta)=(x+y)/(1-xy)$, then take $\tan^{-1}$ of both sides to recover $\alpha+\beta=\tan^{-1}[(x+y)/(1-xy)]$, valid when $xy<1$ so that $\alpha+\beta$ stays within the correct principal range.

What double-angle formula follows from this sum formula?

Setting $y=x$ gives $2\tan^{-1}x = \tan^{-1}(2x/(1-x^2))$, obtained simply by substituting $y=x$ directly into the general sum formula and simplifying the resulting expression.

How is this formula applied to solve an equation like $\tan^{-1}[(x-1)/(x+1)]+\tan^{-1}[2x/(x+1)]=\pi/4$?

The two inverse-tangent terms are combined into a single $\tan^{-1}$ expression using the sum formula, the combined argument is set equal to $\tan(\pi/4)=1$, and the resulting algebraic equation is solved for x, checking that the solution keeps both original arguments within a valid principal range.

Why is checking the domain condition important after solving such an equation?

Because the sum formula can introduce an extraneous solution or require a $\pm\pi$ adjustment when $xy\geq1$, so any candidate solution for x must be substituted back into the original expression to confirm it produces a genuine sum equal to $\pi/4$ and not merely a solution of the simplified algebraic equation.

Multiple Choice Questions (MCQs)

The domain of $\sin^{-1}x$ is:

A. $(-\infty,\infty)$

B. $[-1,1]$

C. $[0,\pi]$

D. $[-\pi/2,\pi/2]$

Answer: B. $[-1,1]$

Explanation: $\sin^{-1}x$ is defined only for inputs between -1 and 1 inclusive.

The range of $\cos^{-1}x$ is:

A. $[-\pi/2,\pi/2]$

B. $(-\pi/2,\pi/2)$

C. $[0,\pi]$

D. $(0,\pi)$

Answer: C. $[0,\pi]$

Explanation: $\cos x$ is restricted to $[0,\pi]$ to make it invertible, so this is the range of $\cos^{-1}x$.

The graph of $y=\tan^{-1}x$ is obtained from $y=\tan x$ (restricted) by:

A. Shifting it right

B. Reflecting about the x-axis

C. Reflecting about the line $y=x$

D. Stretching vertically

Answer: C. Reflecting about the line $y=x$

Explanation: Every inverse function's graph is the reflection of the original (restricted) function's graph about $y=x$.

Which is the correct reciprocal identity?

A. $\csc^{-1}x=\cos^{-1}(1/x)$

B. $\sec^{-1}x=\cos^{-1}(1/x)$

C. $\cot^{-1}x=\sin^{-1}(1/x)$

D. $\csc^{-1}x=\tan^{-1}(1/x)$

Answer: B. $\sec^{-1}x=\cos^{-1}(1/x)$

Explanation: $\sec^{-1}x=\cos^{-1}(1/x)$ is the correct reciprocal identity linking secant's inverse to cosine's inverse.

$\sin^{-1}x + \cos^{-1}x$ equals:

A. π

B. $\pi/2$

C. 0

D. 2π

Answer: B. $\pi/2$

Explanation: This is the standard complementary identity, valid for all x in [-1,1].

The principal value of $\tan^{-1}(-1/\sqrt{3})$ is:

A. $\pi/6$

B. $-\pi/6$

C. $5\pi/6$

D. $-5\pi/6$

Answer: B. $-\pi/6$

Explanation: $\tan \theta=-1/\sqrt{3}$ with θ in $(-\pi/2,\pi/2)$ gives $\theta=-\pi/6$.

The domain of $\cos^{-1}(3x-1)$ is:

A. [-1,1]

B. [0,2/3]

C. [0,1]

D. $(-\infty,\infty)$

Answer: B. [0,2/3]

Explanation: Solving $-1\leq 3x-1\leq 1$ gives $0\leq x\leq 2/3$.

The sum formula $\tan^{-1}x+\tan^{-1}y=\tan^{-1}[(x+y)/(1-xy)]$ requires:

A. $x=y$

B. $xy<1$

C. $x,y>0$

D. $x+y=0$

Answer: B. $xy<1$

Explanation: The formula applies directly only when $xy<1$; otherwise a $\pm\pi$ correction is needed.

Setting y=x in the $\sin^{-1}$ sum formula gives:

A. $2\sin^{-1}x=\sin^{-1}(2x)$

B. $2\sin^{-1}x=\sin^{-1}(2x\sqrt{(1-x^2)})$

C. $2\sin^{-1}x=\sin^{-1}(x^2)$

D. $2\sin^{-1}x=2\sin^{-1}x$

Answer: B. $2\sin^{-1}x=\sin^{-1}(2x\sqrt{(1-x^2)})$

Explanation: This is the double-angle corollary obtained by substituting $y=x$ into the $\sin^{-1}$ sum formula.

The domain of $\cot^{-1}x$ is:

A. $[-1,1]$

B. $(-\infty,\infty)$

C. $(0,\pi)$

D. $[0,\pi]$

Answer: B. $(-\infty,\infty)$

Explanation: $\cot x$ is restricted to $(0,\pi)$ to make it one-to-one, but $\cot^{-1}x$ itself accepts all real number inputs.

Quick Revision Summary

- An inverse trig function exists only after restricting the original function's domain to a principal interval where it is one-to-one
- $\sin^{-1}x$: domain $[-1,1]$, range $[-\pi/2,\pi/2]$ | $\cos^{-1}x$: domain $[-1,1]$, range $[0,\pi]$ | $\tan^{-1}x$: domain $(-\infty,\infty)$, range $(-\pi/2,\pi/2)$
- $\csc^{-1}x$ and $\sec^{-1}x$: domain $(-\infty,-1]\cup[1,\infty)$ | $\cot^{-1}x$: domain $(-\infty,\infty)$, range $(0,\pi)$
- Every inverse trig graph is the reflection of its restricted trig graph about the line $y=x$
- Reciprocal identities: $\csc^{-1}x=\sin^{-1}(1/x)$, $\sec^{-1}x=\cos^{-1}(1/x)$, $\cot^{-1}x=\tan^{-1}(1/x)$
- Complementary identities: $\sin^{-1}x+\cos^{-1}x=\pi/2$, $\tan^{-1}x+\cot^{-1}x=\pi/2$, $\sec^{-1}x+\csc^{-1}x=\pi/2$
- Domain of a composite inverse trig function is found by requiring the inner expression to lie within the outer function's domain
- Sum formula: $\sin^{-1}x+\sin^{-1}y=\sin^{-1}[x\sqrt{(1-y^2)}+y\sqrt{(1-x^2)}]$ | $\tan^{-1}x+\tan^{-1}y=\tan^{-1}[(x+y)/(1-xy)]$, $xy<1$
- Double-angle corollaries: $2\tan^{-1}x=\tan^{-1}(2x/(1-x^2))$, $2\sin^{-1}x=\sin^{-1}(2x\sqrt{(1-x^2)})$, $2\cos^{-1}x=\cos^{-1}(2x^2-1)$
- Reference right triangles are the fastest way to evaluate composite expressions like $\sec[\tan^{-1}(-\sqrt{3})]$ or $\sin[\frac{1}{2}\tan^{-1}(4/3)]$

Exam Tips

- Always confirm your final angle answer lies inside the correct principal range for that inverse function -- an otherwise correct calculation with the wrong-range angle is marked incorrect
- When evaluating a composite expression like $\sec[\tan^{-1}(-\sqrt{3})]$, sketch a quick reference right triangle first -- it is faster and less error-prone than working purely algebraically
- For domain questions on composite inverse trig functions, always start from the inequality $-1 \le (\text{inner expression}) \le 1$ (for $\sin^{-1}/\cos^{-1}$) or the matching condition for the other four functions
- Memorize the six principal domain/range pairs as a table -- confusing $\cos^{-1}$'s range $[0,\pi]$ with $\sin^{-1}$'s range $[-\pi/2,\pi/2]$ is the single most common exam mistake
- When applying a sum/difference formula, always check the $xy<1$ (or equivalent) condition before trusting the raw formula output -- a $\pm\pi$ correction may be needed otherwise
- Practice deriving the double-angle corollaries by substituting $y=x$ into the general sum formulas rather than memorizing them separately -- this reduces the total number of formulas you need to recall under exam pressure

Unit 10: Solution of Trigonometric Equations

A trigonometric equation involves one or more trigonometric functions (or their inverses) of an unknown angle, such as $\sin x = 1/\sqrt{2}$ or $2\cos^2\theta - 5\cos\theta + 1 = 0$. Because trigonometric functions are periodic, these equations have infinitely many solutions -- the solutions lying within one period, usually $[0, 2\pi]$, are called the principal solutions, and adding integer multiples of the period to a principal solution generates the full general solution.

This unit develops the algebraic technique for finding principal and general solutions using quadrant analysis and reference angles, including equations that require squaring both sides (and therefore checking for extraneous roots). It then introduces the graphical method for equations too complex to solve algebraically, by plotting both sides of the equation and reading off intersection points, and closes with real-life applications in engineering and architecture -- subtended-angle problems, projectile range, and angle-of-elevation problems -- solved using inverse trigonometric functions.

Learning Objectives

- Define a trigonometric equation, and distinguish between a principal solution and a general solution
- Use quadrant analysis and reference angles to find all principal solutions of $\sin\theta = k$, $\cos\theta = k$, or $\tan\theta = k$ in $[0, 2\pi]$
- Write the general solution of a trigonometric equation using the function's period
- Solve trigonometric equations that require squaring both sides, and correctly identify and discard extraneous roots
- Solve a trigonometric equation graphically by plotting both sides and reading off intersection points
- Apply general and graphical solution methods to real-life problems involving subtended angles, projectile motion, and angles of elevation/depression

Key Concepts

10.1 Trigonometric Equations, Principal and General Solutions

A trigonometric equation is an equation involving trigonometric functions (sine, cosine, tangent, etc.) or their inverses, such as $\sin x=1/\sqrt{2}$ or $2\tan^{-1}x=\pi/2$. An unknown angle satisfying the equation is called a solution; solutions lying in $[0,2\pi]$ are the principal solutions.

Because trigonometric functions repeat every period (2π for sin/cos, π for tan), once a principal solution is found, adding or subtracting integer multiples n of the period generates every other solution -- this full solution set is the general solution, e.g. for $\cos x=1/2$, the general solution is $x=\{\pi/3+2n\pi\}\cup\{5\pi/3+2n\pi\}$, $n\in\mathbb{Z}$.

10.2 Quadrant Analysis and Reference Angles

To solve $\sin\theta=k$, $\cos\theta=k$, or $\tan\theta=k$ algebraically, first determine which quadrants give the correct sign for the function (positive/negative), then use the reference angle (the related acute angle) to write both principal solutions in $[0,2\pi]$ -- e.g. for $\cos x=\sqrt{3}/2$ (positive), x lies in quadrants I and IV with reference angle $\pi/6$, giving $x=\pi/6$ and $x=2\pi-\pi/6=11\pi/6$.

More complex equations are first reduced to this basic form using algebraic manipulation (factoring, using identities to write everything in terms of one function) before quadrant analysis is applied -- e.g. $2\sec^2x-8/3=0$ reduces to $\cos x=\pm\sqrt{3}/2$, each case then solved separately by quadrant analysis.

10.3 Squaring and Extraneous Roots

Some equations, such as $\sqrt{3}\cot x-\csc x-1=0$, cannot be solved by direct quadrant analysis and instead require isolating a term and squaring both sides to eliminate a reciprocal or radical function. Squaring can introduce extraneous roots -- values that satisfy the squared equation but not the original one -- because squaring can make a true statement out of what was originally a false one (e.g. it discards sign information).

Every candidate solution obtained after squaring must therefore be substituted back into the ORIGINAL (unsquared) equation to verify it actually satisfies it; any candidate that fails this check is discarded as extraneous, and only the verified solutions are reported in the final general solution.

10.4 Graphical Solution of Trigonometric Equations

When an equation $f(x)=g(x)$ is too complex for algebraic methods (e.g. $\sin x=x/2$ or $\cos x=x/8$), the graphical method plots $y=f(x)$ and $y=g(x)$ on the same axes over a given interval; each point where the two curves intersect gives an x-value that is a solution of the equation.

This method does not give exact values but provides accurate decimal approximations and also reveals the total number of solutions in the interval at a glance -- e.g. $\sin x=x/2$ on $[-\pi,\pi]$ has exactly three intersection points, giving solutions $x\approx-1.89$, 0, and 1.89 radians.

10.5 Real-Life Applications

Subtended-angle problems (e.g. a lighthouse lantern viewed from a ship) use the difference formula for $\tan^{-1}$ to relate two angles of elevation measured from the same observation point to different heights on a structure, solving the resulting equation for an unknown height.

Projectile range problems use $R=(v_0^2/g)\sin 2\theta$ to relate launch angle to horizontal range, solved using $\sin^{-1}$ and the identity $\sin\theta=\sin(180°-\theta)$ to find both possible launch angles for a given target distance, while the maximum range is found by recognizing $\sin 2\theta$ is maximized (equal to 1) when $2\theta=90°$, i.e. $\theta=45°$.

Important Definitions

What is a trigonometric equation?

An equation involving one or more trigonometric functions (or their inverses) of an unknown angle, such as $\sin x=1/\sqrt{2}$.

What is a principal solution?

A solution of a trigonometric equation that lies within the interval $[0,2\pi]$.

What is a general solution?

The complete set of all solutions of a trigonometric equation, obtained by adding integer multiples of the function's period to each principal solution.

What is a reference angle?
The acute angle between the terminal side of an angle and the x-axis, used to find related angles with the same trigonometric ratio magnitude in other quadrants.

What is an extraneous root?
A value obtained while solving an equation (often after squaring both sides) that satisfies the transformed equation but does NOT satisfy the original equation.

What is the period of sinx and cosx?
2π.

What is the period of tanx?
π.

What is the graphical method for solving a trigonometric equation?
Plotting $y=f(x)$ and $y=g(x)$ for an equation $f(x)=g(x)$ and reading off the x-coordinates of their points of intersection as the solutions.

At what launch angle is projectile range $R=(v_0^2/g)\sin 2\theta$ maximized?
$\theta=45°$, since $\sin 2\theta$ reaches its maximum value of 1 when $2\theta=90°$.

Why must candidate roots be checked after squaring a trigonometric equation?
Because squaring can introduce extraneous roots that satisfy the squared equation but not the original one, so each candidate must be substituted back into the original equation to confirm validity.

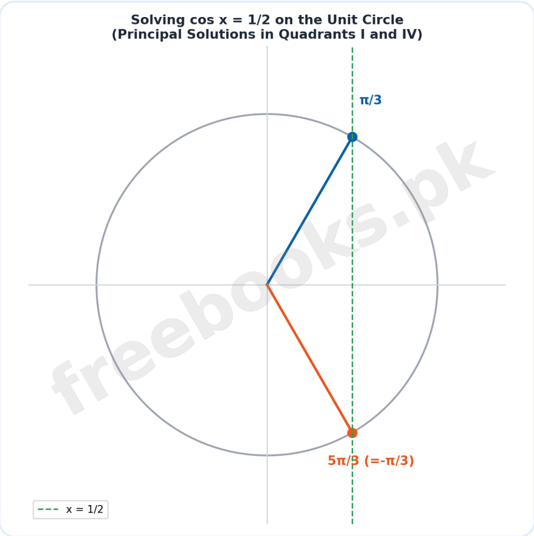
Key Facts and Relations

Topic	Key Fact / Relation
General solution of cosx=k	$x = \{\theta_r + 2n\pi\} \cup \{-\theta_r + 2n\pi\}$, $n\in\mathbb{Z}$, θ_r = reference angle
General solution of sinx=k	$x = \{\theta_r + 2n\pi\} \cup \{(\pi-\theta_r) + 2n\pi\}$, $n\in\mathbb{Z}$
General solution of tanx=k	$x = \theta_r + n\pi$, $n\in\mathbb{Z}$ (period π)

Topic	Key Fact / Relation
Reference-angle sign rule	Sign of the trig function determines the quadrant pair; the reference angle gives the acute angle in each
Difference formula for $\tan^{-1}$	$\tan^{-1}\alpha - \tan^{-1}\beta = \tan^{-1}[(\alpha-\beta)/(1+\alpha\beta)]$
Projectile range formula	$R = (v_0^2/g) \sin 2\theta$
Maximum projectile range	$R_{max} = v_0^2/g$, achieved at $\theta=45°$
Sine supplementary-angle identity	$\sin\theta = \sin(180°-\theta)$
Pythagorean identity (used after squaring)	$\sin^2 x + \cos^2 x = 1$
Angle of elevation relation	$\tan\theta$ = (height)/(horizontal distance)

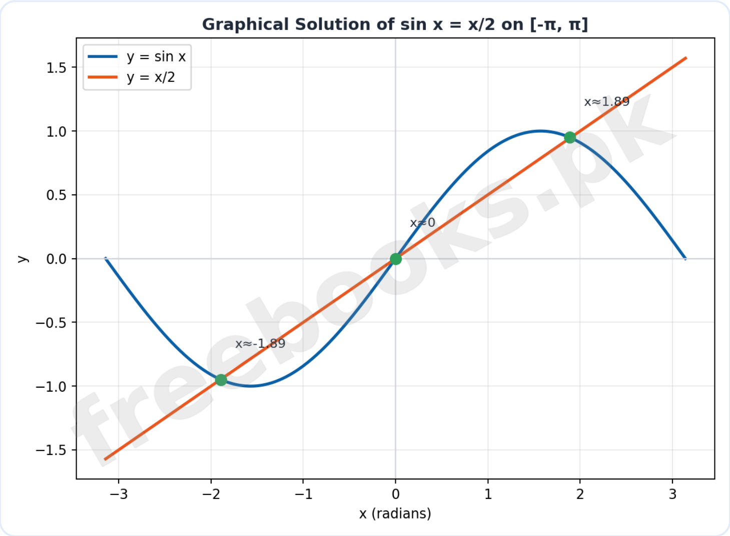
Diagrams

Solving cos x = 1/2 on the Unit Circle: The unit circle showing the two principal solutions $x=\pi/3$ and $x=5\pi/3$ for $\cos x=1/2$, found in quadrants I and IV using the reference angle $\pi/3$



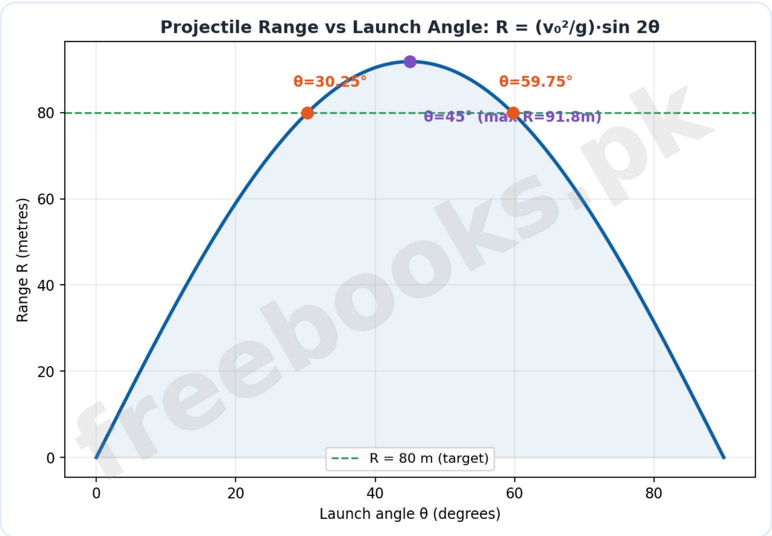
Solving cos x = 1/2 on the Unit Circle

Graphical Solution of sin x = x/2: The graphs of $y=\sin x$ and $y=x/2$ plotted together on $[-\pi,\pi]$, showing all three intersection points that give the equation's approximate solutions



Graphical Solution of sin x = x/2

Projectile Range vs Launch Angle: The range $R=(v_0^2/g)\sin 2\theta$ plotted against launch angle θ, showing the two angles that give a range of 80m and the single angle (45°) that gives the maximum range



Projectile Range vs Launch Angle

Solved Examples

Example 1: Solving a Basic Sine Equation

Problem: Solve $\sin x = -\sqrt{3}/2$.

1. Since sine is negative, x lies in quadrants III and IV, with reference angle $\pi/3$ (since $\sin(\pi/3)=\sqrt{3}/2$).

2. Quadrant III solution: $x = \pi + \pi/3 = 4\pi/3$.
3. Quadrant IV solution: $x = 2\pi - \pi/3 = 5\pi/3$.
4. Since 2π is the period of $\sin x$, the general solution adds $2n\pi$ to each principal solution.
5. Final answer: general solution = $\{4\pi/3 + 2n\pi\} \cup \{5\pi/3 + 2n\pi\}$, $n \in \mathbb{Z}$.

Example 2: Solving a Squared-Function Equation

Problem: Solve $2\sec^2 x - 8/3 = 0$.

1. Isolate $\sec^2 x$: $\sec^2 x = 4/3$, so $\sec x = \pm 2/\sqrt{3}$, giving $\cos x = \sqrt{3}/2$ or $\cos x = -\sqrt{3}/2$.
2. For $\cos x=\sqrt{3}/2$ (positive, quadrants I and IV, reference angle $\pi/6$): $x=\pi/6$ and $x=11\pi/6$.
3. For $\cos x=-\sqrt{3}/2$ (negative, quadrants II and III, reference angle $\pi/6$): $x=5\pi/6$ and $x=7\pi/6$.
4. Since 2π is the period of $\cos x$, add $2n\pi$ to each of the four principal solutions.
5. Final answer: general solution = $\{\pi/6+2n\pi\} \cup \{5\pi/6+2n\pi\} \cup \{7\pi/6+2n\pi\} \cup \{11\pi/6+2n\pi\}$, $n \in \mathbb{Z}$.

Example 3: Solving an Equation Requiring Squaring, with Extraneous Root Check

Problem: Solve $\sqrt{3} \cot x - \csc x - 1 = 0$.

1. Rewrite in terms of sin and cos, isolate the radical-free term: $\sqrt{3}\cos x - 1 = \sin x$, then square both sides: $(\sqrt{3}\cos x-1)^2 = \sin^2 x$.
2. Expand and use $\sin^2 x=1-\cos^2 x$ to reduce to $4\cos^2 x - 2\sqrt{3}\cos x = 0$, which factors as $2\cos x(2\cos x-\sqrt{3})=0$, giving $\cos x=0$ or $\cos x=\sqrt{3}/2$.
3. For $\cos x=0$: candidates $x=\pi/2$ and $x=3\pi/2$. Checking in the ORIGINAL equation: $x=\pi/2$ gives LHS$=-2\neq 0$ (extraneous); $x=3\pi/2$ gives LHS$=0=$RHS (valid).

4. For $\cos x=\sqrt{3}/2$: candidates $x=\pi/6$ and $x=11\pi/6$. Checking in the original equation: $x=\pi/6$ gives LHS=0=RHS (valid); $x=11\pi/6$ gives LHS=$-2\neq0$ (extraneous).
5. Final answer: general solution = $\{3\pi/2 + 2n\pi\} \cup \{\pi/6 + 2n\pi\}$, $n\in\mathbb{Z}$ (the two extraneous roots are discarded).

Example 4: Graphical Solution of an Equation

Problem: Solve $\sin x = x/2$ graphically over $[-\pi, \pi]$.

1. Recognize this equation cannot be solved algebraically in closed form since x appears both inside and outside the sine function.
2. Plot $y=\sin x$ and $y=x/2$ on the same axes over the interval $[-\pi,\pi]$.
3. Identify each point where the two curves cross -- reading the graph shows three intersection points.
4. Read off the approximate x-coordinates of these three intersection points.
5. Final answer: $x \approx -1.89$, 0, or 1.89 radians.

Example 5: Graphical Solution with a Wider Search Interval

Problem: Solve $\cos x = x/8$ graphically.

1. Since $-1\leq\cos x\leq1$, any solution must satisfy $-1\leq x/8\leq1$, i.e. $-8\leq x\leq8$, so this is the natural search interval.
2. Plot $y=\cos x$ and $y=x/8$ together over $[-8,8]$.
3. Count and locate each intersection point of the two curves -- the graph shows six intersection points in this interval.
4. Read off the approximate x-coordinate of each intersection point.
5. Final answer: $x \approx -4.16$, -1.8, 0, 1.39, 5.46, and 6.82 radians.

Example 6: Real-Life Application: Subtended Angle of a Lighthouse Lantern

Problem: A 15m lantern sits atop a cliff. From a ship 120m offshore, the lantern subtends the same angle as a 1.5m buoy at the base of the cliff. Find the height of the cliff.

1. Let h be the cliff height. From the ship, the angle to the cliff top is $\theta_1=\tan^{-1}(h/120)$, and the angle to the lantern top is $\theta=\tan^{-1}[(h+15)/120]$.
2. The lantern's subtended angle is $\theta_2=\theta-\theta_1 = \tan^{-1}[(h+15)/120] - \tan^{-1}(h/120)$. The buoy's subtended angle is $\theta_3=\tan^{-1}(1.5/120)$.
3. Since $\theta_2=\theta_3$, apply the $\tan^{-1}$ difference formula to combine the left side into a single inverse tangent, then equate arguments: 15/120 divided by $(14400+h^2+15h)/14400 = 1.5/120$.
4. Simplify to the quadratic $h^2 + 15h - 129600 = 0$, then solve with the quadratic formula: $h = [-15 + \sqrt{(225+518400)}]/2 = [-15+720.2]/2$.
5. Final answer: the cliff height is $h \approx 352.6$ metres.

Example 7: Real-Life Application: Projectile Launch Angle

Problem: A projectile launched at $v_0=30$ m/s must hit a target 80m away. Find the launch angle θ, using $R=(v_0^2/g)\sin 2\theta$ with $g=9.8$ m/s².

1. Substitute the known values: $80 = (30^2/9.8)\sin 2\theta = (900/9.8)\sin 2\theta$, so $\sin 2\theta = 80\times 9.8/900 = 784/900 \approx 0.8711$.
2. Take $\sin^{-1}$ of both sides: $2\theta = \sin^{-1}(0.8711) \approx 60.5°$, giving $\theta = 30.25°$.
3. Since $\sin\theta=\sin(180°-\theta)$, a second solution exists: $2\theta = 180°-60.5° = 119.5°$, giving $\theta = 59.75°$.
4. Both angles are valid launch angles that produce the same 80m range (a low, flat trajectory and a high, steep trajectory).
5. Final answer: $\theta \approx 30.25°$ or $\theta \approx 59.75°$.

Example 8: Real-Life Application: Maximum Projectile Range

Problem: For the same projectile (v_0=30 m/s, g=9.8 m/s^2), find the launch angle that gives the maximum range, and compute that maximum range.

1. The range $R=(v_0^2/g)\sin2\theta$ is maximized when $\sin2\theta$ is at its maximum possible value.
2. Since the maximum value of the sine function is 1, set $\sin2\theta=1$, giving $2\theta=\sin^{-1}(1)=90°$.
3. Solve for θ: $\theta = 90°/2 = 45°$.
4. Substitute θ=45° (or directly use $\sin2\theta=1$) into the range formula: $R_{max} = v_0^2/g = 900/9.8$.
5. Final answer: the maximum range occurs at θ=45°, giving $R_{max} \approx 91.84$ metres.

Short Questions & Answers

What distinguishes a principal solution from a general solution?

A principal solution lies only within [0,2π], while the general solution includes every solution obtained by adding integer multiples of the period to each principal solution.

Why does squaring both sides of a trigonometric equation risk introducing extraneous roots?

Because squaring discards sign information, so a value that makes the squared equation true might not actually satisfy the original (unsquared) equation.

How is an extraneous root identified and handled?

By substituting the candidate solution back into the ORIGINAL equation; if it does not satisfy the original equation, it is discarded as extraneous.

What is the main advantage of the graphical method for solving trigonometric equations?

It can find approximate solutions for equations too complex for algebraic methods, and it also reveals the total number of solutions in a given interval at a glance.

In quadrant analysis, what determines which two quadrants contain the solutions?

The sign (positive or negative) of the trigonometric function value in the equation, since each function is positive in two quadrants and negative in the other two.

At what angle is projectile range $R=(v_0^2/g)\sin 2\theta$ maximized, and why?

$\theta=45°$, because $\sin 2\theta$ reaches its maximum possible value of 1 when $2\theta=90°$.

Why can a target range typically be hit at two different launch angles?

Because $\sin\theta=\sin(180°-\theta)$, so the equation $\sin 2\theta=k$ generally has two solutions for 2θ in $[0°,180°]$, giving two valid values of θ.

Long Questions & Answers

Explain the complete process for finding the general solution of a trigonometric equation like cos x = k, including how quadrant analysis, reference angles, and the function's period are used.

How is the correct pair of quadrants identified for cosx=k?

The sign of k determines the quadrants: if k is positive, cosx is positive in quadrants I and IV; if k is negative, cosx is negative in quadrants II and III -- this narrows the search to exactly two quadrants within $[0,2\pi]$.

How is the reference angle used to find both principal solutions?

The reference angle θ_r is the acute angle satisfying $\cos(\theta_r)=|k|$; the two principal solutions are then obtained by placing θ_r correctly in each of the two identified quadrants, e.g. for quadrant I the solution is simply θ_r, and for quadrant IV it is $2\pi-\theta_r$.

How is the period of cosx used to extend the principal solutions to a general solution?

Since cosx repeats every 2π, adding any integer multiple $2n\pi$ ($n\in\mathbb{Z}$) to each principal solution produces another valid solution, so the general solution is written as the union of $\{\theta_r+2n\pi\}$ and $\{-\theta_r+2n\pi\}$ (or equivalently $\{2\pi-\theta_r+2n\pi\}$), covering every possible solution.

How does this process change for sinx=k or tanx=k?

For sinx=k, the two quadrants are I/II (positive) or III/IV (negative), and the period is also 2π; for tanx=k, the two quadrants are I/III (positive) or II/IV (negative), but since tanx has period π (not 2π), only a single principal solution plus nπ is needed to generate the full general solution.

Explain how the graphical method solves trigonometric equations that resist algebraic methods, and how it was applied to find the launch angles and maximum range of a projectile.

Why is the graphical method needed for an equation like sinx=x/2?

Because x appears both inside the trigonometric function and as a plain algebraic term, there is no algebraic manipulation that isolates x exactly -- the equation is transcendental, so no closed-form solution exists, making a visual/numerical approach necessary.

What is the general procedure for solving f(x)=g(x) graphically?

Plot y=f(x) and y=g(x) on the same set of axes over the interval of interest, then identify every point where the two curves cross; the x-coordinate of each crossing point is an approximate solution to the equation, and the total number of crossings gives the number of solutions in that interval.

How does the projectile range formula $R=(v_0^2/g)\sin 2\theta$ connect to solving for a launch angle?

Setting R equal to a target distance and rearranging gives $\sin 2\theta = (Rg)/v_0^2$, an equation solved not by plotting but by taking $\sin^{-1}$ of both sides directly, since it reduces to the standard form sin(angle)=k; because sinθ=sin(180°-θ), this generally yields two valid launch angles for the same target range.

How is the maximum-range angle found, and why is it unique?

Since sin2θ can never exceed 1, the range is maximized exactly when sin2θ=1, which happens only at 2θ=90°, i.e. θ=45° within the physically meaningful range of launch angles (0° to 90°) -- making 45° the single unique angle that achieves the maximum possible range for a given launch speed.

Multiple Choice Questions (MCQs)

Solutions of a trigonometric equation lying within $[0,2\pi]$ are called:

A. General solutions

B. Principal solutions

C. Extraneous solutions

D. Reference solutions

Answer: B. Principal solutions

Explanation: Principal solutions are, by definition, the solutions restricted to one full period $[0,2\pi]$.

The general solution of $\cos x=1/2$ is:

A. $x=\pi/3+n\pi$

B. $x=\{\pi/3+2n\pi\}\cup\{5\pi/3+2n\pi\}$

C. $x=\pi/3+2n\pi$ only

D. $x=5\pi/3-n\pi$

Answer: B. $x=\{\pi/3+2n\pi\}\cup\{5\pi/3+2n\pi\}$

Explanation: Both principal solutions $\pi/3$ and $5\pi/3$ must be extended by adding $2n\pi$ (the period of cosine) to each.

An extraneous root arises most often from which operation?

A. Adding a constant

B. Squaring both sides of an equation

C. Taking a reciprocal

D. Multiplying by a positive constant

Answer: B. Squaring both sides of an equation

Explanation: Squaring can turn a false statement into a true one, introducing solutions that don't satisfy the original equation.

How should a suspected extraneous root be checked?

A. Substitute it into the squared equation

B. Substitute it into the original (unsquared) equation

C. Ignore it if it's negative

D. Round it to the nearest integer

Answer: B. Substitute it into the original (unsquared) equation

Explanation: Only substitution into the ORIGINAL equation correctly confirms or rejects a candidate root.

The graphical method finds solutions of f(x)=g(x) by locating:

A. The maximum of f(x)

B. Points where the two graphs intersect

C. The y-intercepts of both graphs

D. The period of f(x)

Answer: B. Points where the two graphs intersect

Explanation: Each intersection point's x-coordinate is a solution of f(x)=g(x).

The period of tanx used when writing its general solution is:

A. 2π

B. π

C. π/2

D. 4π

Answer: B. π

Explanation: Unlike sine and cosine, tangent repeats every π, so only nπ (not 2nπ) is added to its principal solution.

The projectile range formula $R=(v_0^2/g)\sin 2\theta$ is maximized when:

A. θ=0°

B. θ=90°

C. θ=45°

D. θ=60°

Answer: C. θ=45°

Explanation: sin2θ reaches its maximum value of 1 when 2θ=90°, i.e. θ=45°.

Why can a projectile hit a target at two different launch angles?

A. Because g is not constant

B. Because sinθ=sin(180°-θ) gives two valid angles for the same sin2θ value

C. Because v_0 changes with angle

D. It cannot -- only one angle is ever possible

Answer: B. Because sinθ=sin(180°-θ) gives two valid angles for the same sin2θ value

Explanation: The supplementary-angle sine identity means sin2θ=k has two solutions for 2θ in [0°,180°].

For sinx=k with k negative, the solutions lie in which quadrants?

A. I and II

B. I and IV

C. III and IV

D. II and III

Answer: C. III and IV

Explanation: Sine is negative in quadrants III and IV.

In the equation $\sqrt{3}\cot x - \csc x - 1 = 0$, why is squaring used?

A. To simplify the coefficients

B. To eliminate the reciprocal trig functions and reduce to a solvable polynomial in cosx

C. To find the period

D. Squaring is not used in this example

Answer: B. To eliminate the reciprocal trig functions and reduce to a solvable polynomial in cosx

Explanation: Squaring both sides after isolating terms converts the equation into a polynomial equation in cosx that can be factored and solved.

Quick Revision Summary

- Trigonometric equation: an equation in one or more trig functions or their inverses; solution = angle satisfying it
- Principal solution: lies in $[0,2\pi]$ | General solution: principal solution(s) + integer multiples of the period
- Quadrant analysis: sign of k determines the quadrant pair; reference angle gives the acute angle within each
- Periods used in general solutions: sinx and cosx use $2n\pi$; tanx uses $n\pi$
- Squaring both sides can introduce extraneous roots -- always verify every candidate in the ORIGINAL equation
- Graphical method: plot both sides of $f(x)=g(x)$; intersection points' x-coordinates are the approximate solutions
- Subtended-angle problems use the $\tan^{-1}$ difference formula to relate two angles of elevation from one observation point
- Projectile range: $R=(v_0^2/g)\sin 2\theta$; two launch angles usually give the same range ($\sin\theta=\sin(180°-\theta)$)
- Maximum projectile range occurs at $\theta=45°$, since $\sin 2\theta$'s maximum value of 1 occurs at $2\theta=90°$

- Angle of elevation/depression problems use $\tan\theta = (\text{height})/(\text{horizontal distance})$ as the basic relation

Exam Tips

- Always determine the correct pair of quadrants FIRST (from the sign of k) before computing the reference angle -- working in the wrong quadrant pair is the most common error
- When an equation involves squaring, always circle or list every candidate solution before checking, then substitute each one individually into the ORIGINAL equation -- do not skip this verification step
- For graphical-method questions, sketch both curves as accurately as possible and count intersection points carefully -- missing an intersection near the edge of the given interval is a common mistake
- Remember $\tan x$ has period π, not 2π -- using $2n\pi$ instead of $n\pi$ for a tangent equation's general solution will produce an incomplete or incorrect answer
- For real-life subtended-angle problems, draw a clear diagram labeling every triangle and angle before writing any equation -- it prevents mixing up which angle belongs to which triangle
- For projectile problems, remember the supplementary-angle identity $\sin\theta = \sin(180° - \theta)$ whenever an inverse-sine step is involved -- it is the source of the second (usually higher) launch angle

Unit 11: Vector Valued Functions and Their Differentiations

A scalar function assigns a single real number to each input, like temperature at a point. A vector-valued function instead assigns a vector -- a quantity with both magnitude and direction -- making it the natural tool for modelling position, velocity, or force in two or three dimensions. Written as $r(t)=f(t)\hat{i}+g(t)\hat{j}+h(t)\hat{k}$, each component is itself an ordinary scalar function of the parameter t (often time), and together the three components trace out a curve in space as t varies.

This unit defines scalar and vector-valued functions, shows how to construct a vector-valued function from a real-world description, finds the domain and range by intersecting the domains of the individual components, and develops the derivative $r'(t)$ (found by differentiating each component separately) as the key tool for analyzing motion. Building on this, it defines velocity as $r'(t)$, acceleration as $r''(t)$, speed as the magnitude of velocity, and direction of motion as the unit tangent vector, applying these to real-life problems in robotics, drones, projectiles, and orbital motion.

Learning Objectives

- Distinguish between a scalar function and a vector-valued function, with examples of each
- Construct a vector-valued function $r(t)$ from a description of how each coordinate changes with t
- Find the domain of a vector-valued function by intersecting the domains of its component functions
- Find the range of a vector-valued function given its component functions
- Differentiate a vector-valued function by differentiating each component separately
- Define and compute the velocity vector $v(t)=r'(t)$ and acceleration vector $a(t)=r''(t)$ of a moving particle
- Compute the speed of a particle as the magnitude of its velocity vector, and its direction of motion as the unit tangent vector

- Apply vector-valued functions and their derivatives to real-life motion problems in robotics, drones, projectiles, and orbits

Key Concepts

11.1 Scalar Functions and Vector-Valued Functions

A scalar function assigns a single real number to each value of its input variable, e.g. $f(t)=3t+2\cos t$; it is typically denoted by a lowercase letter and describes quantities like temperature or potential energy that have magnitude only, no direction.

A vector-valued function of a single variable takes one input (usually t) and returns a vector, written $r(t)=\langle f(t),g(t),h(t)\rangle=f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$, where each component $f(t)$, $g(t)$, $h(t)$ is itself a scalar function. It is typically denoted by a boldface or arrow-topped letter and describes quantities like position, velocity, or force that have both magnitude and direction.

11.2 Constructing a Vector-Valued Function

To build a vector-valued function from a real-world description: first determine how each coordinate (x, y, and if needed z) changes with respect to t; then combine these component functions into a single vector using either angle-bracket notation $\langle f(t),g(t),h(t)\rangle$ or unit-vector notation $f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$; finally specify the domain of t appropriate to the physical situation (e.g. $t\geq 0$ for time-dependent motion starting at $t=0$).

Common physical examples include projectile motion, $r(t)=(v_0\cos\theta)t\hat{\imath}+[(v_0\sin\theta)t-\frac{1}{2}gt^2]\hat{\jmath}$, where the x-component models constant horizontal speed and the y-component models vertical motion under gravity; and helical motion, $r(t)=R\cos t\,\hat{\imath}+R\sin t\,\hat{\jmath}+ct\,\hat{k}$, where the x,y-components trace a circle of radius R and the z-component models steady linear ascent.

11.3 Domain and Range of a Vector-Valued Function

The domain of $r(t)=f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$ is the set of all t for which every component function is simultaneously defined -- found by determining the domain of each component separately, then taking the intersection of all these domains.

The range of r(t) is the set of all possible output vectors as t varies over the domain; it is described component-wise by finding the range of each individual component function (given the domain restriction) and expressing the overall output as an ordered tuple ⟨x,y⟩ or ⟨x,y,z⟩ subject to those component ranges.

11.4 The Derivative of a Vector-Valued Function

The derivative of r(t) is defined by the same limit-of-difference-quotient idea as for scalar functions: $r'(t)=\lim_{\Delta t\to 0}[r(t+\Delta t)-r(t)]/\Delta t$. When this limit exists, it can be computed simply by differentiating each component separately: $r'(t)=f_1'(t)\hat{i}+f_2'(t)\hat{j}+f_3'(t)\hat{k}$.

This componentwise rule means every differentiation technique already known for scalar functions (product rule, chain rule, quotient rule, trig/exponential/log derivatives) applies directly to each component of a vector-valued function -- no new differentiation rules are needed, only careful bookkeeping across the components.

11.5 Velocity, Acceleration, Speed, and Direction of Motion

If r(t) is a particle's position vector, its velocity vector is $v(t)=r'(t)$, which is tangent to the path of motion at every point; its acceleration vector is $a(t)=v'(t)=r''(t)$, the second derivative of position. Both are themselves vector-valued functions of t.

The particle's speed at a given instant is the magnitude of its velocity vector, $|v(t)|$, a scalar; its direction of motion is given by the unit tangent vector $\hat{v}(t)=v(t)/|v(t)|$, a vector of magnitude 1 pointing in the instantaneous direction of travel. Together these four quantities -- velocity, acceleration, speed, and direction -- give a complete instantaneous description of a particle's motion along its path.

Important Definitions

What is a scalar function?

A function that assigns a single real number to each value of the input variable, such as $f(t)=3t+2\cos t$.

What is a vector-valued function?

A function of a single variable that returns a vector as output, written $r(t)=f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$, where each component is a scalar function of t.

What is the domain of a vector-valued function r(t)?

The set of all t for which every component function of r(t) is simultaneously defined -- the intersection of each component's individual domain.

What is the range of a vector-valued function?

The set of all possible output vectors r(t) as t varies over the domain.

How is the derivative of a vector-valued function computed?

By differentiating each component function separately: $r'(t)=f_1'(t)\hat{\imath}+f_2'(t)\hat{\jmath}+f_3'(t)\hat{k}$.

What is the velocity vector of a moving particle?

$v(t)=r'(t)$, the derivative of the position vector, tangent to the particle's path of motion.

What is the acceleration vector of a moving particle?

$a(t)=v'(t)=r''(t)$, the derivative of the velocity vector (second derivative of position).

What is the speed of a particle?

The magnitude of its velocity vector, $|v(t)|$, a scalar quantity.

What is the unit tangent vector, and what does it represent?

$\hat{v}(t)=v(t)/|v(t)|$, a vector of magnitude 1 that gives the particle's instantaneous direction of motion.

Give an example of a helical vector-valued function.

$r(t)=R\cos t\ \hat{\imath} + R\sin t\ \hat{\jmath} + ct\ \hat{k}$, which traces a circle of radius R in the xy-plane while climbing steadily along the z-axis at rate c.

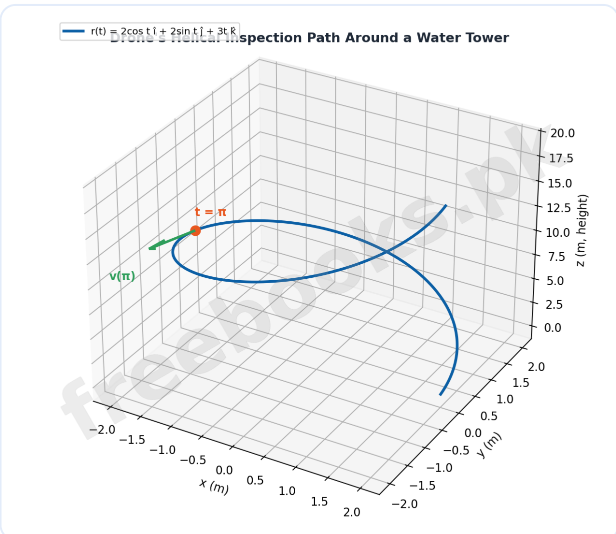
Key Facts and Relations

Topic	Key Fact / Relation
Vector-valued function (component form)	$r(t) = \langle f(t), g(t), h(t)\rangle = f(t)\hat{\imath} + g(t)\hat{\jmath} + h(t)\hat{k}$
	$r'(t) = f_1'(t)\hat{\imath} + f_2'(t)\hat{\jmath} + f_3'(t)\hat{k}$

Topic	Key Fact / Relation
Derivative of a vector-valued function	
Derivative as a limit	$r'(t) = \lim(\Delta t \to 0)\ [r(t+\Delta t) - r(t)] / \Delta t$
Velocity vector	$v(t) = r'(t)$
Acceleration vector	$a(t) = v'(t) = r''(t)$
Speed	$\lvert v(t)\rvert = \sqrt{[(f_1'(t))^2 + (f_2'(t))^2 + (f_3'(t))^2]}$
Unit tangent vector (direction of motion)	$\hat{v}(t) = v(t) / \lvert v(t)\rvert$
Projectile motion position vector	$r(t) = (v_0\cos\theta)t\ \hat{\imath} + [(v_0\sin\theta)t - \frac{1}{2}gt^2]\ \hat{\jmath}$
Helical motion position vector	$r(t) = R\cos t\ \hat{\imath} + R\sin t\ \hat{\jmath} + ct\ \hat{k}$
Domain of a vector-valued function	Domain[r(t)] = Domain[f(t)] ∩ Domain[g(t)] ∩ Domain[h(t)]

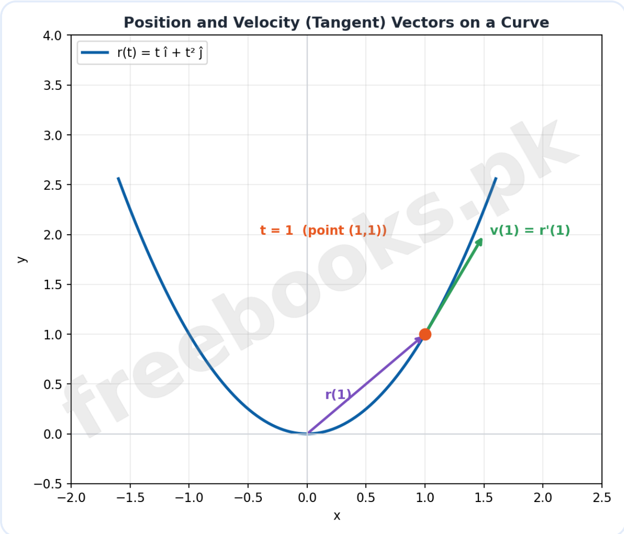
Diagrams

Drone's Helical Inspection Path Around a Water Tower: The 3D helix traced by $r(t)=2\cos t\ \hat{\imath}+2\sin t\ \hat{\jmath}+3t\ \hat{k}$, with the velocity vector shown tangent to the path at $t=\pi$



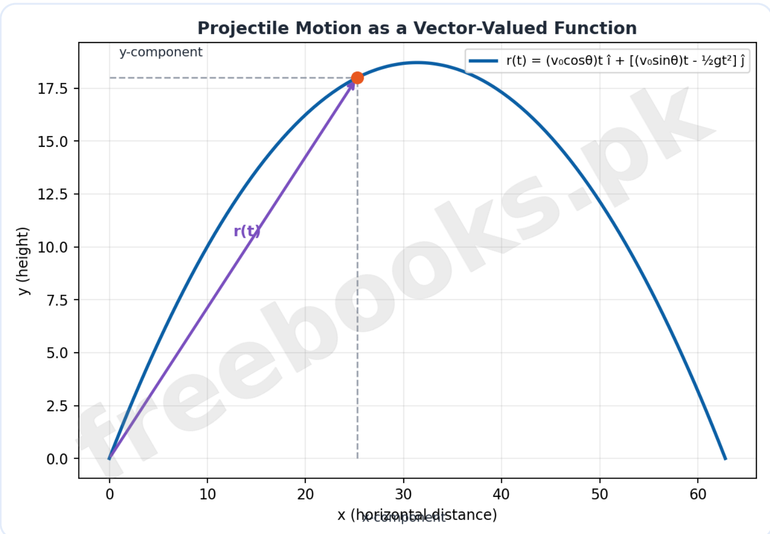
Drone's Helical Inspection Path Around a Water Tower

Position and Velocity (Tangent) Vectors on a Curve: The curve $r(t)=t\,\hat{i}+t^2\,\hat{j}$ with the position vector r(1) from the origin and the tangent velocity vector v(1)=r'(1) shown at the point (1,1)



Position and Velocity (Tangent) Vectors on a Curve

Projectile Motion as a Vector-Valued Function: The parabolic trajectory of a projectile with its position vector r(t) shown decomposed into horizontal (x) and vertical (y) components at a point along the path



Projectile Motion as a Vector-Valued Function

Solved Examples

Example 1: Constructing and Evaluating a Vector-Valued Function

Problem: For the vector-valued function $r(t)=\langle \cos t, \sin t, t\rangle$, find $r(\pi/2)$ and describe the type of curve it traces.

1. Substitute $t=\pi/2$ into each component: $\cos(\pi/2)=0$, $\sin(\pi/2)=1$, and the third component is simply $\pi/2$.
2. Combine the results: $r(\pi/2) = \langle 0, 1, \pi/2\rangle$.
3. This vector represents a specific point in 3D space -- the drone or particle's location at that instant.
4. Since the x,y-components trace a unit circle while the z-component increases linearly and steadily with t, the overall curve spirals upward at a constant rate.
5. Final answer: $r(\pi/2)=\langle 0,1,\pi/2\rangle$, and the function traces a helix (a 3D spiral).

Example 2: Finding the Domain and Range of a Vector-Valued Function

Problem: Find the domain and range of $r(t) = \sqrt{4-t^2}\,\hat{\imath} + 1/(t-1)\,\hat{\jmath}$.

1. For the first component $\sqrt{4-t^2}$, require $4-t^2\geq 0$, i.e. $t^2\leq 4$, giving $t\in[-2,2]$.
2. For the second component $1/(t-1)$, require the denominator nonzero: $t-1\neq 0$, i.e. $t\neq 1$.
3. Intersect both conditions: domain = $[-2,1) \cup (1,2]$.
4. For the range: since $4-t^2$ ranges over $[0,4]$ on this domain, $\sqrt{4-t^2}$ ranges over $[0,2]$, so $x\in[0,2]$; since $1/(t-1)$ can equal any nonzero real number as t varies (excluding $t=1$), y can be any real number except 0.
5. Final answer: domain = $[-2,1)\cup(1,2]$; range = $\{\langle x,y\rangle \mid 0\leq x\leq 2, y\in\mathbb{R}, y\neq 0\}$.

Example 3: Differentiating a Vector-Valued Function

Problem: Find $r'(t)$ for $r(t) = t^2\hat{\imath} + \sin t\,\hat{\jmath} + e^t\hat{k}$.

1. Differentiate the first component: $d/dt(t^2) = 2t$.
2. Differentiate the second component: $d/dt(\sin t) = \cos t$.
3. Differentiate the third component: $d/dt(e^t) = e^t$.
4. Combine the three derivatives into a single vector-valued derivative.
5. Final answer: $r'(t) = 2t\,\hat{\imath} + \cos t\,\hat{\jmath} + e^t\,\hat{k}$.

Example 4: Evaluating a Derivative at a Specific Point

Problem: Find $r'(\pi/4)$ for $r(t) = \sec t\,\hat{\imath} + \tan t\,\hat{\jmath} - \cos t\,\hat{k}$.

1. Differentiate each component symbolically: $d/dt(\sec t)=\sec t \tan t$, $d/dt(\tan t)=\sec^2 t$, $d/dt(-\cos t)=\sin t$.
2. Evaluate the first at $t=\pi/4$: $\sec(\pi/4)\tan(\pi/4) = \sqrt{2} \times 1 = \sqrt{2}$.
3. Evaluate the second at $t=\pi/4$: $\sec^2(\pi/4) = (\sqrt{2})^2 = 2$.
4. Evaluate the third at $t=\pi/4$: $\sin(\pi/4) = 1/\sqrt{2}$.
5. Final answer: $r'(\pi/4) = \sqrt{2}\,\hat{\imath} + 2\,\hat{\jmath} + (1/\sqrt{2})\,\hat{k}$.

Example 5: Velocity, Acceleration, Speed, and Direction on a Conveyor Belt

Problem: A part's position on a vibrating conveyor belt is $r(t)=3t^2\hat{\imath}+2\sin t\,\hat{\jmath}$ (metres). Find its velocity, acceleration, speed, and direction of motion at $t=\pi/2$.

1. Velocity: $v(t)=r'(t)=6t\hat{\imath}+2\cos t\,\hat{\jmath}$; at $t=\pi/2$, $v(\pi/2)=6(\pi/2)\hat{\imath}+2\cos(\pi/2)\hat{\jmath} = 3\pi\hat{\imath} + 0\hat{\jmath}$.
2. Acceleration: $a(t)=v'(t)=6\hat{\imath}-2\sin t\,\hat{\jmath}$; at $t=\pi/2$, $a(\pi/2)=6\hat{\imath}-2\sin(\pi/2)\hat{\jmath} = 6\hat{\imath} - 2\hat{\jmath}$.
3. Speed: $|v(t)|=\sqrt{36t^2+4\cos^2 t}$; at $t=\pi/2$, $|v(\pi/2)|=\sqrt{(3\pi)^2+0}=3\pi \approx 9.42$ m/s.

4. Direction: $\hat{v}(\pi/2)=v(\pi/2)/|v(\pi/2)| = (3\pi\hat{\imath}+0\hat{\jmath})/(3\pi) = \hat{\imath} + 0\hat{\jmath}$.
5. Final answer: at $t=\pi/2$ the part moves due east at $3\pi\approx 9.42$ m/s, is accelerating eastward and starting to accelerate southward, though its instantaneous velocity has no north-south component.

Example 6: Velocity, Acceleration, Speed, and Direction for a Drone's Helical Path

Problem: A drone inspecting a water tower follows $r(t)=2\cos t\,\hat{\imath}+2\sin t\,\hat{\jmath}+3t\,\hat{k}$. Find its velocity, acceleration, speed, and direction of motion at $t=\pi$.

1. Velocity: $v(t)=r'(t)=-2\sin t\,\hat{\imath}+2\cos t\,\hat{\jmath}+3\hat{k}$; at $t=\pi$, $v(\pi)=-2\sin(\pi)\hat{\imath}+2\cos(\pi)\hat{\jmath}+3\hat{k} = 0\hat{\imath} - 2\hat{\jmath} + 3\hat{k}$.
2. Acceleration: $a(t)=v'(t)=-2\cos t\,\hat{\imath}-2\sin t\,\hat{\jmath}$; at $t=\pi$, $a(\pi)=-2\cos(\pi)\hat{\imath}-2\sin(\pi)\hat{\jmath} = 2\hat{\imath} + 0\hat{\jmath} + 0\hat{k}$.
3. Speed: $|v(t)|=\sqrt{4\cos^2 t+4\sin^2 t+9}=\sqrt{13}$ (constant for all t, since the trig terms simplify to 4).
4. Direction: $\hat{v}(\pi) = (0\hat{\imath}-2\hat{\jmath}+3\hat{k})/\sqrt{13} = 0\hat{\imath} - (2/\sqrt{13})\hat{\jmath} + (3/\sqrt{13})\hat{k}$.
5. Final answer: at $t=\pi$ the drone has no east-west motion, moves south and climbs simultaneously at constant speed $\sqrt{13}\approx 3.61$ m/s, with acceleration pointing eastward toward the tower's central axis (centripetal acceleration).

Example 7: Real-Life Application: Cannonball Projectile Motion

Problem: A cannonball is fired at 30° with initial velocity 100 m/s ($g=9.8$ m/s²). Formulate $r(t)$ and find the velocity vector at $t=5$ seconds.

1. Use the projectile position formula: $r(t)=(v_0\cos\theta)t\hat{\imath}+[(v_0\sin\theta)t-\frac{1}{2}gt^2]\hat{\jmath}$, with $v_0=100$, $\theta=30°$, $g=9.8$.
2. Substitute: $r(t) = (100\cos 30°)t\hat{\imath} + [(100\sin 30°)t - 4.9t^2]\hat{\jmath} = (86.6t)\hat{\imath} + (50t - 4.9t^2)\hat{\jmath}$.
3. Differentiate to get velocity: $v(t) = r'(t) = 86.6\hat{\imath} + (50 - 9.8t)\hat{\jmath}$.

4. Evaluate at $t=5$: $v(5) = 86.6\hat{\imath} + (50 - 9.8\times5)\hat{\jmath} = 86.6\hat{\imath} + (50-49)\hat{\jmath} = 86.6\hat{\imath} + 1\hat{\jmath}$.
5. Final answer: $r(t)=86.6t\hat{\imath}+(50t-4.9t^2)\hat{\jmath}$, and $v(5)\approx86.6\hat{\imath}+1\hat{\jmath}$ (m/s) -- the cannonball is nearly at the peak of its trajectory at t=5s, moving almost purely horizontally.

Example 8: Real-Life Application: Comet's Elliptical Orbit

Problem: A comet's position (in AU) is $x=2\cos t$, $y=\sin t$. Write $r(t)$, and find its velocity and speed at $t=\pi/2$.

1. Combine the components into a vector-valued function: $r(t) = 2\cos t\,\hat{\imath} + \sin t\,\hat{\jmath}$.
2. Differentiate to find velocity: $v(t) = r'(t) = -2\sin t\,\hat{\imath} + \cos t\,\hat{\jmath}$.
3. Evaluate at $t=\pi/2$: $v(\pi/2) = -2\sin(\pi/2)\hat{\imath} + \cos(\pi/2)\hat{\jmath} = -2\hat{\imath} + 0\hat{\jmath}$.
4. Compute speed: $|v(\pi/2)| = \sqrt{(-2)^2+0^2} = \sqrt{4} = 2$ AU/year.
5. Final answer: $r(t)=2\cos t\,\hat{\imath}+\sin t\,\hat{\jmath}$; at $t=\pi/2$, $v(\pi/2)=-2\hat{\imath}$ (AU/year), speed = 2 AU/year, moving due west (negative x-direction) at that instant.

Short Questions & Answers

What is the key difference between a scalar function and a vector-valued function?

A scalar function outputs a single real number, while a vector-valued function outputs a vector with both magnitude and direction.

How is the derivative of $r(t)=f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$ computed?

By differentiating each component function separately: $r'(t)=f'(t)\hat{\imath}+g'(t)\hat{\jmath}+h'(t)\hat{k}$.

How is the domain of a vector-valued function found?

By finding the domain of each component function individually, then taking the intersection of all these domains.

What does the velocity vector $v(t)=r'(t)$ represent geometrically?

A vector tangent to the particle's path of motion at each point, pointing in the instantaneous direction of travel.

How is the speed of a particle related to its velocity vector?

Speed is the magnitude (length) of the velocity vector, $|v(t)|$, a scalar quantity.

How is the direction of motion found from the velocity vector?

By computing the unit tangent vector $\hat{v}(t)=v(t)/|v(t)|$, which has magnitude 1 and points in the direction of travel.

What real-world quantity does the acceleration vector $a(t)=r''(t)$ represent?

The instantaneous rate of change of velocity -- how quickly and in what direction the particle's velocity itself is changing.

Long Questions & Answers

Explain how a vector-valued function is constructed from a real-world description, and how its domain and range are determined, using a concrete example.

What are the three steps for constructing a vector-valued function from a description?

First, determine how each coordinate (x, and y, and z if 3D) changes with respect to the parameter t; second, combine these component functions into a single vector using either angle-bracket notation $\langle f(t),g(t),h(t)\rangle$ or unit-vector notation $f(t)\hat{\imath}+g(t)\hat{\jmath}+h(t)\hat{k}$; third, specify the domain of t appropriate to the physical situation, such as $t\geq 0$ for motion starting at time zero.

How is the domain of a constructed vector-valued function found?

The domain of each component function is determined separately (e.g. requiring a square root's argument to be non-negative, or a denominator to be nonzero), and then the overall domain of $r(t)$ is the intersection of all these individual component domains, since every component must be defined simultaneously.

How is the range of a vector-valued function described?

The range is described by finding the range of each component function given the domain restriction, then expressing the set of all possible output vectors as an ordered tuple $\langle x,y \rangle$ or $\langle x,y,z \rangle$ subject to those individual component ranges and any relationships between them.

How does this process apply to $r(t)=\sqrt{4-t^2}\,\hat{i}+1/(t-1)\,\hat{j}$?

The first component requires $4-t^2 \geq 0$ giving $t \in [-2,2]$, and the second requires $t \neq 1$, so the domain is $[-2,1) \cup (1,2]$; on this domain, $\sqrt{4-t^2}$ ranges over $[0,2]$ while $1/(t-1)$ can take any nonzero real value, giving the range $\{\langle x,y \rangle \mid 0 \leq x \leq 2,\ y \in \mathbb{R},\ y \neq 0\}$.

Explain how velocity, acceleration, speed, and direction of motion are all derived from a position vector r(t), and how they are interpreted for a particle moving along a curved path.

How is the velocity vector obtained from the position vector, and what does it represent?

The velocity vector is $v(t)=r'(t)$, found by differentiating each component of the position vector; geometrically it is tangent to the particle's path at every point and indicates the instantaneous direction and rate of change of position.

How is the acceleration vector obtained, and what does it represent?

The acceleration vector is $a(t)=v'(t)=r''(t)$, the derivative of velocity (second derivative of position); it captures how the velocity itself -- both its magnitude and direction -- is changing at each instant, such as speeding up, slowing down, or curving.

How is speed obtained from the velocity vector, and why is it a scalar rather than a vector?

Speed is the magnitude $|v(t)|$ of the velocity vector, computed as the square root of the sum of the squares of its components; it is a scalar because it measures only how fast the particle is moving, discarding all directional information.

How is the direction of motion found, and why is it useful separately from speed?

The direction of motion is the unit tangent vector $\hat{v}(t)=v(t)/|v(t)|$, obtained by dividing the velocity vector by its own magnitude so the result has length 1; separating direction from speed lets an analyst describe exactly which way a particle is heading independent of how fast it is currently travelling, which is essential in applications like drone navigation or orbital tracking.

Multiple Choice Questions (MCQs)

A vector-valued function differs from a scalar function because it:

A. Always uses time as input
B. Outputs a vector with magnitude and direction, not just a number
C. Cannot be differentiated
D. Has no domain restrictions
Answer: B. Outputs a vector with magnitude and direction, not just a number
Explanation: The defining feature of a vector-valued function is that its output is a vector, not a single real number.

The derivative of $r(t)=f(t)\hat{i}+g(t)\hat{j}+h(t)\hat{k}$ is found by:

A. Differentiating only f(t)
B. Multiplying all components together then differentiating
C. Differentiating each component separately
D. Taking the magnitude first, then differentiating
Answer: C. Differentiating each component separately
Explanation: Vector-valued differentiation is done componentwise: $r'(t)=f'(t)\hat{i}+g'(t)\hat{j}+h'(t)\hat{k}$.

The domain of a vector-valued function r(t) is:

A. The union of each component's domain
B. The intersection of each component's domain
C. Always all real numbers
D. The domain of only the first component
Answer: B. The intersection of each component's domain
Explanation: Every component must be defined simultaneously, so the overall domain is the intersection of the individual component domains.

The velocity vector of a moving particle is defined as:

A. v(t) = r(t)

B. v(t) = r'(t)

C. v(t) = r''(t)

D. v(t) = |r(t)|

Answer: B. v(t) = r'(t)

Explanation: Velocity is the first derivative of the position vector.

The acceleration vector is defined as:

A. a(t) = r(t)

B. a(t) = v(t)

C. a(t) = v'(t) = r''(t)

D. a(t) = |v(t)|

Answer: C. a(t) = v'(t) = r''(t)

Explanation: Acceleration is the derivative of velocity, equivalently the second derivative of position.

Speed is best described as:

A. A vector pointing in the direction of motion

B. The magnitude of the velocity vector

C. The magnitude of the acceleration vector

D. The same thing as velocity

Answer: B. The magnitude of the velocity vector

Explanation: Speed is a scalar equal to |v(t)|, the length of the velocity vector.

The unit tangent vector $\hat{v}(t)=v(t)/|v(t)|$ represents:

A. The particle's speed

B. The particle's acceleration

C. The particle's direction of motion

D. The particle's position

Answer: C. The particle's direction of motion

Explanation: Dividing the velocity vector by its own magnitude produces a vector of length 1 pointing in the direction of travel.

For $r(t)=R\cos t\,\hat{\imath}+R\sin t\,\hat{\jmath}+ct\,\hat{k}$, the motion described is:

A. Purely linear

B. Purely circular

C. A helix (circular motion plus steady linear climb)

D. Random

Answer: C. A helix (circular motion plus steady linear climb)

Explanation: The x,y-components trace a circle of radius R while the z-component climbs steadily, together tracing a helix.

If a particle's velocity has zero y-component at some instant, this means:

A. The particle has stopped moving entirely

B. The particle has no north-south motion at that instant

C. The particle's acceleration is also zero

D. The position vector is undefined

Answer: B. The particle has no north-south motion at that instant

Explanation: A zero component in the velocity vector means no motion in that particular direction at that instant, not that the particle has stopped altogether.

Every differentiation technique used for scalar functions (product rule, chain rule, etc.) applies to a vector-valued function's components because:

A. Vector-valued functions have their own separate differentiation rules

B. Each component of a vector-valued function is itself an ordinary scalar function of t

C. Only the first component needs to be differentiated

D. Vector-valued functions cannot actually be differentiated

Answer: B. Each component of a vector-valued function is itself an ordinary scalar function of t

Explanation: Since f(t), g(t), h(t) are each ordinary scalar functions, all standard differentiation rules apply directly to each one.

Quick Revision Summary

- Scalar function: outputs a single real number | Vector-valued function: outputs a vector, $r(t)=f(t)\hat{i}+g(t)\hat{j}+h(t)\hat{k}$
- Constructing r(t): determine each component's dependence on t, combine into vector form, specify domain
- Domain of r(t) = intersection of the domains of f(t), g(t), h(t) | Range described component-wise

- Derivative: $r'(t)=f'(t)\hat{\imath}+g'(t)\hat{\jmath}+h'(t)\hat{k}$ -- differentiate each component separately, using ordinary scalar rules
- Velocity: $v(t)=r'(t)$ | Acceleration: $a(t)=v'(t)=r''(t)$
- Speed: $|v(t)|$ (scalar, magnitude of velocity) | Direction of motion: $\hat{v}(t)=v(t)/|v(t)|$ (unit vector)
- Projectile motion: $r(t)=(v_0\cos\theta)t\hat{\imath}+[(v_0\sin\theta)t-\frac{1}{2}gt^2]\hat{\jmath}$
- Helical motion: $r(t)=R\cos t\,\hat{\imath}+R\sin t\,\hat{\jmath}+ct\,\hat{k}$ -- circular motion in xy-plane plus steady linear climb along z
- Every scalar differentiation rule (product, chain, quotient, trig/exp/log derivatives) applies componentwise to vector-valued functions
- Applications: robotics (arm trajectories), drones (helical inspection paths), projectiles (cannonballs, footballs), orbits (comets, satellites, Ferris wheels)

Exam Tips

- Always differentiate each component of a vector-valued function completely separately -- treat $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ as fixed labels, not variables, and never mix terms across components
- When finding the domain of r(t), work out each component's domain restriction on its own first (square roots need non-negative arguments, denominators need to be nonzero, logs need positive arguments), then intersect all of them at the end
- Remember speed is always a non-negative scalar (a magnitude), while velocity is a vector -- do not confuse the two when a question asks specifically for one or the other
- For direction-of-motion questions, always divide the velocity vector by its OWN magnitude at that same instant -- using a magnitude from a different value of t gives a meaningless result
- For real-life motion problems, first identify what physical quantity each component represents (horizontal, vertical, or height) before differentiating -- this makes interpreting the final velocity/acceleration vector far more intuitive
- When acceleration is constant (like -g in projectile motion) but velocity's other components are non-constant, differentiate each term individually rather than assuming the whole acceleration vector is zero or constant without checking