MH2100 Calculus III – Revision Notes

Quantitative Research Society @NTU

Academic Year 2026–2027, Semester 1

Abstract

Revision notes for MH2100 Calculus III through the midterm. The course sequence covers parametric curves, multivariable limits and continuity, partial derivatives and extrema, tangent spaces and differentiability, the chain rule and directional derivatives, Lagrange multipliers, and double integrals. The scope ends at double integrals.

Chapter	Planned Week	Deck Week	Key Concepts
Ch. 1: Parametric Curves	1	W1	Parametrisations, velocity, acceleration, tangent lines, and singularities.
Ch. 2: Limits, Continuity, and Partial Derivatives	2	W2	Multivariable limits, continuity, and partial derivatives.
Ch. 3: Extreme Points	3	W4	Critical points, the Hessian test, and local and global extrema.
Ch. 4: Tangent Spaces and Differentiability	4	W3	Tangent planes, differentiability, and linear approximation.
Ch. 5: Chain Rule and Directional Derivatives	5	W3–W4	Multivariable chain rule, gradients, directional derivatives, and implicit differentiation.
Ch. 6: Lagrange Multipliers	6	W5	Constrained optimisation and geometric interpretation of constraints.
Ch. 7: Double Integrals	7	W5–W6	Iterated integrals, regions, order changes, and polar coordinates.
Ch. 8: Midterm Summary	–	W1–W6	Comparison tables, method patterns, traps, integrated examples, and a memory sheet.

"This revision guide follows the planned MH2100 topic sequence through the midterm, one chapter per topic plus a cumulative midterm summary."

For academic reference only. This document is provided for educational purposes and does not constitute an official question paper, answer key, or any formally authorised academic material. Its content is not guaranteed to be complete or accurate.

Contents

How to Use This Document 4

I Main Notes 5

1 Topic 1: Parametric Curves 5
- 1.1 Vectors in the plane and in space . . . 5
- 1.2 Vector-valued functions and parametrizations . . . 6
- 1.3 Standard parametrizations and parameter ranges . . . 8
- 1.4 Tangents from a parametrization . . . 11
- 1.5 Singular parameters and the cusp . . . 13
- 1.6 Functions of two variables: a short preview . . . 16
- 1.7 Chapter Summary . . . 18

2 Topic 2: Limits, Continuity, and Partial Derivatives 19
- 2.1 Functions, domains, level sets, and distance . . . 19
- 2.2 Finite limits on a relative domain . . . 22
- 2.3 Vector limits and continuity . . . 26
- 2.4 Limits at infinity, open sets, and closed sets . . . 28
- 2.5 Partial derivatives . . . 30
- 2.6 Chapter Summary . . . 31

3 Topic 3: Extreme Points and Their Classification 32
- 3.1 Local and absolute extrema . . . 32
- 3.2 Second partials and the quadratic approximation . . . 34
- 3.3 Reading the quadratic form before using a table . . . 39
- 3.4 Absolute extrema on compact domains . . . 41
- 3.5 Chapter Summary . . . 43

4 Topic 4: Tangent Spaces and Differentiability 44
- 4.1 Planes and the vector product . . . 44
- 4.2 Tangent plane functions . . . 46
- 4.3 Total differentiability and the total differential . . . 49
- 4.4 Consequences of the total-differentiability definition . . . 50
- 4.5 When partial derivatives are insufficient . . . 52
- 4.6 A usable sufficient criterion . . . 53
- 4.7 Linear approximation and error propagation . . . 54
- 4.8 Chapter Summary . . . 56

5 Topic 5: Chain Rule and Directional Derivatives 57
- 5.1 The chain rule . . . 57
- 5.2 Implicit curves and surfaces . . . 59
- 5.3 Directional derivatives and gradients . . . 61
- 5.4 Proof details and diagnostic counterexamples . . . 63
- 5.5 A scalar implicit-function proof in detail . . . 64
- 5.6 Chapter Summary . . . 68

6 Topic 6: Lagrange Multipliers **69**
6.1 Constraint sets and their geometry . . . 69
6.2 Candidates, compactness, and singular points . . . 71
6.3 A three-variable constraint and a noncompact feasible set . . . 73
6.4 Two constraints in three variables . . . 75
6.5 The Lagrangian and safe algebra . . . 76
6.6 Domain boundaries and constraint boundaries . . . 79
6.7 Chapter Summary . . . 80

7 Topic 7: Double Integrals **81**
7.1 Rectangles, partitions, and Riemann sums . . . 81
7.2 General regions and the integrability criterion . . . 83
7.3 Applications of general double integrals . . . 85
7.4 Fubini's theorem and iterated integrals . . . 85
7.5 Changing the order of integration . . . 89
7.6 Linearity, restriction, and additivity . . . 90
7.7 Polar coordinates and the area factor . . . 93
7.8 Chapter Summary . . . 95

8 Topic 8: Midterm Summary **97**
8.1 Comparison table: what each question is asking . . . 97
8.2 Method patterns . . . 98
8.3 Traps that change a correct-looking answer into a wrong one . . . 100
8.4 Integrated examples . . . 101
8.5 Checklist before handing in a solution . . . 107
8.6 Memory sheet . . . 108

A Formula & Key Results Reference **109**

B Notation **112**

C Exam & Admin Information **115**

QRS@NTU

How to Use This Document

Structure of the Notes

Each topic chapter develops definitions, theorems, derivations, worked examples, common mistakes, and exam-focused checks. Where available, examples identify their lecture source; otherwise they are worked illustrations of the relevant method. The planned syllabus weeks differ from the supplied slide-deck weeks, as the topic map shows.

Suggested Revision Flow

- First pass: read the core theory and Key Formula boxes.
- Second pass: rederive the central results and attempt Worked Examples before reading solutions.
- Final pass: use the midterm summary, Common Mistake boxes, and Exam Focus boxes to diagnose weak methods.

Colour and Box System (Legend)

Colour	Box Name	Use This For
Green	**Definition**	Formal definitions and assumptions (numbered).
Blue (band)	**Theorem**	Formal results; a proof or derivation follows immediately.
Blue	**Key Formula**	Equations to memorise or rederive quickly.
Brown	**Worked Example**	Fully worked questions. Try on your own first.
Red	**Common Mistake**	Typical errors and conceptual traps.
Yellow	**Policy Application**	Applied context where it supports mathematical interpretation.
Purple	**Exam Focus**	High-yield checklists and exam strategy.
Cyan	**Extension**	Beyond-syllabus depth (non-examinable unless stated).

QRS@NTU

Part I
Main Notes

1 Topic 1: Parametric Curves

Policy Application

This chapter follows the Week 1 curve material through the regular and singular tangent discussion. The domain and level-set material at the end is a short preview only; limits, continuity, and partial derivatives are developed in Topic 2. No arc-length formula is asserted here because it is outside the displayed Week 1 examinable scope.

1.1 Vectors in the plane and in space

Definition 1.1: Vector coordinates

A vector in $\mathbb{R}^n$ is an ordered list of n real components. Write $\boldsymbol{a} = (a_1, \dots, a_n)$, and use its components to define

$$\boldsymbol{a} + \boldsymbol{b} = (a_1 + b_1, \dots, a_n + b_n), \qquad c\boldsymbol{a} = (ca_1, \dots, ca_n).$$

In $\mathbb{R}^2$, the standard basis is $\boldsymbol{i} = (1, 0)$, $\boldsymbol{j} = (0, 1)$. In $\mathbb{R}^3$, it is $\boldsymbol{i} = (1, 0, 0)$, $\boldsymbol{j} = (0, 1, 0)$, $\boldsymbol{k} = (0, 0, 1)$.

Definition 1.2: Dot product, norm, and angle

For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$, define

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{j=1}^{n} a_j b_j, \qquad \|\boldsymbol{a}\| = \sqrt{\langle \boldsymbol{a}, \boldsymbol{a} \rangle}.$$

If both vectors are nonzero, let $\theta \in [0, \pi]$ be their usual Euclidean angle.

Theorem 1.1: Dot-angle identity

For nonzero $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ with angle θ,

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos\theta.$$

Consequently, $\boldsymbol{a} \perp \boldsymbol{b}$ exactly when $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$.

Proof. By the cosine law applied to the triangle with side vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{a} - \boldsymbol{b}$,

$$\|\boldsymbol{a} - \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\,\|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos\theta.$$

Expanding the left side with the component dot product gives

$$\|\boldsymbol{a} - \boldsymbol{b}\|^2 = \langle \boldsymbol{a} - \boldsymbol{b}, \boldsymbol{a} - \boldsymbol{b} \rangle = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\,\langle \boldsymbol{a}, \boldsymbol{b} \rangle .$$

Equating the two expressions and cancelling common terms proves the identity. For the second, $\boldsymbol{a} \perp \boldsymbol{b}$ means $\theta = \pi/2$, hence $\cos\theta = 0$, so the dot product is zero. Conversely, a zero dot product makes $\cos\theta = 0$; since $\theta \in [0, \pi]$, this gives $\theta = \pi/2$. □

Theorem 1.2: Cauchy–Schwarz and the triangle inequality

For all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$,

$$|\langle \boldsymbol{a}, \boldsymbol{b} \rangle| \leq \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| , \qquad \|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\| .$$

Proof. If either vector is zero, Cauchy–Schwarz is immediate. Otherwise the dot-angle identity and $|\cos\theta| \leq 1$ give it. For the triangle inequality, expand and use Cauchy–Schwarz:

$$\begin{aligned} \|\boldsymbol{a} + \boldsymbol{b}\|^2 &= \|\boldsymbol{a}\|^2 + 2\,\langle \boldsymbol{a}, \boldsymbol{b} \rangle + \|\boldsymbol{b}\|^2 \\ &\leq \|\boldsymbol{a}\|^2 + 2\,\|\boldsymbol{a}\|\,\|\boldsymbol{b}\| + \|\boldsymbol{b}\|^2 \\ &= (\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)^2. \end{aligned}$$

Both sides are nonnegative, so taking square roots proves the claim. □

QRS check: angle and perpendicular direction. Let $\boldsymbol{a} = (2, -1, 2)$ and $\boldsymbol{b} = (1, 2, 0)$. Then

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 2 - 2 + 0 = 0.$$

Thus the vectors are perpendicular. Also $\|\boldsymbol{a}\| = 3$ and $\|\boldsymbol{b}\| = \sqrt{5}$, so the angle formula gives

$$\boxed{\theta = \frac{\pi}{2}.}$$

Common Mistake

Do not divide by $\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|$ before checking that both vectors are nonzero. The angle formula is not defined for the zero vector.

1.2 Vector-valued functions and parametrizations

Definition 1.3: Parametric curve

A continuous vector-valued function on an interval I,

$$\boldsymbol{r} : I \longrightarrow \mathbb{R}^n, \qquad \boldsymbol{r}(t) = (x_1(t), \ldots, x_n(t)),$$

parametrizes its image

$$C = \{\boldsymbol{r}(t) : t \in I\}.$$

The variable t is a parameter. A single geometric curve can have many parametrizations, and a parametrization can visit one point more than once.

Definition 1.4: Derivative and regular parameter

If every component is differentiable at t_0, define

$$\boldsymbol{r}'(t_0) = \lim_{h \to 0} \frac{\boldsymbol{r}(t_0 + h) - \boldsymbol{r}(t_0)}{h}.$$

A parameter t_0 is *regular* when $\boldsymbol{r}'(t_0) \neq \boldsymbol{0}$. It is *singular* when the derivative is zero or does not exist.

Theorem 1.3: Componentwise differentiation

If $\boldsymbol{r}(t) = (x_1(t), \ldots, x_n(t))$ and all components are differentiable at t_0, then

$$\boldsymbol{r}'(t_0) = \big(x_1'(t_0), \ldots, x_n'(t_0)\big).$$

Proof. For $h \neq 0$, the difference quotient is

$$\frac{\boldsymbol{r}(t_0+h) - \boldsymbol{r}(t_0)}{h} = \left(\frac{x_1(t_0+h) - x_1(t_0)}{h}, \ldots, \frac{x_n(t_0+h) - x_n(t_0)}{h}\right).$$

A vector limit exists precisely when its finitely many component limits exist, and the limit is taken componentwise. Taking $h \to 0$ therefore yields the stated vector. □

Definition 1.5: Velocity, speed, and acceleration

For a position curve $\boldsymbol{r}(t)$, define

$$\boldsymbol{v}(t) = \boldsymbol{r}'(t), \qquad \text{speed} = \|\boldsymbol{v}(t)\|, \qquad \boldsymbol{a}(t) = \boldsymbol{r}''(t).$$

Velocity has direction; speed is its nonnegative magnitude.

Theorem 1.4: Affine parametrization of a line

Let $A \in \mathbb{R}^n$ and let $\boldsymbol{d} \neq \boldsymbol{0}$. Then

$$\boldsymbol{r}(t) = A + t\boldsymbol{d}, \qquad t \in \mathbb{R},$$

parametrizes exactly the line through A with direction $\boldsymbol{d}$. Moreover, $\boldsymbol{r}'(t) = \boldsymbol{d}$, so every parameter is regular.

Proof. Every point of the displayed form differs from A by a scalar multiple of $\boldsymbol{d}$, and hence lies on that line. Conversely, the definition of the line says that each of its points has this form for some real scalar t. Componentwise differentiation gives $\boldsymbol{r}'(t) = \boldsymbol{d} \neq \boldsymbol{0}$. □

QRS check: a line and a non-affine reparametrization. The line through $A = (1, -2, 3)$ in direction $\boldsymbol{d} = (2, 1, -1)$ is

$$\boldsymbol{r}(t) = (1, -2, 3) + t(2, 1, -1), \qquad t \in \mathbb{R}.$$

It is also traced by

$$\boldsymbol{s}(u) = (1, -2, 3) + (u^3 + u)(2, 1, -1), \qquad u \in \mathbb{R},$$

because $u^3 + u$ takes every real value. Both have the same image, while

$$\boxed{\boldsymbol{r}'(t) = (2, 1, -1), \qquad \boldsymbol{s}'(u) = (3u^2 + 1)(2, 1, -1).}$$

Common Mistake

Replacing t by t^2 in $A + t\boldsymbol{d}$ does *not* usually trace the whole line: it only obtains coefficients $t^2 \geq 0$, hence a ray from A. Always verify the range of the new scalar parameter.

Peano-type space-filling curves

The Week 1 deck briefly notes that a continuous map can have an unexpectedly large image, such as a Peano curve filling a square. This is a caution that a curve refers to a parametrized image; it does not authorize treating every such map as a regular smooth curve.

1.3 Standard parametrizations and parameter ranges

Definition 1.6: Circle, cycloid, and helix

For a circle with centre (a, b) and radius $R > 0$, and for the other constants indicated below, common parametrizations are

$$\begin{aligned} \text{circle:} \quad & \boldsymbol{c}(t) = (a + R\cos t, b + R\sin t), && 0 \le t < 2\pi, \\ \text{cycloid:} \quad & \boldsymbol{q}(t) = (R(t - \sin t), R(1 - \cos t)), && t \in \mathbb{R}, \\ \text{helix:} \quad & \boldsymbol{h}(t) = (R\cos t, R\sin t, ct), && t \in \mathbb{R}, \end{aligned}$$

where $c \neq 0$ is the vertical rise per radian for the helix. A stated parameter interval is part of a parametrization, not decoration.

The cycloid formula comes from a circle rolling to the right without slipping. After a clockwise rotation through angle t, its centre is (Rt, R): the horizontal displacement equals the rolled arc length Rt. The radius vector to a marked point initially at ground contact is then $(-R\sin t, -R\cos t)$. Adding this vector to the centre gives $(R(t - \sin t), R(1 - \cos t))$, as asserted.

Theorem 1.5: Verification of the standard circle

The curve $\boldsymbol{c}(t) = (a + R\cos t, b + R\sin t)$, $0 \le t < 2\pi$, has image

$$(x - a)^2 + (y - b)^2 = R^2.$$

Every point of that circle is obtained once on the half-open interval, except that the endpoint convention prevents repetition of $(a + R, b)$.

Proof. For every allowed t,

$$(a + R\cos t - a)^2 + (b + R\sin t - b)^2 = R^2,$$

so the image lies on the circle. Conversely, if (x, y) lies on that circle, then $((x - a)/R, (y - b)/R)$ lies on the unit circle. There is an angle $t \in [0, 2\pi)$ whose cosine and sine are these two coordinates, so $(x, y) = \boldsymbol{c}(t)$. The interval contains exactly one representative of each angle modulo 2π. □

QRS: eliminate a parameter

Let

$$\boldsymbol{r}(t) = (3\cos t, 2\sin t), \qquad 0 \le t < 2\pi.$$

Method 1: Pythagorean identity with a converse check. Write $x = 3\cos t$, $y = 2\sin t$, so $x/3 = \cos t$ and $y/2 = \sin t$. Square and add:

$$\frac{x^2}{9} + \frac{y^2}{4} = \cos^2 t + \sin^2 t = 1.$$

Conversely, if $x^2/9 + y^2/4 = 1$, then $(x/3, y/2)$ is on the unit circle, so by Theorem 1.5 (with $R = 1$) exactly one $t \in [0, 2\pi)$ has $(\cos t, \sin t) = (x/3, y/2)$, giving $\boldsymbol{r}(t) = (x, y)$. As $\boldsymbol{r}(0) = (3, 0)$ and $\boldsymbol{r}(\pi/2) = (0, 2)$, the ellipse is traced once counterclockwise from $(3, 0)$.

$$\boxed{\frac{x^2}{9} + \frac{y^2}{4} = 1, \quad 0 \le t < 2\pi.}$$

For the second Week 1 prompt, let

$$\boldsymbol{p}(t) = (t^2 + t, t^2 - t), \qquad t \in \mathbb{R}.$$

Method 1: Sum and difference of the components. With $x = t^2 + t$ and $y = t^2 - t$,

$$x - y = (t^2 + t) - (t^2 - t) = 2t, \qquad x + y = (t^2 + t) + (t^2 - t) = 2t^2,$$
$$(x - y)^2 = 4t^2 = 2 \cdot 2t^2 = 2(x + y),$$

so the image lies in C. Conversely, let $(x - y)^2 = 2(x + y)$ and put $t := (x - y)/2$. Then

$$t^2 = \frac{(x-y)^2}{4} = \frac{2(x+y)}{4} = \frac{x+y}{2},$$
$$t^2 + t = \frac{x+y}{2} + \frac{x-y}{2} = x, \qquad t^2 - t = \frac{x+y}{2} - \frac{x-y}{2} = y,$$

so $(x, y) = \boldsymbol{p}(t)$ and C lies in the image.

$$\boxed{C = \{(x, y) \in \mathbb{R}^2 : (x - y)^2 - 2(x + y) = 0\}.}$$

Method 2: Rotated orthonormal coordinates. Along the orthonormal vectors $\frac{1}{\sqrt{2}}(1, -1)$, $\frac{1}{\sqrt{2}}(1, 1)$ the coordinates are $u = (x - y)/\sqrt{2}$, $w = (x + y)/\sqrt{2}$; on the curve,

$$u = \frac{2t}{\sqrt{2}} = \sqrt{2}\, t, \qquad w = \frac{2t^2}{\sqrt{2}} = \sqrt{2}\, t^2 = \sqrt{2}\Big(\frac{u}{\sqrt{2}}\Big)^2 = \frac{u^2}{\sqrt{2}}.$$

As t runs over $\mathbb{R}$, so does $u = \sqrt{2}\, t$; the image is the whole parabola $w = u^2/\sqrt{2}$ (axis along $y = x$). Multiplying $(x + y)/\sqrt{2} = (x - y)^2/(2\sqrt{2})$ by $2\sqrt{2}$ gives $2(x + y) = (x - y)^2$.

$$\boxed{C = \{(x, y) \in \mathbb{R}^2 : (x - y)^2 - 2(x + y) = 0\}.}$$

QRS: helix motion

For

$$\boldsymbol{h}(t) = (2\cos t, 2\sin t, 3t),$$

Method 1: Componentwise differentiation. By Theorem 1.3, differentiate each component twice:

$$\boldsymbol{v}(t) = \left(\tfrac{d}{dt}(2\cos t),\ \tfrac{d}{dt}(2\sin t),\ \tfrac{d}{dt}(3t)\right) = (-2\sin t, 2\cos t, 3),$$
$$\boldsymbol{a}(t) = \left(\tfrac{d}{dt}(-2\sin t),\ \tfrac{d}{dt}(2\cos t),\ \tfrac{d}{dt}(3)\right) = (-2\cos t, -2\sin t, 0).$$

Hence

$$\|\boldsymbol{v}(t)\| = \sqrt{(-2\sin t)^2 + (2\cos t)^2 + 3^2} = \sqrt{4(\sin^2 t + \cos^2 t) + 9}$$
$$= \sqrt{4+9} = \boxed{\sqrt{13}.}$$

Method 2: Orthogonal horizontal and vertical motion. Split $\boldsymbol{h}(t) = (2\cos t, 2\sin t, 0) + (0, 0, 3t)$; the two parts have velocities

$$\boldsymbol{v}_1 = (-2\sin t, 2\cos t, 0), \qquad \boldsymbol{v}_2 = (0, 0, 3).$$

Then $\|\boldsymbol{v}_1\| = \sqrt{4\sin^2 t + 4\cos^2 t} = 2$ (the circle speed $R = 2$), $\|\boldsymbol{v}_2\| = 3$, and $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0+0+0 = 0$. Expanding the squared norm,

$$\|\boldsymbol{v}_1 + \boldsymbol{v}_2\|^2 = \|\boldsymbol{v}_1\|^2 + 2\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle + \|\boldsymbol{v}_2\|^2 = 4 + 0 + 9 = 13,$$

$$\|\boldsymbol{v}(t)\| = \boxed{\sqrt{13}.}$$

Method 3: Constant speed from $\langle \boldsymbol{v}, \boldsymbol{a} \rangle = 0$. Using the derivatives of Method 1,

$$\langle \boldsymbol{v}, \boldsymbol{a} \rangle = (-2\sin t)(-2\cos t) + (2\cos t)(-2\sin t) + 3 \cdot 0 = 0.$$

By the product rule for dot products, $\frac{d}{dt}\|\boldsymbol{v}\|^2 = 2\langle \boldsymbol{v}, \boldsymbol{a} \rangle = 0$ on $\mathbb{R}$, so $\|\boldsymbol{v}\|^2$ is constant by the Mean Value Theorem. At $t = 0$, $\boldsymbol{v}(0) = (0, 2, 3)$, hence

$$\|\boldsymbol{v}(t)\| = \|\boldsymbol{v}(0)\| = \sqrt{0+4+9} = \boxed{\sqrt{13}.}$$

Theorem 1.6: Regularity of the standard circle, cycloid, and helix

The circle and helix above are regular at every parameter. The cycloid is singular exactly at $t = 2\pi k$, where $k \in \mathbb{Z}$.

Proof. Their derivatives are

$$\boldsymbol{c}'(t) = (-R\sin t, R\cos t),$$
$$\boldsymbol{q}'(t) = (R(1-\cos t), R\sin t),$$

and

$$\boldsymbol{h}'(t) = (-R\sin t, R\cos t, c).$$

For the circle, $\|\boldsymbol{c}'(t)\| = R > 0$. For the helix, its third component is $c \neq 0$. For the cycloid, both derivative components vanish exactly when $\sin t = 0$ and $1 - \cos t = 0$, equivalently $t = 2\pi k$. □

Common Mistake

An implicit equation identifies an image but can lose direction, repetition, and the selected portion. For example, $x^2+y^2=R^2$ does not say whether the circle is traversed clockwise, once, or repeatedly. Keep the parameter interval with the formula.

1.4 Tangents from a parametrization

Definition 1.7: Tangent line at a regular parameter

If $\boldsymbol{r}(t_0)=P$ and $\boldsymbol{r}'(t_0)\neq\mathbf{0}$, the tangent line at the parameter value t_0 is

$$\ell(s)=P+s\boldsymbol{r}'(t_0), \qquad s\in\mathbb{R}.$$

The vector $\boldsymbol{r}'(t_0)$ is a tangent vector. The unit tangent, oriented with increasing parameter, is

$$\boldsymbol{T}(t_0)=\frac{\boldsymbol{r}'(t_0)}{\|\boldsymbol{r}'(t_0)\|}.$$

It has norm one because $\|cv\|=|c|\,\|v\|$, and dividing by the positive speed preserves the derivative's orientation.

Theorem 1.7: Secant directions approach the derivative direction

Suppose $\boldsymbol{r}$ is differentiable at t_0 and $\boldsymbol{r}'(t_0)\neq\mathbf{0}$. Then the unoriented direction of the secant from $\boldsymbol{r}(t_0)$ to $\boldsymbol{r}(t_0+h)$ approaches the direction of $\boldsymbol{r}'(t_0)$ as $h\to 0$. Thus the displayed line in Definition 1.7 is the induced tangent line.

Proof. For $h\neq 0$, the secant displacement is

$$\boldsymbol{r}(t_0+h)-\boldsymbol{r}(t_0)=h\left[\frac{\boldsymbol{r}(t_0+h)-\boldsymbol{r}(t_0)}{h}\right].$$

The bracket tends to $\boldsymbol{r}'(t_0)$ by differentiability. Since that limit is nonzero, normalizing the bracket is continuous near the limit; its unit direction tends to $\boldsymbol{r}'(t_0)/\|\boldsymbol{r}'(t_0)\|$. Changing the sign of h only reverses an unoriented line direction, so the limiting line has direction $\boldsymbol{r}'(t_0)$. □

QRS: recovered Week 1 tangent calculation

The rendered Week 1 slide gives

$$\boldsymbol{r}(t) = (4t^3 + 5, t^2 + 2t^4, -2\ln(2t)), \qquad t > 0.$$

Method 1: Componentwise derivative and the tangent-line definition. At $t = 1$, $P = \boldsymbol{r}(1) = (4 + 5,\ 1 + 2,\ -2\ln 2) = (9, 3, -2\ln 2)$. Differentiate each component; the third uses the Chain Rule with inner function $2t$, which is positive on the domain $t > 0$:

$$\tfrac{d}{dt}(4t^3 + 5) = 12t^2, \qquad \tfrac{d}{dt}(t^2 + 2t^4) = 2t + 8t^3,$$
$$\tfrac{d}{dt}\big(-2\ln(2t)\big) = -2 \cdot \frac{1}{2t} \cdot 2 = -\frac{2}{t}.$$

Hence

$$\boldsymbol{r}'(t) = (12t^2,\ 2t + 8t^3,\ -2/t),$$
$$\boldsymbol{r}'(1) = (12,\ 2 + 8,\ -2) = (12, 10, -2) \neq \mathbf{0}.$$

So $t = 1$ is regular and Definition 1.7 gives

$$\boxed{\ell(s) = (9, 3, -2\ln 2) + s(12, 10, -2), \qquad s \in \mathbb{R}.}$$

Method 2: Secant limit from first principles. For $h \neq 0$ with $1 + h > 0$, expand each component of $\boldsymbol{r}(1 + h) - \boldsymbol{r}(1)$:

$$4(1 + h)^3 + 5 - 9 = 12h + 12h^2 + 4h^3,$$
$$(1 + h)^2 + 2(1 + h)^4 - 3 = (2h + h^2) + (8h + 12h^2 + 8h^3 + 2h^4) = 10h + 13h^2 + 8h^3 + 2h^4,$$
$$-2\ln(2 + 2h) + 2\ln 2 = -2\ln(1 + h).$$

Divide by h and let $h \to 0$, using $\lim_{h\to 0} \ln(1 + h)/h = 1$:

$$\frac{\boldsymbol{r}(1 + h) - \boldsymbol{r}(1)}{h} = \Big(12 + 12h + 4h^2,\ 10 + 13h + 8h^2 + 2h^3,\ -2\,\frac{\ln(1 + h)}{h}\Big)$$
$$\longrightarrow (12, 10, -2) \neq \mathbf{0}.$$

By Theorem 1.7, the secant lines through P tend to the line through P with this direction:

$$\boxed{\ell(s) = (9, 3, -2\ln 2) + s(12, 10, -2), \qquad s \in \mathbb{R}.}$$

QRS: one image point, two tangents

Consider

$$\boldsymbol{r}(t) = (t^2 - 1, t(t^2 - 1)).$$

Method 1: Solve for every parameter, then differentiate. The second component is t times the first, so

$$\boldsymbol{r}(t) = (0,0) \iff t^2 - 1 = 0 \text{ and } t(t^2-1) = 0 \iff t = \pm 1.$$

With $y(t) = t^3 - t$, componentwise differentiation gives

$$\boldsymbol{r}'(t) = \left(\tfrac{d}{dt}(t^2-1),\ \tfrac{d}{dt}(t^3-t)\right) = (2t,\ 3t^2-1),$$
$$\boldsymbol{r}'(1) = (2,\ 3-1) = (2,2) \neq \mathbf{0}, \qquad \boldsymbol{r}'(-1) = (-2,\ 3-1) = (-2,2) \neq \mathbf{0}.$$

Both parameters are regular, so Definition 1.7 gives two lines through the origin:

$$s(2,2) = (2s, 2s) \implies y = x, \qquad s(-2,2) = (-2s, 2s) \implies y = -x.$$

$$\boxed{y = x \quad \text{and} \quad y = -x.}$$

Method 2: Limiting secant slopes. For $t \neq \pm 1$, $x = t^2 - 1 \neq 0$, and the secant from $(0,0)$ to $\boldsymbol{r}(t)$ has slope

$$\frac{y}{x} = \frac{t(t^2-1)}{t^2-1} = t.$$

Since $\boldsymbol{r}'(\pm 1) \neq \mathbf{0}$, Theorem 1.7 applies at each parameter: the slope tends to 1 as $t \to 1$ and to -1 as $t \to -1$.

$$\boxed{y = x \quad \text{and} \quad y = -x.}$$

Method 3: Cartesian equation and branch derivatives. For $x \neq 0$, $t = y/x$ and $t^2 = x + 1$, so the image lies on $y^2 = x^2(x+1)$. The parameters near $t = 1$ give the branch $y = x\sqrt{1+x}$ ($t = \sqrt{1+x}$), and those near $t = -1$ give $y = -x\sqrt{1+x}$. By the Product Rule,

$$\frac{d}{dx}(x\sqrt{1+x})\Big|_{x=0} = \left(\sqrt{1+x} + \frac{x}{2\sqrt{1+x}}\right)\Big|_{x=0} = 1 + 0 = 1,$$

so the branches have slopes 1 and -1 at the origin.

$$\boxed{y = x \quad \text{and} \quad y = -x.}$$

Common Mistake

At a self-intersection, the question point may have several parameter values. Solving $\boldsymbol{r}(t) = P$ is a required first step; using only one root can omit a tangent branch.

1.5 Singular parameters and the cusp

Definition 1.8: One-sided and semi-regular tangent

At a singular parameter, the ordinary derivative may fail to choose a direction. If one-sided derivative directions exist and determine the same line, that line is a natural tangent candidate. A parametrization is semi-regular at a joining point when its one-sided derivatives are nonzero, even though it need not be differentiable there.

Theorem 1.8: The algebraic cusp is singular

Let $\boldsymbol{r}(t) = (x(t), y(t))$ be differentiable at an interior parameter t_0, with $\boldsymbol{r}(t_0) = (0, 0)$, and suppose its nearby image lies on

$$x^2 = y^3.$$

Then $\boldsymbol{r}'(t_0) = \mathbf{0}$. Thus the cusp point of $x^2 = y^3$ cannot be regular under such a differentiable parametrization.

Proof. The equation forces $y \geq 0$. Since $y(t_0) = 0$, the function $y(t)$ has a local minimum at the interior point t_0, so $y'(t_0) = 0$. Writing $h = t - t_0$, this gives $y(t) = o(h)$. Also $|x(t)| = y(t)^{3/2}$, hence

$$\left|\frac{x(t)}{h}\right| = \left|\frac{y(t)}{h}\right|^{3/2} |h|^{1/2} \longrightarrow 0.$$

Therefore $x'(t_0) = 0$ as well, and the derivative vector is zero. □

QRS@NTU — qrsntu.org

QRS: the cusp and its unique tangent line

Method 1: One-sided derivatives of a semi-regular parametrization. The rendered Week 1 recovery uses the semi-regular parametrization

$$\boldsymbol{v}(t) = \begin{cases} (-\sqrt{(-t)^3}, -t), & t \le 0, \\ (\sqrt{t^3}, t), & t \ge 0. \end{cases}$$

For $t \ge 0$, $x^2 = t^3 = y^3$; for $t \le 0$, $x^2 = (-t)^3 = y^3$. So its image satisfies $x^2 = y^3$, with $\boldsymbol{v}(0) = (0,0)$. For $h > 0$ (right side) and for $-k$ with $k > 0$ (left side), the difference quotients are

$$\frac{\boldsymbol{v}(h) - \boldsymbol{v}(0)}{h} = \left(\frac{h^{3/2}}{h}, \frac{h}{h}\right) = (h^{1/2}, 1) \longrightarrow (0,1) = \boldsymbol{v}'_+(0),$$
$$\frac{\boldsymbol{v}(-k) - \boldsymbol{v}(0)}{-k} = \left(\frac{-k^{3/2}}{-k}, \frac{k}{-k}\right) = (k^{1/2}, -1) \longrightarrow (0,-1) = \boldsymbol{v}'_-(0).$$

Both are nonzero and $\boldsymbol{v}'_-(0) = -\boldsymbol{v}'_+(0)$, so they span one line through the origin with direction $(0,1)$. Since they differ, $\boldsymbol{v}'(0)$ does not exist and $t = 0$ is singular.

$$\boxed{x = 0, \quad \text{the } y\text{-axis.}}$$

Method 2: Limiting secant direction from $x^2 = y^3$. Let $(x,y) \neq (0,0)$ satisfy $x^2 = y^3$. Then $y^3 = x^2 \ge 0$ forces $y \ge 0$, and $y = 0$ would give $x = 0$; hence $y > 0$ and

$$|x| = y^{3/2} \implies \frac{x}{y} = \pm\sqrt{y} \longrightarrow 0 \qquad \text{as } (x,y) \to (0,0).$$

The secant from the cusp to (x,y) has direction (x,y), parallel to $(x/y, 1) \to (0,1)$, so every secant tends to the y-axis.

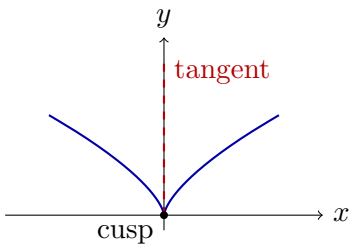
$$\boxed{x = 0, \quad \text{the } y\text{-axis.}}$$

Method 3: Limit of unit tangents of $\boldsymbol{w}(t) = (t^3, t^2)$. This curve (plotted below) satisfies $x^2 = t^6 = y^3$, and $\boldsymbol{w}'(t) = (3t^2, 2t)$, so $\boldsymbol{w}'(0) = \mathbf{0}$, in line with Theorem 1.8. For $t \neq 0$, $\|\boldsymbol{w}'(t)\| = \sqrt{9t^4 + 4t^2} = |t|\sqrt{9t^2 + 4}$, hence

$$\boldsymbol{T}(t) = \frac{(3t^2, 2t)}{|t|\sqrt{9t^2+4}} = \frac{t}{|t|} \cdot \frac{(3t, 2)}{\sqrt{9t^2+4}} \longrightarrow \pm\frac{(0,2)}{2} = \pm(0,1) \qquad (t \to 0^{\pm}).$$

Both limits lie on the line $x = 0$.

$$\boxed{x = 0, \quad \text{the } y\text{-axis.}}$$



Exam Focus

For a singular point, do not automatically write a tangent line using $\boldsymbol{r}'(t_0)$. First test whether a one-sided or an implicit-curve argument identifies a line. A zero derivative alone neither supplies a direction nor rules out a geometric tangent.

1.6 Functions of two variables: a short preview

Definition 1.9: Domain and graph

For a real-valued function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$, the domain is the set D of input pairs where the formula is defined. Its graph is the subset of $\mathbb{R}^3$

$$\text{graph}(f) = \{(x, y, z) : (x, y) \in D,\ z = f(x, y)\}.$$

Domain restrictions are often obtained from denominators, even roots, and logarithms.

QRS check: domain from two restrictions. The Week 1 rendered slide asks for the domain of

$$f(x, y) = \frac{x}{\sqrt{y^2 - x}}.$$

The square root must be real and, because it is in the denominator, nonzero. Thus $y^2 - x > 0$, or equivalently $x < y^2$. The numerator imposes no further restriction. Therefore

$$\boxed{D = \{(x, y) \in \mathbb{R}^2 : x < y^2\}.}$$

Definition 1.10: Level set

For $c \in \mathbb{R}$, the level set of f at level c is

$$L_c = \{(x, y) \in D : f(x, y) = c\}.$$

A contour plot draws several such plane curves together. Level sets describe horizontal slices of the graph at heights $z = c$.

Theorem 1.9: Level sets are horizontal graph sections

For $f : D \to \mathbb{R}$, the points of the graph at height c project precisely onto L_c.

Proof. A graph point at height c has the form (x, y, c) and satisfies $c = f(x, y)$. Its projection (x, y) is therefore in L_c. Conversely, if $(x, y) \in L_c$, then $(x, y, f(x, y)) = (x, y, c)$ is a graph point at height c. □

QRS: level curves of a hemisphere

The recovered Week 1 slide uses

$$f(x,y)=\sqrt{36-x^2-y^2}, \qquad D=\{(x,y): x^2+y^2\le 36\}.$$

Method 1: Solve $f=c$ with sign and range checks. A square root is nonnegative, so every level has $c\ge 0$; since $36-x^2-y^2\le 36$, also $c\le 6$. For $0\le c\le 6$, both sides of $\sqrt{36-x^2-y^2}=c$ are nonnegative, so squaring is reversible:

$$f(x,y)=c \iff 36-x^2-y^2=c^2 \iff x^2+y^2=36-c^2.$$

These are concentric circles of radius $\sqrt{36-c^2}$; for $c=0,1,\dots,6$ the radii are

$$6,\ \sqrt{35},\ \sqrt{32},\ \sqrt{27},\ \sqrt{20},\ \sqrt{11},\ 0.$$

At the two extreme levels,

$$c=6:\ x^2+y^2=0 \iff (x,y)=(0,0), \qquad c=0:\ x^2+y^2=36.$$

$$\boxed{L_6=\{(0,0)\}, \qquad L_0=\{x^2+y^2=36\}.}$$

Method 2: Horizontal sections of the hemisphere. For $(x,y)\in D$, $z=f(x,y) \iff \big(z\ge 0$ and $x^2+y^2+z^2=36\big)$, so the graph is the upper hemisphere of radius 6 centred at the origin. By Theorem 1.9, L_c is the projection of the section by the plane $z=c$:

$$x^2+y^2+c^2=36, \quad z=c \implies x^2+y^2=36-c^2.$$

The plane $z=6$ meets the hemisphere only at the top point $(0,0,6)$, and $z=0$ cuts it in the equator $x^2+y^2=36$.

$$\boxed{L_6=\{(0,0)\}, \qquad L_0=\{x^2+y^2=36\}.}$$

Method 3: Radial profile in polar coordinates. With $r=\sqrt{x^2+y^2}\in[0,6]$, $f=g(r):=\sqrt{36-r^2}$, and for $0<r<6$

$$g'(r)=\frac{-2r}{2\sqrt{36-r^2}}=-\frac{r}{\sqrt{36-r^2}}<0.$$

So g is continuous and strictly decreasing on $[0,6]$ from $g(0)=6$ to $g(6)=0$; by the Intermediate Value Theorem each $c\in[0,6]$ is taken at exactly one radius, $r=\sqrt{36-c^2}$. Level $c=6$ gives $r=0$, and level $c=0$ gives $r=6$.

$$\boxed{L_6=\{(0,0)\}, \qquad L_0=\{x^2+y^2=36\}.}$$

Common Mistake

For $f(x,y)=\sqrt{36-x^2-y^2}$, the equation after squaring is only valid for levels $c\ge 0$. Do not report circles for negative levels: the square root output cannot be negative.

Policy Application

This preview only identifies domains, graphs, and level sets from Week 1. The limit and continuity tools needed to analyze multivariable functions begin in Topic 2.

1.7 Chapter Summary

Key Formula

For a parametrized curve $\boldsymbol{r}(t)$, keep the parameter range visible:

$$\boxed{\boldsymbol{v}(t) = \boldsymbol{r}'(t), \qquad \text{speed} = \|\boldsymbol{r}'(t)\|, \qquad \boldsymbol{a}(t) = \boldsymbol{r}''(t), \qquad \ell(s) = \boldsymbol{r}(t_0) + s\boldsymbol{r}'(t_0)}$$

when $\boldsymbol{r}'(t_0) \neq \boldsymbol{0}$. For vectors, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos\theta$.

Exam Focus

A reliable curve workflow is: state the parameter interval; locate every parameter value for the requested point; differentiate componentwise; test regularity; then write the tangent line from the appropriate branch. At a cusp or another singular parameter, replace the regular tangent formula with a justified one-sided or implicit argument.

QRS@NTU — qrsntu.org

2 Topic 2: Limits, Continuity, and Partial Derivatives

This chapter develops the local language used throughout multivariable calculus. A limit asks what outputs approach when inputs approach a point through the actual domain. Continuity makes substitution legitimate. Partial derivatives measure coordinate-wise change, but do not by themselves give a linear model; that stronger question belongs to Chapter 4.

2.1 Functions, domains, level sets, and distance

A real-valued function of two variables is a rule $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^2$. A vector-valued function is a rule $F : D \to \mathbb{R}^m$. The domain is part of the function: an expression can be meaningful at some points and excluded at others by a denominator, square root, logarithm, or an explicit restriction.

Definition 2.1: Euclidean norm and distance

For $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$, define

$$\|u\| = \sqrt{u_1^2 + \cdots + u_m^2}.$$

For $p, q \in \mathbb{R}^m$, their Euclidean distance is

$$\operatorname{dist}(p, q) = \|p - q\|.$$

The norm is a length, so it is nonnegative and vanishes only at the zero vector. It also has the homogeneity and triangle properties

$$\|cu\| = |c| \, \|u\|, \qquad \|u + v\| \le \|u\| + \|v\|.$$

These inequalities turn algebra into quantitative control of an error.

Theorem 2.1: Reverse triangle inequality

For every $u, v \in \mathbb{R}^m$,

$$\big|\|u\| - \|v\|\big| \le \|u - v\|.$$

Consequently, the norm map $u \mapsto \|u\|$ is continuous.

Proof. The triangle inequality applied to $u = (u - v) + v$ gives

$$\|u\| \le \|u - v\| + \|v\|,$$

so $\|u\| - \|v\| \le \|u - v\|$. Interchanging u and v gives $\|v\| - \|u\| \le \|u - v\|$. Combining the two bounds proves the displayed inequality.

If $u \to a$, then the inequality gives

$$\big|\|u\| - \|a\|\big| \le \|u - a\| \to 0.$$

Thus $\|u\| \to \|a\|$, which is continuity of the norm map. □

Definition 2.2: Balls, limit points, and level curves

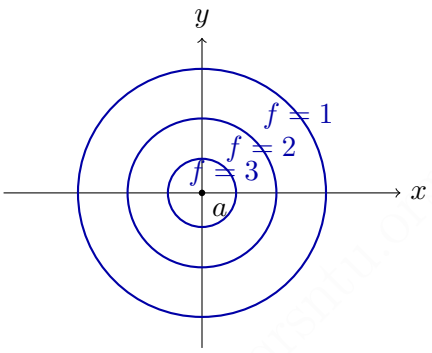
For $a \in \mathbb{R}^m$ and $r > 0$, the open ball is

$$B_r(a) = \{x \in \mathbb{R}^m : \|x - a\| < r\}.$$

A point a is a *limit point* of $D \subseteq \mathbb{R}^m$ if every $B_r(a)$ contains a point of D different from a. For $f : D \to \mathbb{R}$, the level set at height k is

$$\{(x, y) \in D : f(x, y) = k\}.$$

At a non-limit point there is no nearby domain point from which to test a limit. In particular, a singleton domain has no limit points. This is why continuity at an isolated point can hold without placing any restriction on the assigned value there.



Nested level curves: height changes while the input stays in the domain.

QRS@NTU · Revision Notes

Worked Example (Week 1): a domain and its level curves

Let

$$f(x,y)=\frac{x}{\sqrt{y^2-x}}.$$

Method 1: Defining conditions, then squaring and completing the square. The square root needs $y^2-x\geq 0$, and the denominator needs $\sqrt{y^2-x}\neq 0$, that is $y^2-x\neq 0$. Together,

$$y^2-x>0 \iff x<y^2.$$

For the level $f=0$ the numerator must vanish, so $x=0$; then $x<y^2$ reads $0<y^2$. Hence

$$f=0 \iff x=0,\ y\neq 0.$$

For a level $f=k\neq 0$, write $x=k\sqrt{y^2-x}$. Since $\sqrt{y^2-x}>0$, x has the sign of k, so $x/k>0$. Squaring,

$$\begin{aligned} x^2&=k^2(y^2-x),\\ x^2+k^2x-k^2y^2&=0,\\ \left(x+\frac{k^2}{2}\right)^2-k^2y^2&=\frac{k^4}{4},\\ \frac{(x+k^2/2)^2}{(k^2/2)^2}-\frac{y^2}{(|k|/2)^2}&=1. \end{aligned}$$

This is a hyperbola with centre $(-k^2/2,0)$ and vertices $(0,0)$, $(-k^2,0)$. The sign condition keeps the branch with $x>0$ when $k>0$, and the branch with $x\leq -k^2$ when $k<0$. On that branch $k^2(y^2-x)=x^2>0$, so it lies in the domain.

$$\boxed{\mathrm{Dom}(f)=\{(x,y):x<y^2\}}$$

Method 2: Boundary parabola with test points; level sets solved for y. The function y^2-x is continuous and vanishes exactly on the parabola $x=y^2$, so it keeps one sign on each of the two regions that the parabola cuts out. Test one point in each region:

$$(-1,0):\ y^2-x=1>0\ \text{(defined)},\qquad (1,0):\ y^2-x=-1<0\ \text{(undefined)}.$$

The parabola itself is excluded because there the denominator is 0. So the domain is the region containing $(-1,0)$, namely $x<y^2$. For $f=k\neq 0$ with $x/k>0$, solve $x^2=k^2(y^2-x)$ for y:

$$y^2=x+\frac{x^2}{k^2}=\frac{x(x+k^2)}{k^2},\qquad y=\pm\frac{\sqrt{x(x+k^2)}}{|k|}.$$

With $x>0$ (for $k>0$) the radicand is positive; with $x<0$ (for $k<0$) it is nonnegative exactly when $x\leq -k^2$. These are the two hyperbola branches of Method 1, and $f=0$ is again the punctured line $x=0$, $y\neq 0$.

$$\boxed{\mathrm{Dom}(f)=\{(x,y):x<y^2\}}$$

Domain first

Do not test a multivariable limit along paths that leave the domain. Also, do not cancel a factor and silently enlarge a punctured domain: the cancelled formula may describe the limiting behaviour, but it is not automatically the original function at the missing point.

2.2 Finite limits on a relative domain

Definition 2.3: Limit at a limit point

Let $f : D \to \mathbb{R}^m$, and let a be a limit point of D. We write

$$\lim_{x \to a,\ x \in D} f(x) = L$$

when, for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$x \in D, \quad 0 < \|x - a\| < \delta \quad \Longrightarrow \quad \|f(x) - L\| < \epsilon.$$

The phrase "$x \in D$" makes this a relative-domain definition.

The deleted condition $x \neq a$ means that the value $f(a)$, if it exists, does not affect the limit. A limit is a statement about arbitrarily close nearby values, not a substitution rule.

Theorem 2.2: Uniqueness of limits

If a is a limit point of D and

$$\lim_{x \to a,\ x \in D} f(x) = L \quad \text{and} \quad \lim_{x \to a,\ x \in D} f(x) = M,$$

then $L = M$.

Proof. Suppose, for contradiction, that $L \neq M$. Set

$$\epsilon = \frac{\|L - M\|}{3} > 0.$$

Choose δ_L and δ_M from the two limit definitions for this ϵ, and put $\delta = \min\{\delta_L, \delta_M\}$. Because a is a limit point of D, choose $x \in D$ with $0 < \|x - a\| < \delta$. The triangle inequality then gives

$$\|L - M\| \leq \|L - f(x)\| + \|f(x) - M\| < \epsilon + \epsilon = \frac{2}{3}\|L - M\|,$$

which is impossible. Therefore $L = M$. □

Theorem 2.3: Limit laws

Let $f, g : D \to \mathbb{R}$ have limits A, B at the same limit point a. Then

$$\lim(f + g) = A + B, \quad \lim(cf) = cA, \quad \lim(fg) = AB.$$

If $B \neq 0$, then $\lim(f/g) = A/B$ wherever g is nonzero sufficiently near a. The same statements hold componentwise for vector-valued maps.

Proof. For the sum, choose δ so that both $|f(x) - A| < \epsilon/2$ and $|g(x) - B| < \epsilon/2$ hold. Then

$$|(f(x) + g(x)) - (A + B)| \leq |f(x) - A| + |g(x) - B| < \epsilon.$$

Scalar multiplication follows by using $\epsilon/|c|$ when $c \neq 0$; the case $c = 0$ is immediate.

For products, first choose a neighbourhood in which $|f(x) - A| < 1$. There $|f(x)| \leq |A| + 1$. Next choose the neighbourhood small enough that

$$|f(x) - A| < \frac{\epsilon}{2(|B| + 1)}, \qquad |g(x) - B| < \frac{\epsilon}{2(|A| + 1)}.$$

Using $fg - AB = f(g - B) + B(f - A)$ gives

$$|fg - AB| \le (|A| + 1)|g - B| + |B||f - A| < \epsilon.$$

For quotients, choose a punctured neighbourhood where $|g(x) - B| < |B|/2$, so $|g(x)| > |B|/2$. There

$$\left|\frac{1}{g(x)} - \frac{1}{B}\right| = \frac{|g(x) - B|}{|g(x)||B|} \le \frac{2|g(x) - B|}{|B|^2}.$$

Making $|g(x)-B| < \epsilon|B|^2/2$ proves $1/g \to 1/B$. The already proved product rule now yields $f/g \to A/B$. Componentwise laws follow by applying the scalar proof to each coordinate; componentwise division requires every denominator limit to be nonzero. □

Theorem 2.4: Squeeze theorem

Suppose $u(x) \le f(x) \le v(x)$ for all domain points sufficiently close to a limit point a, except possibly a itself. If $u(x) \to L$ and $v(x) \to L$, then $f(x) \to L$.

Proof. For $\epsilon > 0$, choose a punctured neighbourhood small enough for both $u(x) > L-\epsilon$ and $v(x) < L+\epsilon$, and for the assumed inequalities. Then $L-\epsilon < f(x) < L+\epsilon$, so $|f(x)-L| < \epsilon$. This is the limit definition. In particular, a bound $|f(x) - L| \le h(x)$ with $h(x) \to 0$ proves the desired limit by squeezing between $L - h(x)$ and $L + h(x)$. □

A proof menu for finite limits

algebraic simplification $\longrightarrow$ limit laws; $\quad |f(x) - L| \le h(x) \to 0 \quad \longrightarrow \quad f(x) \to L.$

Two paths with different limiting outputs prove non-existence. Agreement on some paths is evidence only; it is never a proof that a two-variable limit exists.

Worked Example (Week 2): a curved-path counterexample

Consider

$$f(x,y)=\frac{xy^2}{x^2+y^4},\qquad (x,y)\neq(0,0).$$

Method 1: Two paths with different limits. Along the x-axis, the points $(x,0)$ with $x\neq 0$ tend to $(0,0)$ as $x\to 0$, and

$$f(x,0)=\frac{x\cdot 0^2}{x^2+0^4}=\frac{0}{x^2}=0 \longrightarrow 0.$$

Along the parabola $x=y^2$, the points (y^2,y) with $y\neq 0$ tend to $(0,0)$ as $y\to 0$, and

$$f(y^2,y)=\frac{y^2\cdot y^2}{(y^2)^2+y^4}=\frac{y^4}{2y^4}=\frac{1}{2} \longrightarrow \frac{1}{2}.$$

Both paths lie in the domain and approach $(0,0)$, but $0\neq\frac{1}{2}$. If the limit were L, every path limit would equal L by uniqueness of limits (Theorem 2.2), forcing $L=0$ and $L=\frac{1}{2}$.

$$\boxed{\lim_{(x,y)\to(0,0)}\frac{xy^2}{x^2+y^4}\text{ does not exist}}$$

Method 2: The parabola family $x=my^2$. Straight lines do not separate the values: for $y=mx$ with $x\neq 0$,

$$f(x,mx)=\frac{m^2x^3}{x^2+m^4x^4}=\frac{m^2x}{1+m^4x^2} \longrightarrow 0.$$

On the remaining line $x=0$, $f(0,y)=0/y^4=0\to 0$ for $y\neq 0$. For the parabolas $x=my^2$ with $y\neq 0$,

$$f(my^2,y)=\frac{my^2\cdot y^2}{m^2y^4+y^4}=\frac{m\,y^4}{(m^2+1)y^4}=\frac{m}{m^2+1}.$$

This constant depends on m: $m=0$ gives 0, $m=1$ gives $\frac{1}{2}$, $m=-1$ gives $-\frac{1}{2}$. Distinct path limits rule out a limit.

$$\boxed{\lim_{(x,y)\to(0,0)}\frac{xy^2}{x^2+y^4}\text{ does not exist}}$$

Method 3: Substitution $v=y^2$ and polar angle. Put $u=x$, $v=y^2\geq 0$; then $(u,v)\to(0,0)$ as $(x,y)\to(0,0)$ and $f=uv/(u^2+v^2)$. Write $u=\rho\cos\varphi$, $v=\rho\sin\varphi$ with $\rho>0$, $0\leq\varphi\leq\pi$:

$$f=\frac{\rho^2\cos\varphi\sin\varphi}{\rho^2}=\frac{1}{2}\sin 2\varphi.$$

The value depends only on the angle. Every ball about the origin contains points with $\varphi=0$ ($y=0$, $x>0$, value 0) and $\varphi=\pi/4$ ($x=y^2$, value $\frac{1}{2}$).

$$\boxed{\lim_{(x,y)\to(0,0)}\frac{xy^2}{x^2+y^4}\text{ does not exist}}$$

Worked Example (Week 2): squeeze with an epsilon–delta conclusion

For $(x, y) \neq (0, 0)$, set

$$g(x, y) = \frac{3x^2 y}{x^2 + y^2}.$$

Method 1: Squeeze bound and an explicit δ. Because $0 \leq x^2 \leq x^2 + y^2$ and $x^2 + y^2 > 0$,

$$0 \leq \frac{x^2}{x^2 + y^2} \leq 1.$$

Also $|y| = \sqrt{y^2} \leq \sqrt{x^2 + y^2}$. Hence

$$|g(x, y) - 0| = 3|y| \cdot \frac{x^2}{x^2 + y^2} \leq 3|y| \cdot 1 \leq 3\sqrt{x^2 + y^2} = 3 \|(x, y)\|.$$

Since $3 \|(x, y)\| \to 0$, the squeeze theorem (Theorem 2.4) gives $g \to 0$. To write the epsilon–delta conclusion, let $\epsilon > 0$ and set $\delta = \epsilon/3$. Then

$$0 < \|(x, y)\| < \delta \implies |g(x, y) - 0| \leq 3 \|(x, y)\| < 3\delta = \epsilon.$$

$$\boxed{\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^2 + y^2} = 0}$$

Method 2: Polar coordinates. Write $x = r\cos\theta$, $y = r\sin\theta$ with $r = \|(x, y)\| > 0$. Then $x^2 + y^2 = r^2$ and

$$g = \frac{3(r^2\cos^2\theta)(r\sin\theta)}{r^2} = 3r\cos^2\theta\sin\theta.$$

Since $|\cos^2\theta\sin\theta| \leq 1$ for every θ,

$$|g - 0| \leq 3r \longrightarrow 0 \qquad (r \to 0^+).$$

The bound $3r$ does not depend on θ, so it controls every approach to the origin at once, and $r < \epsilon/3$ forces $|g| < \epsilon$.

$$\boxed{\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^2 + y^2} = 0}$$

Method 3: AM–GM bound on the denominator. On the axes $g = 0$. For $x \neq 0$, $y \neq 0$, AM–GM gives $x^2 + y^2 \geq 2|x||y| > 0$, so

$$|g(x, y)| = \frac{3x^2|y|}{x^2 + y^2} \leq \frac{3x^2|y|}{2|x||y|} = \frac{3}{2}|x| \leq \frac{3}{2} \|(x, y)\|.$$

Thus $|g - 0| \leq \frac{3}{2} \|(x, y)\| \to 0$ at every point near the origin, and the squeeze theorem gives the limit.

$$\boxed{\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^2 + y^2} = 0}$$

2.3 Vector limits and continuity

Theorem 2.5: Component criterion for vector limits

Let $F = (f_1, \ldots, f_m) : D \to \mathbb{R}^m$. Then

$$\lim_{x \to a} F(x) = L = (L_1, \ldots, L_m)$$

if and only if $\lim_{x \to a} f_j(x) = L_j$ for every j.

Proof. Suppose first that $F(x) \to L$. Given $\epsilon > 0$, choose $\delta > 0$ from the vector-limit definition. For $x \in D$ with $0 < \|x - a\| < \delta$,

$$|f_j(x) - L_j| \leq \|F(x) - L\| < \epsilon,$$

so every component has the stated limit. Conversely, let $\epsilon > 0$. For each j, choose δ_j so that $|f_j(x) - L_j| < \epsilon/\sqrt{m}$ when $0 < \|x - a\| < \delta_j$. With $\delta = \min_j \delta_j$,

$$\|F(x) - L\|^2 = \sum_{j=1}^{m} |f_j(x) - L_j|^2 < m \left(\frac{\epsilon}{\sqrt{m}} \right)^2 = \epsilon^2.$$

Taking square roots proves $F(x) \to L$. □

Definition 2.4: Continuity relative to the domain

Let $f : D \to \mathbb{R}^m$ and let $a \in D$. The map f is continuous at a if, for every $\epsilon > 0$, there is $\delta > 0$ such that

$$x \in D, \quad \|x - a\| < \delta \quad \Longrightarrow \quad \|f(x) - f(a)\| < \epsilon.$$

The definition includes $x = a$. If a is a limit point of D, continuity is equivalent to $\lim_{x \to a,\ x \in D} f(x) = f(a)$. If a is isolated, choose δ so that $B_\delta(a) \cap D = \{a\}$; continuity follows immediately.

Theorem 2.6: Continuity laws and composition

If $f, g : D \to \mathbb{R}^m$ are continuous at a, then their componentwise sums, scalar multiples, and products are continuous there. If $h : E \to \mathbb{R}^k$ is continuous at $f(a)$ and $f(D) \subseteq E$, then $h \circ f$ is continuous at a.

Proof. If a is isolated in D, all the stated algebraic combinations are automatically continuous there. Otherwise substitute the limits $f(x) \to f(a)$ and $g(x) \to g(a)$ into the limit laws of Theorem 2.3. For composition, let $\epsilon > 0$. Continuity of h supplies $\eta > 0$ such that

$$\|u - f(a)\| < \eta \Longrightarrow \|h(u) - h(f(a))\| < \epsilon.$$

Continuity of f supplies $\delta > 0$ such that

$$\|x - a\| < \delta \Longrightarrow \|f(x) - f(a)\| < \eta.$$

Putting $u = f(x)$ in the first implication proves the claim. □

Worked Example (Week 2): a continuous piecewise map and its partials

Define

$$p(x,y)=\begin{cases}\dfrac{x^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Method 1: Squeeze for continuity; difference quotients at the origin. For $x\neq 0$, $x^2+y^2\geq x^2>0$, so

$$|p(x,y)|=\frac{|x|^3}{x^2+y^2}\leq\frac{|x|^3}{x^2}=|x|.$$

For $x=0$, $y\neq 0$, $p(0,y)=0=|x|$. So for every $(x,y)\neq(0,0)$,

$$|p(x,y)-0|\leq|x|\leq\sqrt{x^2+y^2}=\|(x,y)\|\ \longrightarrow\ 0,$$

and the squeeze theorem gives $\lim_{(x,y)\to(0,0)}p(x,y)=0=p(0,0)$, which is continuity at $(0,0)$. The origin is the join of the two formulas, so the partials there come from the defining limits. For $h\neq 0$, $p(h,0)=h^3/h^2=h$ and $p(0,h)=0/h^2=0$, hence

$$p_x(0,0)=\lim_{h\to0}\frac{p(h,0)-p(0,0)}{h}=\lim_{h\to0}\frac{h-0}{h}=\lim_{h\to0}1=1,$$
$$p_y(0,0)=\lim_{h\to0}\frac{p(0,h)-p(0,0)}{h}=\lim_{h\to0}\frac{0-0}{h}=\lim_{h\to0}0=0.$$

Away from the origin p is a quotient with nonzero denominator, so the quotient rule applies:

$$p_x=\frac{3x^2(x^2+y^2)-x^3\cdot 2x}{(x^2+y^2)^2}=\frac{3x^4+3x^2y^2-2x^4}{(x^2+y^2)^2}=\frac{x^4+3x^2y^2}{(x^2+y^2)^2},$$
$$p_y=\frac{0\cdot(x^2+y^2)-x^3\cdot 2y}{(x^2+y^2)^2}=-\frac{2x^3y}{(x^2+y^2)^2}.$$

$$\boxed{p\text{ is continuous at }(0,0),\quad p_x(0,0)=1,\quad p_y(0,0)=0}$$

Method 2: Polar bound; one-variable slices. With $x=r\cos\theta$, $y=r\sin\theta$, $r>0$,

$$p=\frac{r^3\cos^3\theta}{r^2}=r\cos^3\theta,\qquad |p-0|=r|\cos\theta|^3\leq r\ \longrightarrow\ 0,$$

uniformly in θ, so $p\to 0=p(0,0)$. For the partials use the slice theorem (Theorem 2.9). The horizontal slice $\varphi(x)=p(x,0)$ equals $x^3/x^2=x$ for $x\neq 0$ and 0 at $x=0$, so $\varphi(x)=x$ for every x. The vertical slice $\psi(y)=p(0,y)$ equals $0/y^2=0$ for $y\neq 0$ and 0 at $y=0$, so $\psi\equiv 0$. Therefore

$$p_x(0,0)=\varphi'(0)=1,\qquad p_y(0,0)=\psi'(0)=0.$$

$$\boxed{p\text{ is continuous at }(0,0),\quad p_x(0,0)=1,\quad p_y(0,0)=0}$$

Partials are not a differentiability test

The statements "$f_x(a,b)$ and $f_y(a,b)$ exist" and "f is continuous at (a,b)" do not, by themselves, imply differentiability. Chapter 4 gives the linear-approximation definition and sufficient hypotheses. Here, use partial derivatives only as coordinate limits with their stated domains.

2.4 Limits at infinity, open sets, and closed sets

Definition 2.5: A vector limit at infinity

Let D be unbounded. For $F : D \to \mathbb{R}^m$, write $\lim_{x\to\infty} F(x) = L$ when, for every $\epsilon > 0$, there is $N > 0$ such that

$$x \in D, \quad \|x\| > N \quad \Longrightarrow \quad \|F(x) - L\| < \epsilon.$$

The condition is radial: $\|x\| \to \infty$ means escaping every ball about the origin. It is not an iterated instruction such as "first let $x \to \infty$, then let $y \to \infty$".

Theorem 2.7: A radial decay test

If $F : D \to \mathbb{R}^m$ satisfies

$$\|F(x) - L\| \le \phi(\|x\|)$$

outside some ball, where $\phi(r) \to 0$ as $r \to \infty$, then $F(x) \to L$ as $x \to \infty$.

Proof. Given $\epsilon > 0$, choose N such that $r > N$ implies $\phi(r) < \epsilon$. For $\|x\| > N$, the displayed bound gives

$$\|F(x) - L\| \le \phi(\|x\|) < \epsilon.$$

This is precisely the definition of the limit at infinity. □

QRS@NTU – qrs.ntu.org

Worked Example (Week 2): iterated versus radial infinity

For $q(x,y) = \exp(-(x+y))$, each iterated limit with $x \to \infty$ and y fixed is 0, and then the remaining limit is also 0. Now let

$$F(x,y) = \left(\frac{x}{x^2+y^2}, -1\right), \qquad (x,y) \neq (0,0).$$

Method 1: Two escape paths; radial decay bound. For fixed y, $\lim_{x\to\infty} e^{-(x+y)} = e^{-y}\lim_{x\to\infty} e^{-x} = e^{-y}\cdot 0 = 0$, and then $\lim_{y\to\infty} 0 = 0$; the other order gives 0 in the same way because $x+y$ is symmetric. This iterated value does not control $\|(x,y)\| \to \infty$, so test two escape paths:

$$\begin{aligned} q(x,-x) &= e^{-(x-x)} = e^0 = 1 \to 1 && (x \to +\infty,\ \|(x,-x)\| = \sqrt{2}|x| \to \infty), \\ q(t,0) &= e^{-(t+0)} = e^{-t} \to 0 && (t \to +\infty,\ \|(t,0)\| = t \to \infty). \end{aligned}$$

Since $1 \neq 0$, no single L meets the definition at infinity. For F, the second component is constant, so

$$F(x,y) - (0,-1) = \left(\frac{x}{x^2+y^2}, 0\right), \qquad \|F(x,y) - (0,-1)\| = \frac{|x|}{x^2+y^2}.$$

Using $|x| \le \sqrt{x^2+y^2}$,

$$\frac{|x|}{x^2+y^2} \le \frac{\sqrt{x^2+y^2}}{x^2+y^2} = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{\|(x,y)\|}.$$

The radial decay test (Theorem 2.7) applies with $\phi(r) = 1/r \to 0$ as $r \to \infty$.

$$\boxed{\lim_{(x,y)\to\infty} F(x,y) = (0,-1)}$$

$$\boxed{\lim_{(x,y)\to\infty} \exp(-(x+y)) \text{ does not exist}}$$

Method 2: Level-line family; component criterion in polar form. The function q is constant on each line $x+y=c$. Parametrise it by $(t, c-t)$:

$$q(t,c-t) = e^{-(t+c-t)} = e^{-c} \text{ for all } t, \qquad \|(t,c-t)\| \ge |t| \to \infty \quad (t \to +\infty).$$

So each line escapes to infinity with path limit e^{-c}; $c=0$ gives 1 and $c=1$ gives $e^{-1} \neq 1$. For F, the component criterion (Theorem 2.5, whose proof applies with $\|x\| > N$ in place of $0 < \|x-a\| < \delta$) reduces the claim to the two components. The second is the constant -1. For the first, write $x = r\cos\theta$, $y = r\sin\theta$ with $r = \|(x,y)\|$:

$$\left|\frac{x}{x^2+y^2} - 0\right| = \frac{r|\cos\theta|}{r^2} = \frac{|\cos\theta|}{r} \le \frac{1}{r}.$$

Given $\epsilon > 0$, take $N = 1/\epsilon$; then $r > N$ gives a value below ϵ, for every θ.

$$\boxed{\lim_{(x,y)\to\infty} F(x,y) = (0,-1)}$$

$$\boxed{\lim_{(x,y)\to\infty} \exp(-(x+y)) \text{ does not exist}}$$

Definition 2.6: Open and closed subsets of Euclidean space

A set $U \subseteq \mathbb{R}^m$ is open if every $a \in U$ has some $r > 0$ with $B_r(a) \subseteq U$. A set $C \subseteq \mathbb{R}^m$ is closed if it contains all of its limit points.

Theorem 2.8: Complement criterion

A set $C \subseteq \mathbb{R}^m$ is closed if and only if its complement C^c is open.

Proof. Suppose C is closed and take $a \in C^c$. Since a is not a limit point of C, some ball $B_r(a)$ contains no point of C other than possibly a. But $a \notin C$, so $B_r(a) \subseteq C^c$. Thus C^c is open.

Conversely, suppose C^c is open. If a is a limit point of C but $a \notin C$, then $a \in C^c$, so an open ball around a lies in C^c. That ball cannot contain a point of C distinct from a, contradicting that a is a limit point. Hence $a \in C$, and C is closed. □

Open, closed, and bounded are different ideas

$$B_1(0) \text{ is open and bounded,}$$
$$\{x : \|x\| \leq 1\} \text{ is closed and bounded,}$$
$$\mathbb{R}^m \text{ is open and closed but unbounded.}$$

Closed means "contains limit points"; bounded means "fits inside a ball." Neither word means the other.

2.5 Partial derivatives

Definition 2.7: Partial derivatives on an interior domain

Let (a, b) be an interior point of $D \subseteq \mathbb{R}^2$, and let $f : D \to \mathbb{R}$. The first partial derivatives, when the limits exist, are

$$f_x(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}, \qquad f_y(a,b) = \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

The interior condition ensures that the required horizontal and vertical nearby points remain in the domain for all sufficiently small h. At a boundary point, one-sided coordinate derivatives can be defined separately, but they are not the default objects in this chapter.

Theorem 2.9: Geometric coordinate-slice interpretation

Suppose $f_x(a, b)$ exists. The curve obtained by fixing $y = b$ has local secant slopes

$$\frac{f(a+h,b) - f(a,b)}{h},$$

so its tangent slope is $f_x(a, b)$. The analogous statement for the slice $x = a$ has tangent slope $f_y(a, b)$.

Proof. Define the one-variable slice $\varphi(x) = f(x, b)$ on its induced domain. Then the ordinary derivative

definition gives

$$\varphi'(a) = \lim_{h\to 0} \frac{\varphi(a+h) - \varphi(a)}{h} = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} = f_x(a,b).$$

The statement about its tangent slope is the one-variable tangent-line theorem. Replacing φ by $\psi(y) = f(a,y)$ proves the y claim. □

Partial-derivative workflow

state the domain $\longrightarrow$
differentiate while holding the other coordinate fixed
$\longrightarrow$ use the difference quotient at exceptional points.

For a formula valid away from a point, substitute that point only after the partial derivative has been justified there by its defining limit.

Exam Focus: a reliable limit-and-partials checklist

- Write the domain and identify whether the target is a limit point.
- To disprove a limit, give two admissible paths with unequal outputs.
- To prove a zero limit, seek a bound by a constant times $\|x-a\|$.
- For infinity, test all escape directions using $\|x\| \to \infty$.
- For a piecewise partial at its join, start from the difference quotient.
- Do not use partial derivatives as a substitute for the differentiability definition developed in Chapter 4.

2.6 Chapter Summary

Key Formula: Chapter 2 recap

$$\lim_{x\to a} F(x) = L \iff \forall \epsilon > 0\ \exists \delta > 0 : x \in D, \quad 0 < \|x-a\| < \delta \Rightarrow \|F(x) - L\| < \epsilon,$$

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h}, \qquad f_y(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

Exam Focus: Chapter 2 takeaways

Limits are unique and relative to the domain. Component limits characterize vector limits. Continuity permits substitution and survives legal algebra and composition. Infinity is radial, not iterated. Partial derivatives are slopes of coordinate slices; their existence does not settle differentiability.

3 Topic 3: Extreme Points and Their Classification

This chapter classifies unconstrained extrema of real-valued functions. All comparisons are relative to the stated domain. The gradient geometry and directional derivatives from the first part of Week 4 are developed in Chapter 5; constrained extrema wait for Chapter 6.

3.1 Local and absolute extrema

Definition 3.1: Relative local and absolute extrema

Let $f : D \to \mathbb{R}$ and $p \in D$. The point p is an absolute minimum if

$$f(p) \leq f(x) \quad \text{for every } x \in D.$$

It is an absolute maximum if the reverse inequality holds for every $x \in D$. It is a local minimum if some $r > 0$ satisfies

$$x \in D, \quad \|x - p\| < r \quad \Longrightarrow \quad f(p) \leq f(x),$$

and a local maximum if the reverse inequality holds.

The word local never permits us to ignore the domain boundary. A boundary point can be an absolute extremum, but it is not eligible for the interior critical-point theorem below.

Definition 3.2: Critical and saddle points

An interior point p of D is critical when either at least one first partial derivative does not exist at p, or every first partial exists and

$$\nabla f(p) = 0.$$

A saddle point is a critical point with existing first partials that is neither a local maximum nor a local minimum.

Theorem 3.1: Fermat's necessary condition

Let $p = (a, b)$ be an interior point of D. If f has a local extremum at p and $f_x(p), f_y(p)$ exist, then

$$\nabla f(p) = (0, 0).$$

Hence every interior local extremum is a critical point.

Proof. Fix $y = b$ and define the one-variable slice $g(x) = f(x, b)$. Because p is interior, a is an interior point of the slice domain. Because p is a local extremum of f, a is a local extremum of g. The one-variable Fermat theorem gives

$$f_x(a, b) = g'(a) = 0.$$

Repeating the argument for $h(y) = f(a, y)$ gives $f_y(a, b) = 0$. Thus the gradient is zero. □

Worked Example (Week 4): a saddle point

Let $f(x, y) = y^2 - x^2$, a polynomial on all of $\mathbb{R}^2$.
Method 1: Stationary point and comparison along the axes
The first partials are

$$f_x = \frac{\partial}{\partial x}\left(y^2 - x^2\right) = -2x, \qquad f_y = \frac{\partial}{\partial y}\left(y^2 - x^2\right) = 2y.$$

Both exist everywhere, so the critical points are the solutions of

$$-2x = 0 \implies x = 0, \qquad 2y = 0 \implies y = 0.$$

Thus $(0, 0)$ is the only critical point, and $f(0, 0) = 0$. Every disc around the origin contains nonzero points of both axes, where

$$f(x, 0) = -x^2 < 0 = f(0, 0) \qquad (x \neq 0 : \text{ not a local min})$$
$$f(0, y) = y^2 > 0 = f(0, 0) \qquad (y \neq 0 : \text{ not a local max})$$

$$\therefore \boxed{(0, 0) \text{ is a saddle point of } y^2 - x^2}$$

Method 2: Hessian discriminant test
The second partials are constants, hence continuous on $\mathbb{R}^2$, so the Hessian discriminant test applies at the stationary point $(0, 0)$:

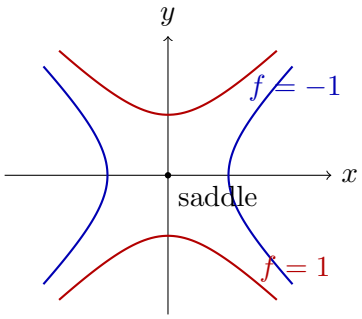
$$f_{xx} = \frac{\partial}{\partial x}(-2x) = -2, \qquad f_{yy} = \frac{\partial}{\partial y}(2y) = 2, \qquad f_{xy} = \frac{\partial}{\partial y}(-2x) = 0.$$

Therefore

$$D(0, 0) = f_{xx} f_{yy} - f_{xy}^2 = (-2)(2) - 0^2 = -4 < 0,$$

and $D < 0$ gives a saddle:

$$\therefore \boxed{(0, 0) \text{ is a saddle point of } y^2 - x^2}$$



Worked Example: a nondifferentiable critical point

For $f(x,y) = |x| + y^2$, the point $(0,0)$ is interior to $\mathbb{R}^2$ and $f(0,0) = 0$.
Method 1: Difference quotients and direct comparison
The difference quotient for $f_x(0,0)$ is

$$\frac{f(h,0) - f(0,0)}{h} = \frac{|h| + 0 - 0}{h} = \frac{|h|}{h},$$

with one-sided limits

$$\lim_{h\to 0^+} \frac{|h| - 0}{h} = \lim_{h\to 0^+} \frac{h}{h} = 1$$
$$\lim_{h\to 0^-} \frac{|h| - 0}{h} = \lim_{h\to 0^-} \frac{-h}{h} = -1 \qquad \Longrightarrow \ f_x(0,0) \text{ does not exist}$$

The other partial exists:

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = \lim_{k\to 0} \frac{k^2}{k} = \lim_{k\to 0} k = 0.$$

Hence $(0,0)$ is critical (undefined partial). For $(x,y) \neq (0,0)$, either $x \neq 0$, so $|x| > 0$, or $x = 0$ and $y \neq 0$, so $y^2 > 0$. In both cases

$$|x| + y^2 > 0 = f(0,0) \qquad (|x| \geq 0,\ y^2 \geq 0,\ \text{not both } 0)$$

$$\boxed{(0,0) \text{ is the strict absolute minimum of } |x| + y^2}$$

Method 2: Separable sum of one-variable functions
Write $f(x,y) = g(x) + k(y)$ with $g(x) = |x|$ and $k(y) = y^2$. By definition of the partial derivative, $f_x(0,0)$ exists exactly when $g'(0)$ does, and $g(x) = |x|$ has a corner at 0 (slope -1 on the left, $+1$ on the right), so $f_x(0,0)$ does not exist and $(0,0)$ is critical. Each summand is minimised separately:

$$g(x) \geq 0 = g(0),\ \text{equality iff } x = 0, \qquad k(y) \geq 0 = k(0),\ \text{equality iff } y = 0.$$

Adding the two inequalities,

$$f(x,y) = g(x) + k(y) \geq g(0) + k(0) = 0 = f(0,0),$$

with equality only when $x = 0$ and $y = 0$ simultaneously. Therefore

$$\boxed{(0,0) \text{ is the strict absolute minimum of } |x| + y^2}$$

Stationary is only one kind of critical

Solving $f_x = f_y = 0$ finds stationary critical points, not all critical points. It also gives candidates only: Fermat's theorem is necessary, never sufficient. Boundary points require a separate comparison on their boundary pieces.

3.2 Second partials and the quadratic approximation

For a function with suitable derivatives, write

$$f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

The Hessian matrix is

$$\operatorname{Hess} f(a,b) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}_{(a,b)}.$$

Theorem 3.2: Clairaut's theorem

If f_{xy} and f_{yx} are continuous on a neighbourhood of (a,b), then

$$f_{xy}(a,b) = f_{yx}(a,b).$$

Proof. For small nonzero h,k, form the rectangular increment

$$Q(h,k) = f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b).$$

For the order x then y, apply the mean value theorem in x to

$$\phi(x) = f(x,b+k) - f(x,b).$$

For some ξ between a and $a+h$, apply it in y to $f_x(\xi,y)$. This gives some η between b and $b+k$ with

$$\frac{Q(h,k)}{hk} = f_{xy}(\xi,\eta).$$

For the reverse order, apply the theorem in y to $\psi(y) = f(a+h,y) - f(a,y)$, then in x to $f_y(x,\eta')$. Thus there are η' between b and $b+k$ and ξ' between a and $a+h$ with

$$\frac{Q(h,k)}{hk} = f_{yx}(\xi',\eta').$$

Letting $(h,k) \to (0,0)$ and using both continuities yields the same limit, which proves the equality. □

Theorem 3.3: Second-order Taylor expansion at a stationary point

Suppose f has continuous second partials on a neighbourhood of $p = (a,b)$, and $\nabla f(p) = 0$. For $u = (h,k) \to 0$,

$$f(p+u) - f(p) = \frac{1}{2} u^{\mathsf{T}} \operatorname{Hess} f(p) u + o(\|u\|^2).$$

Proof. Define $g(t) = f(a+th, b+tk)$ for $0 \le t \le 1$. One-variable Taylor's theorem gives

$$g(1) = g(0) + g'(0) + \frac{1}{2} g''(\theta)$$

for some $0 < \theta < 1$. The chain rule gives

$$g'(0) = \nabla f(p) \cdot u = 0,$$

and

$$g''(t) = u^{\mathsf{T}} \operatorname{Hess} f(p+tu) u.$$

Add and subtract $\operatorname{Hess} f(p)$ in the last expression. Continuity of the Hessian makes the difference $o(\|u\|^2)$, proving the formula. □

Theorem 3.4: Hessian discriminant test

Let p be a stationary point, assume the hypotheses of Theorem 3.3, and set

$$D = f_{xx}(p) f_{yy}(p) - f_{xy}(p)^2.$$

If $D > 0$ and $f_{xx}(p) > 0$, p is a strict local minimum. If $D > 0$ and $f_{xx}(p) < 0$, p is a strict local maximum. If $D < 0$, p is a saddle point. If $D = 0$, this test is inconclusive.

Proof. By Clairaut, the quadratic term is

$$q(h,k) = f_{xx}h^2 + 2f_{xy}hk + f_{yy}k^2.$$

If $f_{xx} \neq 0$, complete the square:

$$q = f_{xx}\left(h + \frac{f_{xy}}{f_{xx}}k\right)^2 + \frac{D}{f_{xx}}k^2.$$

Thus $D > 0$ makes q positive definite when $f_{xx} > 0$ and negative definite when $f_{xx} < 0$. The Taylor remainder is smaller than, say, half the nonzero quadratic bound for sufficiently small u, so it cannot change the sign; this proves the minimum and maximum statements.

If $D < 0$ and $f_{xx} \neq 0$, the displayed square form takes both signs by choosing $k \neq 0$ and then choosing the square term respectively small or large. The Taylor remainder again preserves these signs along the corresponding fixed rays close to p, proving a saddle. If $f_{xx} = 0$, then $D = -f_{xy}^2 < 0$, so $f_{xy} \neq 0$. In this case $q(h,k) = 2f_{xy}hk + f_{yy}k^2$. Setting $k = 1$ and choosing h large positive or negative gives both signs; scaling these two fixed vectors towards zero proves the same conclusion. When $D = 0$, the quadratic term is semidefinite or zero and has directions where it vanishes; higher-order terms decide the result. □

The two-variable classification table

$$D = f_{xx}f_{yy} - f_{xy}^2,$$
$$D > 0,\ f_{xx} > 0 \Rightarrow \text{strict local minimum},$$
$$D > 0,\ f_{xx} < 0 \Rightarrow \text{strict local maximum},$$
$$D < 0 \Rightarrow \text{saddle},$$
$$D = 0 \Rightarrow \text{inconclusive: inspect higher-order behaviour.}$$

QRS@NTU

Worked Example: degenerate points can behave differently

Let $f(x,y) = x^4 + y^4$ and $g(x,y) = x^4 - y^4$, both polynomials on $\mathbb{R}^2$.
Method 1: Second-order data, then direct comparison
The first and second partials are

$$\begin{aligned} f_x &= 4x^3, & f_y &= 4y^3, & f_{xx} &= 12x^2, & f_{yy} &= 12y^2, & f_{xy} &= 0, \\ g_x &= 4x^3, & g_y &= -4y^3, & g_{xx} &= 12x^2, & g_{yy} &= -12y^2, & g_{xy} &= 0. \end{aligned}$$

For both functions, $4x^3 = 0$ and $\pm 4y^3 = 0$ force $x = y = 0$, so the origin is the only stationary point, and there

$$D_f(0,0) = 0 \cdot 0 - 0^2 = 0, \qquad D_g(0,0) = 0 \cdot 0 - 0^2 = 0.$$

The Hessian test is inconclusive for both, so compare values directly:

$$x^4 + y^4 > 0 = f(0,0) \qquad (x^4 \geq 0,\ y^4 \geq 0,\ \text{not both } 0)$$

$$\boxed{(0,0) \text{ is a strict absolute minimum of } x^4 + y^4}$$

For g, every disc around the origin meets both axes, where

$$\begin{aligned} g(x,0) &= x^4 > 0 = g(0,0) & (x \neq 0:&\ \text{not a local max}) \\ g(0,y) &= -y^4 < 0 = g(0,0) & (y \neq 0:&\ \text{not a local min}) \end{aligned}$$

$$\boxed{(0,0) \text{ is a saddle point of } x^4 - y^4}$$

Method 2: Polar coordinates
Let $x = r\cos\theta$, $y = r\sin\theta$ with $r > 0$. Using $\cos^2\theta + \sin^2\theta = 1$ and $2\sin\theta\cos\theta = \sin 2\theta$,

$$\begin{aligned} x^4 + y^4 &= r^4\left(\cos^4\theta + \sin^4\theta\right) = r^4\left(\left(\cos^2\theta + \sin^2\theta\right)^2 - 2\cos^2\theta\sin^2\theta\right) \\ &= r^4\left(1 - \tfrac{1}{2}\sin^2 2\theta\right) \ \geq\ \tfrac{1}{2}r^4 > 0 = f(0,0) \qquad (\sin^2 2\theta \leq 1). \end{aligned}$$

Every point other than the origin has $r > 0$, so

$$\boxed{(0,0) \text{ is a strict absolute minimum of } x^4 + y^4}$$

For g, factor the difference of squares:

$$x^4 - y^4 = r^4\left(\cos^2\theta - \sin^2\theta\right)\left(\cos^2\theta + \sin^2\theta\right) = r^4\cos 2\theta.$$

On every circle $r > 0$, $\theta = 0$ gives $g = r^4 > 0$ and $\theta = \frac{\pi}{2}$ gives $g = -r^4 < 0$, so every disc around the origin contains values above and below $g(0,0) = 0$:

$$\boxed{(0,0) \text{ is a saddle point of } x^4 - y^4}$$

Worked Example (Week 4): $\sin(xy)$ and global non-strict extrema

Let $f(x,y)=\sin(xy)$. By the Chain Rule,

$$f_x=\cos(xy)\cdot\frac{\partial}{\partial x}(xy)=y\cos(xy),\qquad f_y=\cos(xy)\cdot\frac{\partial}{\partial y}(xy)=x\cos(xy).$$

Method 1: Stationary set, diagonal comparison, range bound
Both partials vanish when $y\cos(xy)=0$ and $x\cos(xy)=0$. If $\cos(xy)\neq 0$, then $y=0$ and $x=0$, giving $(0,0)$, where $\cos 0=1\neq 0$. If $\cos(xy)=0$, both equations hold, which happens exactly on the hyperbolas $xy=\left(k+\frac{1}{2}\right)\pi$, $k\in\mathbb{Z}$. At the origin $f(0,0)=0$, and for $0<|x|<\sqrt{\pi}$ we have $0<x^2<\pi$, so

$$f(x,x)=\sin(x^2)>0=f(0,0),\qquad f(x,-x)=\sin(-x^2)=-\sin(x^2)<0=f(0,0).$$

Both points lie at distance $\sqrt{2}|x|$ from the origin, so every disc around it contains values on both sides: $\therefore$ the origin is a saddle.
With $n\in\mathbb{Z}$ and $-1\le\sin(xy)\le 1$:

$$\begin{aligned}
&xy=\left(2n+\tfrac{1}{2}\right)\pi\\
\Longrightarrow\ &f=\sin\left(2n\pi+\tfrac{\pi}{2}\right)\\
&\quad=1 && \text{(global, non-strict max)}\\
&xy=\left(2n+\tfrac{3}{2}\right)\pi\\
\Longrightarrow\ &f=\sin\left(2n\pi+\tfrac{3\pi}{2}\right)\\
&\quad=-1 && \text{(global, non-strict min)}
\end{aligned}$$

Each value is attained along a whole hyperbola, so no extremum is strict.

$$\boxed{\max\sin(xy)=1,\quad \min\sin(xy)=-1}$$

Method 2: Hessian test and the substitution $u=xy$
Differentiating f_x and f_y with the product and Chain Rules,

$$f_{xx}=-y^2\sin(xy),\qquad f_{yy}=-x^2\sin(xy),\qquad f_{xy}=\cos(xy)-xy\sin(xy).$$

These are continuous on $\mathbb{R}^2$, so Clairaut gives $f_{yx}=f_{xy}$ and the Hessian test applies. At $(0,0)$: $f_{xx}=0$, $f_{yy}=0$, $f_{xy}=1$, and $D=0\cdot 0-1^2=-1<0$, a saddle. On $xy=\left(k+\frac{1}{2}\right)\pi$, write $s=\sin(xy)=\pm 1$ and $\cos(xy)=0$:

$$D=\left(-y^2s\right)\left(-x^2s\right)-(-xys)^2=x^2y^2s^2-x^2y^2s^2=0,$$

so the test is inconclusive there. For the range, set $u=xy$. The map $(x,y)\mapsto xy$ takes every real value ($u=1\cdot u$ at $(1,u)$), so

$$\{\sin(xy):(x,y)\in\mathbb{R}^2\}=\{\sin u:u\in\mathbb{R}\}=[-1,1],$$

with $\sin u=1$ iff $u=\frac{\pi}{2}+2n\pi$ and $\sin u=-1$ iff $u=\frac{3\pi}{2}+2n\pi$, $n\in\mathbb{Z}$:

$$\boxed{\max\sin(xy)=1,\quad \min\sin(xy)=-1}$$

$D = 0$ is not a classification

The conclusion $D = 0$ means only that the quadratic approximation is unable to decide. It does not mean saddle, and it does not mean extremum. Compare nearby values, factor the expression, use higher-order terms, or use a global range bound when one is available.

3.3 Reading the quadratic form before using a table

The discriminant table is a compressed statement about the sign of a quadratic form. Writing that form explicitly is useful when an exercise asks for a derivation rather than a memorised classification.

Let

$$A = f_{xx}(p), \qquad B = f_{xy}(p), \qquad C = f_{yy}(p).$$

Then the second-order change is controlled by

$$q(h, k) = Ah^2 + 2Bhk + Ck^2.$$

If $A \neq 0$, completing the square gives

$$q(h, k) = A\left(h + \frac{B}{A}k\right)^2 + \left(C - \frac{B^2}{A}\right)k^2.$$

The coefficient in the second term equals D/A.

If $A = 0$ but $D > 0$, then $-B^2 > 0$, which is impossible. Thus the $D > 0$ cases always have $A \neq 0$, so the completed-square argument loses no cases. If $D < 0$ and $A = C = 0$, then $B \neq 0$ and $q(h, k) = 2Bhk$ still has opposite signs in suitable directions.

Theorem 3.5: Sign control from a positive quadratic form

If $q(h, k)$ is positive definite, then there is $c > 0$ such that

$$q(h, k) \geq c(h^2 + k^2)$$

for every (h, k). If it is negative definite, the reversed inequality holds for some $c > 0$ after replacing q by $-q$.

Proof. Restrict q to the unit circle $h^2 + k^2 = 1$. This circle is compact and q is continuous. Positive definiteness makes $q > 0$ there, so the Extreme Value Theorem gives a positive minimum c. For a nonzero vector u, homogeneity gives

$$q(u) = \|u\|^2 q\left(\frac{u}{\|u\|}\right) \geq c\|u\|^2.$$

The negative-definite statement applies the same argument to $-q$. □

This theorem supplies the omitted quantitative step in the Hessian proof: the Taylor remainder is $o(\|u\|^2)$, so for sufficiently small u it has magnitude below $c\|u\|^2/4$. It cannot overturn the sign supplied by the quadratic part.

What higher-order partials are for

When $D = 0$, record the direction on which the quadratic term vanishes. Then inspect the first nonzero higher-order term along that direction or compare the original function on carefully chosen curves. A higher derivative is evidence only after its sign has been linked to nearby function values.

Worked Example (Week 4): completing squares proves a global minimum

Let

$$r(x,y) = x^2 - xy + y^2 - 2x + y.$$

Method 1: Stationary point, then completing the square

$$\begin{aligned} r_x &= 2x - y - 2 = 0 \\ r_y &= -x + 2y + 1 = 0 \\ \implies y &= 2x - 2 && (r_x = 0) \\ -x + 2(2x-2) + 1 &= 0 && (r_y = 0) \\ \implies 3x - 3 = 0 &\implies x = 1, \quad y = 2(1) - 2 = 0 \\ \implies (x,y) &= (1,0) \end{aligned}$$

Completing squares, first in $x-1$ (expand $(x-1)^2 - y(x-1) + y^2 - 1 = x^2 - 2x + 1 - xy + y + y^2 - 1 = r(x,y)$), then in $x - 1 - \frac{y}{2}$:

$$\begin{aligned} r(x,y) &= (x-1)^2 - y(x-1) + y^2 - 1 \\ &= \left(x - 1 - \frac{y}{2}\right)^2 - \frac{y^2}{4} + y^2 - 1 \\ &= \left(x - 1 - \frac{y}{2}\right)^2 + \frac{3}{4}y^2 - 1 \\ &\geq -1 && (\text{squares} \geq 0;\ \text{equality iff } x - 1 - \tfrac{y}{2} = 0 = y) \\ r(1,0) &= 1 - 0 + 0 - 2 + 0 = -1 \\ r(x,0) &= x^2 - 2x \to \infty && (|x| \to \infty :\ \text{no absolute max}) \end{aligned}$$

$$\boxed{\min_{\mathbb{R}^2} r = -1 \text{ at } (1,0)}$$

Method 2: Hessian test, then an exact shift to the stationary point
The second partials are $r_{xx} = 2$, $r_{xy} = \frac{\partial}{\partial y}(2x - y - 2) = -1$, $r_{yy} = 2$, constants and hence continuous, so the Hessian test applies:

$$D(1,0) = (2)(2) - (-1)^2 = 3 > 0, \qquad r_{xx} = 2 > 0,$$

and $(1,0)$ is a strict local minimum. To make this global, put $x = 1 + h$, $y = k$. Since r is quadratic, the expansion has no remainder:

$$\begin{aligned} r(1+h,k) &= (1+h)^2 - (1+h)k + k^2 - 2(1+h) + k \\ &= 1 + 2h + h^2 - k - hk + k^2 - 2 - 2h + k = h^2 - hk + k^2 - 1. \end{aligned}$$

The quadratic form satisfies

$$h^2 - hk + k^2 = \tfrac{1}{2}\left(h^2 + k^2\right) + \tfrac{1}{2}(h-k)^2 \geq \tfrac{1}{2}\left(h^2 + k^2\right),$$

so $r(1+h,k) \geq -1 + \frac{1}{2}(h^2 + k^2)$, with equality only at $h = k = 0$:

$$\boxed{\min_{\mathbb{R}^2} r = -1 \text{ at } (1,0)}$$

Exam Focus: local proof versus global proof

The Hessian test makes a local conclusion. A completed-square identity can make a global conclusion because it bounds the function at every point of the domain. On an unbounded domain, do not claim a global maximum or minimum just because a stationary point has a favourable Hessian.

3.4 Absolute extrema on compact domains

Definition 3.3: Compactness in Euclidean space

A subset of $\mathbb{R}^m$ is compact when every open cover has a finite subcover. In Euclidean space this is equivalent to being closed and bounded. For calculations, "closed and bounded" is the usable compactness test.

Theorem 3.6: Extreme Value Theorem

If $D \subseteq \mathbb{R}^m$ is nonempty and compact and $f : D \to \mathbb{R}$ is continuous, then f attains both an absolute maximum and an absolute minimum on D.

Proof. Continuity maps compact sets to compact sets: if $\{V_\alpha\}$ covers $f(D)$, then $\{f^{-1}(V_\alpha)\}$ is an open cover of D, and a finite subcover pulls forward to a finite subcover. Thus $f(D)$ is compact in $\mathbb{R}$.

A nonempty compact subset of $\mathbb{R}$ is closed and bounded. Let $M = \sup f(D)$ and $m = \inf f(D)$. Closedness puts both endpoints M, m back in $f(D)$. There are points $p, q \in D$ with $f(p) = M$ and $f(q) = m$, which are the required absolute extrema. □

The theorem is an existence result, not a candidate list. On a region with boundary, a complete calculation combines interior critical points with every boundary piece, including corners and endpoints.

Theorem 3.7: Boundary candidate strategy

Let D be a closed rectangle and let f be continuous on D. Every absolute extremum occurs either at an interior critical point, at a relative interior point of an edge where the edge restriction has derivative zero or has no derivative, or at a corner.

Proof. The Extreme Value Theorem gives an absolute extremum $p \in D$. If p is interior, a missing first partial makes it critical by definition; otherwise Fermat's theorem makes its gradient zero. If p lies in the relative interior of an edge, restrict f to that closed interval. If the one-variable derivative exists, Fermat's theorem makes it zero; if not, p is already one of the listed nondifferentiable edge candidates. The remaining endpoints are corners. These cases exhaust D. □

Under these hypotheses the candidate set need not be finite. A finite comparison is justified only after showing that the interior and edge stationary or nondifferentiable candidates are finite, or reducing them to finitely many objective values by a separate argument.

Worked Example (Week 4): absolute extrema on a rectangle

On $D = [0,3] \times [0,2]$, let

$$f(x,y) = x^2 - xy + y^2 - 2x + y.$$

Method 1: Interior critical points, edge restrictions, comparison

$$\begin{aligned} f_x &= 2x - y - 2 = 0 \implies y = 2x - 2 \\ f_y &= 2y - x + 1 = 0 \implies 2(2x-2) - x + 1 = 3x - 3 = 0 \qquad \implies (1,0) \in \partial D \end{aligned}$$

The only stationary point has $y = 0$, so it is on the edge, and f is a polynomial, so no partial fails to exist. No interior critical point, so by the Extreme Value Theorem (D compact, f continuous) the extrema lie on the edges. Substituting each edge into f:

$$\begin{aligned} y = 0: &\quad f = x^2 - 0 + 0 - 2x + 0 = x^2 - 2x,\ f' = 2x - 2 = 0 \iff x = 1 \\ &\quad f(1,0) = -1,\ f(0,0) = 0,\ f(3,0) = 9 - 6 = 3 \\ y = 2: &\quad f = x^2 - 2x + 4 - 2x + 2 = x^2 - 4x + 6,\ f' = 2x - 4 = 0 \iff x = 2 \\ &\quad f(2,2) = 4 - 8 + 6 = 2,\ f(0,2) = 6,\ f(3,2) = 9 - 12 + 6 = 3 \\ x = 0: &\quad f = 0 - 0 + y^2 - 0 + y = y^2 + y,\ f' = 2y + 1 > 0 \text{ on } [0,2] \\ &\quad f(0,0) = 0,\ f(0,2) = 4 + 2 = 6 \\ x = 3: &\quad f = 9 - 3y + y^2 - 6 + y = y^2 - 2y + 3,\ f' = 2y - 2 = 0 \iff y = 1 \\ &\quad f(3,1) = 1 - 2 + 3 = 2,\ f(3,0) = 3,\ f(3,2) = 4 - 4 + 3 = 3 \end{aligned}$$

Comparing the seven candidate values $-1, 0, 3, 2, 6, 3, 2$:

$$\begin{aligned} \min_D f &= f(1,0) = -1 \\ \max_D f &= f(0,2) = 6 \end{aligned}$$

$$\boxed{\min_D f = -1 \text{ at } (1,0), \qquad \max_D f = 6 \text{ at } (0,2)}$$

Method 2: Completed square for the minimum, edgewise convexity for the maximum
The previous example gives, for every $(x,y) \in \mathbb{R}^2$,

$$f(x,y) = \left(x - 1 - \frac{y}{2}\right)^2 + \frac{3}{4}y^2 - 1 \geq -1,$$

with equality only at $(1,0)$, which lies in D; so $\min_D f = -1$. For the maximum, fix $y \in [0,2]$. Then $x \mapsto f(x,y)$ has second derivative $f_{xx} = 2 > 0$, an upward parabola, which on $[0,3]$ is largest at an endpoint:

$$\begin{aligned} &f(x,y) \leq \max\{f(0,y),\ f(3,y)\}, \\ &f(0,y) = y^2 + y \leq 4 + 2 = 6 && (2y + 1 > 0: \text{ increasing on } [0,2]), \\ &f(3,y) = (y-1)^2 + 2 \leq 1 + 2 = 3 && ((y-1)^2 \leq 1 \text{ on } [0,2]). \end{aligned}$$

Hence $f \leq 6$ on D, and $f(0,2) = 6$ attains it:

$$\boxed{\min_D f = -1 \text{ at } (1,0), \qquad \max_D f = 6 \text{ at } (0,2)}$$

Exam Focus: complete extrema workflow

- State the domain and decide whether absolute extrema are guaranteed.
- Find all interior critical points, including undefined-partial points.
- Use the Hessian test only when its hypotheses hold and $D \neq 0$.
- Restrict to every boundary curve as a one-variable problem.
- Compare values at interior candidates, boundary candidates, and corners.
- Do not introduce a constraint multiplier unless the problem imposes a separate constraint equation; that method belongs to Chapter 6.

3.5 Chapter Summary

Key Formula: Chapter 3 recap

$$\nabla f(p) = 0 \text{ is necessary for an interior smooth extremum,} \qquad D = f_{xx}f_{yy} - f_{xy}^2.$$

Positive-definite quadratic behaviour gives a minimum, negative-definite behaviour gives a maximum, and indefinite behaviour gives a saddle. On a closed bounded domain, continuity guarantees absolute extrema.

Exam Focus: Chapter 3 takeaways

Find every critical point, including undefined partials, and investigate $D = 0$. For global extrema, use compactness and every boundary restriction; on an unbounded domain, prove the global bound directly. For saddles, exhibit values on both sides; for strict extrema, rule out nearby equality. Record all attaining points when useful and name the theorem or comparison used.

QRS@NTU

4 Topic 4: Tangent Spaces and Differentiability

This chapter turns coordinate slopes into a genuine local linear model. For a graph in three-dimensional space, the model is a tangent plane. For a function of two variables, total differentiability says that the error made by that plane is negligible compared with the input displacement. The chain rule, implicit differentiation, and directional derivatives use this model later in Chapter 5; they are deliberately not developed here.

4.1 Planes and the vector product

A plane through a point needs two independent directions, or equivalently one nonzero normal vector. The vector product supplies that normal from two directions in $\mathbb{R}^3$. We write points as $P = (x, y, z)$ and identify a displacement with its vector.

Definition 4.1: Cross product

For $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ in $\mathbb{R}^3$, define

$$u \times v = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \big(u_2v_3 - u_3v_2,\ u_3v_1 - u_1v_3,\ u_1v_2 - u_2v_1\big).$$

The determinant notation is a compact memory aid. Its first row contains the standard basis vectors, and expansion along that row gives the component formula on the right.

Theorem 4.1: Cross product gives a normal and an area

For $u, v \in \mathbb{R}^3$,

$$\langle u, u \times v\rangle = 0, \qquad \langle v, u \times v\rangle = 0, \qquad \|u \times v\|^2 = \|u\|^2 \|v\|^2 - \langle u, v\rangle^2 .$$

Consequently, if u and v are nonzero and nonparallel, then $u \times v$ is a nonzero vector perpendicular to both, and

$$\|u \times v\| = \|u\| \|v\| \sin\theta,$$

where $\theta \in [0, \pi]$ is their angle.

Proof. Put $w = u \times v$. Expanding the first dot product by components gives

$$\begin{aligned} \langle u, w\rangle &= u_1(u_2v_3 - u_3v_2) + u_2(u_3v_1 - u_1v_3) + u_3(u_1v_2 - u_2v_1) \\ &= u_1u_2v_3 - u_1u_3v_2 + u_2u_3v_1 - u_1u_2v_3 + u_1u_3v_2 - u_2u_3v_1 = 0. \end{aligned}$$

The same expansion with v in place of u has cancelling pairs, so $\langle v, w\rangle = 0$.

For the length, expand the three squares in $\|w\|^2$ and collect the diagonal and mixed terms:

$$\begin{aligned} \|w\|^2 &= (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 \\ &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) \\ &\quad - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\ &= \|u\|^2 \|v\|^2 - \langle u, v\rangle^2 . \end{aligned}$$

If $u, v \neq 0$, the dot-product angle formula gives $\langle u, v\rangle = \|u\| \|v\| \cos\theta$. Substitution yields

$$\|w\|^2 = \|u\|^2 \|v\|^2 (1 - \cos^2\theta) = \|u\|^2 \|v\|^2 \sin^2\theta.$$

All factors are nonnegative, so taking square roots proves the formula. For nonparallel vectors, $\sin\theta \neq 0$, hence $w \neq 0$. □

The order matters: $u \times v = -(v \times u)$. It follows from the component formula that the product distributes over addition and is compatible with scalar multiplication, but it is not associative.

Definition 4.2: Plane in normal and direction form

Let $P_0 = (x_0, y_0, z_0)$. If $n = (A, B, C) \neq 0$, the plane through P_0 with normal n is

$$\Pi_n(P_0) = \{(x, y, z) \in \mathbb{R}^3 : A(x - x_0) + B(y - y_0) + C(z - z_0) = 0\}.$$

If u, v are nonparallel nonzero vectors, the plane through P_0 in their directions is

$$\Pi_{u,v}(P_0) = \{P_0 + su + tv : s, t \in \mathbb{R}\}.$$

Theorem 4.2: Equivalent descriptions of a plane

Let u, v be nonparallel nonzero vectors and put $n = u \times v$. Then $n \neq 0$, and

$$\Pi_{u,v}(P_0) = \Pi_n(P_0).$$

Thus a point and a nonzero normal determine one plane, while a point and two nonparallel directions determine the same plane uniquely.

Proof. The preceding theorem gives $n \neq 0$ and $\langle n, u\rangle = \langle n, v\rangle = 0$. If $P = P_0 + su + tv$, then

$$\langle n, P - P_0\rangle = s\langle n, u\rangle + t\langle n, v\rangle = 0,$$

so $\Pi_{u,v}(P_0) \subseteq \Pi_n(P_0)$.

Conversely, u, v, n are linearly independent because their scalar triple product is

$$\det[u\ v\ n] = \langle u \times v, n\rangle = \|n\|^2 \neq 0.$$

Hence u, v, n form a basis of $\mathbb{R}^3$. Write $P - P_0 = su + tv + rn$ in this basis. If $P \in \Pi_n(P_0)$, then

$$0 = \langle n, P - P_0\rangle = r\|n\|^2,$$

so $r = 0$ and $P = P_0 + su + tv$. This proves the reverse inclusion.

Finally, any plane through P_0 perpendicular to n must contain exactly the vectors q satisfying $\langle n, q\rangle = 0$, which is the displayed normal equation. Therefore it is unique. □

Plane equations to recognize immediately

$$\boxed{A(x - x_0) + B(y - y_0) + C(z - z_0) = 0} \iff \boxed{P = P_0 + su + tv}.$$

For the second form, use $n = u \times v$ to move to the first form. For the first form, choose any two independent vectors orthogonal to n when direction vectors are needed.

Worked Example (geometry): two directions to a plane

Let $P_0 = (1, -1, 2)$, $u = (1, 1, 0)$, and $v = (1, 0, 1)$.
Method 1: Cross-product normal. Expand the determinant of Definition 4.1 along its first row:

$$u \times v = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = \big(1 \cdot 1 - 0 \cdot 0,\ 0 \cdot 1 - 1 \cdot 1,\ 1 \cdot 0 - 1 \cdot 1\big)$$
$$= (1, -1, -1) \neq 0.$$

The product is nonzero, so u, v are nonparallel and Theorem 4.2 gives $\Pi_{u,v}(P_0) = \Pi_n(P_0)$ with $n = (1, -1, -1)$. The normal equation through P_0 is

$$0 = 1 \cdot (x - 1) + (-1) \cdot \big(y - (-1)\big) + (-1) \cdot (z - 2)$$
$$= x - 1 - y - 1 - z + 2 = x - y - z.$$

The direction form uses the given point and directions directly:

$$\boxed{x - y - z = 0}, \qquad \boxed{(x, y, z) = (1, -1, 2) + s(1, 1, 0) + t(1, 0, 1)}.$$

Method 2: Eliminating the parameters. Write $P_0 + su + tv$ by components:

$$x = 1 + s + t, \qquad y = -1 + s, \qquad z = 2 + t.$$

The second and third equations give $s = y + 1$ and $t = z - 2$. Substitute both into the first:

$$x = 1 + (y + 1) + (z - 2) = y + z \iff x - y - z = 0.$$

Conversely, any (x, y, z) with $x = y + z$ arises from $s = y + 1$, $t = z - 2$, so both descriptions give the same set:

$$\boxed{x - y - z = 0}, \qquad \boxed{(x, y, z) = (1, -1, 2) + s(1, 1, 0) + t(1, 0, 1)}.$$

Cross-product order and plane data

The normal $v \times u$ is the negative of $u \times v$. This changes the signs in a normal equation but not the plane. More seriously, two parallel directions do *not* determine a plane: their cross product is zero, so there is no nonzero normal from them.

4.2 Tangent plane functions

Let $f : D \to \mathbb{R}$ have open domain $D \subseteq \mathbb{R}^2$. Its graph is the surface

$$\Gamma_f = \{(x, y, f(x, y)) : (x, y) \in D\} \subseteq \mathbb{R}^3.$$

Fix $p = (a, b) \in D$ and set $P = (a, b, f(a, b))$. The vertical coordinate slices through P have parametrizations

$$v_1(t) = (t, b, f(t, b)), \qquad v_2(t) = (a, t, f(a, t)).$$

When the partial derivatives exist, their tangent vectors are

$$v_1'(a) = (1, 0, f_x(a, b)), \qquad v_2'(b) = (0, 1, f_y(a, b)).$$

They cannot be parallel because their first two components are $(1, 0)$ and $(0, 1)$.

Definition 4.3: Tangent plane function

If $f_x(a,b)$ and $f_y(a,b)$ exist, the tangent plane function of f at $p=(a,b)$ is

$$L_{f,p}(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Its graph is the plane through $(a,b,f(a,b))$ parallel to the two coordinate slice tangents.

Theorem 4.3: Tangent graph basis and normal

Under the hypotheses of Definition 4.3, the graph of $L_{f,p}$ has direction basis

$$t_x = (1,0,f_x(a,b)), \qquad t_y = (0,1,f_y(a,b)),$$

and normal vector

$$n = t_x \times t_y = (-f_x(a,b), -f_y(a,b), 1).$$

Consequently, its equation is exactly $z = L_{f,p}(x,y)$.

Proof. Starting at $P=(a,b,f(a,b))$, a point of the plane generated by the slice tangents is

$$P + st_x + tt_y = (a+s, b+t, f(a,b) + sf_x(a,b) + tf_y(a,b)).$$

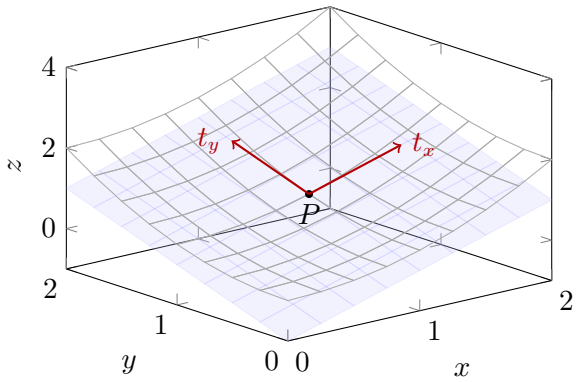
Set $x = a+s$ and $y = b+t$. Then $s = x-a$ and $t = y-b$, so its third coordinate is

$$f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) = L_{f,p}(x,y).$$

The cross-product component formula gives

$$t_x \times t_y = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 0 & f_x(a,b) \\ 0 & 1 & f_y(a,b) \end{vmatrix} = (-f_x(a,b), -f_y(a,b), 1).$$

The last component is 1, so this normal is nonzero. By Theorem 4.2, it is normal to precisely the plane just parametrized. □



For $f = (x^2+y^2)/2$, the mesh and blue tangent plane $z = x+y-1$ touch at $P=(1,1,1)$. Their tangent vectors are $(1,0,1)$ and $(0,1,1)$, and their normal is $(-1,-1,1)$. At this stage, existence of the two partials constructs a plane, but it does not yet say that the surface is well approximated by that plane in every nearby direction.

Worked Example (Week 3, p. 15): a smooth tangent plane

For $f(x,y) = 2x^2 + y^2$ at $p = (1,1)$:
Method 1: Graph formula $z = L_{f,p}(x,y)$. Differentiate in each variable with the other held fixed, then evaluate:

$$f_x(x,y) = 4x, \qquad f_y(x,y) = 2y,$$
$$f(1,1) = 2\cdot 1^2 + 1^2 = 3, \qquad f_x(1,1) = 4, \quad f_y(1,1) = 2.$$

Substitute into Definition 4.3:

$$L_{f,(1,1)}(x,y) = 3 + 4(x-1) + 2(y-1) = 3 + 4x - 4 + 2y - 2 = 4x + 2y - 3.$$

By Theorem 4.3, the graph normal is $t_x \times t_y = (1,0,4)\times(0,1,2) = (0\cdot 2 - 4\cdot 1,\ 4\cdot 0 - 1\cdot 2,\ 1\cdot 1 - 0\cdot 0) = (-4,-2,1)$. Both partials are polynomials, hence continuous on $\mathbb{R}^2$, so Theorem 4.7 makes f totally differentiable at $(1,1)$ and this plane is its first-order approximation:

$$\boxed{L_{f,(1,1)}(x,y) = 4x + 2y - 3}$$

Method 2: Level-surface normal. (Uses the gradient normal to a level surface, Chapter 5.) The graph is the level set $F = 0$ of $F(x,y,z) = 2x^2 + y^2 - z$, with $F_x = 4x$, $F_y = 2y$, $F_z = -1$. At $P = (1,1,3)$, $\nabla F(1,1,3) = (4,2,-1) \neq 0$, so the plane is

$$4(x-1) + 2(y-1) - (z-3) = 0 \iff z = 4x - 4 + 2y - 2 + 3 = 4x + 2y - 3.$$

$$\boxed{L_{f,(1,1)}(x,y) = 4x + 2y - 3}$$

Method 3: Expansion about p and its linear part. Put $x = 1 + h$, $y = 1 + k$ and expand:

$$f(1+h, 1+k) = 2(1 + 2h + h^2) + (1 + 2k + k^2) = 3 + 4h + 2k + (2h^2 + k^2).$$

The remainder $r(h,k) = 2h^2 + k^2$ satisfies, for $(h,k) \neq 0$,

$$\frac{|r(h,k)|}{\|(h,k)\|} \leq \frac{2(h^2+k^2)}{\|(h,k)\|} = 2\,\|(h,k)\| \longrightarrow 0.$$

So $df_{(1,1)}(h,k) = 4h + 2k$ by Definition 4.4, unique by Theorem 4.4. With $h = x - 1$, $k = y - 1$:

$$L_{f,(1,1)}(x,y) = 3 + 4(x-1) + 2(y-1) = 4x + 2y - 3.$$

$$\boxed{L_{f,(1,1)}(x,y) = 4x + 2y - 3}$$

4.3 Total differentiability and the total differential

Definition 4.4: Total differentiability

Let $f : D \to \mathbb{R}$ have open domain and let $p = (a, b) \in D$. The function is *totally differentiable* at p if there is a linear map $A : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{f(p+h) - f(p) - A(h)}{\|h\|} = 0.$$

The associated affine approximation is

$$L(x) = f(p) + A(x - p).$$

When it exists, A is called the total differential and is denoted df_p.

Equivalently, there is a remainder $r(h)$ satisfying

$$f(p+h) = f(p) + df_p(h) + r(h), \qquad \frac{r(h)}{\|h\|} \to 0.$$

The notation $r(h) = o(\|h\|)$ means precisely this limiting relation.

QRS@NTU – qrsntu.org

Worked Example (Week 3, pp. 20–21): explicit remainder

Let $f(x,y) = 2x^2 - xy$, and fix $p = (a,b)$.
Method 1: Expanding the increment (definition). With $h = x-a$, $k = y-b$, expand the two products:

$$2(a+h)^2 = 2a^2 + 4ah + 2h^2, \qquad (a+h)(b+k) = ab + ak + bh + hk.$$

Hence

$$\begin{aligned} f(a+h,b+k) - f(a,b) &= 2a^2 + 4ah + 2h^2 - ab - ak - bh - hk - 2a^2 + ab \\ &= (4a-b)h - ak + 2h^2 - hk. \end{aligned}$$

Take the linear candidate $df_{(a,b)}(h,k) = (4a-b)h - ak$ and the remainder $r(h,k) = 2h^2 - hk$. Since $|h| \le \|(h,k)\|$ and $|k| \le \|(h,k)\|$, for $(h,k) \neq 0$,

$$\frac{|r(h,k)|}{\|(h,k)\|} \le \frac{2h^2 + |h||k|}{\|(h,k)\|} \le \frac{3\,\|(h,k)\|^2}{\|(h,k)\|} = 3\,\|(h,k)\| \longrightarrow 0.$$

By Definition 4.4, f is totally differentiable at every (a,b), and $L = f(a,b) + df_{(a,b)}(x-a, y-b)$:

$$\boxed{L_{f,(a,b)}(x,y) = 2a^2 - ab + (4a-b)(x-a) - a(y-b).}$$

Method 2: Continuous-partials criterion. Differentiate in each variable with the other held fixed:

$$f_x(x,y) = 4x - y, \qquad f_y(x,y) = -x.$$

Both are polynomials, hence continuous on all of $\mathbb{R}^2$, so every disk about (a,b) meets the hypothesis of Theorem 4.7 and f is totally differentiable at (a,b). Theorem 4.5 then gives $L_{f,(a,b)}$ from

$$f(a,b) = 2a^2 - ab, \qquad f_x(a,b) = 4a - b, \qquad f_y(a,b) = -a :$$

$$\boxed{L_{f,(a,b)}(x,y) = 2a^2 - ab + (4a-b)(x-a) - a(y-b).}$$

4.4 Consequences of the total-differentiability definition

The total-differentiability definition makes an affine function a valid local model only when its error is negligible per unit of input distance. We now derive the properties that identify that linear map with the tangent-plane coefficients.

Theorem 4.4: Uniqueness of the linear derivative

If A and B both satisfy the total-differentiability definition for f at p, then $A = B$.

Proof. Subtract the two remainder relations to obtain, for $h \neq 0$,

$$\frac{(A-B)(h)}{\|h\|} = \frac{f(p+h) - f(p) - B(h)}{\|h\|} - \frac{f(p+h) - f(p) - A(h)}{\|h\|} \longrightarrow 0.$$

Fix any $q \in \mathbb{R}^2$. For $t \neq 0$, use linearity with $h = tq$:

$$\frac{(A-B)(tq)}{\|tq\|} = \frac{t(A-B)(q)}{|t|\,\|q\|} = \operatorname{sgn}(t)\frac{(A-B)(q)}{\|q\|}$$

when $q \neq 0$. The left side tends to zero as $t \to 0$. Taking $t > 0$ shows $(A-B)(q)/\|q\| = 0$, hence $(A-B)(q) = 0$. This holds for every nonzero q, and also for $q = 0$ by linearity. Therefore $A = B$. □

Theorem 4.5: Differentiability determines the partial derivatives

If f is totally differentiable at $p = (a, b)$ with linear derivative A, then $f_x(a, b)$ and $f_y(a, b)$ exist and

$$A(h, k) = f_x(a, b)h + f_y(a, b)k.$$

Thus the unique affine approximation is the tangent plane function

$$L_{f,p}(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Proof. Let $e_1 = (1, 0)$. For nonzero t, differentiability gives

$$\frac{f(p + te_1) - f(p)}{t} = A(e_1) + \frac{r(te_1)}{t}.$$

Since $\|te_1\| = |t|$,

$$\left|\frac{r(te_1)}{t}\right| = \frac{|r(te_1)|}{\|te_1\|} \longrightarrow 0.$$

The left side therefore has limit $A(e_1)$, which is $f_x(a, b)$. Repeating the argument with $e_2 = (0, 1)$ gives $f_y(a, b) = A(e_2)$.

Every (h, k) equals $he_1 + ke_2$. By linearity,

$$A(h, k) = hA(e_1) + kA(e_2) = f_x(a, b)h + f_y(a, b)k.$$

Substitute $(h, k) = (x - a, y - b)$ into $f(p) + A(h, k)$ to get the displayed tangent-plane formula. □

Theorem 4.6: Total differentiability implies continuity

If f is totally differentiable at p, then f is continuous at p. Equivalently, failure of continuity at p proves failure of total differentiability there.

Proof. Write

$$f(p + h) - f(p) = A(h) + r(h), \qquad \frac{r(h)}{\|h\|} \to 0.$$

Let $A(e_1) = \alpha$ and $A(e_2) = \beta$. For $h = (h_1, h_2)$, the Cauchy–Schwarz inequality gives

$$|A(h)| = |\alpha h_1 + \beta h_2| \le \sqrt{\alpha^2 + \beta^2}\, \|h\|.$$

Also, choose $\delta_1 > 0$ so that $0 < \|h\| < \delta_1$ implies $|r(h)| \le \|h\|$. Then

$$|f(p + h) - f(p)| \le \left(\sqrt{\alpha^2 + \beta^2} + 1\right) \|h\| \longrightarrow 0.$$

This is continuity at p. The contrapositive gives the final assertion. □

Total differential and first-order error

If f is totally differentiable at $p = (a, b)$, then

$$\boxed{df_p(h, k) = f_x(a, b)h + f_y(a, b)k}$$

and, with $\Delta x = x - a$, $\Delta y = y - b$,

$$\boxed{f(x, y) = f(a, b) + f_x(a, b)\Delta x + f_y(a, b)\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right).}$$

The approximation error is smaller than every fixed multiple of the input distance once the input is sufficiently close to p.

4.5 When partial derivatives are insufficient

The implication from total differentiability to partial derivatives is one-way. Partials are necessary for total differentiability, but they need not be sufficient because they inspect only horizontal and vertical changes.

Worked Example (Week 3, pp. 23–24): partials but no differentiability

Define

$$f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Method 1: Partials by definition, then the forced remainder on $y=x$. For $t\neq 0$, $f(t,0)=\dfrac{t\cdot 0}{t^2}=0$ and $f(0,t)=\dfrac{0\cdot t}{t^2}=0$, so

$$f_x(0,0)=\lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t}=\lim_{t\to 0}\frac{0}{t}=0,\qquad f_y(0,0)=\lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t}=0.$$

If f were totally differentiable, Theorem 4.5 would force $df_{(0,0)}=0$ and $L_{f,(0,0)}=0$, so the remainder quotient $|f(x,y)-0|/\|(x,y)\|$ would tend to 0. But for $x\neq 0$, $f(x,x)=x^2/(2x^2)=1/2$ and $\|(x,x)\|=\sqrt{2}|x|$, so

$$\frac{|f(x,x)-0|}{\|(x,x)\|}=\frac{1/2}{\sqrt{2}|x|}=\frac{1}{2\sqrt{2}\,|x|}\longrightarrow\infty\qquad(x\to 0).$$

$$\boxed{f_x(0,0)=f_y(0,0)=0,\quad f\text{ is not totally differentiable at }(0,0).}$$

Method 2: Polar remainder. Put $x=\rho\cos\theta$, $y=\rho\sin\theta$ with $\rho>0$:

$$f(\rho\cos\theta,\rho\sin\theta)=\frac{\rho^2\cos\theta\sin\theta}{\rho^2}=\tfrac{1}{2}\sin 2\theta.$$

On the axes ($\theta\in\{0,\frac{\pi}{2},\pi,\frac{3\pi}{2}\}$), $\sin 2\theta=0$, so $f(t,0)=f(0,t)=0$ for $t\neq 0$; both difference quotients are $0/t$ and $f_x(0,0)=f_y(0,0)=0$. The forced plane is $L_{f,(0,0)}=0$, with remainder quotient

$$\frac{|f-0|}{\rho}=\frac{|\sin 2\theta|}{2\rho},\qquad \theta=\tfrac{\pi}{4}:\ \frac{1}{2\rho}\longrightarrow\infty\quad(\rho\to 0^+),$$

which does not tend to 0:

$$\boxed{f_x(0,0)=f_y(0,0)=0,\quad f\text{ is not totally differentiable at }(0,0).}$$

This example also exposes a quicker obstruction. Along $y=x$, $f(x,x)=1/2$, while $f(0,0)=0$. The function is discontinuous at the origin, so Theorem 4.6 already rules out total differentiability.

Worked Example (continuity obstruction): a noncontinuous function

Define

$$g(x,y)=\begin{cases}\dfrac{x^2-y^2}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Method 1: Two axis paths and the continuity obstruction. Along each axis, away from the origin,

$$g(x,0)=\frac{x^2-0}{x^2+0}=1\quad(x\neq 0),\qquad g(0,y)=\frac{0-y^2}{0+y^2}=-1\quad(y\neq 0).$$

Hence $\lim_{x\to 0} g(x,0) = 1$ and $\lim_{y\to 0} g(0,y) = -1$. Two paths give different limits, so $\lim_{(x,y)\to(0,0)} g(x,y)$ does not exist and g is not continuous at $(0,0)$. By the contrapositive of Theorem 4.6:

$$\boxed{g \text{ is not totally differentiable at } (0,0).}$$

Method 2: Polar form. Put $x=\rho\cos\theta$, $y=\rho\sin\theta$ with $\rho>0$:

$$g(\rho\cos\theta,\rho\sin\theta)=\frac{\rho^2(\cos^2\theta-\sin^2\theta)}{\rho^2}=\cos 2\theta.$$

For every $\rho>0$, the value is 1 at $\theta=0$ and -1 at $\theta=\frac{\pi}{2}$. A limit ℓ would need (take $\varepsilon=1$) $|1-\ell|<1$ and $|-1-\ell|<1$, impossible since $|1-\ell|+|-1-\ell|\geq|1-(-1)|=2$. So g is discontinuous at $(0,0)$, and Theorem 4.6 gives

$$\boxed{g \text{ is not totally differentiable at } (0,0).}$$

Method 3: A missing partial derivative. The x-difference quotient at the origin is

$$\frac{g(t,0)-g(0,0)}{t}=\frac{1-0}{t}=\frac{1}{t}\longrightarrow\pm\infty\quad(t\to 0^{\pm}),$$

so $g_x(0,0)$ does not exist. Theorem 4.5 says total differentiability forces both partials to exist; its contrapositive gives

$$\boxed{g \text{ is not totally differentiable at } (0,0).}$$

Existence of partials is not enough

The existence of f_x and f_y does not establish differentiability. The Week 3 counterexample has both partials equal to zero at the origin and still has no first-order planar approximation there. Use the remainder definition, a valid sufficient criterion, or a direct obstruction such as discontinuity.

4.6 A usable sufficient criterion

Checking the remainder directly is effective for polynomials and carefully chosen functions. For routine formulas, continuous first partial derivatives provide a more efficient criterion.

Theorem 4.7: Continuous partials imply total differentiability

Let $f: D\to\mathbb{R}$ have open domain, let $p=(a,b)\in D$, and suppose there is an open disk about p contained in D on which f_x and f_y exist and are continuous. Then f is totally differentiable at p.

Proof. Take (x,y) sufficiently close to (a,b) that the horizontal and vertical segments used below remain

in the disk. Coordinate telescoping gives

$$f(x,y)-f(a,b)=\big[f(x,y)-f(a,y)\big]+\big[f(a,y)-f(a,b)\big].$$

Apply the one-variable Mean Value Theorem to the first bracket as a function of its first coordinate. There is ξ between a and x such that

$$f(x,y)-f(a,y)=f_x(\xi,y)(x-a).$$

Apply it to the second bracket as a function of its second coordinate. There is η between b and y such that

$$f(a,y)-f(a,b)=f_y(a,\eta)(y-b).$$

Subtract the proposed linear term to obtain

$$\begin{aligned}R(x,y)&=f(x,y)-f(a,b)-f_x(a,b)(x-a)-f_y(a,b)(y-b)\\&=\big[f_x(\xi,y)-f_x(a,b)\big](x-a)\\&\quad+\big[f_y(a,\eta)-f_y(a,b)\big](y-b).\end{aligned}$$

Put $\rho=\|(x-a,y-b)\|$. As $(x,y)\to(a,b)$, both (ξ,y) and (a,η) tend to (a,b). Continuity of the two partials therefore implies

$$\varepsilon_x(x,y)=f_x(\xi,y)-f_x(a,b)\to 0,\qquad \varepsilon_y(x,y)=f_y(a,\eta)-f_y(a,b)\to 0.$$

Using $|x-a|\le\rho$ and $|y-b|\le\rho$ gives

$$\frac{|R(x,y)|}{\rho}\le|\varepsilon_x(x,y)|+|\varepsilon_y(x,y)|\longrightarrow 0.$$

This is total differentiability at p. □

The hypothesis is local: it requires continuity on a neighbourhood, not merely at the point. It is sufficient rather than necessary; a differentiable function can have partial derivatives that fail to be continuous.

Exam Focus: choosing a differentiability argument

- To prove differentiability of a polynomial, rational expression away from a zero denominator, verify that both first partials are continuous nearby and cite Theorem 4.7.
- To prove it at an exceptional point, expand $f(p+h)-f(p)$ and bound the remainder divided by $\|h\|$.
- To disprove it, test the forced tangent plane along one suitable path, or first show that the function is discontinuous.

4.7 Linear approximation and error propagation

At a totally differentiable point, set

$$\Delta x=x-a,\qquad \Delta y=y-b,\qquad \Delta f=f(x,y)-f(a,b).$$

The total differential gives the first-order estimate

$$\Delta f\approx df_{(a,b)}(\Delta x,\Delta y)=f_x(a,b)\Delta x+f_y(a,b)\Delta y.$$

For absolute measurement bounds $|\Delta x|\le E_x$ and $|\Delta y|\le E_y$, the triangle inequality gives the conservative bound

$$|\Delta f|\approx|df_p(\Delta x,\Delta y)|,\qquad |df_p(\Delta x,\Delta y)|\le|f_x(p)|E_x+|f_y(p)|E_y.$$

It is a first-order estimate, not an exact error bound unless additional control of the remainder is supplied.

Theorem 4.8: Linear-approximation formula

If f is totally differentiable at $p = (a, b)$, then

$$f(x, y) - f(a, b) = f_x(a, b)(x - a) + f_y(a, b)(y - b) + o(\|(x - a, y - b)\|) .$$

Hence its tangent plane is the unique first-order affine approximation.

Proof. By Definition 4.4,

$$f(p + h) - f(p) = df_p(h) + r(h), \qquad \frac{r(h)}{\|h\|} \to 0.$$

Theorem 4.5 identifies $df_p(h_1, h_2) = f_x(a, b)h_1 + f_y(a, b)h_2$. Set $h = (x - a, y - b)$ to obtain the displayed formula. Uniqueness follows from Theorem 4.4. □

Worked Example (Week 3, p. 26): cone-volume measurement error

A cone has measured radius $r = 10$ and height $h = 25$. Each measurement has absolute error at most 0.1. Its volume is $V(r, h) = \frac{\pi}{3}r^2h$.

Method 1: Total differential with absolute errors. Differentiate in each variable with the other held fixed:

$$V_r = \frac{\pi}{3} \cdot 2r \cdot h = \frac{2\pi}{3}rh, \qquad V_h = \frac{\pi}{3}r^2.$$

Both are polynomials, continuous on $\mathbb{R}^2$, so V is totally differentiable at $(10, 25)$ (Theorem 4.7) and Theorem 4.8 gives $\Delta V \approx V_r\Delta r + V_h\Delta h$. At $(10, 25)$,

$$V_r(10, 25) = \frac{2\pi \cdot 10 \cdot 25}{3} = \frac{500\pi}{3}, \qquad V_h(10, 25) = \frac{\pi \cdot 10^2}{3} = \frac{100\pi}{3}.$$

With $|\Delta r|, |\Delta h| \le 0.1$, the triangle inequality gives

$$|\Delta V| \lesssim |V_r(10, 25)|\,(0.1) + |V_h(10, 25)|\,(0.1) = \frac{500\pi}{30} + \frac{100\pi}{30} = \frac{600\pi}{30} = 20\pi.$$

$$\boxed{|\Delta V| \lesssim 20\pi}.$$

Method 2: Relative-error form of the differential. Divide the differential by $V = \frac{\pi}{3}r^2h > 0$:

$$\frac{dV}{V} = \frac{\frac{2\pi}{3}rh}{\frac{\pi}{3}r^2h}\,dr + \frac{\frac{\pi}{3}r^2}{\frac{\pi}{3}r^2h}\,dh = 2\,\frac{dr}{r} + \frac{dh}{h}.$$

At $(10, 25)$, $V = \frac{\pi}{3} \cdot 100 \cdot 25 = \frac{2500\pi}{3}$ and, by the triangle inequality with $|dr|, |dh| \le 0.1$,

$$\frac{|dV|}{V} \le 2 \cdot \frac{0.1}{10} + \frac{0.1}{25} = 0.02 + 0.004 = 0.024, \qquad |\Delta V| \lesssim 0.024 \cdot \frac{2500\pi}{3} = \frac{60\pi}{3} = 20\pi.$$

$$\boxed{|\Delta V| \lesssim 20\pi}.$$

Interpreting an error estimate

The cone calculation combines worst-case input errors by the triangle inequality, so it is intentionally conservative. It does not assert that every measurement produces error 20π, nor does it replace a statistical uncertainty model when errors are random.

Beyond this chapter: what comes next

The total differential is the linear map used in the multivariable chain rule and in directional-derivative formulas. Those applications, together with implicit differentiation, belong to Chapter 5; this chapter establishes only the local plane and its error order.

4.8 Chapter Summary

Key Formula: Chapter 4 recap

For $p = (a, b)$ and a totally differentiable f,

$$\begin{aligned} L_{f,p}(x, y) &= f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b), \\ \text{tangent basis:} &\quad (1, 0, f_x(p)),\ (0, 1, f_y(p)), \\ \text{normal:} &\quad (-f_x(p), -f_y(p), 1), \\ \text{error condition:} &\quad \frac{f(p + h) - f(p) - df_p(h)}{\|h\|} \to 0. \end{aligned}$$

Exam Focus: Chapter 4 takeaways

Build a plane from a point and a normal, or two nonparallel directions and their cross product. Partials give a candidate plane; total differentiability requires the stronger remainder condition. It implies partial derivatives and continuity, while neither converse holds without extra hypotheses. Continuous first partials nearby suffice. Use the total differential for first-order error estimates.

5 Topic 5: Chain Rule and Directional Derivatives

This chapter explains change through a map, implicit curves, and rates in arbitrary directions. Differentiability is the prerequisite supplied by Chapter 4; a directional derivative alone is not a substitute for it.

5.1 The chain rule

Theorem 5.1: Chain rule from $\mathbb{R}^2$ to $\mathbb{R}$

Let $z = f(x, y)$ have continuous first partials near (a, b), and let $x = x(t), y = y(t)$ be differentiable at t_0, with $(x(t_0), y(t_0)) = (a, b)$. Then

$$z'(t_0) = f_x(a, b)x'(t_0) + f_y(a, b)y'(t_0).$$

Proof. Differentiability gives

$$f(a + \Delta x, b + \Delta y) - f(a, b) = f_x(a, b)\Delta x + f_y(a, b)\Delta y + r,$$

where $r/\sqrt{\Delta x^2 + \Delta y^2} \to 0$. Divide by Δt. Since $\Delta x/\Delta t \to x'$ and $\Delta y/\Delta t \to y'$, while $\sqrt{\Delta x^2 + \Delta y^2}/|\Delta t|$ stays bounded, $r/\Delta t \to 0$. Taking the limit proves the formula. □

Theorem 5.2: Chain rule in Jacobian form

Let $G : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at a, and let $F : \mathbb{R}^m \to \mathbb{R}^q$ be differentiable at $G(a)$. Then

$$J_{F\circ G}(a) = J_F(G(a))J_G(a).$$

Proof. Write $G(a+h) = G(a)+Ah+r_G(h)$, with $r_G(h) = o(\|h\|)$, and $F(G(a)+k) = F(G(a))+Bk+r_F(k)$, with $r_F(k) = o(\|k\|)$. Substitution gives

$$F(G(a + h)) = F(G(a)) + BAh + Br_G(h) + r_F(Ah + r_G(h)).$$

The two final terms are $o(\|h\|)$ because linear maps are bounded. Thus the derivative is BA, exactly the displayed matrix product. □

Chain-rule bookkeeping

For $z = f(x, y)$ with $x = x(s, t)$ and $y = y(s, t)$,

$$z_s = f_x x_s + f_y y_s, \qquad z_t = f_x x_t + f_y y_t,$$

and every f_x, f_y is evaluated at the transformed point $(x(s, t), y(s, t))$.

Worked Example (Week 3): a $(2,1)$ chain rule

Let $z = x^2y + \sin y$, where $x = t^2 - 1$ and $y = \sin t$.
Method 1: Chain Rule. Write $f(x,y) = x^2y + \sin y$. Its first partials are

$$f_x = 2xy, \qquad f_y = x^2 + \cos y,$$

both continuous on $\mathbb{R}^2$, and $x(t), y(t)$ are differentiable. Hence the Chain Rule (Theorem 5.1) applies:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = f_x\, x'(t) + f_y\, y'(t) = (2xy)\, x'(t) + (x^2 + \cos y)\, y'(t).$$

The inner derivatives are $x'(t) = 2t$ and $y'(t) = \cos t$. At $t = \pi/2$,

$$x = \frac{\pi^2}{4} - 1, \qquad y = \sin\frac{\pi}{2} = 1, \qquad x' = 2\cdot\frac{\pi}{2} = \pi, \qquad y' = \cos\frac{\pi}{2} = 0.$$

The outer partials at $(x,y) = (\pi^2/4 - 1,\, 1)$ are

$$f_x = 2\left(\frac{\pi^2}{4} - 1\right)(1), \qquad f_y = \left(\frac{\pi^2}{4} - 1\right)^2 + \cos 1.$$

Therefore

$$z'(\pi/2) = 2\left(\frac{\pi^2}{4} - 1\right)\pi + \left[\left(\frac{\pi^2}{4} - 1\right)^2 + \cos 1\right]\cdot 0 = 2\pi\left(\frac{\pi^2}{4} - 1\right).$$

$$\boxed{z'(\pi/2) = 2\pi\left(\pi^2/4 - 1\right)}$$

Method 2: Direct substitution. Substitute $x = t^2 - 1$ and $y = \sin t$ before differentiating:

$$z(t) = (t^2 - 1)^2 \sin t + \sin(\sin t).$$

The product rule on the first term and the one-variable chain rule on both terms give

$$z'(t) = 2(t^2 - 1)(2t)\sin t + (t^2 - 1)^2 \cos t + \cos(\sin t)\cos t.$$

At $t = \pi/2$, $\sin t = 1$ and $\cos t = 0$, so

$$z'(\pi/2) = 4\cdot\frac{\pi}{2}\left(\frac{\pi^2}{4} - 1\right)(1) + 0 + 0 = 2\pi\left(\frac{\pi^2}{4} - 1\right).$$

$$\boxed{z'(\pi/2) = 2\pi\left(\pi^2/4 - 1\right)}$$

Theorem 5.3: Second mixed chain derivative

If the displayed functions are twice continuously differentiable, then

$$z_{st} = f_{xx}x_s x_t + f_{xy}(x_s y_t + y_s x_t) + f_{yy}y_s y_t + f_x x_{st} + f_y y_{st}.$$

Proof. Differentiate $z_s = f_x x_s + f_y y_s$ with respect to t. Apply the first chain rule to f_x and f_y, then use Clairaut's theorem to replace f_{yx} by f_{xy}. Collect the four quadratic terms and the two second-derivative terms. □

Checking the mixed formula by expansion

Let $z = x^2y$ with $x = s + t$ and $y = st$. Here

$$f_x = 2xy, \quad f_y = x^2, \quad f_{xx} = 2y, \quad f_{xy} = 2x, \quad f_{yy} = 0.$$

As $x_s = x_t = 1$, $y_s = t$, $y_t = s$, and $x_{st} = 0$, $y_{st} = 1$,

$$z_{st} = 2st + 2(s + t)(s + t) + (s + t)^2.$$

Thus

$$\boxed{z_{st} = 3(s + t)^2 + 2st}.$$

Direct expansion of $(s + t)^2st$ gives the same derivative.

Do not freeze the wrong variables

When differentiating a composite expression, f_x is not a constant unless its input is fixed. Write the intermediate variables, then apply the chain rule. Missing an x_t or evaluating f_x at (s, t) instead of (x, y) is the standard error.

5.2 Implicit curves and surfaces

Theorem 5.4: Implicit-function theorem, scalar form

Let F be C^1 on an open neighbourhood of (a, b, c), with

$$F(a, b, c) = 0, \qquad F_z(a, b, c) \neq 0.$$

Then near (a, b) there is a unique C^1 function $z = g(x, y)$ with $g(a, b) = c$ and $F(x, y, g(x, y)) = 0$. Moreover

$$g_x = -\frac{F_x}{F_z}, \qquad g_y = -\frac{F_y}{F_z}.$$

Proof. This is the scalar implicit-function theorem. For completeness, continuity of F_z makes it have one nonzero sign in a small box. Along each vertical line, endpoint values therefore have opposite signs after shrinking the box; the intermediate value theorem supplies a root. The mean value theorem and a positive lower bound for $|F_z|$ make that root unique and locally Lipschitz in (x, y). Applying the first-order expansion of F to two nearby roots gives differentiability of g; comparing the x and y increments yields the two displayed derivatives. Continuity of the quotient follows from C^1 regularity and the nonvanishing denominator. □

The implicit formula on a circle

For $F(x, y) = x^2 + y^2 - 1$, the unit circle is $F = 0$. At $(0, 1)$, $F_y = 2 \neq 0$, so locally $y = g(x)$ and

$$g'(0) = -\frac{F_x(0, 1)}{F_y(0, 1)} = 0.$$

The tangent line through $(0, 1)$ is therefore

$$\boxed{y = 1}.$$

The calculation is local: it selects the upper semicircle near the point.

Worked Example: an implicit surface derivative

Let

$$F(x,y,z) = x^2 + y^2 + z^2 - 9.$$

Method 1: Implicit Function Theorem formula. The first partials $F_x = 2x$, $F_y = 2y$, $F_z = 2z$ are continuous, so $F \in C^1(\mathbb{R}^3)$. The remaining IFT hypotheses at $(1,2,2)$ are

$$F(1,2,2) = 1 + 4 + 4 - 9 = 0, \qquad F_z(1,2,2) = 2 \cdot 2 = 4 \neq 0.$$

By Theorem 5.4 there is a C^1 function $z = g(x,y)$ near $(1,2)$ with $g(1,2) = 2$, and

$$g_x = -\frac{F_x}{F_z} = -\frac{2x}{2z} = -\frac{x}{z}, \qquad g_y = -\frac{F_y}{F_z} = -\frac{2y}{2z} = -\frac{y}{z}.$$

At $(x,y,z) = (1,2,2)$ these give $g_x(1,2) = -\frac{1}{2}$ and $g_y(1,2) = -\frac{2}{2} = -1$:

$$\boxed{g_x(1,2) = -\frac{1}{2}, \qquad g_y(1,2) = -1}$$

Method 2: Implicit differentiation. The IFT check of Method 1 ($F \in C^1$, $F(1,2,2) = 0$, $F_z(1,2,2) \neq 0$) gives a differentiable g with $g(1,2) = 2$ and

$$x^2 + y^2 + g(x,y)^2 = 9 \qquad \text{near } (1,2).$$

Differentiate with respect to x, holding y fixed; the chain rule on g^2 gives $2g\,g_x$. Then differentiate with respect to y:

$$2x + 2g\,g_x = 0 \implies g_x = -\frac{x}{g}, \qquad 2y + 2g\,g_y = 0 \implies g_y = -\frac{y}{g}.$$

With $g(1,2) = 2$, $g_x(1,2) = -\frac{1}{2}$ and $g_y(1,2) = -\frac{2}{2} = -1$:

$$\boxed{g_x(1,2) = -\frac{1}{2}, \qquad g_y(1,2) = -1}$$

Method 3: Explicit solution for z. Solving gives $z = \pm\sqrt{9 - x^2 - y^2}$. The point has $z = 2 > 0$, and $9 - x^2 - y^2 = 4 > 0$ at $(1,2)$, so near $(1,2)$ the surface is the graph of

$$g(x,y) = \sqrt{9 - x^2 - y^2}, \qquad g(1,2) = \sqrt{9 - 1 - 4} = 2.$$

The chain rule on the square root gives

$$g_x = \frac{-2x}{2\sqrt{9 - x^2 - y^2}} = -\frac{x}{\sqrt{9 - x^2 - y^2}}, \qquad g_y = \frac{-2y}{2\sqrt{9 - x^2 - y^2}} = -\frac{y}{\sqrt{9 - x^2 - y^2}}.$$

At $(1,2)$ the root equals 2, so $g_x(1,2) = -\frac{1}{2}$, $g_y(1,2) = -\frac{2}{2} = -1$:

$$\boxed{g_x(1,2) = -\frac{1}{2}, \qquad g_y(1,2) = -1}$$

Proof scope for the implicit-function theorem

The complete IFT is a structural theorem whose standard proof uses the one-variable IVT and MVT as outlined above. The examinable skill here is to check openness, C^1 regularity, $F = 0$, and a nonzero F_z, then derive the implicit derivatives by the chain rule.

5.3 Directional derivatives and gradients

Definition 5.1: Directional derivative

For a unit vector $u \in \mathbb{R}^n$, the directional derivative of f at an interior point p, when it exists, is

$$D_u f(p) = \lim_{h \to 0} \frac{f(p + hu) - f(p)}{h}.$$

Theorem 5.5: Differentiability gives gradient dot direction

If f is differentiable at p, then for every unit vector u,

$$D_u f(p) = \nabla f(p) \cdot u.$$

Proof. Differentiability says

$$f(p + hu) - f(p) = \nabla f(p) \cdot (hu) + o(|h|).$$

Divide by h and let $h \to 0$. The remainder divided by h tends to zero, leaving the dot product. □

Theorem 5.6: Fastest ascent and descent

If f is differentiable at p, then for every unit u,

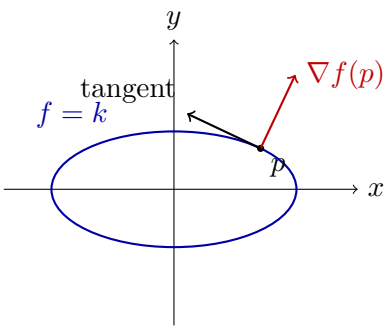
$$D_u f(p) = \|\nabla f(p)\| \cos\theta,$$

when $\nabla f(p) \neq 0$, where θ is the angle from $\nabla f(p)$ to u. Thus a nonzero gradient points in fastest ascent and its negative points in fastest descent. If $\nabla f(p) = 0$, every unit directional rate is zero.

Proof. The preceding theorem and the dot-product identity give the formula. Cauchy–Schwarz gives

$$|\nabla f(p) \cdot u| \leq \|\nabla f(p)\| \, \|u\| = \|\nabla f(p)\| \, .$$

Equality holds precisely when u is parallel to the gradient, with its sign choosing ascent or descent. If the gradient is zero, the dot product is zero. □



Here $f(x,y) = x^2/(1.8)^2 + y^2/(0.85)^2$ and the ellipse is $f = 1$.

Theorem 5.7: Gradient is normal to a regular level curve

If f is differentiable at p, $\nabla f(p) \neq 0$, and $r(t)$ is a differentiable curve in the level set $f(r(t)) = k$ with $r(t_0) = p$, then

$$\nabla f(p) \cdot r'(t_0) = 0.$$

Proof. Differentiate the identity $f(r(t)) = k$ at t_0 by the chain rule. The derivative of the constant is zero, giving the claimed orthogonality. □

Worked Example (Week 4): fastest rate for xe^y

For $f(x,y) = xe^y$,
Method 1: Gradient and Cauchy–Schwarz. The first partials $f_x = e^y$ and $f_y = xe^y$ are continuous on $\mathbb{R}^2$, so f is differentiable and Theorem 5.5 gives $D_u f(2,0) = \nabla f(2,0) \cdot u$. At $(2,0)$,

$$f_x(2,0) = e^0 = 1, \qquad f_y(2,0) = 2e^0 = 2, \qquad \nabla f(2,0) = (1,2), \qquad \|\nabla f(2,0)\| = \sqrt{1^2+2^2} = \sqrt{5}.$$

For a unit vector u, Cauchy–Schwarz gives

$$D_u f(2,0) = (1,2) \cdot u \leq \|(1,2)\|\,\|u\| = \sqrt{5},$$

with equality exactly when u is a positive multiple of $(1,2)$, that is $u = (1,2)/\sqrt{5}$. For this u,

$$D_u f(2,0) = (1,2) \cdot \frac{(1,2)}{\sqrt{5}} = \frac{1+4}{\sqrt{5}} = \sqrt{5}.$$

$$\boxed{u = \frac{(1,2)}{\sqrt{5}}, \qquad \max D_u f(2,0) = \sqrt{5}}$$

Method 2: Angle parametrisation of the unit direction. Every unit vector is $u = (\cos\theta, \sin\theta)$ with $\theta \in [0, 2\pi]$. With $\nabla f(2,0) = (1,2)$ from the partials above,

$$\phi(\theta) = D_u f(2,0) = \cos\theta + 2\sin\theta.$$

ϕ is continuous on the closed interval $[0, 2\pi]$, so by the EVT it attains a maximum, at an endpoint or where

$$\phi'(\theta) = -\sin\theta + 2\cos\theta = 0 \iff \sin\theta = 2\cos\theta.$$

If $\cos\theta = 0$ this forces $\sin\theta = 0$, impossible; so $\tan\theta = 2$, $\cos^2\theta = 1/(1+\tan^2\theta) = 1/5$, and $(\cos\theta, \sin\theta) = \pm(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$. The candidate values are

$$\phi = \frac{1}{\sqrt{5}} + \frac{4}{\sqrt{5}} = \sqrt{5} \ \ (\text{sign } +), \qquad \phi = -\sqrt{5} \ \ (\text{sign } -), \qquad \phi(0) = \phi(2\pi) = 1.$$

The largest is $\sqrt{5}$, at $u = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$:

$$\boxed{u = \frac{(1,2)}{\sqrt{5}}, \qquad \max D_u f(2,0) = \sqrt{5}}$$

The converse is false

Existence of directional derivatives in many, or even every, direction does not imply differentiability. The theorem proved here has one direction only: differentiability implies the gradient-dot-direction formula.

5.4 Proof details and diagnostic counterexamples

For the Jacobian theorem, the estimate hidden by matrix notation is important. If $A = J_G(a)$, then

$$G(a+h) - G(a) = Ah + r_G(h).$$

For a matrix norm compatible with the vector norm, there is K with

$$\|Ah\| \le K \|h\|.$$

Since $r_G(h) = o(\|h\|)$, the inner increment satisfies

$$\|Ah + r_G(h)\| = O(\|h\|).$$

The outer remainder therefore obeys

$$r_F(Ah + r_G(h)) = o(\|Ah + r_G(h)\|) = o(\|h\|).$$

Also $Br_G(h) = o(\|h\|)$. These two estimates justify the product $J_F(G(a))J_G(a)$.

Dependency diagram for a composite map

$$(s,t) \longmapsto (x(s,t), y(s,t)) \longmapsto f(x,y).$$

An arrow carries a derivative. A node tells us where to evaluate it:

$$f_x, f_y, f_{xx}, f_{xy}, f_{yy} \text{ are evaluated at } (x(s,t), y(s,t)),$$

while x_s, y_s, x_{st}, y_{st} are evaluated at (s,t).

The mixed formula can be checked without memory. Start with

$$z_s = f_x x_s + f_y y_s.$$

Product differentiation gives

$$z_{st} = (f_x)_t x_s + f_x x_{st} + (f_y)_t y_s + f_y y_{st}.$$

The chain rule for the two outer partials gives

$$(f_x)_t = f_{xx} x_t + f_{xy} y_t,$$

and

$$(f_y)_t = f_{yx} x_t + f_{yy} y_t.$$

Substitution, then $f_{yx} = f_{xy}$, produces exactly the theorem's formula. This order also shows why the terms $f_x x_{st}$ and $f_y y_{st}$ cannot be dropped when the inner variables have genuine mixed dependence.

Worked Example: evaluate a nontrivial mixed composite

Let

$$z = \exp(xy), \qquad x = s^2 + t, \qquad y = s + t^2.$$

Method 1: Second-order Chain Rule formula. The outer function $f(x, y) = e^{xy}$ is C^2 on $\mathbb{R}^2$, so Clairaut gives $f_{yx} = f_{xy}$ and Theorem 5.3 applies. Its partials are

$$f_x = ye^{xy}, \qquad f_y = xe^{xy}, \qquad f_{xx} = y^2e^{xy}, \qquad f_{yy} = x^2e^{xy},$$

and the product rule on $f_x = ye^{xy}$ gives

$$f_{xy} = \frac{\partial}{\partial y}(ye^{xy}) = e^{xy} + xye^{xy} = (1 + xy)e^{xy}.$$

The inner derivatives are

$$x_s = 2s, \qquad x_t = 1, \qquad x_{st} = 0, \qquad y_s = 1, \qquad y_t = 2t, \qquad y_{st} = 0.$$

At $(s, t) = (1, 0)$, $(x, y) = (1^2 + 0,\ 1 + 0^2) = (1, 1)$ and $xy = 1$, so

$$\begin{aligned} &x_s = 2, \qquad x_t = 1, \qquad y_s = 1, \qquad y_t = 0, \\ &f_x = e, \qquad f_y = e, \qquad f_{xx} = e, \qquad f_{yy} = e, \qquad f_{xy} = (1 + 1)e = 2e. \end{aligned}$$

Substitute into the mixed formula:

$$\begin{aligned} z_{st} &= f_{xx}x_sx_t + f_{xy}(x_sy_t + y_sx_t) + f_{yy}y_sy_t + f_xx_{st} + f_yy_{st} \\ &= e(2)(1) + 2e\big((2)(0) + (1)(1)\big) + e(1)(0) + e(0) + e(0) \\ &= 2e + 2e = 4e. \end{aligned}$$

$$\boxed{z_{st}(1, 0) = 4e}$$

Method 2: Direct substitution. Write $z = e^{g(s,t)}$ with

$$g(s, t) = (s^2 + t)(s + t^2) = s^3 + s^2t^2 + st + t^3.$$

Its partials are

$$g_s = 3s^2 + 2st^2 + t, \qquad g_t = 2s^2t + s + 3t^2, \qquad g_{st} = \frac{\partial}{\partial t}(3s^2 + 2st^2 + t) = 4st + 1.$$

Differentiating $z = e^g$ first in s, then the product g_se^g in t,

$$z_s = g_se^g, \qquad z_{st} = g_{st}e^g + g_s\, g_te^g = (g_{st} + g_sg_t)e^g.$$

At $(1, 0)$: $g = 1 + 0 + 0 + 0 = 1$, $g_s = 3 + 0 + 0 = 3$, $g_t = 0 + 1 + 0 = 1$, $g_{st} = 0 + 1 = 1$. Hence

$$z_{st}(1, 0) = (1 + 3 \cdot 1)e^1 = 4e.$$

$$\boxed{z_{st}(1, 0) = 4e}$$

5.5 A scalar implicit-function proof in detail

Here is the local argument behind the scalar IFT. Continuity of F_z and $F_z(a, b, c) \neq 0$ permit a closed box about (a, b, c) on which either

$$F_z \geq m > 0$$

or $F_z \le -m < 0$. Reverse the sign of F if necessary, so assume the first case. Choose a small vertical half-width ρ. The mean value theorem on the central vertical line gives

$$F(a,b,c+\rho) > 0, \qquad F(a,b,c-\rho) < 0.$$

Continuity of F preserves these strict endpoint signs for all (x,y) in a sufficiently small horizontal neighbourhood U of (a,b).

For each $(x,y) \in U$, the function $z \mapsto F(x,y,z)$ is continuous on $[c-\rho, c+\rho]$ and has opposite endpoint signs. The intermediate value theorem supplies $g(x,y)$ in that interval with

$$F(x,y,g(x,y)) = 0.$$

If two roots z_1, z_2 existed, the one-variable MVT would yield

$$0 = F(x,y,z_1) - F(x,y,z_2) = F_z(x,y,\zeta)(z_1 - z_2).$$

The lower bound $F_z \ge m$ forces $z_1 = z_2$. Thus the root is unique.

To prove continuity, compare roots at (x,y) and (x',y'). Insert and subtract $F(x',y',g(x,y))$ to get

$$m|g(x,y) - g(x',y')| \le |F(x',y',g(x,y)) - F(x,y,g(x,y))|.$$

Uniform continuity of F on the small closed box makes the right side tend to zero as $(x',y') \to (x,y)$. Thus g is locally continuous, indeed locally Lipschitz when the horizontal partials are bounded.

Finally put $\Delta x = x - a$, $\Delta y = y - b$, and $\Delta z = g(x,y) - c$. The differentiability expansion of F at (a,b,c) is

$$0 = F_x \Delta x + F_y \Delta y + F_z \Delta z + o\left(\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}\right).$$

The local Lipschitz bound makes $\Delta z = O(\sqrt{\Delta x^2 + \Delta y^2})$. Divide by the horizontal norm and solve for Δz; this gives

$$g(x,y) - c = -\frac{F_x}{F_z}\Delta x - \frac{F_y}{F_z}\Delta y + o(\sqrt{\Delta x^2 + \Delta y^2}).$$

The same argument applies at every root in a smaller open box. Hence g is differentiable there with the asserted partial derivatives. Their continuity follows because the derivatives of F are continuous and F_z stays away from zero. This completes the scalar C^1 conclusion.

A singular level set need not have one tangent

For $F(x,y) = y^2 - x^3$, the level set $F = 0$ has a cusp at $(0,0)$ and

$$\nabla F(0,0) = (0,0).$$

The IFT hypothesis fails. Dividing by $F_y = 2y$ would be illegal at the cusp; there is no regular implicit graph through the point. Always check the nonzero derivative before quoting an implicit slope or normal vector.

Worked Example: normalize a velocity direction

For $f(x,y) = x^2 + y^2$ at $(1,2)$, the velocity $v = (3,4)$ is not a unit vector.
Method 1: Gradient formula and Chain Rule. Normalise v:

$$\|v\| = \sqrt{3^2+4^2} = \sqrt{25} = 5, \qquad u = \frac{v}{\|v\|} = \left(\frac{3}{5}, \frac{4}{5}\right).$$

The partials $f_x = 2x$, $f_y = 2y$ are continuous, so f is differentiable and Theorem 5.5 applies with $\nabla f(1,2) = (2\cdot 1,\ 2\cdot 2) = (2,4)$:

$$D_u f(1,2) = (2,4)\cdot\left(\frac{3}{5}, \frac{4}{5}\right) = \frac{6}{5} + \frac{16}{5} = \frac{22}{5}.$$

For the velocity rate, move along $r(t) = (1,2) + tv = (1+3t,\ 2+4t)$, with $r(0) = (1,2)$ and $r'(0) = (3,4)$. The Chain Rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t} f(r(t))\Big|_{t=0} = \nabla f(1,2)\cdot r'(0) = (2,4)\cdot(3,4) = 6 + 16 = 22.$$

$$\boxed{\text{unit directional rate } = 22/5, \quad \text{velocity rate } = 22}$$

Method 2: Limit definition and direct substitution. Along the unit direction, $(1,2) + hu = \left(1 + \frac{3h}{5},\ 2 + \frac{4h}{5}\right)$, so

$$\begin{aligned} f\big((1,2)+hu\big) &= \left(1+\frac{3h}{5}\right)^2 + \left(2+\frac{4h}{5}\right)^2 = 1 + \frac{6h}{5} + \frac{9h^2}{25} + 4 + \frac{16h}{5} + \frac{16h^2}{25} \\ &= 5 + \frac{22h}{5} + h^2. \end{aligned}$$

With $f(1,2) = 5$, Definition 5.1 gives

$$D_u f(1,2) = \lim_{h\to 0} \frac{\frac{22h}{5} + h^2}{h} = \lim_{h\to 0}\left(\frac{22}{5} + h\right) = \frac{22}{5}.$$

Along the velocity path, let $\varphi(t) = f(1+3t, 2+4t)$:

$$\varphi(t) = (1+3t)^2 + (2+4t)^2 = 1 + 6t + 9t^2 + 4 + 16t + 16t^2 = 5 + 22t + 25t^2,$$

so $\varphi'(t) = 22 + 50t$ and $\varphi'(0) = 22$.

$$\boxed{\text{unit directional rate } = 22/5, \quad \text{velocity rate } = 22}$$

Worked Example: directional derivatives without differentiability

Define

$$h(x,y)=\begin{cases}\dfrac{x^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Method 1: Limit definition, then the gradient formula. Let $u=(u_1,u_2)$ be a unit direction, so $u_1^2+u_2^2=1$. For $t\neq 0$, $tu\neq(0,0)$ and

$$h(tu)=\frac{t^3u_1^3}{t^2u_1^2+t^2u_2^2}=\frac{t^3u_1^3}{t^2}=tu_1^3.$$

Definition 5.1 gives

$$D_uh(0,0)=\lim_{t\to0}\frac{h(tu)-h(0,0)}{t}=\lim_{t\to0}\frac{tu_1^3-0}{t}=u_1^3,$$

so every directional derivative exists. In particular $h_x(0,0)=D_{e_1}h(0,0)=1^3=1$ and $h_y(0,0)=D_{e_2}h(0,0)=0^3=0$. Suppose h were differentiable at $(0,0)$. Theorem 5.5 would give

$$D_uh(0,0)=\nabla h(0,0)\cdot u=(1,0)\cdot(u_1,u_2)=u_1.$$

For $u=\frac{1}{\sqrt2}(1,1)$ this would require $u_1^3=\frac{1}{2\sqrt2}$ to equal $u_1=\frac{1}{\sqrt2}$, a contradiction.

$$\boxed{h \text{ has every directional derivative at 0 but is not differentiable there}}$$

Method 2: Homogeneity and the differentiability remainder. For $\lambda\neq0$ and $(x,y)\neq(0,0)$,

$$h(\lambda x,\lambda y)=\frac{\lambda^3x^3}{\lambda^2(x^2+y^2)}=\lambda\,h(x,y),$$

and $h(0,0)=0$. For a unit u, $h(u)=u_1^3/(u_1^2+u_2^2)=u_1^3$, hence $h(tu)=t\,h(u)=tu_1^3$ and $D_uh(0,0)=\lim_{t\to0}tu_1^3/t=u_1^3$ exists. The partials are $h_x(0,0)=1$, $h_y(0,0)=0$, so the only candidate linear map is $L(x,y)=x$. Differentiability at $(0,0)$ means $R(x,y)=\dfrac{h(x,y)-h(0,0)-x}{\sqrt{x^2+y^2}}\to0$. For $(x,y)\neq(0,0)$,

$$R(x,y)=\frac{x^3-x(x^2+y^2)}{(x^2+y^2)^{3/2}}=-\frac{xy^2}{(x^2+y^2)^{3/2}}.$$

Along $y=0$, $R(x,0)=0\to0$. Along $y=x$ with $x>0$,

$$R(x,x)=-\frac{x^3}{(2x^2)^{3/2}}=-\frac{x^3}{2\sqrt2\,x^3}=-\frac{1}{2\sqrt2}\to-\frac{1}{2\sqrt2}\neq0.$$

The two path limits differ, so R has no limit at $(0,0)$ and h is not differentiable there.

$$\boxed{h \text{ has every directional derivative at 0 but is not differentiable there}}$$

When $\nabla f(p)=0$ and f is differentiable, the directional theorem says

$$D_uf(p)=0$$

for every unit u. This does not classify p: Chapter 3 supplies examples where a zero gradient is a saddle, a minimum, or a maximum.

The unit-vector condition makes a directional derivative a rate per unit distance. A path velocity has direction and speed, so its derivative is a rate per unit parameter; normalize a nonzero velocity only when a direction rather than a velocity rate is requested.

For a regular level curve, both ∇f and $-\nabla f$ are normal. The first points toward local increase when the gradient is nonzero.

Exam Focus: Chapter 5 workflow

- Draw the dependency arrows before differentiating a composite map.
- Evaluate outer partials at the transformed point.
- For implicit differentiation, verify the nonzero derivative required by IFT.
- Normalize a direction before using a directional derivative.
- Use Cauchy–Schwarz for fastest ascent, and state the zero-gradient case.

5.6 Chapter Summary

Key Formula: Chapter 5 recap

$$J_{F\circ G} = J_F(G)J_G, \qquad z_x = -\frac{F_x}{F_z}, \quad z_y = -\frac{F_y}{F_z}, \qquad D_u f = \nabla f \cdot u.$$

These identities connect maps, local graphs, and directional rates.

Exam Focus: Chapter 5 takeaways

Use correct points and hypotheses. IFT is local, not global; a nonzero gradient is normal to a level curve and gives fastest change.

QRS@NTU

6 Topic 6: Lagrange Multipliers

Lagrange multipliers turn a constrained extremum problem into equations. The equations are necessary conditions at a regular constrained extremum; they are not an existence theorem and they are not, by themselves, a classification test. This distinction controls every calculation in this chapter.

6.1 Constraint sets and their geometry

Let f be the objective function and let g describe a constraint. In two variables, the set $g(x, y) = k$ is usually a curve. In three variables it is usually a surface. At a regular point, the gradient of g is normal to that curve or surface, by Chapter 5.

Definition 6.1: Equality-constrained optimization

Let $U \subseteq \mathbb{R}^n$ be open, and let $f, g : U \to \mathbb{R}$. For $k \in \mathbb{R}$, set

$$C = \{x \in U : g(x) = k\}.$$

The set C is the *constraint set*. An absolute maximum of f on C is a point $p \in C$ for which $f(p) \geq f(x)$ for every $x \in C$; an absolute minimum reverses the inequality. A point is *regular for the constraint* when $\nabla g(p) \neq 0$.

The word "on" matters. We compare only feasible points, namely points of C. An unconstrained critical point of f might not even lie on C. Conversely, a constrained extremum can have $\nabla f(p) \neq 0$, because moving in the direction of increase may violate the constraint.

Near a regular point, the implicit-function theorem makes C into a graph. For example, if $g_y(p) \neq 0$ in two variables, then $g(x, y) = k$ can locally be written $y = \phi(x)$. A feasible change is then a change along that graph, rather than an arbitrary change in the plane.

Theorem 6.1: Regular single-constraint multiplier theorem

Let $f, g \in C^1(U)$, let $p \in C = \{x \in U : g(x) = k\}$, and suppose that p is a local maximum or local minimum of f relative to C. If $\nabla g(p) \neq 0$, then there is a real number λ such that

$$\nabla f(p) = \lambda \nabla g(p).$$

This covers the two-variable and three-variable single-constraint theorems in the Week 5 slides.

Proof. Write $p = (p_1, \ldots, p_n)$. Because $\nabla g(p) \neq 0$, after relabelling coordinates if necessary, $g_n(p) \neq 0$. The scalar implicit-function theorem proved in Chapter 5 gives a neighbourhood of $p' = (p_1, \ldots, p_{n-1})$ and a C^1 function ϕ such that, near p,

$$g(x_1, \ldots, x_{n-1}, x_n) = k \quad \Longleftrightarrow \quad x_n = \phi(x_1, \ldots, x_{n-1}).$$

Define the restricted objective

$$F(x') = f(x', \phi(x')).$$

Since p is a relative local extremum on C, p' is an unconstrained local extremum of F. The one-variable partial-derivative necessary condition, applied in each coordinate direction, gives $F_i(p') = 0$ for $i = 1, \ldots, n-1$.

Differentiating the identity $g(x', \phi(x')) = k$ with respect to x_i gives

$$g_i(p) + g_n(p)\phi_i(p') = 0, \qquad \phi_i(p') = -\frac{g_i(p)}{g_n(p)}.$$

The chain rule for F now gives

$$\begin{aligned} 0 = F_i(p') &= f_i(p) + f_n(p)\phi_i(p') \\ &= f_i(p) - f_n(p)\frac{g_i(p)}{g_n(p)}. \end{aligned}$$

Thus, for $i < n$,

$$f_i(p) = \frac{f_n(p)}{g_n(p)} g_i(p).$$

Put $\lambda = f_n(p)/g_n(p)$. The displayed equation supplies the first $n-1$ components of $\nabla f(p) = \lambda \nabla g(p)$, and its last component holds from the definition of λ. This proves the result. □

The proof explains the geometry. At a constrained extremum the derivative of f is zero in every tangent direction to the constraint. Both gradients are normal to that tangent space, so they are parallel for one constraint.

Single-constraint equations

For a regular equality constraint $g(x) = k$, solve

$$\boxed{g(x) = k, \qquad \nabla f(x) = \lambda \nabla g(x).}$$

In coordinates this is one constraint equation plus one equation for each component of the gradient. The multiplier λ is an auxiliary unknown; discard it only after solving the full system.

Worked Example (Week 5, pp. 9–11): quadratic on the unit circle

Find the absolute extrema of

$$f(x, y) = x^2 + 2y^2$$

on $C = \{(x, y) : x^2 + y^2 = 1\}$.

Method 1: Lagrange multipliers with EVT comparison. Take $g(x, y) = x^2 + y^2$, so $C = \{g = 1\}$. The partial derivatives are

$$f_x = 2x, \qquad f_y = 4y, \qquad g_x = 2x, \qquad g_y = 2y.$$

Existence: g is continuous, so $C = g^{-1}(\{1\})$ is closed; every point of C has $|x| \le 1$ and $|y| \le 1$, so C is bounded. Hence C is compact, f is a polynomial and so continuous, and the Extreme Value Theorem gives an absolute maximum and an absolute minimum of f on C. Regularity: $\nabla g = (2x, 2y) = (0, 0)$ only at $(0, 0)$, and $0^2 + 0^2 = 0 \neq 1$, so

$$\nabla g = (2x, 2y) \neq (0, 0) \text{ on } C \quad \text{(regular)}.$$

By Theorem 6.2, both extrema are among the solutions of the Lagrange condition $\nabla f = \lambda \nabla g$ with $g = 1$:

$$\begin{aligned} 2x = 2\lambda x &\iff 2x(1-\lambda) = 0 \iff x = 0 \text{ or } \lambda = 1 \\ 4y = 2\lambda y &\iff 2y(2-\lambda) = 0 \iff y = 0 \text{ or } \lambda = 2 \\ x^2 + y^2 &= 1 \end{aligned}$$

Case $x = 0$. The constraint gives $y^2 = 1$, so $y = \pm 1$. The second equation then reads $\pm 4 = \pm 2\lambda$, so $\lambda = 2$. Candidates: $(0, 1)$, $(0, -1)$.
Case $x \neq 0$. The first equation forces $\lambda = 1$, and the second becomes $4y = 2y$, so $y = 0$. The constraint gives $x^2 = 1$, so $x = \pm 1$. Candidates: $(1, 0)$, $(-1, 0)$.
The two cases exhaust the first equation ($x, y \neq 0$ would need $\lambda = 1 = 2$). Evaluate f at every candidate:

$$\begin{aligned} f(0, \pm 1) &= 0^2 + 2(\pm 1)^2 = 2, \\ f(\pm 1, 0) &= (\pm 1)^2 + 2 \cdot 0^2 = 1. \end{aligned}$$

Comparing the four values gives

$$\boxed{\max_C f = 2 \text{ at } (0, \pm 1), \qquad \min_C f = 1 \text{ at } (\pm 1, 0).}$$

Method 2: Eliminating x^2 with the constraint. On C, $x^2 = 1 - y^2$, and a real x exists exactly when $1 - y^2 \geq 0$, that is, $-1 \leq y \leq 1$. Each such y gives the feasible points $(\pm\sqrt{1-y^2}, y)$. Substituting,

$$f(x, y) = (1 - y^2) + 2y^2 = 1 + y^2, \qquad -1 \leq y \leq 1.$$

Since $0 \leq y^2 \leq 1$ on $[-1, 1]$,

$$1 \leq f(x, y) \leq 2 \quad \text{on } C.$$

Equality $f = 1$ holds iff $y = 0$, and then $x^2 = 1$, so $(x, y) = (\pm 1, 0)$. Equality $f = 2$ holds iff $y^2 = 1$, and then $x^2 = 0$, so $(x, y) = (0, \pm 1)$. Therefore

$$\boxed{\max_C f = 2 \text{ at } (0, \pm 1), \qquad \min_C f = 1 \text{ at } (\pm 1, 0).}$$

Method 3: Parametrising the circle. The map $t \mapsto (\cos t, \sin t)$, $t \in [0, 2\pi]$, traces all of C. Let

$$F(t) = f(\cos t, \sin t) = \cos^2 t + 2\sin^2 t = (\cos^2 t + \sin^2 t) + \sin^2 t = 1 + \sin^2 t.$$

F is continuous on the closed interval $[0, 2\pi]$, so its extrema occur at critical points or endpoints. Differentiate:

$$F'(t) = 2\sin t \cos t = \sin 2t = 0 \iff t \in \{0, \tfrac{\pi}{2}, \pi, \tfrac{3\pi}{2}, 2\pi\}.$$

Evaluate F at each of these points:

$$F(0) = F(\pi) = F(2\pi) = 1 + 0 = 1, \qquad (\cos t, \sin t) = (\pm 1, 0),$$
$$F(\tfrac{\pi}{2}) = F(\tfrac{3\pi}{2}) = 1 + 1 = 2, \qquad (\cos t, \sin t) = (0, \pm 1).$$

Hence

$$\boxed{\max_C f = 2 \text{ at } (0, \pm 1), \qquad \min_C f = 1 \text{ at } (\pm 1, 0).}$$

6.2 Candidates, compactness, and singular points

The preceding theorem tells us where a regular local extremum must occur. It does not say that every solution of the multiplier equations is an extremum, nor that an extremum exists on a noncompact feasible set.

Theorem 6.2: Candidate comparison on a compact regular constraint

Let $C = \{x \in U : g(x) = k\}$ be nonempty and compact, let $f, g \in C^1(U)$, and suppose $\nabla g(x) \neq 0$ for every $x \in C$. Then f has an absolute maximum and an absolute minimum on C. Each of them occurs among the solutions of

$$g(x) = k, \qquad \nabla f(x) = \lambda \nabla g(x).$$

Consequently, compare f at every feasible solution of this system.

Proof. The Extreme Value Theorem applies because f is continuous and C is compact, so an absolute maximum $p_{\max}$ and an absolute minimum $p_{\min}$ exist. Each is also a local extremum relative to C. The regularity hypothesis lets us apply Theorem 6.1 to each point, so both are solutions of the multiplier system. Thus the largest and smallest objective values among all feasible multiplier solutions are the

global extrema. The theorem does not claim that every multiplier solution is a maximum or minimum; comparison is still required. □

For a closed and bounded constraint in $\mathbb{R}^2$ or $\mathbb{R}^3$, compactness follows from Heine–Borel. This is the Week 5 "bounded and closed" case. If the constraint is not closed or not bounded, a multiplier calculation still gives regular candidates, but an additional argument is needed to prove a global claim.

Definition 6.2: Singular feasible point

A feasible point $p \in C$ is *singular* when $\nabla g(p) = 0$. The regular multiplier theorem makes no assertion at such a point.

Worked Example: a singular point cannot be discarded

Minimize and maximize $f(x, y) = x$ subject to

$$g(x, y) = x^2 + y^2 = 0.$$

Method 1: Lagrange system plus singular points. The partial derivatives are

$$\nabla f = (f_x, f_y) = (1, 0), \qquad \nabla g = (g_x, g_y) = (2x, 2y).$$

The Lagrange condition $\nabla f = \lambda \nabla g$ together with $g = 0$ reads

$$1 = 2\lambda x, \qquad 0 = 2\lambda y, \qquad x^2 + y^2 = 0.$$

Since $x^2 \geq 0$ and $y^2 \geq 0$, the third equation forces $x = y = 0$. The first equation then reads $1 = 2\lambda \cdot 0 = 0$, which is false for every λ. Hence the system has no solution: there is no regular candidate. Singular feasible points: $\nabla g = (2x, 2y) = (0, 0)$ iff $(x, y) = (0, 0)$, and $0^2 + 0^2 = 0$, so $(0, 0)$ is feasible, with

$$\nabla g(0, 0) = (0, 0), \qquad \nabla f(0, 0) = (1, 0) \neq \lambda \nabla g(0, 0) \ \ \forall \lambda$$

so the only feasible point is singular and the Lagrange condition has no solution. Existence: the solution of the third equation shows $C = \{(0, 0)\}$, a nonempty finite (hence compact) set, and f is continuous, so the Extreme Value Theorem gives a maximum and a minimum. Each is a regular solution of the system (there are none) or a singular feasible point, so both occur at $(0, 0)$, where

$$f(0, 0) = 0.$$

Therefore

$$\boxed{\min_C f = \max_C f = 0 \text{ at } (0, 0).}$$

Method 2: Solving the constraint directly. For real x, y, both $x^2 \geq 0$ and $y^2 \geq 0$, so

$$x^2 + y^2 = 0 \iff x^2 = 0 \text{ and } y^2 = 0 \iff (x, y) = (0, 0).$$

Thus $C = \{(0, 0)\}$. On a one-point set the only value of f is $f(0, 0) = 0$, which is both the largest and the smallest value:

$$\boxed{\min_C f = \max_C f = 0 \text{ at } (0, 0).}$$

Lagrange equations are necessary, not sufficient

Do not write "therefore maximum" after solving $\nabla f = \lambda \nabla g$. First verify the point is feasible. Then decide whether the feasible set is compact. On a compact regular constraint, compare all candidate values. On a noncompact or nonclosed constraint, prove the required bound or analyse limits. Also include every singular feasible point, where the theorem does not apply.

6.3 A three-variable constraint and a noncompact feasible set

The same equations work in three variables. The lidless-box example in the slides also shows why compactness cannot be silently assumed.

Worked Example (Week 5, pp. 14–17): lidless box of area 12

Let $x, y, z > 0$ be the length, width, and height. A lidless box has

$$V(x,y,z) = xyz, \qquad A(x,y,z) = xy + 2xz + 2yz.$$

Maximize V subject to $A = 12$.
Method 1: Lagrange multipliers with an AM–GM global bound. The partial derivatives are

$$\nabla V = (V_x, V_y, V_z) = (yz, xz, xy), \qquad \nabla A = (A_x, A_y, A_z) = (y + 2z,\ x + 2z,\ 2x + 2y).$$

On the open set $\{x, y, z > 0\}$ all entries of ∇A are positive, so $\nabla A \neq 0$ and the constraint is regular. Lagrange condition $\nabla V = \lambda \nabla A$:

$$\begin{aligned} yz &= \lambda(y + 2z) && \text{(i)} \\ xz &= \lambda(x + 2z) && \text{(ii)} \\ xy &= 2\lambda(x + y) && \text{(iii)} \\ xy + 2xz + 2yz &= 12 && \text{(iv)} \end{aligned}$$

If $\lambda = 0$, then (i) gives $yz = 0$, impossible for $y, z > 0$; so $\lambda \neq 0$. Multiply (i), (ii), (iii) by x, y, z respectively; each left side becomes xyz:

$$xyz = \lambda x(y + 2z) = \lambda y(x + 2z) = 2\lambda z(x + y).$$

Divide by $\lambda \neq 0$ and compare in pairs:

$$\begin{aligned} x(y + 2z) = y(x + 2z) &\iff xy + 2xz = xy + 2yz \iff 2z(x - y) = 0 \iff x = y && (z > 0) \\ x(x + 2z) = 2z(x + x) &\iff x^2 + 2xz = 4xz \iff x(x - 2z) = 0 \iff x = 2z && (x > 0) \end{aligned}$$

Substitute $y = x$, $z = x/2$ into (iv):

$$x^2 + 2x\left(\tfrac{x}{2}\right) + 2x\left(\tfrac{x}{2}\right) = 3x^2 = 12 \implies x = 2 \quad (x > 0).$$

So $(x, y, z) = (2, 2, 1)$. From (i), $2 \cdot 1 = \lambda(2 + 2)$ gives $\lambda = \frac{1}{2}$, and then (ii) reads $2 = \frac{1}{2} \cdot 4$ and (iii) reads $4 = 2 \cdot \frac{1}{2} \cdot 4$, so all four equations hold. The only candidate has

$$V(2, 2, 1) = 2 \cdot 2 \cdot 1 = 4.$$

The feasible set $(x, y, z > 0)$ is neither closed nor bounded, so compactness is unavailable and a single candidate is not yet a maximum. Global bound by AM–GM on the three positive terms of A:

$$\begin{aligned} \sqrt[3]{(xy)(2xz)(2yz)} \le \frac{xy + 2xz + 2yz}{3} = \frac{12}{3} = 4 && \text{(AM–GM)} \\ \implies (xy)(2xz)(2yz) = 4x^2y^2z^2 \le 4^3 = 64 \\ \implies (xyz)^2 \le 16 \implies xyz \le 4 && (xyz > 0) \end{aligned}$$

Equality holds iff $xy = 2xz = 2yz$; $xy = 2xz$ gives $y = 2z$ and $2xz = 2yz$ gives $x = y$, so $x = y = 2z$, and $A = 12$ gives $(2, 2, 1)$ as above.
Minimum: for $0 < \varepsilon < 1$ take $z = \varepsilon$ and $x = y = -2\varepsilon + \sqrt{4\varepsilon^2 + 12}$, the positive root of $x^2 + 4\varepsilon x - 12 = 0$; it is positive because $\sqrt{4\varepsilon^2 + 12} > \sqrt{4\varepsilon^2} = 2\varepsilon$.

$$A = x^2 + 2x\varepsilon + 2x\varepsilon = x^2 + 4x\varepsilon = 12 \qquad \text{(feasible)}$$
$$V = x^2\varepsilon = (12 - 4x\varepsilon)\varepsilon \le 12\varepsilon \to 0 \quad (\varepsilon \downarrow 0)$$

Every feasible $V > 0$, so 0 is an unattained infimum (no minimum).

$$\boxed{\max V = 4, \text{ uniquely at dimensions } 2 \times 2 \times 1.}$$

Method 2: Eliminating z and maximizing in one variable. Solve the constraint for z:

$$xy + 2z(x + y) = 12 \iff z = \frac{12 - xy}{2(x + y)},$$

and $z > 0$ iff $xy < 12$. Conversely, every $x, y > 0$ with $xy < 12$ gives a feasible box. Hence

$$V = xyz = \frac{xy(12 - xy)}{2(x + y)}, \qquad x, y > 0,\ xy < 12.$$

Two-term AM–GM: $x + y - 2\sqrt{xy} = (\sqrt{x} - \sqrt{y})^2 \ge 0$, with equality iff $x = y$. Since $xy(12 - xy) > 0$, a larger denominator gives a smaller fraction:

$$V \le \frac{xy(12 - xy)}{4\sqrt{xy}} = \frac{\sqrt{xy}\,(12 - xy)}{4}.$$

Put $t = \sqrt{xy} \in (0, 2\sqrt{3})$ and $h(t) = \dfrac{12t - t^3}{4}$. Then

$$h'(t) = \frac{12 - 3t^2}{4} = \frac{3(2 - t)(2 + t)}{4},$$

so $h' > 0$ on $(0, 2)$ and $h' < 0$ on $(2, 2\sqrt{3})$: h increases and then decreases, and

$$V \le h(t) \le h(2) = \frac{24 - 8}{4} = 4.$$

Equality needs $x = y$ (AM–GM) and $t = 2$, i.e. $xy = 4$; so $x = y = 2$ and

$$z = \frac{12 - 4}{2(2 + 2)} = 1.$$

Therefore

$$\boxed{\max V = 4, \text{ uniquely at dimensions } 2 \times 2 \times 1.}$$

Exam Focus: one equality constraint

For $g = k$, work in this order:

1. State the feasible domain and check $\nabla g \neq 0$ at its regular points.
2. Solve $g = k$ and $\nabla f = \lambda \nabla g$ simultaneously.
3. Add feasible singular points and any endpoints or boundary pieces caused by the domain restrictions.
4. Compare values if compact; otherwise justify a maximum/minimum separately.

6.4 Two constraints in three variables

With two independent equality constraints, their simultaneous solution is usually a curve in $\mathbb{R}^3$. At a constrained extremum, ∇f is normal to that curve. Both constraint gradients are normal to the curve too, so ∇f belongs to their span.

Theorem 6.3: Regular two-constraint multiplier theorem

Let $f, g, h \in C^1(U)$ for an open $U \subseteq \mathbb{R}^n$, and set

$$C = \{x \in U : g(x) = k,\ h(x) = \ell\}.$$

If $p \in C$ is a relative local extremum of f and $\nabla g(p)$ and $\nabla h(p)$ are linearly independent, then real multipliers λ, μ exist such that

$$\nabla f(p) = \lambda \nabla g(p) + \mu \nabla h(p).$$

For the Week 5 three-variable case, this is the stated formula. Linear independence is the regularity condition needed for the conclusion.

Proof. Independence of the two gradient rows means that some 2×2 minor of their Jacobian is invertible. Relabel coordinates so the minor in the final two coordinates is invertible, and interchange these coordinates or the constraints so that $g_n(p) \neq 0$. The scalar IFT first solves $g = k$ as $x_n = \phi(s, x_{n-1})$. Set $H(s, x_{n-1}) = h(s, x_{n-1}, \phi(s, x_{n-1})) - \ell$. Its derivative in the final free coordinate is

$$H_{n-1} = h_{n-1} - h_n \frac{g_{n-1}}{g_n} = \frac{h_{n-1} g_n - h_n g_{n-1}}{g_n} \neq 0$$

at p, by the invertible minor. A second application of the scalar IFT solves $H = 0$ for $x_{n-1} = u(s)$; then $v(s) = \phi(s, u(s))$ is C^1 by the chain rule. Thus the two constraints are locally equivalent to

$$x_{n-1} = u(s), \qquad x_n = v(s), \qquad s = (x_1, \ldots, x_{n-2}).$$

At the parameter s_0 corresponding to p, the restricted function

$$F(s) = f(s, u(s), v(s))$$

has an unconstrained local extremum. Hence, for every $i \leq n - 2$,

$$0 = F_i = f_i + f_{n-1} u_i + f_n v_i. \tag{1}$$

Differentiating both constraint identities gives

$$0 = g_i + g_{n-1} u_i + g_n v_i, \tag{2}$$
$$0 = h_i + h_{n-1} u_i + h_n v_i. \tag{3}$$

All derivatives in (1)–(3) are evaluated at the appropriate point p.

The invertible minor means the two equations

$$f_{n-1} = \lambda g_{n-1} + \mu h_{n-1}, \qquad f_n = \lambda g_n + \mu h_n$$

have a unique solution (λ, μ). Substitute these identities into (1), then use (2) and (3):

$$\begin{aligned} 0 &= f_i + \lambda(g_{n-1}u_i + g_n v_i) + \mu(h_{n-1}u_i + h_n v_i) \\ &= f_i - \lambda g_i - \mu h_i. \end{aligned}$$

Thus the first $n-2$ gradient components also obey $f_i = \lambda g_i + \mu h_i$. Together with the last two defining equations, this is the required vector identity. □

Two-constraint equations in $\mathbb{R}^3$

For $g(x,y,z) = k$ and $h(x,y,z) = \ell$, solve

$$\boxed{g = k, \qquad h = \ell, \qquad \nabla f = \lambda \nabla g + \mu \nabla h.}$$

There are five scalar equations for x, y, z, λ, μ. Check that ∇g and ∇h are independent at each candidate before treating the system as a complete regular-candidate list.

6.5 The Lagrangian and safe algebra

It is convenient to package the equations into one function. This notation does not replace the constraint equation: differentiating with respect to a multiplier is precisely what restores that equation.

Definition 6.3: Lagrangian

For the problem of optimizing $f(x)$ subject to $g(x) = k$, define

$$\mathcal{L}(x, \lambda) = f(x) - \lambda\big(g(x) - k\big).$$

For two constraints $g(x) = k$ and $h(x) = \ell$, define

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \lambda\big(g(x) - k\big) - \mu\big(h(x) - \ell\big).$$

The sign convention is optional: changing the signs merely replaces a multiplier by its negative.

Theorem 6.4: Stationarity equations are the multiplier equations

A stationary point of the single-constraint Lagrangian satisfies exactly

$$g(x) = k, \qquad \nabla f(x) = \lambda \nabla g(x).$$

A stationary point of the two-constraint Lagrangian satisfies exactly

$$g(x) = k, \qquad h(x) = \ell, \qquad \nabla f(x) = \lambda \nabla g(x) + \mu \nabla h(x).$$

Proof. For one constraint, differentiate $\mathcal{L}$ with respect to every coordinate of x and with respect to λ:

$$\frac{\partial \mathcal{L}}{\partial x_i} = f_i - \lambda g_i, \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = -(g - k).$$

Setting these $n+1$ derivatives equal to zero gives the displayed constraint and all components of the gradient equation. For two constraints, the same calculation adds

$$\frac{\partial \mathcal{L}}{\partial \mu} = -(h - \ell), \qquad \frac{\partial \mathcal{L}}{\partial x_i} = f_i - \lambda g_i - \mu h_i.$$

Their simultaneous vanishing is exactly the stated system. □

This shorthand is useful when a question asks for a Lagrangian, but the underlying logic remains Theorems 6.1 and 6.3: regular constrained extrema must be stationary in this augmented-variable sense.

Exam Focus: algebra in a multiplier system

Before dividing an equation by x, y, z, or λ, record the case where that quantity is zero. A zero coordinate is often a valid candidate, as on the unit circle. If the stated physical domain makes all variables positive, say so before the division, as in the lidless-box solution. After eliminating the multipliers, substitute candidates back into every original constraint; transformed equations can otherwise create extraneous points.

Worked Example (Week 5, p. 19): squared distance on cone and plane

Optimize

$$f(x,y,z) = x^2 + y^2 + z^2$$

subject to

$$g(x,y,z) = z^2 - x^2 - y^2 = 0, \qquad h(x,y,z) = x + y - z + 1 = 0.$$

Method 1: Two multipliers with a global bound in $s = x + y$. The partial derivatives are

$$\nabla f = (2x, 2y, 2z), \qquad \nabla g = (-2x, -2y, 2z), \qquad \nabla h = (1, 1, -1).$$

Regularity: $\nabla h \neq 0$, so the two gradients are dependent only if $\nabla g = t\nabla h$, i.e. $(-2x, -2y, 2z) = (t, t, -t)$, which gives $x = y = z = -t/2$. The cone then reads $z^2 = 2z^2$, so $z = 0$ and $x = y = z = 0$, but $h(0,0,0) = 1 \neq 0$ contradicts the plane. So $\nabla g, \nabla h$ are independent on the feasible set, and Theorem 6.3 applies. Lagrange condition $\nabla f = \lambda\nabla g + \mu\nabla h$:

$$\begin{aligned} 2x &= -2\lambda x + \mu && && \text{(i)} \\ 2y &= -2\lambda y + \mu && && \text{(ii)} \\ 2z &= 2\lambda z - \mu && && \text{(iii)} \\ z^2 &= x^2 + y^2, && x + y - z + 1 = 0 && \text{(iv), (v)} \end{aligned}$$

Subtract (ii) from (i):

$$2(x-y) = -2\lambda(x-y) \iff 2(1+\lambda)(x-y) = 0 \iff \lambda = -1 \text{ or } x = y.$$

Case $x \neq y$, so $\lambda = -1$. Then (i) reads $2x = 2x + \mu$, so $\mu = 0$, and (iii) reads $2z = -2z$, so $z = 0$. The cone (iv) gives $x^2 + y^2 = 0$, so $x = y = 0$, contradicting $x \neq y$. No candidate.
Case $x = y =: a$.

$$\begin{aligned} z &= 2a + 1 && \text{(plane)} \\ (2a+1)^2 &= 2a^2 && \text{(cone)} \end{aligned}$$

$$\iff 4a^2 + 4a + 1 = 2a^2 \iff 2a^2 + 4a + 1 = 0 \iff a = \frac{-4 \pm \sqrt{8}}{4} = -1 \pm \tfrac{\sqrt{2}}{2}$$

Multipliers exist for both roots: adding (i) and (iii) with $x = a$, $z = 2a+1$ gives $2(a+z) = 2\lambda(z-a)$, so $\lambda = \frac{3a+1}{a+1}$, where $a + 1 = \pm\frac{\sqrt{2}}{2} \neq 0$; then (i) gives $\mu = 2a(1+\lambda)$, and (ii) is the same equation as (i). On the cone $z^2 = x^2 + y^2 = 2a^2$, so $f = 2a^2 + z^2 = 4a^2$, and $a^2 = 1 \mp \sqrt{2} + \frac{1}{2} = \frac{3}{2} \mp \sqrt{2}$:

$$\begin{aligned} p_+ &= \left(-1 + \tfrac{\sqrt{2}}{2}, -1 + \tfrac{\sqrt{2}}{2}, \sqrt{2} - 1\right), && f(p_+) = 4\left(\tfrac{3}{2} - \sqrt{2}\right) = 6 - 4\sqrt{2} \\ p_- &= \left(-1 - \tfrac{\sqrt{2}}{2}, -1 - \tfrac{\sqrt{2}}{2}, -1 - \sqrt{2}\right), && f(p_-) = 4\left(\tfrac{3}{2} + \sqrt{2}\right) = 6 + 4\sqrt{2} \end{aligned}$$

Global: the feasible set is unbounded (not compact), so let $s = x + y$. Plane: $z = s + 1$; cone: $x^2 + y^2 = (s+1)^2$. Since $x^2 + y^2 - \frac{s^2}{2} = \frac{(x-y)^2}{2} \geq 0$, real x, y with sum s need

$$\begin{aligned} & (s+1)^2 \geq \frac{s^2}{2} && (x^2 + y^2 \geq \tfrac{s^2}{2}) \\ \iff\ & s^2 + 4s + 2 \geq 0 \\ \iff\ & s \leq -2 - \sqrt{2} \text{ or } s \geq -2 + \sqrt{2} \end{aligned}$$

Conversely, for such s the numbers x, y are the roots of $u^2 - su + \frac{1}{2}\big(s^2 - (s+1)^2\big) = 0$, whose discriminant $2(s+1)^2 - s^2$ is ≥ 0; so every such s occurs. On the cone, $f = 2z^2 = 2(s+1)^2$:

$$\begin{aligned} s \geq -2 + \sqrt{2} &\implies s + 1 \geq \sqrt{2} - 1 > 0 \implies f \geq 2(\sqrt{2} - 1)^2 = 6 - 4\sqrt{2}, \\ s \leq -2 - \sqrt{2} &\implies s + 1 \leq -1 - \sqrt{2} < 0 \implies f \geq 2(1 + \sqrt{2})^2 = 6 + 4\sqrt{2}. \end{aligned}$$

Equality in the first line needs $s = -2 + \sqrt{2}$ and $x = y$, i.e. the point p_+. On $s \leq -2 - \sqrt{2}$, $f \to \infty$ as $s \to -\infty$, so there is no maximum.

$$\boxed{\min f = 6 - 4\sqrt{2} \text{ at } p_+; \qquad f \text{ has no maximum.}}$$

Method 2: Parametrising the cone. A feasible point has $z \neq 0$: $z = 0$ forces $x = y = 0$ on the cone, and $h(0,0,0) = 1 \neq 0$. For $z \neq 0$, $(x/z)^2 + (y/z)^2 = 1$, so

$$x = z\cos t, \qquad y = z\sin t \qquad \text{for some } t \in [0, 2\pi).$$

The plane becomes

$$z\cos t + z\sin t - z + 1 = 0 \iff z\,D(t) = 1, \qquad D(t) = 1 - \cos t - \sin t,$$

so $D(t) \neq 0$ and $z = 1/D(t)$. On the cone $f = x^2 + y^2 + z^2 = 2z^2$, so

$$f = \frac{2}{D(t)^2}.$$

Since $\cos t + \sin t = \sqrt{2}\sin\left(t + \frac{\pi}{4}\right) \in [-\sqrt{2}, \sqrt{2}]$,

$$1 - \sqrt{2} \leq D(t) \leq 1 + \sqrt{2}.$$

If $D > 0$ then $|D| \leq 1 + \sqrt{2}$; if $D < 0$ then $|D| \leq \sqrt{2} - 1 < 1 + \sqrt{2}$. Hence

$$f = \frac{2}{D^2} \geq \frac{2}{(1+\sqrt{2})^2} = \frac{2}{3 + 2\sqrt{2}} = \frac{2(3 - 2\sqrt{2})}{9 - 8} = 6 - 4\sqrt{2},$$

with equality iff $D = 1 + \sqrt{2}$, iff $\sin(t + \frac{\pi}{4}) = -1$, iff $t = \frac{5\pi}{4}$. There

$$z = \frac{1}{1 + \sqrt{2}} = \sqrt{2} - 1, \qquad x = y = (\sqrt{2} - 1)\left(-\tfrac{\sqrt{2}}{2}\right) = -1 + \tfrac{\sqrt{2}}{2},$$

which is p_+. For the maximum, as $t \downarrow 0$ we have $D(t) \to 1 - 1 - 0 = 0$, with $D(t) \neq 0$ for small $t > 0$ because $\sin t > 1 - \cos t$ there; so $f = 2/D(t)^2 \to \infty$ along feasible points and f has no maximum. (The point $t = \frac{\pi}{4}$, where $D = 1 - \sqrt{2}$, gives $f = \frac{2}{(1-\sqrt{2})^2} = 6 + 4\sqrt{2}$ at p_-.)

$$\boxed{\min f = 6 - 4\sqrt{2} \text{ at } p_+; \qquad f \text{ has no maximum.}}$$

6.6 Domain boundaries and constraint boundaries

The multiplier theorem describes a relative local extremum only where the feasible set is locally a regular equality-constraint set inside an open domain. Inequalities such as $x \geq 0$ can create edges, endpoints, or corners. Points on those pieces require their own one-variable or endpoint analysis.

Worked Example: boundary edges on a quarter-circle

Find the extrema of $f(x, y) = x + y$ on

$$D = \{(x, y) : x^2 + y^2 = 1,\ x \geq 0,\ y \geq 0\}.$$

Method 1: Lagrange on the open arc plus endpoints. D is closed (an intersection of preimages of closed sets under continuous maps) and bounded ($|x|, |y| \leq 1$), hence compact, and f is continuous, so by the Extreme Value Theorem f attains a maximum and a minimum on D. Split D into the open arc in $U = \{x, y > 0\}$ and the two endpoints $(1, 0)$, $(0, 1)$. On the open set $U = \{x, y > 0\}$ take $g = x^2 + y^2$ (regular: $\nabla g \neq 0$); Lagrange condition:

$$\nabla f = (1, 1), \qquad \nabla g = (2x, 2y) \neq (0, 0) \text{ on } U,$$

$$1 = 2\lambda x, \qquad 1 = 2\lambda y, \qquad x^2 + y^2 = 1.$$

The first equation gives $\lambda \neq 0$, so $x = \frac{1}{2\lambda} = y$. The constraint then gives $2x^2 = 1$, so $x = \pm\frac{1}{\sqrt{2}}$, and $x > 0$ in U leaves $x = y = \frac{1}{\sqrt{2}}$ with

$$f\left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) = \tfrac{2}{\sqrt{2}} = \sqrt{2}.$$

Endpoints $(1, 0), (0, 1)$ lie outside the interior calculation:

$$f(1, 0) = 1 + 0 = 1, \qquad f(0, 1) = 0 + 1 = 1.$$

Comparing $\sqrt{2}$ with 1 gives

$$\boxed{\max_D (x + y) = \sqrt{2} \text{ at } (1/\sqrt{2}, 1/\sqrt{2}), \quad \min_D (x + y) = 1 \text{ at } (1, 0), (0, 1).}$$

Method 2: Parametrising the arc. The map $t \mapsto (\cos t, \sin t)$, $t \in [0, \frac{\pi}{2}]$, traces D exactly once, so

$$F(t) = \cos t + \sin t = \sqrt{2}\left(\tfrac{\sqrt{2}}{2}\cos t + \tfrac{\sqrt{2}}{2}\sin t\right) = \sqrt{2}\sin\left(t + \tfrac{\pi}{4}\right).$$

For $t \in [0, \frac{\pi}{2}]$, $t + \frac{\pi}{4} \in [\frac{\pi}{4}, \frac{3\pi}{4}]$, where sin ranges over $[\frac{\sqrt{2}}{2}, 1]$. Hence

$$1 = \sqrt{2} \cdot \tfrac{\sqrt{2}}{2} \leq F(t) \leq \sqrt{2},$$

with $F = \sqrt{2}$ iff $t + \frac{\pi}{4} = \frac{\pi}{2}$, i.e. $t = \frac{\pi}{4}$, the point $(1/\sqrt{2}, 1/\sqrt{2})$; and $F = 1$ iff $t + \frac{\pi}{4} \in \{\frac{\pi}{4}, \frac{3\pi}{4}\}$, i.e. $t = 0$ or $t = \frac{\pi}{2}$, the points $(1, 0)$, $(0, 1)$.

$$\boxed{\max_D (x + y) = \sqrt{2} \text{ at } (1/\sqrt{2}, 1/\sqrt{2}), \quad \min_D (x + y) = 1 \text{ at } (1, 0), (0, 1).}$$

Method 3: Squaring and two inequalities. On D, $x + y \geq 0$, so it suffices to bound $(x + y)^2$:

$$(x + y)^2 = x^2 + y^2 + 2xy = 1 + 2xy.$$

Lower bound: $xy \geq 0$ on D, so $(x + y)^2 \geq 1$, with equality iff $xy = 0$, i.e. $(x, y) = (1, 0)$ or $(0, 1)$. Upper bound: $2xy \leq x^2 + y^2$ because $(x - y)^2 \geq 0$, so $(x + y)^2 \leq 1 + (x^2 + y^2) = 2$, with equality iff $x = y$, i.e. $x = y = \frac{1}{\sqrt{2}}$. Taking nonnegative square roots, $1 \leq x + y \leq \sqrt{2}$:

$$\boxed{\max_D (x + y) = \sqrt{2} \text{ at } (1/\sqrt{2}, 1/\sqrt{2}), \quad \min_D (x + y) = 1 \text{ at } (1, 0), (0, 1).}$$

A constraint equation does not erase domain restrictions

Writing $g(x) = k$ does not replace the original domain of f and g. For example, the lidless-box variables satisfy $x, y, z > 0$; allowing zero or negative dimensions changes the mathematical problem. Likewise, a circle with $x, y \geq 0$ is an arc with endpoints, not the complete circle. Keep every stated inequality visible until the final comparison.

6.7 Chapter Summary

Key Formula: Chapter 6 recap

At a regular local constrained extremum,

$$\boxed{\nabla f = \lambda \nabla g \quad (g = k),} \qquad \boxed{\nabla f = \lambda \nabla g + \mu \nabla h \quad (g = k,\ h = \ell).}$$

For global extrema, combine these necessary equations with singular points, domain boundaries, and an existence or bounding argument.

Exam Focus: Chapter 6 takeaways

The multiplier equation is a candidate generator. State the regularity condition, solve all equations simultaneously, and remove only the auxiliary multipliers after the solution set is known. Closed and bounded feasible sets permit value comparison through the Extreme Value Theorem. On an open or unbounded feasible set, prove the requested global conclusion separately; the AM–GM argument in the lidless-box example is the model from Week 5.

QRS@NTU

7 Topic 7: Double Integrals

Single-variable integration accumulates a quantity along an interval. A double integral accumulates it over a planar region. For $f \geq 0$, the integral is the volume below $z = f(x, y)$; for signed f, it is the signed volume, with portions below the plane counted negatively. This chapter develops the Riemann definition, the conditions that guarantee existence, iterated integration, order reversal, additivity, and polar form.

7.1 Rectangles, partitions, and Riemann sums

Let $R = [a, b] \times [c, d]$ be a closed standard rectangle. Its area is $(b - a)(d - c)$. A partition P of R is made by choosing

$$a = x_0 < x_1 < \cdots < x_m = b, \qquad c = y_0 < y_1 < \cdots < y_n = d.$$

The resulting closed subrectangles are

$$R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], \qquad 1 \leq i \leq m, \quad 1 \leq j \leq n.$$

Write

$$\Delta x_i = x_i - x_{i-1}, \qquad \Delta y_j = y_j - y_{j-1}, \qquad \Delta A_{ij} = \Delta x_i \Delta y_j.$$

The mesh of P is the largest diameter of one of its subrectangles:

$$\|P\| = \max_{i,j} \sqrt{(\Delta x_i)^2 + (\Delta y_j)^2}.$$

Definition 7.1: Riemann sum and double integral

Let $f : R \to \mathbb{R}$ be bounded. Choose a sample point $(x^*_{ij}, y^*_{ij}) \in R_{ij}$ for every cell. The associated Riemann sum is

$$S(f; P, *) = \sum_{i=1}^{m} \sum_{j=1}^{n} f(x^*_{ij}, y^*_{ij}) \Delta A_{ij}.$$

We say that f is *Riemann integrable* on R with integral I if, for every $\varepsilon > 0$, there is $\delta > 0$ such that

$$\|P\| < \delta \quad \Longrightarrow \quad |S(f; P, *) - I| < \varepsilon$$

for every choice of sample points. We then write

$$I = \iint_R f(x, y) \, \mathrm{d}A.$$

The definition requires one common limit for all increasingly fine grids and all choices of tags. This independence is important: a formula based on a fortunate choice of sample points is not yet an integral.

Definition 7.2: Upper and lower Darboux sums

For bounded f and a partition P, set

$$M_{ij} = \sup_{R_{ij}} f, \qquad m_{ij} = \inf_{R_{ij}} f.$$

The upper and lower sums are

$$U(f,P) = \sum_{i,j} M_{ij}\Delta A_{ij}, \qquad L(f,P) = \sum_{i,j} m_{ij}\Delta A_{ij}.$$

The oscillation of f on a set E is

$$\omega_f(E) = \sup_E f - \inf_E f.$$

Thus

$$U(f,P) - L(f,P) = \sum_{i,j} \omega_f(R_{ij})\Delta A_{ij}.$$

Upper sums use the highest possible height in each column and lower sums use the lowest. Consequently every tagged sum lies between them:

$$L(f,P) \le S(f;P,*) \le U(f,P).$$

Theorem 7.1: Darboux criterion

A bounded function $f : R \to \mathbb{R}$ is Riemann integrable if and only if, for every $\varepsilon > 0$, some partition P satisfies

$$U(f,P) - L(f,P) < \varepsilon.$$

When this holds,

$$\iint_R f \, \mathrm{d}A = \sup_P L(f,P) = \inf_P U(f,P).$$

Proof. Assume first that f has Riemann integral I. Choose a partition whose mesh is small enough that every tagged sum is within $\varepsilon/4$ of I. For each cell, choose tags within $\varepsilon/(4\,\mathrm{area}(R))$ of its supremum and infimum; this is possible from the definition of supremum and infimum. The two resulting tagged sums differ from $U(f,P)$ and $L(f,P)$ by at most $\varepsilon/4$, respectively. They are each within $\varepsilon/4$ of I, so

$$U(f,P) - L(f,P) < \varepsilon.$$

Conversely, suppose such partitions exist. Let

$$\ell = \sup_P L(f,P), \qquad u = \inf_P U(f,P).$$

Refining a partition cannot decrease its lower sum or increase its upper sum. For any partitions P, Q, their common refinement gives $L(f,P) \le U(f,Q)$, hence $\ell \le u$. The assumed inequality gives a P with $u - \ell \le U(f,P) - L(f,P) < \varepsilon$. Thus $u = \ell =: I$.

For a fixed partition P with $U - L < \varepsilon/2$, refine any sufficiently fine partition Q by all grid lines of P. The portions near the finitely many grid lines of P have arbitrarily small area; boundedness of f makes their contribution arbitrarily small. Away from those strips, Q refines cells of P, so its tagged sum remains between the corresponding lower and upper sums. Hence every sufficiently fine tagged sum differs from I by less than ε. This is the Riemann definition. □

Theorem 7.2: Continuous functions are integrable

If $f : R \to \mathbb{R}$ is continuous, then it is bounded and Riemann integrable on R.

Proof. The rectangle R is closed and bounded, hence compact. The Extreme Value Theorem gives finite maximum and minimum values of f, so f is bounded. Continuity on a compact set also gives uniform continuity. For $\varepsilon > 0$, choose $\delta > 0$ such that

$$\|p - q\| < \delta \quad \Longrightarrow \quad |f(p) - f(q)| < \frac{\varepsilon}{\operatorname{area}(R)}.$$

Take a partition with mesh below δ. Any two points of a common cell are less than δ apart, so every cell has oscillation below $\varepsilon/\operatorname{area}(R)$. Therefore

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i,j} \omega_f(R_{ij}) \Delta A_{ij} \\ &< \frac{\varepsilon}{\operatorname{area}(R)} \sum_{i,j} \Delta A_{ij} = \varepsilon. \end{aligned}$$

The Darboux criterion proves integrability. □

Riemann-sum model

For a fine partition, remember

$$\boxed{\iint_R f(x,y)\, \mathrm{d}A \approx \sum_{i,j} f(x_{ij}^*, y_{ij}^*) \Delta x_i \Delta y_j.}$$

The approximation becomes exact in the limit as the mesh tends to zero, provided the tagged sums have a common limit.

7.2 General regions and the integrability criterion

Many planar regions are not rectangles. To define integration over a bounded region D, enclose it in a rectangle R and extend the integrand by zero outside D.

Definition 7.3: Jordan domain and restricted integral

A bounded $D \subseteq \mathbb{R}^2$ is a *Jordan domain* when ∂D has zero area, meaning a finite rectangle cover of arbitrarily small total area. If $R \supseteq D$ is a closed rectangle and

$$\widetilde{f}(p) = \begin{cases} f(p), & p \in D, \\ 0, & p \in R \setminus D, \end{cases}$$

is integrable on R, define $\displaystyle\iint_D f \, \mathrm{d}A = \iint_R \widetilde{f} \, \mathrm{d}A$.

Theorem 7.3: Lebesgue criterion on a rectangle

For a bounded $F : R \to \mathbb{R}$,

$$F \text{ is Riemann integrable} \quad \Longleftrightarrow \quad \operatorname{Disc}(F) \text{ has Lebesgue area zero.}$$

Lebesgue area zero permits a *countable* rectangle cover with arbitrarily small total area.

Proof. Put $\omega_F(p) = \inf_{U \ni p} \omega_F(U \cap R)$ and $E_k = \{p : \omega_F(p) \geq 1/k\}$. Each E_k is closed, hence compact. If F is integrable, choose P with $U(F,P) - L(F,P) < \varepsilon/(2k)$. Away from grid lines, each point of E_k is

in a cell whose oscillation is at least $1/k$. The total area of all such cells is at most

$$k \sum_{\text{selected } Q} \omega_F(Q)\,\text{area}(Q) \le k[U(F,P) - L(F,P)] < \varepsilon/2.$$

Cover the finitely many grid lines by thin rectangles of total area below $\varepsilon/2$. Together these covers include all of E_k. Thus each E_k has finite covers of arbitrarily small area. Covers of E_k of area $\varepsilon/2^k$ produce a countable cover of $\text{Disc}(F) = \bigcup_k E_k$. The union need not have a finite cover, which distinguishes Lebesgue zero area from Jordan zero area.

Conversely let $E = \text{Disc}(F)$ have Lebesgue area zero and take $|F| \le M$ with $M > 0$. Choose k with $\text{area}(R)/k < \varepsilon/2$. The compact $E_k \subseteq E$ has a finite subcover, so it lies in a finite open-rectangle union G of area below $\varepsilon/(8M)$. On compact $R \setminus G$, every point has a neighbourhood of oscillation below $1/k$; finitely many suffice. A fine grid has every cell avoiding G inside one of them. Its cells meeting G lie in G plus thin strips about the finitely many edges of G, chosen with area below $\varepsilon/(8M)$. Hence

$$U(F,P) - L(F,P) < \frac{\varepsilon}{2} + 2M\left(\frac{\varepsilon}{8M} + \frac{\varepsilon}{8M}\right) = \varepsilon.$$

The Darboux criterion completes the proof. □

Theorem 7.4: Area exists exactly when the boundary has zero area

For a bounded $D \subseteq \mathbb{R}^2$,

$$\text{Area}(D) = \iint_D 1\,\mathrm{d}A \text{ exists} \quad \Longleftrightarrow \quad \partial D \text{ has Lebesgue area zero.}$$

For the compact boundary, this is equivalent to a finite rectangle cover of arbitrarily small total area.

Proof. Choose an enclosing rectangle whose interior contains D. The zero extension of 1 on it is $\mathbf{1}_D$. Its discontinuity set is precisely ∂D. Apply Theorem 7.3. □

Theorem 7.5: Consistency and criterion on a Jordan domain

For a Jordan domain D, $f : D \to \mathbb{R}$ is integrable if and only if it is bounded and its relative discontinuity set has Lebesgue area zero. Its integral is independent of the enclosing rectangle.

Proof. The discontinuities of $\widetilde{f}$ lie among ∂D and the relative discontinuities of f, and the converse holds away from ∂D. The rectangle criterion proves the assertion. Enlarging the enclosing rectangle adds finitely many rectangles on which the extension is zero, so it does not change the integral. □

Bounded does not mean integrable

The bounded function that equals 1 on rational-coordinate points and 0 elsewhere is discontinuous at every point of a rectangle. Its upper sum is always the rectangle area and its lower sum is always zero, so it is not Riemann integrable. Boundedness is necessary in this course's Riemann definition, but it is far from sufficient.

7.3 Applications of general double integrals

Theorem 7.6: Volume under a nonnegative graph

Let $D \subseteq \mathbb{R}^2$ be bounded and suppose that $\text{Area}(D)$ exists. If $h : D \to \mathbb{R}$ is bounded, continuous, and nonnegative, the volume between $D \times \{0\}$ and the graph of h is

$$V = \iint_D h(x, y)\ \mathrm{d}A.$$

Proof. The zero extension of h to an enclosing rectangle is Riemann integrable by Theorem 7.5. Over each partition cell, prisms with heights equal to its infimum and supremum bound the solid from below and above. Their total volumes are the lower and upper Darboux sums. Integrability makes their difference arbitrarily small, so the enclosed volume is their common limit, namely the integral. □

Definition 7.4: Mass density and total mass

A *mass-density function* on D is a bounded, continuous, nonnegative function $\rho : D \to \mathbb{R}$. The total mass of the lamina over D is

$$m = \iint_D \rho(x, y)\ \mathrm{d}A.$$

If f is integrable on D and $\text{Area}(D) > 0$, its average value on D is

$$f_{\text{avg}} = \frac{1}{\text{Area}(D)} \iint_D f(x, y)\ \mathrm{d}A.$$

Definition 7.5: Joint probability density

A normalized joint probability density for real-valued random variables X and Y is a nonnegative function $p : \mathbb{R}^2 \to \mathbb{R}$ for which

$$\iint_{\mathbb{R}^2} p(x, y)\ \mathrm{d}A = 1.$$

Because $\mathbb{R}^2$ is unbounded, this total integral is an improper double integral. For a bounded region $E \subseteq \mathbb{R}^2$ for which $\iint_E p(x, y)\ \mathrm{d}A$ exists,

$$\text{Prob}\big((X, Y) \in E\big) = \iint_E p(x, y)\ \mathrm{d}A \leq 1.$$

7.4 Fubini's theorem and iterated integrals

Iterated integrals turn a two-dimensional accumulation into a sequence of one-dimensional accumulations.

Theorem 7.7: Fubini's theorem on a rectangle

If f is continuous on $R = [a, b] \times [c, d]$, then

$$\iint_R f(x, y)\ \mathrm{d}A = \int_a^b \left[\int_c^d f(x, y)\ \mathrm{d}y\right] \mathrm{d}x = \int_c^d \left[\int_a^b f(x, y)\ \mathrm{d}x\right] \mathrm{d}y.$$

Proof. Partition the x-interval and y-interval, and use tags (x_i^*, y_j^*). For a fixed x_i^*, the one-variable

Riemann sum

$$\sum_{j=1}^{n} f(x_i^*, y_j^*)\Delta y_j$$

converges to $g(x_i^*) = \int_c^d f(x_i^*, y)\,\mathrm{d}y$ as the y mesh tends to zero. Uniform continuity of f on R makes this convergence uniform in x_i^*: the error is bounded by $(d-c)$ times the largest within-strip oscillation. Thus the double Riemann sum has the same limit as

$$\sum_i g(x_i^*)\Delta x_i.$$

The same uniform-continuity estimate shows that g is continuous, since

$$|g(x) - g(x')| \le \int_c^d |f(x,y) - f(x',y)|\,\mathrm{d}y.$$

The last sum therefore tends to $\int_a^b g(x)\,\mathrm{d}x$. This proves the first equality. Interchanging the roles of x and y proves the second. □

Theorem 7.8: Riemann Fubini with integrable slices

Let F be Riemann integrable on $[a,b]\times[c,d]$. If every vertical slice $F(x,\cdot)$ is integrable, then

$$\iint_R F\,\mathrm{d}A = \int_a^b \left(\int_c^d F(x,y)\,\mathrm{d}y\right)\mathrm{d}x.$$

The analogous statement holds for horizontal slices.

Proof. Let $g(x) = \int_c^d F(x,y)\,\mathrm{d}y$. For an x-partition and any tags, refine the y-partition. The lower and upper Darboux sums of F squeeze the corresponding tagged sum of g:

$$L(F,P) \le \sum_i g(x_i^*)\Delta x_i \le U(F,P).$$

Indeed, integrate each cell's constant lower and upper bounds along the fixed slice and sum over its y-subintervals. When the two-dimensional mesh is small, $U(F,P) - L(F,P)$ is arbitrarily small and both bounds squeeze the integral of F. Thus the tagged sums of g have that common limit. □

Worked Examples (Week 5): rectangle Fubini

The integrands x^2y on $R_1 = [0,3]\times[1,2]$ and $y\sin(xy)$ on $R_2 = [1,2]\times[0,\pi]$ are continuous on closed rectangles, so Fubini's theorem gives the double integral by either iterated order.
Method 1: Fubini with $\mathrm{d}x$ inside. For R_2, fix y and put $u = xy$, $\mathrm{d}u = y\,\mathrm{d}x$; the new limits are $u = y$ at $x = 1$ and $u = 2y$ at $x = 2$.

$$\iint_{R_1} x^2y\,\mathrm{d}A = \int_1^2\int_0^3 x^2y\,\mathrm{d}x\,\mathrm{d}y = \int_1^2 \Big[\frac{x^3y}{3}\Big]_{x=0}^{x=3}\mathrm{d}y = \int_1^2 9y\,\mathrm{d}y = \Big[\frac{9y^2}{2}\Big]_1^2 = 18-\frac{9}{2} = \boxed{\frac{27}{2}},$$

$$\iint_{R_2} y\sin(xy)\,\mathrm{d}A = \int_0^\pi\int_1^2 y\sin(xy)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\pi\int_y^{2y}\sin u\,\mathrm{d}u\,\mathrm{d}y = \int_0^\pi \big[-\cos u\big]_{u=y}^{u=2y}\mathrm{d}y$$
$$= \int_0^\pi [\cos y - \cos(2y)]\,\mathrm{d}y = \Big[\sin y - \frac{\sin(2y)}{2}\Big]_0^\pi = (0-0)-(0-0) = \boxed{0}.$$

Method 2: Fubini with $\mathrm{d}y$ inside. For R_2, fix $x \geq 1$ and integrate by parts with $u = y$, $\mathrm{d}v = \sin(xy)\,\mathrm{d}y$, $v = -\cos(xy)/x$. The quotient rule gives $\frac{\mathrm{d}}{\mathrm{d}x}\big(-\frac{\sin(\pi x)}{x}\big) = -\frac{\pi\cos(\pi x)}{x} + \frac{\sin(\pi x)}{x^2}$, which is the resulting outer integrand.

$$\iint_{R_1} x^2y\,\mathrm{d}A = \int_0^3\int_1^2 x^2y\,\mathrm{d}y\,\mathrm{d}x = \int_0^3 x^2\Big[\frac{y^2}{2}\Big]_1^2\mathrm{d}x = \int_0^3 \frac{3}{2}x^2\,\mathrm{d}x = \Big[\frac{x^3}{2}\Big]_0^3 = \boxed{\frac{27}{2}},$$

$$\int_0^\pi y\sin(xy)\,\mathrm{d}y = \Big[-\frac{y\cos(xy)}{x}\Big]_{y=0}^{y=\pi} + \frac{1}{x}\int_0^\pi \cos(xy)\,\mathrm{d}y = -\frac{\pi\cos(\pi x)}{x} + \Big[\frac{\sin(xy)}{x^2}\Big]_{y=0}^{y=\pi}$$
$$= -\frac{\pi\cos(\pi x)}{x} + \frac{\sin(\pi x)}{x^2},$$

$$\iint_{R_2} y\sin(xy)\,\mathrm{d}A = \int_1^2\Big(-\frac{\pi\cos(\pi x)}{x} + \frac{\sin(\pi x)}{x^2}\Big)\mathrm{d}x = \Big[-\frac{\sin(\pi x)}{x}\Big]_1^2 = -\frac{\sin(2\pi)}{2} + \frac{\sin\pi}{1} = \boxed{0}.$$

Definition 7.6: Type I and Type II regions

A Type I region has the form

$$D = \{(x,y) : a \leq x \leq b,\ g_1(x) \leq y \leq g_2(x)\},$$

where g_1, g_2 are continuous and $g_1 \leq g_2$. A Type II region has the form

$$D = \{(x,y) : c \leq y \leq d,\ h_1(y) \leq x \leq h_2(y)\},$$

where h_1, h_2 are continuous and $h_1 \leq h_2$.

Theorem 7.9: Fubini on Type I and Type II regions

Let f be continuous on a Type I region D. Then

$$\iint_D f(x,y)\,\mathrm{d}A = \int_a^b\int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

For a Type II region,

$$\iint_D f(x,y)\,\mathrm{d}A = \int_c^d\int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Proof. If $a = b$, both the region area and the iterated integral are zero. Otherwise, for a continuous

$g : [a, b] \to \mathbb{R}$, uniform continuity gives a partition with oscillation below $\varepsilon/(b-a)$ on each subinterval. Rectangles $I_i \times [\inf_{I_i} g, \sup_{I_i} g]$ cover its graph and have total area below ε. Thus both boundary graphs, and the two vertical segments, have zero area. Hence D is a Jordan domain.

Enclose D in a rectangle and use its integrable zero extension F. For each fixed x, $F(x, \cdot)$ is continuous except possibly at the two slice endpoints, so it is integrable; its integral is

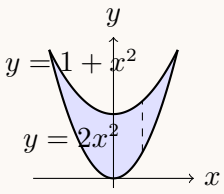
$$\int_{g_1(x)}^{g_2(x)} f(x, y)\, \mathrm{d}y.$$

Theorem 7.8, rather than the continuous-rectangle version, now yields the Type I formula. The Type II proof swaps x and y. □

Worked Example (Week 5): a Type I region

Let

$$D = \{(x, y) : -1 \le x \le 1,\ 2x^2 \le y \le 1 + x^2\}.$$



For $f(x, y) = x + 2y$, continuous on D. The curves meet where $2x^2 = 1 + x^2$, i.e. at $(\pm 1, 2)$.
Method 1: Fubini on the Type I description. In the evaluation the even powers cancel between the limits and the odd powers double.

$$\begin{aligned}
\iint_D (x + 2y)\, \mathrm{d}A &= \int_{-1}^{1} \int_{2x^2}^{1+x^2} (x + 2y)\, \mathrm{d}y\, \mathrm{d}x = \int_{-1}^{1} \left[xy + y^2\right]_{y=2x^2}^{y=1+x^2} \mathrm{d}x \\
&= \int_{-1}^{1} \left[(x + x^3 + 1 + 2x^2 + x^4) - (2x^3 + 4x^4)\right] \mathrm{d}x \\
&= \int_{-1}^{1} (1 + x + 2x^2 - x^3 - 3x^4)\, \mathrm{d}x = \left[x + \frac{x^2}{2} + \frac{2x^3}{3} - \frac{x^4}{4} - \frac{3x^5}{5}\right]_{-1}^{1} \\
&= \left(1 + \frac{1}{2} + \frac{2}{3} - \frac{1}{4} - \frac{3}{5}\right) - \left(-1 + \frac{1}{2} - \frac{2}{3} - \frac{1}{4} + \frac{3}{5}\right) \\
&= 2\left(1 + \frac{2}{3} - \frac{3}{5}\right) = \frac{2(15 + 10 - 9)}{15} \implies \boxed{\iint_D (x + 2y)\, \mathrm{d}A = \frac{32}{15}.}
\end{aligned}$$

Method 2: Symmetry and Fubini on Type II slices. Solving $y = 2x^2$ and $y = 1 + x^2$ for x, the horizontal slice at height y is $|x| \le \sqrt{y/2}$ for $0 \le y \le 1$ and $\sqrt{y-1} \le |x| \le \sqrt{y/2}$ for $1 \le y \le 2$. Each slice is symmetric about $x = 0$, so $\int x\, \mathrm{d}x = 0$ on it, and $\int 2y\, \mathrm{d}x$ is $2y$ times the slice length; here $2y \cdot 2\sqrt{y/2} = 2\sqrt{2}\, y^{3/2}$. In the last integral put $u = y - 1$, $\mathrm{d}u = \mathrm{d}y$, with new limits $u = 0$ and $u = 1$.

$$\begin{aligned}
\iint_D (x + 2y)\, \mathrm{d}A &= \int_0^1 2\sqrt{2}\, y^{3/2}\, \mathrm{d}y + \int_1^2 2y \cdot 2\left(\sqrt{y/2} - \sqrt{y-1}\right) \mathrm{d}y \\
&= \int_0^2 2\sqrt{2}\, y^{3/2}\, \mathrm{d}y - 4\int_1^2 y\sqrt{y-1}\, \mathrm{d}y \\
&= 2\sqrt{2}\left[\frac{2y^{5/2}}{5}\right]_0^2 - 4\int_0^1 (u+1)u^{1/2}\, \mathrm{d}u = 2\sqrt{2} \cdot \frac{8\sqrt{2}}{5} - 4\left[\frac{2u^{5/2}}{5} + \frac{2u^{3/2}}{3}\right]_0^1 \\
&= \frac{32}{5} - \frac{64}{15} \implies \boxed{\iint_D (x + 2y)\, \mathrm{d}A = \frac{32}{15}.}
\end{aligned}$$

7.5 Changing the order of integration

The symbols in an iterated integral do not determine the region by themselves. Read the inner bounds as a moving slice, draw it, then describe the same set using the other orientation.

Theorem 7.10: Order reversal

Suppose a region D is both Type I and Type II, and f is continuous on it. Then either iterated description gives the same double integral. In particular, changing order changes limits, but not the value.

Proof. The Type I formula in Theorem 7.9 and the Type II formula in the same theorem both equal the defining double integral $\iint_D f \, dA$. Equating these two values proves the claim. The continuity hypothesis supplies the integrability required by both uses of Fubini. □

Worked Example (Week 5): choose the Type II order

Let D be bounded by $y = x - 1$ and $y^2 = 2x + 6$. The curves meet where $y + 1 = y^2/2 - 3$, i.e. $y^2 - 2y - 8 = (y - 4)(y + 2) = 0$, at $(-1, -2)$ and $(5, 4)$. The integrand xy is continuous, so Fubini applies.

Method 1: Fubini on the Type II description. For $-2 \le y \le 4$ the parabola $x = y^2/2 - 3$ lies left of the line $x = y + 1$:

$$D = \{(x, y) : -2 \le y \le 4, \quad y^2/2 - 3 \le x \le y + 1\}.$$

$$\begin{aligned}
\iint_D xy \, dA &= \int_{-2}^{4} \int_{y^2/2-3}^{y+1} xy \, dx \, dy = \int_{-2}^{4} \Big[\frac{x^2 y}{2}\Big]_{x=y^2/2-3}^{x=y+1} dy \\
&= \frac{1}{2} \int_{-2}^{4} y\Big[(y^2 + 2y + 1) - \Big(\frac{y^4}{4} - 3y^2 + 9\Big)\Big] \, dy \\
&= \frac{1}{2} \int_{-2}^{4} \Big(-\frac{y^5}{4} + 4y^3 + 2y^2 - 8y\Big) \, dy = \frac{1}{2} \Big[-\frac{y^6}{24} + y^4 + \frac{2y^3}{3} - 4y^2\Big]_{-2}^{4} \\
&= \frac{1}{2}\Big[\Big(-\frac{512}{3} + 256 + \frac{128}{3} - 64\Big) - \Big(-\frac{8}{3} + 16 - \frac{16}{3} - 16\Big)\Big] = \frac{1}{2}[64 - (-8)] = \boxed{36}.
\end{aligned}$$

Method 2: Fubini on Type I slices, split at $x = -1$. Solving $y^2 = 2x + 6$ gives $y = \pm\sqrt{2x + 6}$ for $x \ge -3$. A vertical slice starts on the lower parabola branch for $-3 \le x \le -1$ and on the line for $-1 \le x \le 5$; it always ends on the upper branch:

$$D = \{-3 \le x \le -1, \ -\sqrt{2x + 6} \le y \le \sqrt{2x + 6}\} \cup \{-1 \le x \le 5, \ x - 1 \le y \le \sqrt{2x + 6}\}.$$

On the first piece $\int_{-\sqrt{2x+6}}^{\sqrt{2x+6}} xy \, dy = \frac{x}{2}[(2x + 6) - (2x + 6)] = 0$. Additivity over the two pieces gives

$$\begin{aligned}
\iint_D xy \, dA &= 0 + \int_{-1}^{5} \Big[\frac{xy^2}{2}\Big]_{y=x-1}^{y=\sqrt{2x+6}} dx = \frac{1}{2} \int_{-1}^{5} x\big[(2x + 6) - (x^2 - 2x + 1)\big] \, dx \\
&= \frac{1}{2} \int_{-1}^{5} (-x^3 + 4x^2 + 5x) \, dx \\
&= \frac{1}{2}\Big[-\frac{x^4}{4} + \frac{4x^3}{3} + \frac{5x^2}{2}\Big]_{-1}^{5} = \frac{1}{2}\Big[\Big(-\frac{625}{4} + \frac{500}{3} + \frac{125}{2}\Big) - \Big(-\frac{1}{4} - \frac{4}{3} + \frac{5}{2}\Big)\Big] \\
&= \frac{1}{2}\Big(\frac{875}{12} - \frac{11}{12}\Big) = \boxed{36}.
\end{aligned}$$

Worked Example (Week 5): tetrahedron volume

The tetrahedron bounded by

$$x + 2y + z = 2, \qquad x = 2y, \qquad x = 0, \qquad z = 0$$

lies below $z = 2 - x - 2y \geq 0$ over the triangle T in $z = 0$ with vertices $(0, 0)$, $(0, 1)$, $(1, \frac{1}{2})$ (where $x + 2y = 2$ meets $x = 0$ and $x = 2y$), i.e. $0 \leq x \leq 1$, $x/2 \leq y \leq 1 - x/2$. The height is continuous, so Fubini applies.

Method 1: Fubini on the Type I description. In the last integral put $u = 1 - x$, $\mathrm{d}u = -\,\mathrm{d}x$, with new limits $u = 1$ at $x = 0$ and $u = 0$ at $x = 1$.

$$\begin{aligned} V &= \int_0^1 \int_{x/2}^{1-x/2} (2 - x - 2y)\, \mathrm{d}y\, \mathrm{d}x = \int_0^1 \left[(2-x)y - y^2\right]_{y=x/2}^{y=1-x/2} \mathrm{d}x \\ &= \int_0^1 \left[(2-x)(1-x) - \left(\left(1 - \tfrac{x}{2}\right)^2 - \tfrac{x^2}{4}\right)\right] \mathrm{d}x \\ &= \int_0^1 \left[(2-x)(1-x) - (1-x)\right] \mathrm{d}x = \int_0^1 (1-x)^2\, \mathrm{d}x = \int_0^1 u^2\, \mathrm{d}u = \left[\frac{u^3}{3}\right]_0^1 = \boxed{\frac{1}{3}}. \end{aligned}$$

Method 2: Fubini on Type II slices, split at $y = \frac{1}{2}$. Horizontally, T is $0 \leq x \leq 2y$ for $0 \leq y \leq \frac{1}{2}$ and $0 \leq x \leq 2 - 2y$ for $\frac{1}{2} \leq y \leq 1$. For a slice of length h, $\int_0^h (2 - 2y - x)\, \mathrm{d}x = (2 - 2y)h - \frac{h^2}{2}$; with $h = 2y$ this is $4y - 6y^2$, and with $h = 2 - 2y$ it is $\frac{(2-2y)^2}{2} = 2(1-y)^2$.

$$\begin{aligned} V &= \int_0^{1/2} (4y - 6y^2)\, \mathrm{d}y + \int_{1/2}^1 2(1-y)^2\, \mathrm{d}y = \left[2y^2 - 2y^3\right]_0^{1/2} + \left[-\frac{2(1-y)^3}{3}\right]_{1/2}^1 \\ &= \left(\frac{1}{2} - \frac{1}{4}\right) + \frac{2}{3} \cdot \frac{1}{8} = \frac{1}{4} + \frac{1}{12} = \boxed{\frac{1}{3}}. \end{aligned}$$

Method 3: Pyramid volume formula. The four planes meet in the vertices $(0, 0, 0)$, $(0, 1, 0)$, $(1, \frac{1}{2}, 0)$ and the apex $(0, 0, 2)$ (from $x = 0$, $x = 2y$, $z = 2 - x - 2y$). The base T has the side from $(0, 0)$ to $(0, 1)$ of length 1 and opposite vertex at distance 1, so $\mathrm{Area}(T) = \frac{1}{2}$; the apex is at height 2. Hence $V = \frac{1}{3}\,\mathrm{Area}(T) \cdot 2 = \frac{1}{3} \cdot \frac{1}{2} \cdot 2 = \boxed{\frac{1}{3}}$.

Do not swap symbols and keep the old limits

Writing $\int \int f\, \mathrm{d}y\, \mathrm{d}x$ after merely exchanging $\mathrm{d}x$ and $\mathrm{d}y$ is usually wrong. First solve the inequalities that describe the same shaded region. If one horizontal or vertical slice changes its endpoints partway across the region, split the integral at that change.

7.6 Linearity, restriction, and additivity

The basic algebra of a double integral is inherited from the corresponding algebra of finite sums.

Theorem 7.11: Linearity and comparison

If f, g are integrable over D and $\alpha, \beta \in \mathbb{R}$, then

$$\iint_D (\alpha f + \beta g)\ \mathrm{d}A = \alpha \iint_D f\ \mathrm{d}A + \beta \iint_D g\ \mathrm{d}A.$$

If $f \leq g$ on D, then

$$\iint_D f\ \mathrm{d}A \leq \iint_D g\ \mathrm{d}A.$$

Consequently,

$$\left|\iint_D f\ \mathrm{d}A\right| \leq \iint_D |f|\ \mathrm{d}A.$$

Proof. For every tagged partition sum,

$$S(\alpha f + \beta g; P, *) = \alpha S(f; P, *) + \beta S(g; P, *).$$

Taking the common limits proves linearity. If $f \leq g$, then each term of every common tagged sum for f is at most the corresponding term for g because cell areas are nonnegative. Taking limits proves comparison. Before using comparison, note that $|f|$ is integrable: for any p, q,

$$\big||f(p)| - |f(q)|\big| \leq |f(p) - f(q)|,$$

so every cell oscillation of $|f|$ is at most that of f, and the Darboux criterion applies. Now $-|f| \leq f \leq |f|$ gives the stated bound. □

Theorem 7.12: Additivity over pieces

Let $D = D_1 \cup \cdots \cup D_N$, where each D_k is a Jordan domain and any two distinct pieces intersect only on a set of area zero. If f is integrable on every D_k, it is integrable on D and

$$\iint_D f\ \mathrm{d}A = \sum_{k=1}^{N} \iint_{D_k} f\ \mathrm{d}A.$$

Proof. Use a common enclosing rectangle and write $F = f\mathbf{1}_D$ and $F_k = f\mathbf{1}_{D_k}$, with zero values outside the named sets. By hypothesis each F_k is integrable, and their finite sum is bounded. Let $O = \bigcup_{i<j}(D_i \cap D_j)$. Two piece interiors cannot intersect, since a common open disk has positive area. Thus $O \subseteq \bigcup_k \partial D_k$, a compact zero-area set. The bounded function $\sum_k F_k - F$ is supported there, so its upper absolute Darboux sums can be made arbitrarily small by covering that set with tiny rectangles. Its integral is zero. Hence F is integrable, and linearity gives

$$\iint_D f\ \mathrm{d}A = \int F\ \mathrm{d}A = \sum_{k=1}^{N} \int F_k\ \mathrm{d}A = \sum_{k=1}^{N} \iint_{D_k} f\ \mathrm{d}A.$$

□

Conversely, if f is integrable on D, its restriction to each Jordan piece is integrable: the zero extension F_k can be discontinuous only where F is discontinuous or on ∂D_k; apply the Lebesgue criterion.

Worked Example (Week 6): additivity on three pieces

Let D be the union of

$$D_1=\{(x,y):-2\le x\le 0,\ 0\le y\le (x+2)^2\},\qquad D_2=\{(x,y):0\le y\le 4,\ 0\le x\le y-y^3/16\},$$

and

$$D_3=\{(x,y):-4\le y\le 0,\ -2\le x\le y-y^3/16\}.$$

The overlaps are segments or a point, of area zero. Write $g(y)=y-y^3/16$; $f(x,y)=2x+5y$ is continuous, and every piece is Type I or Type II, so Fubini applies on each.

Method 1: Additivity over D_1, D_2, D_3. On D_1 put $u=x+2$, $\mathrm{d}u=\mathrm{d}x$, $x=u-2$, with new limits $u=0$ and $u=2$:

$$\begin{aligned}\iint_{D_1} f\,\mathrm{d}A&=\int_{-2}^{0}\Big[2xy+\frac{5y^2}{2}\Big]_{y=0}^{y=(x+2)^2}\mathrm{d}x=\int_{-2}^{0}\Big(2x(x+2)^2+\frac{5}{2}(x+2)^4\Big)\mathrm{d}x\\&=\int_0^2\Big(\frac{5u^4}{2}+2u^3-4u^2\Big)\mathrm{d}u=\Big[\frac{u^5}{2}+\frac{u^4}{2}-\frac{4u^3}{3}\Big]_0^2=16+8-\frac{32}{3}=\frac{40}{3}.\end{aligned}$$

On D_2 and D_3 the inner antiderivative is x^2+5xy, and at $x=g(y)$ it equals $g^2+5yg=\big(y^2-\frac{y^4}{8}+\frac{y^6}{256}\big)+\big(5y^2-\frac{5y^4}{16}\big)=6y^2-\frac{7y^4}{16}+\frac{y^6}{256}=:p(y)$, an even polynomial, so $\int_{-4}^{0}p=\int_0^4 p$.

$$\begin{aligned}\iint_{D_2} f\,\mathrm{d}A&=\int_0^4\big[x^2+5xy\big]_{x=0}^{x=g(y)}\mathrm{d}y=\int_0^4 p(y)\,\mathrm{d}y=\Big[2y^3-\frac{7y^5}{80}+\frac{y^7}{1792}\Big]_0^4\\&=128-\frac{448}{5}+\frac{64}{7}=\frac{4480-3136+320}{35}=\frac{1664}{35},\\\iint_{D_3} f\,\mathrm{d}A&=\int_{-4}^{0}\big[x^2+5xy\big]_{x=-2}^{x=g(y)}\mathrm{d}y=\int_{-4}^{0}p(y)\,\mathrm{d}y-\int_{-4}^{0}(4-10y)\,\mathrm{d}y\\&=\frac{1664}{35}-\big[4y-5y^2\big]_{-4}^{0}=\frac{1664}{35}-96=-\frac{1696}{35},\\\iint_D f\,\mathrm{d}A&=\frac{40}{3}+\frac{1664}{35}-\frac{1696}{35}=\frac{40}{3}-\frac{32}{35}=\frac{1400-96}{105}\implies\boxed{\iint_D(2x+5y)\,\mathrm{d}A=\frac{1304}{105}.}\end{aligned}$$

Method 2: Additivity with a removed parabolic cap. Let $E=\{-4\le y\le 4,\ -2\le x\le g(y)\}$ and $A=\{0\le y\le 4,\ -2\le x\le -2+\sqrt{y}\}$, the part of $[-2,0]\times[0,4]$ above $y=(x+2)^2$. Since $g\ge 0$ on $[0,4]$, $E=D\cup A$ with $D\cap A$ an arc of zero area, so $\iint_D f\,\mathrm{d}A=\iint_E f\,\mathrm{d}A-\iint_A f\,\mathrm{d}A$. On E, the inner antiderivative x^2+5xy gives $g^2+5yg-(4-10y)=6y^2-\frac{7y^4}{16}+\frac{y^6}{256}-4+10y$; the even part doubles over $[-4,4]$ and $10y$ integrates to 0:

$$\begin{aligned}\iint_E f\,\mathrm{d}A&=2\Big[2y^3-\frac{7y^5}{80}+\frac{y^7}{1792}\Big]_0^4-4\cdot 8=2\Big(128-\frac{448}{5}+\frac{64}{7}\Big)-32=\frac{3328}{35}-32=\frac{2208}{35},\\\iint_A f\,\mathrm{d}A&=\int_0^4\big[x^2+5xy\big]_{x=-2}^{x=-2+\sqrt{y}}\mathrm{d}y=\int_0^4\big[(4-4\sqrt{y}+y)+(-10y+5y^{3/2})-(4-10y)\big]\mathrm{d}y\\&=\int_0^4\big(y-4y^{1/2}+5y^{3/2}\big)\mathrm{d}y=\Big[\frac{y^2}{2}-\frac{8y^{3/2}}{3}+2y^{5/2}\Big]_0^4=8-\frac{64}{3}+64=\frac{152}{3},\\\iint_D f\,\mathrm{d}A&=\frac{2208}{35}-\frac{152}{3}=\frac{6624-5320}{105}\implies\boxed{\iint_D(2x+5y)\,\mathrm{d}A=\frac{1304}{105}.}\end{aligned}$$

7.7 Polar coordinates and the area factor

Cartesian rectangles are poorly suited to circles, sectors, and radial boundaries. In polar coordinates,

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad r \geq 0.$$

The angle interval should describe each point once, apart from boundary rays.

Theorem 7.13: Polar area element

Under the polar map $T(r,\theta) = (r\cos\theta, r\sin\theta)$,

$$\mathrm{d}A = r\ \mathrm{d}r\ \mathrm{d}\theta.$$

Thus a polar rectangle

$$0 \leq a \leq r \leq b, \qquad \alpha \leq \theta \leq \beta, \qquad 0 \leq \beta - \alpha \leq 2\pi,$$

has area

$$\int_\alpha^\beta \int_a^b r\ \mathrm{d}r\ \mathrm{d}\theta = \frac{1}{2}(b^2 - a^2)(\beta - \alpha).$$

Proof. For an annular sector $[r, r+\Delta r] \times [\theta, \theta + \Delta\theta]$, the exact area is

$$\frac{1}{2}\big((r+\Delta r)^2 - r^2\big)\Delta\theta = \left(r\Delta r + \frac{1}{2}(\Delta r)^2\right)\Delta\theta.$$

Replacing r by any tag r^* changes the cell area by at most $(\Delta r)^2\Delta\theta$. Across a partition of $[a,b]\times[\alpha,\beta]$, the total error is at most $(\max \Delta r)(b-a)(\beta-\alpha)$ and tends uniformly to zero. The limit is therefore the Riemann sum for $r\ \mathrm{d}r\ \mathrm{d}\theta$. □

Theorem 7.14: Polar-coordinate integration

Let

$$D = \{(r\cos\theta, r\sin\theta) : \alpha \leq \theta \leq \beta,\ h_1(\theta) \leq r \leq h_2(\theta)\},$$

where h_1, h_2 are continuous, nonnegative, and $h_1 \leq h_2$. Assume $0 \leq \beta - \alpha \leq 2\pi$, and let f be continuous on D. Then

$$\iint_D f(x,y)\ \mathrm{d}A = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta) r\ \mathrm{d}r\ \mathrm{d}\theta.$$

Proof. Partition the angle interval and use the supremum of h_1 and infimum of h_2 on each part as inner step bounds, discarding a piece if they cross. These sectors stay inside D. Uniform continuity of the radii makes the omitted annular strips have total area tending to zero. Partition the resulting parameter rectangles into small rectangles. The exact sector estimate from Theorem 7.13 makes the difference between the Cartesian-area sum and

$$\sum_{i,j} f(r_{ij}^*\cos\theta_{ij}^*, r_{ij}^*\sin\theta_{ij}^*) r_{ij}^* \Delta r_i \Delta\theta_j$$

tend uniformly to zero; boundedness of f controls the accumulated error. The remaining sum is the Riemann sum of the displayed iterated integral. Boundary rays and the origin have zero area, so repeated descriptions there do not affect the integral. □

Polar setup

$$\boxed{x = r\cos\theta, \quad y = r\sin\theta, \quad \mathrm{d}A = r\ \mathrm{d}r\ \mathrm{d}\theta.}$$

For a polar region with inner and outer radii h_1, h_2,

$$\boxed{\iint_D f\ \mathrm{d}A = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta) r\ \mathrm{d}r\ \mathrm{d}\theta.}$$

The extra r is the area factor. Omitting it changes the problem.

Worked Example (Week 6): a left half-annulus

Let

$$D = \{(x,y) : x \le 0,\ 1 \le x^2 + y^2 \le 4\}.$$

For $f(x,y) = x + y$, continuous on D.
Method 1: Polar coordinates. With $x = r\cos\theta$, $y = r\sin\theta$: $1 \le x^2 + y^2 = r^2 \le 4$ gives $1 \le r \le 2$, and $x \le 0$ means $\cos\theta \le 0$:

$$1 \le r \le 2, \qquad \frac{\pi}{2} \le \theta \le \frac{3\pi}{2}.$$

The integrand times the area factor r is $r^2(\cos\theta + \sin\theta)$, a product of a function of r and one of θ.

$$\begin{aligned}
\iint_D (x+y)\ \mathrm{d}A &= \int_{\pi/2}^{3\pi/2} \int_1^2 (r\cos\theta + r\sin\theta) r\ \mathrm{d}r\ \mathrm{d}\theta \\
&= \left(\int_{\pi/2}^{3\pi/2} (\cos\theta + \sin\theta)\ \mathrm{d}\theta\right)\left(\int_1^2 r^2\ \mathrm{d}r\right) \\
&= \big[\sin\theta - \cos\theta\big]_{\pi/2}^{3\pi/2} \left[\frac{r^3}{3}\right]_1^2 = \big[(-1-0) - (1-0)\big]\left(\frac{8}{3} - \frac{1}{3}\right) \\
&= (-2)\left(\frac{7}{3}\right) \implies \boxed{\iint_D (x+y)\ \mathrm{d}A = -\frac{14}{3}.}
\end{aligned}$$

Method 2: Cartesian half-disks and additivity. For $a > 0$ let $H_a = \{x \le 0,\ x^2 + y^2 \le a^2\} = \{-a \le x \le 0,\ -s_a(x) \le y \le s_a(x)\}$ with $s_a(x) = \sqrt{a^2 - x^2}$. Then $H_2 = D \cup H_1$ with $D \cap H_1$ an arc of zero area, so $\iint_D = \iint_{H_2} - \iint_{H_1}$. By Fubini (Type I), the y-term vanishes on every symmetric slice. In the x-term put $u = a^2 - x^2$, $\mathrm{d}u = -2x\ \mathrm{d}x$, with new limits $u = 0$ at $x = -a$ and $u = a^2$ at $x = 0$:

$$\begin{aligned}
\iint_{H_a} (x+y)\ \mathrm{d}A &= \int_{-a}^0 \left[xy + \frac{y^2}{2}\right]_{y=-s_a(x)}^{y=s_a(x)} \mathrm{d}x = \int_{-a}^0 \left(2x\sqrt{a^2 - x^2} + 0\right) \mathrm{d}x \\
&= -\int_0^{a^2} u^{1/2}\ \mathrm{d}u = -\left[\frac{2u^{3/2}}{3}\right]_0^{a^2} = -\frac{2a^3}{3}, \\
\iint_D (x+y)\ \mathrm{d}A &= -\frac{2\cdot 8}{3} - \left(-\frac{2\cdot 1}{3}\right) = -\frac{16}{3} + \frac{2}{3} \implies \boxed{\iint_D (x+y)\ \mathrm{d}A = -\frac{14}{3}.}
\end{aligned}$$

Week 6 application: an offset disk and volume. Find the volume under $z = 4 - x^2 - y^2$ and above

$$D = \{(x,y) : (x-1)^2 + y^2 \le 1\}.$$

Expanding the boundary gives $r^2 \le 2r\cos\theta$. Thus

$$-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}, \qquad 0 \le r \le 2\cos\theta.$$

The volume is

$$\begin{aligned} V &= \int_{-\pi/2}^{\pi/2} \int_0^{2\cos\theta} (4 - r^2) r \, \mathrm{d}r \, \mathrm{d}\theta \\ &= \int_{-\pi/2}^{\pi/2} \left[2r^2 - \frac{r^4}{4} \right]_0^{2\cos\theta} \mathrm{d}\theta \\ &= \int_{-\pi/2}^{\pi/2} \left(8\cos^2\theta - 4\cos^4\theta \right) \mathrm{d}\theta. \end{aligned}$$

Using

$$\int_{-\pi/2}^{\pi/2} \cos^2\theta \, \mathrm{d}\theta = \frac{\pi}{2}, \qquad \int_{-\pi/2}^{\pi/2} \cos^4\theta \, \mathrm{d}\theta = \frac{3\pi}{8},$$

we obtain

$$V = 8\left(\frac{\pi}{2}\right) - 4\left(\frac{3\pi}{8}\right) = \frac{5\pi}{2}.$$

Hence

$$\boxed{V = \frac{5\pi}{2}.}$$

Polar bounds and the Jacobian

After substituting $x = r\cos\theta$ and $y = r\sin\theta$, replace every occurrence of $x^2 + y^2$ by r^2 and multiply the integrand by r. Also check $h_2(\theta) \ge 0$ on the stated interval. For the offset disk above, $2\cos\theta$ is nonnegative only on $[-\pi/2, \pi/2]$.

7.8 Chapter Summary

Key Formula: Chapter 7 recap

For continuous f on a rectangle,

$$\boxed{\iint_{[a,b]\times[c,d]} f \, \mathrm{d}A = \int_a^b \int_c^d f \, \mathrm{d}y \, \mathrm{d}x = \int_c^d \int_a^b f \, \mathrm{d}x \, \mathrm{d}y.}$$

For Type I regions, integrate vertically; for Type II regions, integrate horizontally. For polar regions, use

$$\boxed{\mathrm{d}A = r \, \mathrm{d}r \, \mathrm{d}\theta.}$$

Bounded f is integrable over a Jordan domain exactly when its discontinuities have area zero.

Exam Focus: Chapter 7 takeaways

- Sketch the region before selecting iterated bounds. The inner integral follows the moving slice; the outer integral follows its parameter.
- To reverse order, solve the same inequalities for the other variable and split the domain wherever the slice description changes.
- Use additivity only when overlaps have area zero; shared curves and shared points are harmless, shared two-dimensional pieces are not.
- Use polar coordinates for circles, annuli, sectors, and radial curves. State the angle range, radial range, substituted integrand, and r.
- A continuity hypothesis is enough for every Fubini formula here. For bounded discontinuous functions, use the discontinuity-set criterion.

QRS@NTU — qrsntu.org

8 Topic 8: Midterm Summary

This is a cumulative method sheet for Topics 1–7. It ends at double integrals. Use it after working through the chapter proofs: the purpose here is to choose a method, check its hypotheses, and connect techniques that are often examined in the same question.

8.1 Comparison table: what each question is asking

Signal in the question	**Primary tool**	**Essential check and output**
A curve is given by $\boldsymbol{r}(t)$ and a point or direction is requested.	Differentiate componentwise; use $\boldsymbol{r}'(t_0)$.	Find every t_0 with $\boldsymbol{r}(t_0) = p$. Check $\boldsymbol{r}'(t_0) \neq \mathbf{0}$ before writing $p + s\boldsymbol{r}'(t_0)$.
Does the limit exist near a point?	Paths can disprove; a radial bound or algebra can prove.	Two admissible paths with different values disprove a limit. Matching paths do not prove it. State the domain.
Continuity or a piecewise value at a point.	Test relative-domain continuity; at an accumulation point, evaluate the limit and compare with the stated value.	For an accumulation point $p \in D$, continuity means $\lim_{q \to p,\ q \in D} f(q) = f(p)$. An isolated domain point is automatically continuous.
Local maximum, minimum, or saddle.	Solve $f_x = f_y = 0$, then use the Hessian discriminant.	Also include interior nondifferentiability and any boundary required by the domain. If $D = 0$, use another argument.
Tangent plane, linearization, or error.	Use total differentiability and $L_p(h) = f(p) + \nabla f(p) \cdot h$.	Compute partials at the base point, not at the displaced point. A C^1 hypothesis near p is a sufficient test.
Variables depend on other variables, or an equation defines one variable locally.	Chain rule or implicit-function theorem.	Track evaluation points. For IFT verify $F(p) = 0$, F is C^1 nearby, and the derivative solved against is nonzero.
Rate in a direction, steepest increase, or a level curve normal.	Normalize first, then use $D_u f = \nabla f \cdot u$.	The direction in a directional derivative is unit length. A nonzero gradient is normal to a regular level set.
Optimise subject to $g = 0$.	Lagrange multipliers: $\nabla f = \lambda \nabla g$.	Where f, g are C^1 nearby, check $\nabla g \neq 0$, solve the constraint simultaneously, then compare all candidate values.
Area, mass, average value, or accumulation over a planar region.	Set up a double integral; change order or use polar coordinates when useful.	Sketch the actual region. In polar include r and translate both bounds and integrand.

One map of the midterm tools

$$\boldsymbol{r}(t) \xrightarrow{\mathrm{d}/\mathrm{d}t} \text{tangent direction},$$
$$f(p+h) = f(p) + \nabla f(p)\cdot h + o(\|h\|),$$
$$f \circ G \xrightarrow{\text{chain rule}} J_f(G)J_G,$$
$$f \text{ on } g = 0 \xrightarrow{\text{Lagrange}} \nabla f = \lambda \nabla g,$$
$$\iint_R f(x,y)\,\mathrm{d}A \xrightarrow{x=r\cos\theta,\ y=r\sin\theta} \iint f(r\cos\theta, r\sin\theta)\, r\,\mathrm{d}r\,\mathrm{d}\theta.$$

The same derivative has several interpretations: $\boldsymbol{r}'(t)$ is a curve velocity, $\nabla f(p)$ records first-order change of a scalar field, and $\nabla F(p)$ is a normal to a regular level set $F = 0$.

8.2 Method patterns

Method Pattern 1: parametric curves and tangents

Start by recording the allowed parameter interval. To find a tangent at a geometric point p, solve $\boldsymbol{r}(t) = p$ before differentiating; one image point can have more than one parameter value. At each solution t_0, compute

$$\boldsymbol{v}(t_0) = \boldsymbol{r}'(t_0), \qquad \ell(s) = \boldsymbol{r}(t_0) + s\boldsymbol{v}(t_0).$$

If $\boldsymbol{v}(t_0) = \boldsymbol{0}$, the regular tangent-line formula is unavailable. Inspect one-sided secants, eliminate the parameter, or use a justified implicit description. Chapter 1 gives the regular and singular cases.

Method Pattern 2: limits and continuity

For a finite limit at p, first state the domain near p. To disprove a limit, choose simple admissible paths through p, commonly axes, lines $y = mx$, or a curve adapted to the powers in the expression. Different path limits give a complete disproof. To prove a limit, transform the expression into a bound in $\|(x,y) - p\|$ and invoke the squeeze theorem from Chapter 2. Only after the limit is known should continuity be tested by comparing it with the defined value.

Method Pattern 3: extrema on a stated domain

For an unconstrained interior problem, solve

$$f_x(x,y) = 0, \qquad f_y(x,y) = 0.$$

Assuming f is C^2 near a stationary point, compute

$$D = f_{xx}f_{yy} - f_{xy}^2.$$

Then $D > 0$ with $f_{xx} > 0$ gives a local minimum, $D > 0$ with $f_{xx} < 0$ gives a local maximum, and $D < 0$ gives a saddle. If $D = 0$, the test says nothing. For an absolute-extrema problem on a closed bounded region, include every interior critical point and reduce each boundary piece to a one-variable problem before comparing values; see Chapter 3.

Method Pattern 4: planes and differentiability

At $p = (a, b)$, compute $f(a, b)$, $f_x(a, b)$, and $f_y(a, b)$. The candidate tangent plane to $z = f(x, y)$ is

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

This formal plane is a first-order approximation only when total differentiability holds. Continuous first partials on a neighbourhood are enough. If the function is piecewise or singular, test the remainder

$$\frac{f(p + h) - f(p) - \nabla f(p) \cdot h}{\|h\|} \longrightarrow 0$$

directly. Existence of partial derivatives alone is not enough.

Method Pattern 5: chain rule, IFT, and directions

Draw the dependency arrows. For $z = f(x, y)$ with $x = x(s, t)$ and $y = y(s, t)$, write

$$z_s = f_x x_s + f_y y_s, \qquad z_t = f_x x_t + f_y y_t,$$

with f_x, f_y evaluated at $(x(s, t), y(s, t))$. For $F(x, y, z) = 0$ locally solved as $z = g(x, y)$, verify the IFT hypotheses before using

$$g_x = -\frac{F_x}{F_z}, \qquad g_y = -\frac{F_y}{F_z}.$$

For a given nonzero vector v, set $u = v/\|v\|$ before calculating a directional derivative. Chapter 5 proves each displayed rule.

Method Pattern 6: constrained extrema

Write the objective, the equality constraint, and the full system

$$\nabla f(x, y) = \lambda \nabla g(x, y), \qquad g(x, y) = 0.$$

Use the regular multiplier theorem only where f and g are C^1 nearby and $\nabla g \neq \mathbf{0}$. Do not divide by a component of ∇g before checking whether it can vanish. Solve all branches created by the equations, retain only points on the constraint, and add singular feasible points or domain endpoints that the regular multiplier theorem cannot detect. On a nonempty compact constraint, a continuous objective attains absolute extrema; comparing all these candidates finds them. The geometric reason is that at a regular constrained extremum both gradients are normal to the same tangent direction, as developed in Chapter 6.

Method Pattern 7: double integrals and polar coordinates

Sketch R before choosing bounds. A type-I description has

$$\iint_R f \, \mathrm{d}A = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, \mathrm{d}y \, \mathrm{d}x;$$

a type-II description reverses the slices. Change order when the existing inner antiderivative is awkward. For disks, annuli, sectors, and expressions in $x^2 + y^2$, use polar coordinates and include the area factor:

$$\mathrm{d}A = r \, \mathrm{d}r \, \mathrm{d}\theta.$$

The factor r is part of the change of variables, not part of the integrand. Chapter 7 supplies the geometric derivation and the conditions for the iterated-integral rules.

8.3 Traps that change a correct-looking answer into a wrong one

A path test is one-sided evidence

If two paths give different limiting values, the multivariable limit fails. If ten paths give the same value, the limit may still fail along an eleventh. Use paths to disprove, and use a uniform estimate to prove. For example, testing only $y = mx$ can miss a failure along $y = x^2$.

The Hessian is a local test, not a global comparison

The discriminant classifies a nondegenerate interior stationary point. It does not inspect the boundary, it does not classify $D = 0$, and it does not by itself prove an absolute extremum on an unbounded region. Keep the words local and absolute visible in the solution.

A tangent plane needs a linear approximation

Writing the formal plane from partial derivatives is algebraically easy, but the plane represents first-order behaviour only under differentiability. For a piecewise function, test the remainder or invoke continuous partials on a neighbourhood. Do not infer differentiability merely from f_x and f_y existing at the point.

Keep base, transformed, and direction points separate

In a chain rule, outer partials are evaluated at the transformed point. In an IFT calculation, all F-partials are evaluated at the point on $F = 0$. In a directional derivative, the gradient is evaluated at the base point and the direction is normalized. A useful annotation is

$$(s, t) \mapsto (x(s, t), y(s, t)) \mapsto f(x, y).$$

Put an evaluation point beside every derivative until the calculation is finished.

Lagrange multipliers need the constraint and regularity

Solving $\nabla f = \lambda \nabla g$ without $g = 0$ creates spurious points. At a point with $\nabla g = 0$, the standard multiplier theorem does not apply; inspect that point separately if it belongs to the constraint. A multiplier equation finds candidates, not an answer until objective values are compared.

Polar coordinates change the whole integral

When $x = r\cos\theta$ and $y = r\sin\theta$, change all three pieces: the integrand, the region, and $\mathrm{d}A$. In particular,

$$x^2 + y^2 = r^2, \qquad \mathrm{d}A = r\ \mathrm{d}r\ \mathrm{d}\theta.$$

Forgetting the final r gives the wrong area even when the bounds are correct. Negative r is usually avoidable: use $r \geq 0$ and choose a suitable interval of angles.

8.4 Integrated examples

Integrated Example 1: a parametric curve meets a level curve

Let

$$\boldsymbol{r}(t) = (t^2 - 1, t^3 - t), \qquad F(x, y) = x^2 + y^2.$$

Find the tangent line to the curve at the image point $p = (0, 0)$, and decide whether the curve has a unique regular tangent there.

Method 1: Preimages of p and the velocity vector

Solve $\boldsymbol{r}(t) = (0, 0)$ one component at a time:

$$\begin{aligned} t^2 - 1 = 0 &\iff t = 1 \text{ or } t = -1, \\ t^3 - t = t(t^2 - 1) = 0 &\iff t = 0,\ t = 1 \text{ or } t = -1. \end{aligned}$$

Both components vanish only for $t = \pm 1$, so the curve visits p twice. Differentiate componentwise and evaluate at each visit:

$$\begin{aligned} \boldsymbol{r}'(t) &= \left(\tfrac{\mathrm{d}}{\mathrm{d}t}(t^2 - 1),\ \tfrac{\mathrm{d}}{\mathrm{d}t}(t^3 - t)\right) = (2t,\ 3t^2 - 1), \\ \boldsymbol{r}'(1) &= (2 \cdot 1,\ 3 \cdot 1^2 - 1) = (2, 2) \neq \mathbf{0}, \\ \boldsymbol{r}'(-1) &= (2 \cdot (-1),\ 3 \cdot (-1)^2 - 1) = (-2, 2) \neq \mathbf{0}. \end{aligned}$$

Both visits are regular, so each has the tangent line $\boldsymbol{r}(t_0) + s\boldsymbol{r}'(t_0)$:

$$\begin{aligned} \ell_+(s) &= (0, 0) + s(2, 2) = (2s, 2s) &&\implies y = x, \\ \ell_-(s) &= (0, 0) + s(-2, 2) = (-2s, 2s) &&\implies y = -x. \end{aligned}$$

The velocities are not parallel, since $\det \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix} = 4 + 4 = 8 \neq 0$, so the two lines differ and no single tangent line describes the image at p.

$$\boxed{\text{the image has two regular tangent lines: } y = x \text{ and } y = -x.}$$

Method 2: Eliminating the parameter into two graph branches

From $x = t^2 - 1$ we get $t^2 = x + 1$, and $y = t(t^2 - 1) = tx$. Near $t = 1$ we have $t = \sqrt{x + 1}$, and near $t = -1$ we have $t = -\sqrt{x + 1}$. Hence near p the image is the union of the two graphs

$$y_+(x) = x\sqrt{x + 1}, \qquad y_-(x) = -x\sqrt{x + 1}, \qquad x > -1,$$

and $y_\pm(0) = 0$. Differentiate with the product rule:

$$\begin{aligned} y'_+(x) &= \sqrt{x + 1} + \frac{x}{2\sqrt{x + 1}}, &&y'_+(0) = 1 + 0 = 1, \\ y'_-(x) &= -\sqrt{x + 1} - \frac{x}{2\sqrt{x + 1}}, &&y'_-(0) = -1 - 0 = -1. \end{aligned}$$

Each branch is a differentiable graph through $(0, 0)$, so its tangent line is $y = 0 + y'_\pm(0)(x - 0)$, that is $y = x$ and $y = -x$; the two slopes differ.

$$\boxed{\text{the image has two regular tangent lines: } y = x \text{ and } y = -x.}$$

Integrated Example 2: a limit, continuity, and a tangent plane

Define

$$f(x,y)=\begin{cases}\dfrac{x^2y}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Determine continuity and differentiability at the origin, then give the tangent plane if it exists.

Method 1: Squeeze theorem, difference quotients, and two remainder paths

Continuity. For $(x,y)\neq(0,0)$, $0\le x^2\le x^2+y^2$ gives $x^2/(x^2+y^2)\le 1$, so

$$|f(x,y)-f(0,0)|=|y|\,\frac{x^2}{x^2+y^2}\le|y|\le\sqrt{x^2+y^2}.$$

Since $\sqrt{x^2+y^2}\to 0$, the squeeze theorem gives $\lim_{(x,y)\to(0,0)} f(x,y)=0=f(0,0)$. Partials: for $h\neq 0$, $f(h,0)=h^2\cdot 0/h^2=0$ and $f(0,h)=0\cdot h/h^2=0$, so

$$f_x(0,0)=\lim_{h\to0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to0}\frac{0-0}{h}=0,\qquad f_y(0,0)=\lim_{h\to0}\frac{f(0,h)-f(0,0)}{h}=0.$$

With $\nabla f(0,0)=\mathbf{0}$, f is differentiable at $(0,0)$ if and only if $E(h,k)\to 0$ as $(h,k)\to(0,0)$, where E is the remainder quotient; test it along $k=h>0$:

$$E(h,k)=\frac{f(h,k)-f(0,0)-0\cdot h-0\cdot k}{\sqrt{h^2+k^2}}=\frac{h^2k}{(h^2+k^2)^{3/2}},$$
$$E(h,h)=\frac{h^2\cdot h}{(2h^2)^{3/2}}=\frac{h^3}{2\sqrt{2}\,h^3}=\frac{1}{2\sqrt{2}}\longrightarrow\frac{1}{2\sqrt{2}}\neq 0.$$

Along $k=0$, $E(h,0)=0\to 0$. Two paths give different limits, so $\lim E$ does not exist and $E\not\to 0$.

f is continuous at $(0,0)$ but not differentiable there;
the formal plane $z=0$ is not a first-order approximation.

Method 2: AM–GM bound and linearity of directional derivatives

Since $2|xy|\le x^2+y^2$, $|f(x,y)|=|x|\,\dfrac{|xy|}{x^2+y^2}\le\dfrac{|x|}{2}\to 0$, so f is continuous at $(0,0)$. For a unit vector $u=(a,b)$, $f(sa,sb)=\dfrac{s^3a^2b}{s^2(a^2+b^2)}=sa^2b$, so

$$D_uf(0,0)=\lim_{s\to0}\frac{f(sa,sb)-f(0,0)}{s}=\lim_{s\to0}\frac{sa^2b}{s}=a^2b.$$

If f were differentiable at $(0,0)$, then $D_uf(0,0)=\nabla f(0,0)\cdot u=a\,D_{(1,0)}f(0,0)+b\,D_{(0,1)}f(0,0)=a\cdot 0+b\cdot 0=0$ for every u. But $u=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$ gives $a^2b=\frac{1}{2\sqrt{2}}\neq 0$, a contradiction.

f is continuous at $(0,0)$ but not differentiable there;
the formal plane $z=0$ is not a first-order approximation.

Integrated Example 3: local classification and absolute extrema

Let

$$f(x,y) = x^3 + y^2 - 3x$$

on the rectangle $R = [-2,2] \times [-1,1]$. Find all absolute extrema on R.

Method 1: Interior critical points, Hessian test, and boundary edges

R is closed and bounded and f is a polynomial, so by the Extreme Value Theorem both absolute extrema exist; they occur at interior critical points or on the boundary. Interior, $(x,y) \in (-2,2) \times (-1,1)$:

$$\begin{aligned} f_x &= 3x^2 - 3 = 0 \iff x^2 = 1 \iff x = \pm 1 \in (-2,2), \\ f_y &= 2y = 0 \iff y = 0 \in (-1,1), \\ f_{xx} &= 6x, \quad f_{yy} = 2, \quad f_{xy} = \tfrac{\partial}{\partial y}(3x^2 - 3) = 0, \quad D = 6x \cdot 2 - 0^2 = 12x. \end{aligned}$$

$$\begin{aligned} (1,0): &\quad D = 12 > 0,\ f_{xx} = 6 > 0 \text{ (local minimum)}, && f(1,0) = 1 + 0 - 3 = -2, \\ (-1,0): &\quad D = -12 < 0 \text{ (saddle)}, && f(-1,0) = -1 + 0 + 3 = 2. \end{aligned}$$

Boundary of R, each edge as a one-variable problem on a closed interval:

$$\begin{aligned} y = \pm 1: &\quad f(x,\pm 1) = x^3 - 3x + 1, \quad \tfrac{\mathrm{d}}{\mathrm{d}x}(x^3 - 3x + 1) = 3x^2 - 3 = 0 \iff x = \pm 1; \\ &\quad x = 1:\ 1 - 3 + 1 = -1, \quad x = -1:\ -1 + 3 + 1 = 3, \\ &\quad x = 2:\ 8 - 6 + 1 = 3, \quad x = -2:\ -8 + 6 + 1 = -1. \\ x = 2: &\quad f(2,y) = 8 + y^2 - 6 = 2 + y^2, \quad \tfrac{\mathrm{d}}{\mathrm{d}y}(2 + y^2) = 2y = 0 \iff y = 0; \\ &\quad y = 0:\ 2, \quad y = \pm 1:\ 3. \\ x = -2: &\quad f(-2,y) = -8 + y^2 + 6 = -2 + y^2, \quad 2y = 0 \iff y = 0; \\ &\quad y = 0:\ -2, \quad y = \pm 1:\ -1. \end{aligned}$$

The complete candidate list:

point	f	point	f
$(1,0)$	-2	$(-1,0)$	2
$(2,0)$	2	$(-2,0)$	-2
$(1,\pm 1)$	-1	$(-1,\pm 1)$	3
$(2,\pm 1)$	3	$(-2,\pm 1)$	-1

The smallest value is -2 and the largest is 3:

$$\boxed{\min_R f = -2 \text{ at } (1,0),(-2,0); \qquad \max_R f = 3 \text{ at } (-1,\pm 1),(2,\pm 1).}$$

Method 2: Separation into a sum of one-variable functions

Write $f(x,y) = \varphi(x) + \psi(y)$ with $\varphi(x) = x^3 - 3x$ on $[-2,2]$ and $\psi(y) = y^2$ on $[-1,1]$. On a rectangle x and y vary independently, so $\varphi(x) + \psi(y) \geq \min \varphi + \min \psi$, with equality exactly when x minimises φ and y minimises ψ; the same holds for maxima. For φ: $\varphi'(x) = 3x^2 - 3 = 0 \iff x = \pm 1$, and

$$\varphi(-2) = -8 + 6 = -2, \quad \varphi(-1) = -1 + 3 = 2, \quad \varphi(1) = 1 - 3 = -2, \quad \varphi(2) = 8 - 6 = 2,$$

so $\min \varphi = -2$ at $x \in \{-2, 1\}$ and $\max \varphi = 2$ at $x \in \{-1, 2\}$. For ψ: $\min \psi = 0$ at $y = 0$ and $\max \psi = 1$ at $y = \pm 1$. Hence $\min f = -2 + 0 = -2$ and $\max f = 2 + 1 = 3$:

$$\boxed{\min_R f = -2 \text{ at } (1,0),(-2,0); \qquad \max_R f = 3 \text{ at } (-1,\pm 1),(2,\pm 1).}$$

Integrated Example 4: IFT, tangent plane, and directional change

The surface

$$F(x,y,z) = x^2 + y^2 + z^2 - 9 = 0$$

passes through $p = (1,2,2)$. Treat it locally as $z = g(x,y)$, find the tangent plane, and find the directional derivative of g at $(1,2)$ in the direction from $(1,2)$ to $(4,6)$.

Method 1: Implicit Function Theorem and the gradient of g

Partials of F: $F_x = 2x$, $F_y = 2y$, $F_z = 2z$. F is a polynomial, hence C^1; $F(1,2,2) = 1+4+4-9 = 0$; and $F_z(1,2,2) = 4 \neq 0$. By the Implicit Function Theorem there is a C^1 function g near $(1,2)$ with $g(1,2) = 2$ and $F(x,y,g(x,y)) = 0$. Differentiating this identity by the Chain Rule:

$$F_x + F_z\, g_x = 0 \implies g_x = -\frac{F_x}{F_z} = -\frac{2x}{2z} = -\frac{x}{z}, \qquad g_x(1,2) = -\frac{1}{2},$$
$$F_y + F_z\, g_y = 0 \implies g_y = -\frac{F_y}{F_z} = -\frac{2y}{2z} = -\frac{y}{z}, \qquad g_y(1,2) = -\frac{2}{2} = -1.$$

Tangent plane $z = g(1,2) + g_x(1,2)(x-1) + g_y(1,2)(y-2)$:

$$z = 2 - \frac{1}{2}(x-1) - 1\cdot(y-2).$$

Direction: $v = (4,6)-(1,2) = (3,4)$, $\|v\| = \sqrt{9+16} = 5$, $u = \left(\frac{3}{5}, \frac{4}{5}\right)$. Since g is C^1, it is differentiable and

$$D_u g(1,2) = \nabla g(1,2)\cdot u = -\frac{1}{2}\cdot\frac{3}{5} - 1\cdot\frac{4}{5} = -\frac{3}{10} - \frac{8}{10} = -\frac{11}{10}.$$

$$\boxed{z - 2 = -\frac{1}{2}(x-1) - (y-2), \qquad D_u g(1,2) = -\frac{11}{10}.}$$

Method 2: Level-surface normal and restriction to the line

$\nabla F(1,2,2) = (2,4,4) \neq \mathbf{0}$ is normal to the level surface $F = 0$ at p, so

$$2(x-1) + 4(y-2) + 4(z-2) = 0 \iff z - 2 = -\frac{1}{2}(x-1) - (y-2).$$

For the rate, solve explicitly: $z^2 = 9 - x^2 - y^2$ and $z = 2 > 0$ at p, so near $(1,2)$, $g(x,y) = \sqrt{9 - x^2 - y^2}$. Restrict g to the line $(1,2) + tu = \left(1 + \frac{3t}{5},\, 2 + \frac{4t}{5}\right)$:

$$9 - \left(1 + \tfrac{3t}{5}\right)^2 - \left(2 + \tfrac{4t}{5}\right)^2 = 9 - \left(1 + \tfrac{6t}{5} + \tfrac{9t^2}{25}\right) - \left(4 + \tfrac{16t}{5} + \tfrac{16t^2}{25}\right) = 4 - \tfrac{22t}{5} - t^2,$$
$$\phi(t) = \sqrt{4 - \tfrac{22t}{5} - t^2}, \qquad \phi'(t) = \frac{-\frac{22}{5} - 2t}{2\sqrt{4 - \frac{22t}{5} - t^2}}, \qquad \phi'(0) = \frac{-\frac{22}{5}}{2\cdot 2} = -\frac{11}{10}.$$

By definition $D_u g(1,2) = \phi'(0)$.

$$\boxed{z - 2 = -\frac{1}{2}(x-1) - (y-2), \qquad D_u g(1,2) = -\frac{11}{10}.}$$

Integrated Example 5: a constrained extremum and gradient geometry

Find the absolute maximum and minimum of

$$f(x,y) = xy$$

subject to the elliptical constraint

$$g(x,y) = x^2 + 4y^2 - 4 = 0.$$

Method 1: Lagrange multipliers with every branch closed
The ellipse $g^{-1}(0)$ is closed, bounded ($x^2 \le 4$, $y^2 \le 1$) and nonempty, and f is continuous, so extrema exist (Extreme Value Theorem). Partials: $\nabla f = (y,x)$, $\nabla g = (2x, 8y)$, both C^1; $\nabla g = \mathbf{0}$ only at $(0,0)$, where $g = -4 \neq 0$, so $\nabla g \neq \mathbf{0}$ on $g = 0$ and $\nabla f = \lambda \nabla g$ applies: (i) $y = 2\lambda x$, (ii) $x = 8\lambda y$, (iii) $x^2 + 4y^2 = 4$. If $x = 0$, (i) gives $y = 0$; if $y = 0$, (ii) gives $x = 0$; either way $(0,0)$ fails (iii) since $0 \neq 4$. Hence $x, y \neq 0$, and multiplying (i) and (ii) gives $xy = 16\lambda^2 xy$, so $16\lambda^2 = 1$ and $\lambda = \pm\frac{1}{4}$. The candidates, whose values we compare, are:

$$\begin{aligned} \lambda = \tfrac{1}{4}: &\quad y = \tfrac{x}{2}, \quad x^2 + 4\cdot\tfrac{x^2}{4} = 2x^2 = 4 \implies x = \pm\sqrt{2},\ y = \pm\tfrac{1}{\sqrt{2}}, &\quad f = x\cdot\tfrac{x}{2} = \tfrac{2}{2} = 1, \\ \lambda = -\tfrac{1}{4}: &\quad y = -\tfrac{x}{2}, \quad 2x^2 = 4 \implies x = \pm\sqrt{2},\ y = \mp\tfrac{1}{\sqrt{2}}, &\quad f = -\tfrac{x^2}{2} = -1. \end{aligned}$$

$$\boxed{\max xy = 1 \text{ at } (\sqrt{2}, 1/\sqrt{2}), (-\sqrt{2}, -1/\sqrt{2});}$$
$$\boxed{\min xy = -1 \text{ at } (\sqrt{2}, -1/\sqrt{2}), (-\sqrt{2}, 1/\sqrt{2}).}$$

Method 2: Parametrising the constraint
Since $(x/2)^2 + y^2 = 1$, the map $\theta \mapsto (2\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$, traces the ellipse exactly once, and

$$h(\theta) = f(2\cos\theta, \sin\theta) = 2\sin\theta\cos\theta = \sin 2\theta, \qquad h'(\theta) = 2\cos 2\theta = 0 \iff \theta \in \left\{\tfrac{\pi}{4}, \tfrac{3\pi}{4}, \tfrac{5\pi}{4}, \tfrac{7\pi}{4}\right\}.$$

Since $-1 \le \sin 2\theta \le 1$, the critical values ± 1 below are the absolute extrema:

$$\begin{aligned} \theta = \tfrac{\pi}{4} &: (\sqrt{2}, \tfrac{1}{\sqrt{2}}),\ h = \sin\tfrac{\pi}{2} = 1; &\quad \theta = \tfrac{5\pi}{4} &: (-\sqrt{2}, -\tfrac{1}{\sqrt{2}}),\ h = \sin\tfrac{5\pi}{2} = 1; \\ \theta = \tfrac{3\pi}{4} &: (-\sqrt{2}, \tfrac{1}{\sqrt{2}}),\ h = \sin\tfrac{3\pi}{2} = -1; &\quad \theta = \tfrac{7\pi}{4} &: (\sqrt{2}, -\tfrac{1}{\sqrt{2}}),\ h = \sin\tfrac{7\pi}{2} = -1. \end{aligned}$$

$$\boxed{\max xy = 1 \text{ at } (\sqrt{2}, 1/\sqrt{2}), (-\sqrt{2}, -1/\sqrt{2});}$$
$$\boxed{\min xy = -1 \text{ at } (\sqrt{2}, -1/\sqrt{2}), (-\sqrt{2}, 1/\sqrt{2}).}$$

Method 3: Completing the square on the constraint
On $x^2 + 4y^2 = 4$, $(x - 2y)^2 = x^2 + 4y^2 - 4xy = 4 - 4xy \ge 0$ gives $xy \le 1$, with equality iff $x = 2y$; then $4y^2 + 4y^2 = 4$, $y = \pm\frac{1}{\sqrt{2}}$, $(x,y) = \pm(\sqrt{2}, \frac{1}{\sqrt{2}})$. Likewise $(x+2y)^2 = x^2 + 4y^2 + 4xy = 4 + 4xy \ge 0$ gives $xy \ge -1$, with equality iff $x = -2y$; then $8y^2 = 4$ and $(x,y) = \pm(\sqrt{2}, -\frac{1}{\sqrt{2}})$.

$$\boxed{\max xy = 1 \text{ at } (\sqrt{2}, 1/\sqrt{2}), (-\sqrt{2}, -1/\sqrt{2});}$$
$$\boxed{\min xy = -1 \text{ at } (\sqrt{2}, -1/\sqrt{2}), (-\sqrt{2}, 1/\sqrt{2}).}$$

Integrated Example 6: a double integral over a polar region

Let R be the upper semicircular annulus

$$R = \{(x, y) : 1 \le x^2 + y^2 \le 4,\ y \ge 0\}.$$

Find its area and the integral of $x^2 + y^2$ over R.

Method 1: Polar coordinates

Put $x = r\cos\theta$, $y = r\sin\theta$, $r \ge 0$. Then $1 \le x^2 + y^2 \le 4 \iff 1 \le r \le 2$, and for $r > 0$, $y \ge 0 \iff \sin\theta \ge 0 \iff 0 \le \theta \le \pi$. So $R = \{1 \le r \le 2,\ 0 \le \theta \le \pi\}$, and the Jacobian is

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r\cos^2\theta + r\sin^2\theta = r, \qquad \mathrm{d}A = r\ \mathrm{d}r\ \mathrm{d}\theta.$$

Both integrands below are continuous on the rectangle $[1, 2] \times [0, \pi]$, so Fubini's theorem allows either order:

$$\mathrm{Area}(R) = \int_0^\pi \int_1^2 r\ \mathrm{d}r\ \mathrm{d}\theta = \int_0^\pi \left[\frac{r^2}{2}\right]_1^2 \mathrm{d}\theta = \int_0^\pi \left(2 - \frac{1}{2}\right) \mathrm{d}\theta = \frac{3}{2}\pi,$$

$$\iint_R (x^2 + y^2)\ \mathrm{d}A = \int_0^\pi \int_1^2 r^2 \cdot r\ \mathrm{d}r\ \mathrm{d}\theta = \int_0^\pi \left[\frac{r^4}{4}\right]_1^2 \mathrm{d}\theta = \int_0^\pi \left(4 - \frac{1}{4}\right) \mathrm{d}\theta = \frac{15}{4}\pi.$$

$$\boxed{\mathrm{Area}(R) = \frac{3\pi}{2}, \qquad \iint_R (x^2 + y^2)\ \mathrm{d}A = \frac{15\pi}{4}.}$$

Method 2: Cartesian half-disks and trigonometric substitution

For $a > 0$ let $H_a = \{x^2 + y^2 \le a^2,\ y \ge 0\} = \{-a \le x \le a,\ 0 \le y \le \sqrt{a^2 - x^2}\}$ (type I, Fubini), and $I(a) = \iint_{H_a} (x^2 + y^2)\ \mathrm{d}A$. R is H_2 minus H_1 up to the arc $x^2 + y^2 = 1$, which has zero area, so every integral over R is the H_2 value minus the H_1 value. Substitute $x = a\sin\varphi$, $\mathrm{d}x = a\cos\varphi\ \mathrm{d}\varphi$, $\varphi : -\frac{\pi}{2} \to \frac{\pi}{2}$, $\sqrt{a^2 - x^2} = a\cos\varphi$ (as $\cos\varphi \ge 0$), and use $\cos^2\varphi = \frac{1+\cos 2\varphi}{2}$, $\sin^2\varphi\cos^2\varphi = \frac{1-\cos 4\varphi}{8}$, $\cos^4\varphi = \frac{3}{8} + \frac{\cos 2\varphi}{2} + \frac{\cos 4\varphi}{8}$; each $\cos k\varphi$ term integrates to $\frac{1}{k}[\sin k\varphi]_{-\pi/2}^{\pi/2} = 0$ for $k = 2, 4$:

$$\mathrm{Area}(H_a) = \int_{-a}^{a} \sqrt{a^2 - x^2}\ \mathrm{d}x = \int_{-\pi/2}^{\pi/2} a\cos\varphi \cdot a\cos\varphi\ \mathrm{d}\varphi = a^2 \int_{-\pi/2}^{\pi/2} \frac{1 + \cos 2\varphi}{2}\ \mathrm{d}\varphi = \frac{\pi a^2}{2},$$

$$\begin{aligned} I(a) &= \int_{-a}^{a} \left[x^2 y + \frac{y^3}{3}\right]_0^{\sqrt{a^2 - x^2}} \mathrm{d}x = \int_{-a}^{a} \left(x^2\sqrt{a^2 - x^2} + \frac{(a^2 - x^2)^{3/2}}{3}\right) \mathrm{d}x \\ &= \int_{-\pi/2}^{\pi/2} \left(a^2\sin^2\varphi \cdot a\cos\varphi + \frac{a^3\cos^3\varphi}{3}\right) a\cos\varphi\ \mathrm{d}\varphi \\ &= a^4 \int_{-\pi/2}^{\pi/2} \left(\sin^2\varphi\cos^2\varphi + \frac{\cos^4\varphi}{3}\right) \mathrm{d}\varphi = a^4 \left(\frac{\pi}{8} + \frac{1}{3} \cdot \frac{3\pi}{8}\right) = \frac{\pi a^4}{4}, \end{aligned}$$

$$\mathrm{Area}(R) = \frac{\pi \cdot 2^2}{2} - \frac{\pi \cdot 1^2}{2} = \frac{3\pi}{2}, \qquad I(2) - I(1) = \frac{\pi \cdot 2^4}{4} - \frac{\pi \cdot 1^4}{4} = \frac{15\pi}{4}.$$

$$\boxed{\mathrm{Area}(R) = \frac{3\pi}{2}, \qquad \iint_R (x^2 + y^2)\ \mathrm{d}A = \frac{15\pi}{4}.}$$

8.5 Checklist before handing in a solution

Two-minute audit

1. Did I state the domain, region, parameter interval, or constraint?
2. Did I identify every candidate: all parameter preimages, all critical points, all relevant boundary points, or all multiplier branches?
3. Did I check the theorem hypotheses before using it: regular velocity, limit estimate, differentiability, nonzero IFT derivative, regular constraint, or valid polar bounds?
4. Did I keep evaluation points correct in every partial derivative and chain-rule term?
5. Did I normalize a direction vector before calling it a directional derivative?
6. Did I distinguish a local claim from an absolute claim?
7. Did I include the polar Jacobian r and integrate over the actual region rather than its bounding rectangle or disk?
8. Did I give a final answer with the requested object and units or geometric interpretation where appropriate?

Minimal formulas to rederive quickly

$$\boldsymbol{v} = \boldsymbol{r}', \quad \text{speed} = \|\boldsymbol{r}'\|, \quad \ell(s) = \boldsymbol{r}(t_0) + s\boldsymbol{r}'(t_0),$$
$$D = f_{xx}f_{yy} - f_{xy}^2,$$
$$L_p(h) = f(p) + \nabla f(p) \cdot h,$$
$$J_{F\circ G}(p) = J_F(G(p))J_G(p),$$
$$D_u f(p) = \nabla f(p) \cdot u \quad (\|u\| = 1),$$
$$\nabla f = \lambda \nabla g, \quad g = 0,$$
$$\iint_R f(x,y)\ \mathrm{d}A = \int\!\!\int f(r\cos\theta, r\sin\theta) r\ \mathrm{d}r\ \mathrm{d}\theta.$$

Each formula is conditional. The conditions are usually where a short solution earns or loses the decisive marks.

8.6 Memory sheet

Read the question, then say the trigger aloud

If I see ...	**I immediately think ...**
curve $\boldsymbol{r}(t)$ at a point	solve for t first; velocity gives direction only when nonzero.
limit of a quotient in x, y	paths may kill it; a radial bound proves it.
piecewise formula at one point	compare the nearby limit with the defined value; then test the linear remainder if needed.
local extrema	critical points, Hessian, and $D = 0$ means investigate.
closed bounded region	EVT: interior plus every boundary piece, then compare values.
surface $F = 0$	gradient is normal; IFT needs the solved derivative nonzero.
steepest or normal	gradient; normalize any supplied direction.
subject to	objective and constraint together; multipliers give candidates.
circle, sector, $x^2 + y^2$	polar bounds and the factor r.

The compact narrative of the course is:

parameter or region $\longrightarrow$ setup and hypothesis check $\longrightarrow$ derivative or integral calculation $\longrightarrow$ interpreted conclusion

When unsure, return to that order. It prevents most omitted cases before any calculation begins.

QRS@NTU – qrs.ntu.edu.sg

A Formula & Key Results Reference

Use this as a setup and hypothesis check, not as a substitute for describing the domain, region, or constraint in a solution. The scope ends at double integrals in polar coordinates.

Topic 1: Parametric Curves

- A parametrized curve is a continuous map $\boldsymbol{r} : I \to \mathbb{R}^n$ on an interval, $\boldsymbol{r}(t) = (x_1(t), \ldots, x_n(t))$. Different parametrizations can have the same image and may trace it with different directions or multiplicities.
- Differentiate componentwise: $\boldsymbol{r}'(t) = (x_1'(t), \ldots, x_n'(t))$. For motion, $\boldsymbol{v} = \boldsymbol{r}'$, speed $= \|\boldsymbol{v}\|$, and $\boldsymbol{a} = \boldsymbol{r}''$.
- t_0 is *regular* when $\boldsymbol{r}'(t_0) \neq \boldsymbol{0}$. It is singular when $\boldsymbol{r}'(t_0) = \boldsymbol{0}$ or the derivative does not exist. At a regular point $P = \boldsymbol{r}(t_0)$, the tangent line is

$$\boxed{\boldsymbol{\ell}(s) = P + s\boldsymbol{r}'(t_0), \qquad s \in \mathbb{R}.}$$

- At a singular parameter, do not automatically use $\boldsymbol{r}'(t_0)$ as a direction. Inspect one-sided secant directions or a local implicit description; a cusp or self-intersection can require separate tangent lines. At regular parameters the unit tangent is $\boldsymbol{T} = \boldsymbol{r}' / \|\boldsymbol{r}'\|$.
- For nonzero $\boldsymbol{a}, \boldsymbol{b}$, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos\theta$, and $\boldsymbol{a} \perp \boldsymbol{b}$ iff $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$.

Topic 2: Limits, Continuity, and Partial Derivatives

- If $f : D \to \mathbb{R}^n$ and a is a limit point of $D \subseteq \mathbb{R}^m$, then

$$\lim_{x \to a,\, x \in D} f(x) = L \iff (\forall \varepsilon > 0)(\exists \delta > 0) \quad x \in D, \quad 0 < \|x - a\| < \delta \Rightarrow \|f(x) - L\| < \varepsilon.$$

  The limit ignores the value at a, but every tested path must remain in the domain. Two admissible paths with different limits prove non-existence.
- At a limit point $a \in D$, f is continuous iff $\lim_{x \to a,\, x \in D} f(x) = f(a)$. Sums, products, quotients with a nonzero limiting denominator, and compositions preserve continuity. Every function is continuous at each isolated point of its domain.
- For an interior point (a, b),

$$f_x(a,b) = \lim_{h \to 0} \frac{f(a+h, b) - f(a,b)}{h}, \qquad f_y(a,b) = \lim_{h \to 0} \frac{f(a, b+h) - f(a,b)}{h}.$$

  Hold the other variable fixed. Existence of partial derivatives alone does not imply continuity or total differentiability.
- If the second partials exist in a neighbourhood of (a, b) and f_{xy}, f_{yx} are continuous there, then $f_{xy}(a,b) = f_{yx}(a,b)$.

Topic 3: Extreme Points

- An interior local extremum of a differentiable f is stationary: $\nabla f(p) = \boldsymbol{0}$. A critical point also includes an interior point where a required first partial derivative does not exist.
- At a stationary point (a, b), assume $f \in C^2$ on a neighbourhood and set

$$H = \operatorname{Hess} f(a,b) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}, \qquad \Delta = f_{xx} f_{yy} - f_{xy}^2.$$

  Then $\Delta > 0, f_{xx} > 0$ gives a strict local minimum; $\Delta > 0, f_{xx} < 0$ a strict local maximum; $\Delta < 0$ a saddle; and $\Delta = 0$ is inconclusive.
- For absolute extrema on a closed bounded region, list interior critical points and analyse every boundary piece (including endpoints), then compare values. A continuous function on a nonempty compact set attains both an absolute maximum and an absolute minimum.

Topic 4: Tangent Spaces and Differentiability

- For $z = f(x, y)$, if $f_x(a, b), f_y(a, b)$ exist, the candidate tangent plane at $(a, b, f(a, b))$ is

$$\boxed{z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).}$$

  It is a genuine first-order approximation when f is differentiable.

- For a graph $z = f(x, y)$, the coordinate slice directions are $(1, 0, f_x)$ and $(0, 1, f_y)$, with cross-product normal $(-f_x, -f_y, 1)$. In general, a plane through P with independent directions u, v has equation

$$(u \times v) \cdot (X - P) = 0, \qquad X = P + su + tv.$$

  Dependent directions do not determine a unique plane.

- $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at p iff a linear map $Df(p)$ satisfies

$$f(p + h) = f(p) + Df(p)h + o(\|h\|).$$

  For a scalar f with partials, $Df(p)h = \nabla f(p) \cdot h$. Differentiability implies continuity and existence of all partials.

- A sufficient test is $f \in C^1$ on an open neighbourhood of p: all first partial derivatives exist and are continuous there. It is not a necessary condition, and existing partials at one point are insufficient.

- The first-order change is $\Delta f \approx f_x \Delta x + f_y \Delta y$. For input error bounds E_x, E_y, the linearized absolute error is at most $|f_x|E_x + |f_y|E_y$; an exact error bound also needs remainder control.

Topic 5: Chain Rule and Directional Derivatives

- If $z = f(x, y)$, $x = x(t)$, and $y = y(t)$, with the required differentiability, then

$$\boxed{\frac{\mathrm{d}z}{\mathrm{d}t} = f_x \frac{\mathrm{d}x}{\mathrm{d}t} + f_y \frac{\mathrm{d}y}{\mathrm{d}t}.}$$

  In general, for differentiable maps F and G, $D(F \circ G)(p) = DF(G(p))DG(p)$.

- If $F(a, b) = 0$, $F \in C^1$ near (a, b), and $F_y(a, b) \neq 0$, then locally $y = y(x)$ and

$$\boxed{\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{F_x}{F_y}.}$$

  If $F_y = 0$, this formula is unavailable; solve for another variable only when the corresponding nonzero partial hypothesis holds. For $F(a, b, c) = 0$, $F \in C^1$ nearby, and $F_z(a, b, c) \neq 0$, the local graph $z = g(x, y)$ satisfies $g_x = -F_x/F_z$ and $g_y = -F_y/F_z$.

- For C^2 composite functions $z = f(x(s, t), y(s, t))$,

$$z_{st} = f_{xx} x_s x_t + f_{xy}(x_s y_t + y_s x_t) + f_{yy} y_s y_t + f_x x_{st} + f_y y_{st}.$$

  Evaluate outer partials at $(x(s, t), y(s, t))$ and inner ones at (s, t).

- For a unit vector $\boldsymbol{u}$,

$$D_{\boldsymbol{u}} f(p) = \lim_{t \to 0} \frac{f(p + t\boldsymbol{u}) - f(p)}{t}.$$

  If f is differentiable at p, then $D_{\boldsymbol{u}} f(p) = \nabla f(p) \cdot \boldsymbol{u}$. The maximum and minimum unit directional rates are $\pm \|\nabla f(p)\|$, attained in the directions $\pm \nabla f(p) / \|\nabla f(p)\|$ when $\nabla f(p) \neq 0$.

- On a regular level curve $f(x, y) = c$, $\nabla f(p) \neq \mathbf{0}$ is normal and a tangent direction $\boldsymbol{v}$ obeys $\nabla f(p) \cdot \boldsymbol{v} = 0$. At $\nabla f(p) = \mathbf{0}$, this regular-level-curve conclusion needs separate analysis.

Topic 6: Lagrange Multipliers

- For a C^1 objective f and one C^1 equality constraint $g(x) = k$, a constrained local extremum at a regular point p, where $\nabla g(p) \neq \mathbf{0}$, satisfies

$$\boxed{\nabla f(p) = \lambda \nabla g(p), \qquad g(p) = k.}$$

  These equations give necessary candidates, not automatic extrema.

- For two C^1 constraints in $\mathbb{R}^3$ (put $m = 2$ below), apply the regular condition that $\nabla g_1(p), \ldots, \nabla g_m(p)$ are linearly independent, then solve

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \cdots + \lambda_m \nabla g_m(p), \qquad g_i(p) = k_i \quad (1 \le i \le m).$$

- Include feasible singular points, where the constraint gradients fail the stated regularity condition, by direct analysis. For an absolute problem on a compact feasible set with continuous f, compare all regular multiplier candidates, singular candidates, and any boundary candidates.

Topic 7: Double Integrals and Polar Coordinates

- On $R = [a, b] \times [c, d]$, a partition and tags give

$$S(f; P, *) = \sum_{i,j} f(x_{ij}^*, y_{ij}^*)\, \Delta x_i \Delta y_j, \qquad \iint_R f \, \mathrm{d}A = \lim_{\|P\| \to 0} S(f; P, *),$$

  when the common limit exists. For bounded f, equivalently: for every $\varepsilon > 0$, some partition has $U(f, P) - L(f, P) < \varepsilon$.

- Continuous functions on a closed rectangle are Riemann integrable. For a Jordan domain D (bounded, with boundary of area zero), define $\iint_D f \, \mathrm{d}A$ by extending f by zero to an enclosing rectangle. If f is bounded on D, then

$$\boxed{f \text{ is Riemann integrable over } D \iff \text{its relative discontinuity set has Lebesgue measure zero.}}$$

  This criterion requires boundedness and the Jordan-domain hypothesis; boundedness alone is not enough.

- If f is continuous on $[a, b] \times [c, d]$, then

$$\iint_R f \, \mathrm{d}A = \int_a^b \int_c^d f(x, y) \, \mathrm{d}y \, \mathrm{d}x = \int_c^d \int_a^b f(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

  For Type I $D = \{(x, y) : a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$, use $\int_a^b \int_{g_1(x)}^{g_2(x)} f \, \mathrm{d}y \, \mathrm{d}x$; for Type II, reverse the variables. Require continuous bounds and continuous f for these course formulas. Sketch and split the region when its slice bounds change.

- For integrable functions, integration is linear. If $D = D_1 \cup \cdots \cup D_N$ is a finite union of Jordan domains, every two distinct pieces overlap only on a set of area zero, and f is integrable on every D_i, then $\iint_D f = \sum_{i=1}^N \iint_{D_i} f$. Area is $\iint_D 1$, mass is $\iint_D \rho$ for an integrable density $\rho \ge 0$, and volume under a nonnegative integrable height f is $\iint_D f$. If f is integrable and $\mathrm{Area}(D) > 0$, then $f_{\mathrm{avg}} = \iint_D f / \mathrm{Area}(D)$.

- A normalized joint probability density is $p : \mathbb{R}^2 \to \mathbb{R}$ with $p \ge 0$ and improper integral $\iint_{\mathbb{R}^2} p \, \mathrm{d}A = 1$. For a bounded region E for which the integral exists, $\mathrm{Prob}((X, Y) \in E) = \iint_E p \, \mathrm{d}A \le 1$.

- Polar coordinates are $x = r\cos\theta$, $y = r\sin\theta$, with

$$\boxed{\mathrm{d}A = r \, \mathrm{d}r \, \mathrm{d}\theta.}$$

  For continuous f on a polar region $\alpha \le \theta \le \beta$, $h_1(\theta) \le r \le h_2(\theta)$, where h_1, h_2 are continuous, nonnegative, and $h_1 \le h_2$, take an angle interval of length at most 2π that covers points once except on boundary rays, and use

$$\iint_D f \, \mathrm{d}A = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta)\, r \, \mathrm{d}r \, \mathrm{d}\theta.$$

B Notation

This reference fixes the notation used in Topics 1–7. It records meanings and conventions; hypotheses for results remain in the relevant topic.

Vectors, curves, and sets (Topics 1–2)

Notation	Meaning	Use or convention
$\mathbb{R}$, $\mathbb{R}^n$	Real numbers; Euclidean n-space.	A point or vector in $\mathbb{R}^n$ has n real components.
$\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$	Vectors, usually written in bold.	$\mathbf{0}$ is the zero vector; $c\boldsymbol{a}$ is scalar multiplication.
$\langle \boldsymbol{a}, \boldsymbol{b} \rangle$	Dot product $\sum_{j=1}^{n} a_j b_j$.	$\boldsymbol{a} \perp \boldsymbol{b}$ means $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = 0$.
$\|\boldsymbol{a}\|$	Euclidean norm $\sqrt{\langle \boldsymbol{a}, \boldsymbol{a} \rangle}$.	A vector is unit when $\|\boldsymbol{a}\| = 1$.
$\boldsymbol{a} \times \boldsymbol{b}$	Cross product of vectors in $\mathbb{R}^3$.	It is perpendicular to both factors; its order affects its sign.
$\boldsymbol{r} : I \to \mathbb{R}^n$	Parametrized curve, $\boldsymbol{r}(t) = (x_1(t), \ldots, x_n(t))$.	I is the stated parameter interval and the image is the geometric curve.
$\boldsymbol{r}'(t)$, $\boldsymbol{v}$, $\boldsymbol{a}$	Curve derivative; velocity $\boldsymbol{v} = \boldsymbol{r}'$; acceleration $\boldsymbol{a} = \boldsymbol{r}''$.	Speed is $\|\boldsymbol{r}'(t)\|$, not the vector $\boldsymbol{r}'(t)$.
$D \subseteq \mathbb{R}^n$	Domain or a set of admissible input points.	In $f : D \to \mathbb{R}^m$, all limits and paths are taken within D.
$B_\rho(p)$	Open ball $\{x : \|x - p\| < \rho\}$.	Used to state local set and limit notions.
p a limit point of D	For every $\rho > 0$, the deleted ball $B_\rho(p) \setminus \{p\}$ meets D.	This permits $x \to p$ through domain points distinct from p.
$p \in \operatorname{int} D$	Some ball $B_\rho(p)$ lies in D.	Coordinate partial derivatives use an interior point.
$p \in \partial D$	Every ball about p meets both D and its complement.	Boundary points require the stated relative-domain interpretation.
$\{x \in D : f(x) = c\}$	Level set (or level curve/surface in the relevant dimension).	For differentiable f and p in the set with $\nabla f(p) \neq \mathbf{0}$, the gradient gives a normal direction at p.

Partials, extrema, and differentiability (Topics 2–4)

Notation	Meaning	Use or convention
$f_x(a,b),\ f_y(a,b)$	First partial derivatives at (a,b).	Hold the other coordinate fixed in the corresponding difference quotient.
$f_{xx}, f_{xy}, f_{yx}, f_{yy}$	Second partial derivatives.	**Course convention:** $f_{xy} = (f_x)_y$ and $f_{yx} = (f_y)_x$.
$\nabla f(a,b)$	Gradient column vector $(f_x(a,b), f_y(a,b))^\mathsf{T}$.	For scalar f, it records first-order directional change.
Hess $f(a,b)$	Hessian matrix $\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$ at (a,b).	Its determinant is $f_{xx}f_{yy} - f_{xy}f_{yx}$; under the C^2 hypothesis, $f_{xy} = f_{yx}$ and the test uses $f_{xx}f_{yy} - f_{xy}^2$.
p critical for f	An interior stationary point, or an interior point where a required first partial does not exist.	A critical point is a candidate for local classification; it is not itself a conclusion.
$L_{f,p}$	Tangent-plane / affine approximation to a scalar f at $p = (a,b)$.	$L_{f,p}(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$.
df_p or $Df(p)$	Total differential: the linear derivative map at p.	For scalar f, $df_p(h) = \nabla f(p) \cdot h$.
h and $o(\|h\|)$	Increment $h \in \mathbb{R}^n$ and a remainder smaller than first order.	$r(h) = o(\|h\|)$ means $r(h)/\|h\| \to 0$ as $h \to \mathbf{0}$.
$f(p+h) = f(p) + df_p(h) + o(\|h\|)$	Total-differentiability expansion.	It is the meaning of a first-order linear approximation.
$\Delta x, \Delta y, \Delta f$	Changes from a base point.	$\Delta f \approx f_x \Delta x + f_y \Delta y$ is the linear estimate.

Maps, implicit descriptions, and directions (Topic 5)

Notation	Meaning	Use or convention
$F : \mathbb{R}^n \to \mathbb{R}^m$	Vector-valued map, $F = (F_1, \ldots, F_m)$.	Its derivative is a linear map from input increments to output increments.
$J_F(a)$	Jacobian (derivative) matrix $\left[\frac{\partial F_i}{\partial x_j}(a)\right]_{i=1,\ldots,m}^{j=1,\ldots,n}$.	It is $m \times n$ and acts on column increments: $DF(a)h = J_F(a)h$.
$J_f(a)$ versus $\nabla f(a)$	For scalar $f : \mathbb{R}^n \to \mathbb{R}$, $J_f(a)$ is a $1 \times n$ row; $\nabla f(a)$ is the corresponding $n \times 1$ column.	Thus $J_f(a) = \nabla f(a)^\mathsf{T}$ under this matrix convention.
$J_{F \circ G}(a)$	Derivative matrix of a composite map.	If G is differentiable at a and F at $G(a)$, then $J_{F\circ G}(a) = J_F(G(a))J_G(a)$.
$F(x,y) = 0$ or $F(x,y,z) = 0$	Implicit relation defining a curve or surface.	If F is C^1 nearby and the partial for the solved variable is nonzero, the corresponding implicit derivative is available.
$y = y(x)$; $z = g(x,y)$	Local implicit descriptions.	Under those IFT hypotheses, $y' = -F_x/F_y$; $g_x = -F_x/F_z$, $g_y = -F_y/F_z$, at the relevant point.
$D_{\boldsymbol{u}} f(p)$	Directional derivative of f at p in direction $\boldsymbol{u}$.	$\boldsymbol{u}$ must be unit; for differentiable f, $D_{\boldsymbol{u}} f(p) = \nabla f(p) \cdot \boldsymbol{u}$.
$\boldsymbol{u} = \boldsymbol{v}/\|\boldsymbol{v}\|$	Unit vector in the direction of nonzero $\boldsymbol{v}$.	Normalize before using $\boldsymbol{v}$ as a directional-derivative direction.

Constraints and double integrals (Topics 6–7)

Notation	Meaning	Use or convention
f; $g(x) = k$	Objective function; one equality constraint.	Feasible points satisfy the constraint as well as the multiplier equations.
λ	Multiplier for one equality constraint.	For C^1 f, g, at a regular constrained local extremum, $\nabla f = \lambda \nabla g$ together with $g = k$.
$g_1 = k_1, \ldots, g_m = k_m$	Multiple C^1 equality constraints.	The regularity condition is that $\nabla g_1, \ldots, \nabla g_m$ are linearly independent.
$\lambda_1, \ldots, \lambda_m$	Multipliers for independent constraints.	$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m$, with every $g_i = k_i$.
R or D	Planar integration region.	State the actual region before writing iterated bounds.
$\mathrm{d}A = \mathrm{d}x\ \mathrm{d}y$	Cartesian area element.	The differential is part of the integral, alongside integrand and bounds.
Type I region	$D = \{(x, y) : a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$, with continuous $g_1 \le g_2$.	For continuous f, $\iint_D f\ \mathrm{d}A = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\ \mathrm{d}y\ \mathrm{d}x$.
Type II region	$D = \{(x, y) : c \le y \le d,\ h_1(y) \le x \le h_2(y)\}$, with continuous $h_1 \le h_2$.	For continuous f, $\iint_D f\ \mathrm{d}A = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\ \mathrm{d}x\ \mathrm{d}y$.
(r, θ)	Polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$, usually $r \ge 0$.	For the course polar regions, $\mathrm{d}A = r\ \mathrm{d}r\ \mathrm{d}\theta$.
Polar integral	Put $f^\star(r, \theta) = f(r\cos\theta, r\sin\theta)$ on $\alpha \le \theta \le \beta$, where continuous $0 \le h_1 \le h_2$ and $0 \le \beta - \alpha \le 2\pi$.	For continuous f, $\iint_D f\ \mathrm{d}A = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f^\star(r, \theta) r\ \mathrm{d}r\ \mathrm{d}\theta$.
Scope note. This appendix ends with Type I/II iterated integrals and polar coordinates. General integral change-of-variables formulas, line integrals, surface integrals, and triple integrals are outside this reference.		

C Exam & Admin Information

Current-source rule

The AY2026–2027 Semester 1 syllabus, `Reference/Admin/Syllabus.pdf`, governs assessment and examination rules. Scope is cross-checked against the released Week 1–7 decks. A prior paper is practice evidence only: it does not set the 2026 examination format or policy.

Assessment Breakdown

Component	Weight	Verified syllabus information
In-lecture quizzes	5%	Individual, point-based; administered on Wooclap in almost every lecture. Some lowest scores will be dropped, with the exact number announced only after Week 12.
Online homework	20%	Individual, point-based; based on Stewart's *Calculus* (metric version, 9th edition), assigned on WebAssign every teaching week except the last.
Midterm Exam	25%	Individual, point-based; assesses ILOs 1–5. Tuesday, 6 October 2026, 16:30–18:20 (110 minutes), Teaching Week 8; venue TBA.
Final Exam	50%	Individual, point-based; assesses ILOs 1–9. Thursday, 26 November 2026, 13:00–15:00; venue TBA.

Midterm Help-Sheet Rule

What the current syllabus permits

You may consult **ONE double-sided A4-sized help-sheet** with **typed notes**. Attachments are prohibited, including additional sheets and sticky notes. The syllabus gives no further help-sheet specifications here; do not add assumed rules about font size, colour, printing, or calculators.

Do not import an older-paper rule

The 2025–2026 midterm paper permitted typed *or handwritten* notes and stated that calculators could be used. Those were instructions for that paper. For the 2026 midterm, follow the current syllabus's typed-notes rule and the instructions printed on the current paper; the syllabus does not state a calculator policy.

Verified Midterm Scope

The syllabus plans seven topic blocks before the midterm. The released decks arrive in a different order, so revise by topic rather than by a teaching-week label. The pack's scope cutoff is the Week 6 polar-coordinate double-integral material: Week 7 begins triple integrals.

Topic	Planned syllabus block	Released-deck evidence / revision status
1	Parametric curves and tangent lines	Week 1; in scope.
2	Limits, continuity, and partial derivatives	Week 2; in scope.
3	Partial derivatives and classification of extreme points	Week 4; in scope.
4	Tangent spaces and differentiability	Week 3; in scope.
5	Chain Rule and directional derivatives	Weeks 3–4; in scope.
6	Lagrange multipliers	Week 5; in scope.
7	Double integrals, including polar coordinates	Weeks 5–6; in scope.

The accelerated deck order is intentional in this map: actual Week 5 starts double integrals and actual Week 6 completes polar coordinates. Actual Week 7 opens with triple integrals, which are outside this double-integral cutoff. Consequently, do not treat triple integrals, three-dimensional substitution, vector fields, line integrals, Green's Theorem, surface integrals, Stokes' Theorem, or the Divergence Theorem as midterm-pack topics.

110-Minute Practice Plan

Practise an evidence-based timing design

The 2025–2026 midterm had four multi-part problems and allowed 110 minutes. Use four-problem, 110-minute practice runs only as a *practice design*, not as a confirmed 2026 paper format. A workable rehearsal allocation is: 5 minutes to scan and mark methods; about 24 minutes per problem (96 minutes); and 9 minutes to check units, domains, constraints, signs, polar Jacobians, and written justification. Reallocate time by marks when a current paper supplies them.

Revision Checklist Before the Midterm

- ☐ For each of Topics 1–7, complete one mixed problem without consulting notes, then use the error to select a targeted reattempt.
- ☐ In a timed practice run, first sketch regions/constraints and state the theorem or test before differentiating or integrating.
- ☐ Check the assumptions that make a method available: regular parametrization or level set, differentiability, unit direction, constraint regularity, and a correctly described integration region.
- ☐ For limits and differentiability, distinguish a disproof by paths from a proof; for extrema, compare all required interior and boundary candidates.
- ☐ For double integrals, draw the region, choose/split the order, and include the factor r in polar coordinates.
- ☐ Prepare the one permitted double-sided typed A4 sheet early and verify it has no attachment or sticky note. On the day, follow the current paper's instructions if they add details not stated in the syllabus.

Sources: syllabus pp. 3–5 (assessment and help sheet); MH2100 digest and released Week 1–7 slides (scope and deck timing); AY2025–2026 midterm question paper (four-problem practice observation).