Multivariable and complex calculus

(Course summary)

MATHS 2101 & MATHS7101, Semester 1 2017

Version of March 10, 2017

Thomas Leistner

Contents

Contents

Vectors in $\mathbb{R}^n$

Recall

$$\mathbb{R}^n = \{\boldsymbol{x} = (x_1, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}.$$

Notational conventions:

- We usually call elements *vectors* if $n > 1$ and *scalars* if $n = 1$.
- In $\mathbb{R}^2$ and $\mathbb{R}^3$ we usually denote vectors as (x, y) or (x, y, z).
- We will tend to write vectors as $\boldsymbol{x}$ in print. (when writing on the board, we might put little arrows over the vector or tildes under the vector)

$\mathbb{R}^n$ with vector addition, scalar multiplication and null vector is a *real vector space*.

In $\mathbb{R}^n$ we have the *Cartesian basis vectors*

$$\boldsymbol{e}_1 := (1, 0, \ldots, 0), \qquad \boldsymbol{e}_2 := (0, 1, \ldots, 0), \qquad \ldots \qquad , \boldsymbol{e}_n := (0, 0, \ldots, 1).$$

- In $\mathbb{R}^2$ we denote $\boldsymbol{i} := (1, 0)$ and $\boldsymbol{j} := (0, 1)$,
- and in $\mathbb{R}^3$ we have $\boldsymbol{i} := (1, 0, 0)$, $\boldsymbol{j} := (0, 1, 0)$ and $\boldsymbol{k} := (0, 0, 1)$.

The dot product and the norm

If $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ we define the *scalar product* (*dot product* or *inner product*) by

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

We define the *norm* or *length* of a vector by

$$||\boldsymbol{u}|| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}.$$

We also ue the notation $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v}$.

These satisfy

$$\begin{aligned} |\langle \boldsymbol{u}, \boldsymbol{v} \rangle| &\leq ||\boldsymbol{u}||\,||\boldsymbol{v}||, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \quad \textbf{Cauchy's inequality}, \\ ||\boldsymbol{u} + \boldsymbol{v}|| &\leq ||\boldsymbol{u}|| + ||\boldsymbol{v}||, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \quad \textbf{Triangle inequality}. \end{aligned}$$

Note that Cauchy's inequality is equivalent to $\langle \boldsymbol{u}, \boldsymbol{v} \rangle \leq ||\boldsymbol{u}||\,||\boldsymbol{v}||$, $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.

Distance and angles

The distance between two points $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ is $\|\boldsymbol{u} - \boldsymbol{v}\|$.

A vector of length 1 is called a *unit vector*. If $\boldsymbol{u} \neq 0$ we define

$$\hat{\boldsymbol{u}} = \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}$$

the *unit vector in the direction of u*.

If $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ and $\boldsymbol{u} \neq \boldsymbol{0} \neq \boldsymbol{v}$ we define the angle $\theta \in [0, \pi]$ between $\boldsymbol{u}$ and $\boldsymbol{v}$ by

$$\theta = \arccos(\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{v}}).$$

If $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ and $\boldsymbol{u} \neq \boldsymbol{0} \neq \boldsymbol{v}$ then $\theta = \pi/2$ and we say the two vectors are *orthogonal* or *perpendicular*. It follows that if θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ then $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos(\theta)$.

Cross product

If $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ then the *cross product* or *vector product* is a vector $\boldsymbol{u} \times \boldsymbol{v} \in \mathbb{R}^3$ defined by

$$\boldsymbol{u} \times \boldsymbol{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1)$$

Recall that we often use the convenient notation:

$$\boldsymbol{u} \times \boldsymbol{v} = \det \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix}$$

If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ and $\alpha, \beta \in \mathbb{R}$ then from the definition of cross-product we have:

1. $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$
2. $\boldsymbol{u} \times (\alpha \boldsymbol{v} + \beta \boldsymbol{w}) = \alpha(\boldsymbol{u} \times \boldsymbol{v}) + \beta(\boldsymbol{u} \times \boldsymbol{w})$
3. $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = \langle \boldsymbol{u}, \boldsymbol{w} \rangle \boldsymbol{v} - \langle \boldsymbol{u}, \boldsymbol{v} \rangle \boldsymbol{w}$

Useful identity:

$$\langle \boldsymbol{u}, \boldsymbol{v} \times \boldsymbol{w} \rangle = \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}$$

In particular this shows that $\langle \boldsymbol{v}, \boldsymbol{v} \times \boldsymbol{w} \rangle = 0$ and $\langle \boldsymbol{w}, \boldsymbol{v} \times \boldsymbol{w} \rangle = 0$ so that $\boldsymbol{v} \times \boldsymbol{w}$ is orthogonal to both $\boldsymbol{v}$ and $\boldsymbol{w}$. An ordered triple $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u} \times \boldsymbol{v}$ is called a *right-handed triplet*.

Cross product and norm

We have

$$\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - \langle \boldsymbol{u}, \boldsymbol{v} \rangle^2$$

and

$$\boldsymbol{u} \times \boldsymbol{v} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \sin(\theta) \hat{\boldsymbol{n}}$$

where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ and $\hat{\boldsymbol{n}}$ is a unit vector making $\boldsymbol{u}, \boldsymbol{v}, \hat{\boldsymbol{n}}$ a right-handed triplet.

If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ then $\|\boldsymbol{u} \times \boldsymbol{v}\|$ is the area of the parallelogram spanned by $\boldsymbol{u}$ and $\boldsymbol{v}$ and $|\langle \boldsymbol{u}, \boldsymbol{v} \times \boldsymbol{w} \rangle|$ is the volume of the parallelepiped spanned by $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$.

Open and closed sets in $\mathbb{R}^n$

If $\boldsymbol{a} \in \mathbb{R}^n$ and $\delta > 0$ we define

$$B(\boldsymbol{a}, \delta) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{a}\| < \delta\}.$$

and we call this the *(open) ball around $\boldsymbol{a}$ of radius δ*.

We say a set $U \subseteq \mathbb{R}^n$ is *open* if for every $\boldsymbol{x} \in U$ there is a $\delta > 0$ such that $B(\boldsymbol{x}, \delta) \subseteq U$.

Facts about open sets:

- $\mathbb{R}^n$ and the empty set $\emptyset$ are open.
- If $U_1, U_2 \subseteq \mathbb{R}^n$ are open then $U_1 \cap U_2$ and $U_1 \cup U_2$ are open.

If $S \subseteq \mathbb{R}^n$ we say a point $\boldsymbol{x} \in \mathbb{R}^n$ is a *boundary point of S* if every open ball about $\boldsymbol{x}$ contains at least one point in S and at least one point not in S. We denote by ∂S the set of all boundary points of S and call it the *boundary* of S.

We call $\bar{S} = S \cup \partial S$ the *closure* of S. If $S \subseteq \mathbb{R}^n$ we say it is *closed* if $\partial S \subseteq S$.

Facts about closed sets:

- $\bar{S}$ is closed
- S is closed if and only if $S^c = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \notin S\}$ is open

Scalar and vector functions

A *function*, *map* or *transformation*

$$f\colon A \to B$$

is a rule that assigns to any $a \in A$ a specific $f(a) \in B$. We call A the *domain* of f and B the *codomain* or *target* of f. We call

$$f(A) = \{f(a) \mid a \in A\}$$

the *range* or *image* of f. We don't require that $f(A) = B$.

We are interested in the case $A \subseteq \mathbb{R}^n$ and $B = \mathbb{R}^m$ and will adopt the shorthand notation

$$f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m.$$

If $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$ we say that f is *scalar valued* and if $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ with $m > 1$ we say that f is *vector valued*. We could have bolded vector valued functions as $\boldsymbol{f}$ but we won't. Sometimes we are interested in the case that $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and we think of f as attaching a vector $f(x)$ to the point x. This is called a *vector field* and is important in many physical situations such as velocity of a fluid or velocity of the wind on the surface of the earth. We shall usually write vector fields as bolded, typically $\boldsymbol{u}$, $\boldsymbol{v}$ etc.

Graphs, level sets & sketching functions

Notice that if $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ then we can always write

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \dots, f_m(\boldsymbol{x}))$$

where $f_i\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$ is a scalar valued function called the *i-th component function* of f for each $i = 1, \dots m$.

If $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$ we define the *graph* of f to be

$$\text{graph}(f) = \{(\boldsymbol{x}, f(\boldsymbol{x})) \mid \boldsymbol{x} \in A\} \subseteq \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}.$$

In the case that $n = 1$ or $n = 2$ then we can draw the graph as a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$ as a curve or surface.

If $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$ we define the *level set* of f at c to be

$$f^{-1}(c) = \{\boldsymbol{x} \in A \mid f(\boldsymbol{x}) = c\}.$$

If $n = 2$ these are usually level curves and if $n = 3$ they are usually level surfaces.

If $\boldsymbol{u}\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a vector field for $n = 2$ and $n = 3$ they can be sketched by just drawing the vectors on the region A.

Limits

Definition 2.1

Let $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ for an open set A and let $\boldsymbol{a} \in \bar{A}$. Then we say that $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L} \in \mathbb{R}^m$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that if $\boldsymbol{x} \in A$ and $0 < \|\boldsymbol{x} - \boldsymbol{a}\| < \delta$ then $\|f(\boldsymbol{x}) - \boldsymbol{L}\| < \epsilon$. We say "the limit of f as $\boldsymbol{x}$ approaches $\boldsymbol{a}$ is $\boldsymbol{L}$".

Facts about limits:

- Lots of functions have no limit as $\boldsymbol{x}$ approaches $\boldsymbol{a}$.
- If there is a limit as $\boldsymbol{x}$ approaches $\boldsymbol{a}$ it is unique.
- $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$ if and only if $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \|f(\boldsymbol{x}) - \boldsymbol{L}\| = 0$.

Properties of Limits

Assume $f, g\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, A open, $\alpha, \beta\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{a} \in \overline{A}$ and $c \in \mathbb{R}$. Then:

1. If $f(\boldsymbol{x}) = \boldsymbol{L}$ for all $\boldsymbol{x} \in A$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$.
2. $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \boldsymbol{x} = \boldsymbol{a}$.
3. If $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} cf(\boldsymbol{x}) = c\boldsymbol{L}$.
4. If $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$ and $\lim_{\boldsymbol{x}\to\boldsymbol{a}} g(\boldsymbol{x}) = \boldsymbol{J}$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}}(f(\boldsymbol{x}) + g(\boldsymbol{x})) = \boldsymbol{L} + \boldsymbol{J}$.
5. If $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \alpha(\boldsymbol{x}) = A$ and $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \beta(\boldsymbol{x}) = B$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \alpha(\boldsymbol{x})\beta(\boldsymbol{x}) = AB$.
6. If $\beta(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in A$ and $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \beta(x) = B \neq 0$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} 1/\beta(\boldsymbol{x}) = 1/B$.
7. If $f(\boldsymbol{x}) = (f_1(x), \ldots, f_m(x))$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(x) = \boldsymbol{L} = (L_1, \ldots, L_m)$ if and only if $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f_i(x) = L_i$ for every $i = 1, \ldots, m$.
8. $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$ if and only if $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \|f(\boldsymbol{x}) - \boldsymbol{L}\| = 0$.
9. (Squeeze Lemma) If $0 \le \alpha(\boldsymbol{x}) \le \beta(\boldsymbol{x})$ for all $\boldsymbol{x} \in A$ and $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \beta(\boldsymbol{x}) = 0$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \alpha(\boldsymbol{x}) = 0$.

Note: We often use (8) in the following way. If we have $0 \le \|f(\boldsymbol{x}) - \boldsymbol{L}\| \le \beta(\boldsymbol{x})$ for all $x \in A$ and $\lim_{\boldsymbol{x}\to a} \beta(\boldsymbol{x}) = 0$ then $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \|f(\boldsymbol{x}) - \boldsymbol{L}\| = 0$ so that $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{L}$.

Continuous functions

Definition 2.2

Let $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ where A is either open or the closure of an open set. We say that f is *continuous* at $a \in A$ if

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a}).$$

We say that f is continuous on A if f is continuous at every $\boldsymbol{a}$ in A. If f is not continuous at $\boldsymbol{a}$ we say that f is *discontinuous* at $\boldsymbol{a}$.

Assume $f, g\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, where A is either open or the closure of an open set, $\alpha, \beta\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{a} \in A$ and $c \in \mathbb{R}$. Then the following are true:

1. If f is continuous at $\boldsymbol{a}$ then cf is continuous at $\boldsymbol{a}$.
2. If f and g are continuous at $\boldsymbol{a}$ then $f+g$ is continuous at $\boldsymbol{a}$.
3. If α and β are continuous at $\boldsymbol{a}$ then $\alpha\beta$ is continuous at $\boldsymbol{a}$.
4. If $\beta(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in A$ and β is continuous then $1/\beta$ is continuous.
5. If $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x}))$ then f is continuous at $\boldsymbol{a}$ if and only if f_i is continuous at $\boldsymbol{a}$ for all $i = 1, \ldots, m$.

Here cf is the function whose value at $\boldsymbol{x}$ is $cf(\boldsymbol{x})$, $f+g$ is the function whose value at $\boldsymbol{x}$ is $f(\boldsymbol{x}) + g(\boldsymbol{x})$, $\alpha\beta$ is the function whose value at $\boldsymbol{x}$ is $\alpha(\boldsymbol{x})\beta(\boldsymbol{x})$ and $1/\beta$ is the function whose value at $\boldsymbol{x}$ is $1/\beta(\boldsymbol{x})$.

Component functions and composition of continuous functions

Lemma 2.3

The component function $c_i : \mathbb{R}^n \to \mathbb{R}$ *defined by* $c_i(\boldsymbol{x}) = x_i$ *for any* $i = 1, \ldots, n$ *is continuous.*

Definition 2.4

Assume $g\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $f\colon B \subseteq \mathbb{R}^m \to \mathbb{R}^k$ and $g(A) \subseteq B$. We define the *composition* $f \circ g\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ by $(f \circ g)(\boldsymbol{x}) = f(g(\boldsymbol{x}))$.

Proposition 2.5

Let f *and* g *be as above with* A *and* B *open or closures of open sets. If* g *is continuous at* $\boldsymbol{a} \in A$ *and* f *is continuous at* $g(\boldsymbol{a}) \in B$ *then* $f \circ g$ *is continuous at* $\boldsymbol{a}$.

Existence of global extrema

Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$. A point $\boldsymbol{a} \in A$ is an *global (or absolute)* minimum if $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in A$. Similarly a point $\boldsymbol{a} \in A$ is an *global (or absolute)* maximum if $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in A$. A set A is *bounded* if there is some $R > 0$ such that $A \subset B(\mathbf{0}, R)$.

Theorem 2.6

Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$ with A closed and bounded. If f is continuous then there exist global maxima and minima for f on A.

Partial derivatives and differentiability of scalar functions

Let $U \subseteq \mathbb{R}^n$ be open, $\boldsymbol{a} = (a_1, \ldots, a_n) \in U$ and $f\colon U \to \mathbb{R}$ be a scalar valued function. We define the *j-th partial derivative of f at* $\boldsymbol{a}$ to be

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = \frac{d}{dt} f(\boldsymbol{a} + t\boldsymbol{e}_j)_{|t=0} = \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{e}_j) - f(\boldsymbol{a})}{h}.$$

Define a row vector

$$Df(\boldsymbol{a}) = \left(\tfrac{\partial f}{\partial x_1}(\boldsymbol{a}), \cdots, \tfrac{\partial f}{\partial x_n}(\boldsymbol{a})\right).$$

We say that f is *differentiable at* $\boldsymbol{a}$ if

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) - Df(\boldsymbol{a})\boldsymbol{h}}{\|\boldsymbol{h}\|} = 0.$$

Notation for linear maps

Recall from last year that $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if

$$A(\boldsymbol{x}+\boldsymbol{y}) = A(\boldsymbol{x}) + T(\boldsymbol{y}) \quad \text{and} \quad A(\lambda \boldsymbol{x}) = \lambda A(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \text{ and } \lambda \in \mathbb{R}.$$

Every linear map $A : \mathbb{R}^n \to \mathbb{R}^m$ is determined by an $m \times n$ matrix $[A]$ such that

$$A(\boldsymbol{x}) = \Big(\underbrace{\overset{\substack{\text{matrix}\\ \downarrow}}{[A]} \quad \overset{\substack{\text{column vector}\\ \downarrow}}{\boldsymbol{x}^\top}}_{\text{column vector}} \Big)^\top = \left(\sum_{j=1}^n A_{1j} x_j, \dots, \sum_{j=1}^n A_{mj} x_j \right) \tag{3.1}$$

Or, if $A(\boldsymbol{e}_j) = \sum_{i=1}^m A_{ji} \boldsymbol{e}_i$,

$$[A] = \Big(\overset{\substack{\text{column vector}\\ \downarrow}}{A(\boldsymbol{e}_1)^\top}, \dots, A(\boldsymbol{e}_n)^\top \Big) = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix} = \begin{pmatrix} (A_{1j})_{j=1,\dots,n} \\ \vdots \\ (A_{mj})_{j=1,\dots,n} \end{pmatrix} = (A_{ij})_{i=1,\dots,m, j=1,\dots,n}$$

if $A(\boldsymbol{e}_i) = \sum_{j=1}^m A_{ij} \boldsymbol{e}_j$ and $A_{ij} \in \mathbb{R}$. For simplicity, and with the Cartesian basis fixed, we don't distinguish between the linear map A and the matrix $[A]$. We introduce the notation

$$A \cdot \boldsymbol{x} := A(\boldsymbol{x}) = ([A]\boldsymbol{x}^\top)^\top.$$

Remark: *This notation is a bit clumsy, but this is the price we have to pay for dealing with row vectors instead of column vectors.*

Differentiability and derivative of vector functions

Let $U \subseteq \mathbb{R}^n$ be open, $\boldsymbol{a} = (a_1, \ldots, a_n) \in U$ and $f\colon U \to \mathbb{R}^m$ be a vector valued function. We say that $f = (f_1, \ldots, f_m)$ is *differentiable at $\boldsymbol{a} \in U$* if the partial derivatives

$$\frac{\partial f_i}{\partial x_j}(\boldsymbol{a})$$

all exist and

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{\|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) - Df(\boldsymbol{a}) \cdot \boldsymbol{h}\|}{\|\boldsymbol{h}\|} = 0$$

where

$$Df(\boldsymbol{a}) = \begin{bmatrix} Df_1(\boldsymbol{a}) \\ \vdots \\ Df_m(\boldsymbol{a}) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\boldsymbol{a}) \end{bmatrix}$$

and we define

$$Df(\boldsymbol{a}) \cdot \boldsymbol{h} = (Df_1(\boldsymbol{a}) \cdot \boldsymbol{h}, Df_2(\boldsymbol{a}) \cdot \boldsymbol{h}, \ldots, Df_n(\boldsymbol{a}) \cdot \boldsymbol{h}).$$

The matrix $D(f)(\boldsymbol{a})$ is called the *derivative of f at $\boldsymbol{a}$*, the *matrix of partial derivatives*, or the *Jacobi matrix*.

Derivative as approximation

The function $P_1(f, \boldsymbol{a})$ of $\boldsymbol{x}$ given by

$$P_1(f, \boldsymbol{a})(\boldsymbol{x}) = f(\boldsymbol{a}) + Df(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})$$

is called the *best linear (affine) approximation of f at $\boldsymbol{a}$*. Notice that it is a polynomial of degree one satisfying

$$P_1(f, \boldsymbol{a})(\boldsymbol{a}) = f(\boldsymbol{a}) \qquad \text{and} \qquad D(P_1(f, \boldsymbol{a}))(\boldsymbol{a}) = Df(\boldsymbol{a}).$$

The subspace of $\mathbb{R}^{n+m}$ defined by

$$T_{(\boldsymbol{a}, f(\boldsymbol{a}))} \operatorname{graph}(f) = \{(\boldsymbol{h}, Df(\boldsymbol{a}) \cdot \boldsymbol{h}) \mid \boldsymbol{h} \in \mathbb{R}^n\}$$

is called the *tangent space* to the graph of f. It is the subspace tangent to

$$\operatorname{graph}(f) = \{(\boldsymbol{x}, f(\boldsymbol{x})) \mid \boldsymbol{x} \in U\}$$

at the point $(\boldsymbol{a}, f(\boldsymbol{a}))$.

Differentiability and continuity

Theorem 3.1

If $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, for U open, is differentiable at $\boldsymbol{a} \in U$ then f is continuous at $\boldsymbol{a}$.

Definition 3.2

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ for U open and let $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x}))$. If all the partial derivatives of every f_i exist and are continuous on U we say that *f is C^1 on U*.

Theorem 3.3

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ for U open. If f is C^1 on U then f is differentiable at every $\boldsymbol{a} \in U$.

It holds: If $f\colon U \to \mathbb{R}^m$ and $f(x) = (f_1(x), \ldots, f_m(x))$ where $f_i\colon U \to \mathbb{R}$ for each $i = 1, \ldots, m$ then f is differentiable at $\boldsymbol{a}$ if and only if each f_i is differentiable at $\boldsymbol{a}$ and

$$Df(\boldsymbol{a}) = \begin{bmatrix} Df_1(\boldsymbol{a}) \\ \vdots \\ Df_m(\boldsymbol{a}) \end{bmatrix} \quad \text{where each } Df_i(\boldsymbol{a}) = \left(\frac{\partial f_i}{\partial x_1}(\boldsymbol{a}), \ldots, \frac{\partial f_i}{\partial x_n}(\boldsymbol{a}) \right) \text{ is a row vector of length } n.$$

Rules of differentiation

Let $U \subseteq \mathbb{R}^n$ be open and suppose that $f, g \colon U \to \mathbb{R}^m$ and $\alpha, \beta \colon U \to \mathbb{R}$ are differentiable at $\boldsymbol{a} \in U$ and $c \in \mathbb{R}$. Then the following are true:

1. cf is differentiable at $\boldsymbol{a}$ and $D(cf)(\boldsymbol{a}) = cD(f)(\boldsymbol{a})$,
2. $f + g$ is differentiable at $\boldsymbol{a}$ and $D(f+g)(\boldsymbol{a}) = D(f)(\boldsymbol{a}) + D(g)(\boldsymbol{a})$,
3. The *product rule* says that $\alpha\beta$ is differentiable at $\boldsymbol{a}$ and $D(\alpha\beta)(\boldsymbol{a}) = \alpha(\boldsymbol{a})D(\beta)(\boldsymbol{a}) + \beta(\boldsymbol{a})D(\alpha)(\boldsymbol{a})$,
4. The *quotient rule* says that if β is never zero then α/β is differentiable at $\boldsymbol{a}$ and

$$D\left(\frac{\alpha}{\beta}\right)(\boldsymbol{a}) = \frac{\beta(\boldsymbol{a})D(\alpha)(\boldsymbol{a}) - \alpha(\boldsymbol{a})D(\beta)(\boldsymbol{a})}{\beta(\boldsymbol{a})^2}.$$

Theorem 3.4 (The chain rule)

Let $g \colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $f \colon V \subseteq \mathbb{R}^m \to \mathbb{R}^k$ for U and V open sets and $g(U) \subseteq V$. If g is differentiable at $\boldsymbol{a} \in U$ and f is differentiable at $g(\boldsymbol{a}) \in V$ then:

1. *$f \circ g$ is differentiable at $\boldsymbol{a} \in U$ and*
2. *$D(f \circ g)(\boldsymbol{a}) = D(f)(g(\boldsymbol{a}))D(g)(\boldsymbol{a})$.*

The gradient

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$, for U open, by a C^1 function. The *gradient of f* at a is the vector

$$\operatorname{grad} f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) = D(f)(\boldsymbol{a}) = \left(\tfrac{\partial f}{\partial x_1}(\boldsymbol{a}), \cdots, \tfrac{\partial f}{\partial x_n}(\boldsymbol{a})\right).$$

The gradient is a vector field $\nabla f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^n$. If $\boldsymbol{u}\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a vector field we call it *conservative* if $\boldsymbol{u} = \nabla f$ for some $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ and we say that f is a *scalar potential* for $\boldsymbol{u}$. Not all vector fields are conservative.
If $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^n$ is differentiable then the chain rule tells us that

$$\frac{d}{dt}(f \circ \boldsymbol{\gamma}) = Df(\boldsymbol{\gamma}(t))D\boldsymbol{\gamma}(t) = Df(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t)$$

where

$$\boldsymbol{\gamma}'(t) = ((\gamma^1)'(t), \dots, (\gamma^n)'(t)) \qquad \text{and} \qquad (\gamma^i)'(t) = \tfrac{d\gamma^i}{dt}(t).$$

In particular if $\boldsymbol{\gamma}(t) = \boldsymbol{a} + t\hat{\boldsymbol{u}}$ for $\hat{\boldsymbol{u}}$ a unit vector then $\nabla f(\boldsymbol{a}) \cdot \hat{\boldsymbol{u}}$ is the rate of change of f in the direction $\hat{\boldsymbol{u}}$ or the *directional derivative* of f in the direction $\hat{\boldsymbol{u}}$.

Proposition 3.5

∇f is orthogonal to the level sets of f and points in the direction that f increases most rapidly.

Let $c \in \mathbb{R}$. If $Df(\boldsymbol{a}) \neq 0$ for all $\boldsymbol{a} \in f^{-1}(c)$ then this tells us that the tangent space to the level set at $\boldsymbol{a}$ is given by

$$T_{\boldsymbol{a}} f^{-1}(c) = \nabla f(\boldsymbol{a})^{\perp} = \{\boldsymbol{v} \mid \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} = 0\}.$$

Curl

If $\boldsymbol{u}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field then the *curl of* $\boldsymbol{u}$ is defined by

$$\nabla \times \boldsymbol{u} = \operatorname{curl} \boldsymbol{u} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right) \times (u_1, u_2, u_3) = \left(\frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}, \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}\right).$$

A convenient shorthand is

$$\operatorname{curl} \boldsymbol{u} = \det \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ u_1 & u_2 & u_3 \end{bmatrix}.$$

If $\operatorname{curl} \boldsymbol{u} = 0$ we say that $\boldsymbol{u}$ is *irrotational* or *curl free*.

Divergence

If $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a vector field we define its *divergence* $\nabla \cdot \boldsymbol{u}\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ by

$$\nabla \cdot \boldsymbol{u} = \operatorname{div} \boldsymbol{u} = \frac{\partial u_1}{\partial x_1} + \cdots + \frac{\partial u_n}{\partial x_n}.$$

A vector field $\boldsymbol{u}$ with $\operatorname{div} \boldsymbol{u} = 0$ is called *solenoidal* or *divergence free.*

An important and useful property is that $\operatorname{curl} \circ \operatorname{grad} = \mathbf{0}$ and $\operatorname{div} \circ \operatorname{curl} = 0$ so that if $\boldsymbol{u} = \operatorname{grad} f$ then $\operatorname{curl} \boldsymbol{u} = \operatorname{curl}(\operatorname{grad} f) = \mathbf{0}$ so that u is irrotational. Likewise if $\boldsymbol{u} = \operatorname{curl} A$ then $\operatorname{div}(\boldsymbol{u}) = \operatorname{div}(\operatorname{curl}(A)) = 0$ so that $\boldsymbol{u}$ is divergence free.

It can be useful to think of ∇ in $\mathbb{R}^3$ as a vector of differential operators like

$$\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)$$

Then curl is like a cross-product and divergence like a dot-product. Sometimes people emphasise this by writing $\boldsymbol{\nabla}$ for the vector of differential operators.

Derivative identities for scalar and vector fields

For sufficiently differentiable scalar fields $f, g : \mathbb{R}^n \to \mathbb{R}$ and vector fields $\boldsymbol{u}, \boldsymbol{v} : \mathbb{R}^n \to \mathbb{R}^n$ the following can be shown.

1. $\nabla(fg) = f\nabla g + g\nabla f$
2. $\nabla(f/g) = (1/g)\nabla f - (f/g^2)\nabla g$
3. $\nabla \cdot (f\boldsymbol{v}) = (\nabla f) \cdot \boldsymbol{v} + f(\nabla \cdot \boldsymbol{v})$
4. $\nabla \cdot (\nabla f) = (\nabla \cdot \nabla) f = \nabla^2 f$

In the case that $n = 3$ we also have

1. $\nabla(\langle \boldsymbol{u}, \boldsymbol{v} \rangle) = (\boldsymbol{u} \cdot \nabla)\boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{u} + \boldsymbol{u} \times (\nabla \times \boldsymbol{v}) + \boldsymbol{v} \times (\nabla \times \boldsymbol{u})$
2. $\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{v} \cdot (\nabla \times \boldsymbol{u}) - \boldsymbol{u} \cdot (\nabla \times \boldsymbol{v})$
3. $\nabla \times (f\boldsymbol{v}) = (\nabla f) \times \boldsymbol{v} + f(\nabla \times \boldsymbol{v})$
4. $\nabla \times (\boldsymbol{u} \times \boldsymbol{v}) = (\boldsymbol{v} \cdot \nabla)\boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\boldsymbol{v} + (\nabla \cdot \boldsymbol{v})\boldsymbol{u} - (\nabla \cdot \boldsymbol{u})\boldsymbol{v}$
5. $\nabla \times (\nabla \times \boldsymbol{v}) = \nabla(\nabla \cdot \boldsymbol{v}) - \nabla^2 \boldsymbol{v}$
6. $\nabla \times (\nabla f) = 0$
7. $\nabla \cdot (\nabla \times \boldsymbol{v}) = 0$

Iterated partial derivatives

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. Then the partial derivatives such as

$$\frac{\partial f}{\partial x_i}\colon U \to \mathbb{R}$$

are continuous and we can ask if they are C^1. If they are all C^1 we say that f is C^2 and we can define all the *iterated* partial derivatives like

$$\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

We call these *partial derivatives of order* 2. In a similar way we can define partial derivatives of order k and if all the partial derivatives of order k of f exist and are continuous we say that f is of class C^k.

Theorem 3.6 (Clairault's or Schwarz's Theorem)

If $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ is C^2 then

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

for all $1 \le i, j \le n$.

Similarly if f is C^k any iterated partial derivative of order up to and including k is independent of the order in which the partial derivatives are taken.

Taylor's theorem in one dimension

Theorem 3.7 (1-dimensional Taylor's Theorem)

If $f\colon U \subseteq \mathbb{R} \to \mathbb{R}$ is C^{k+1} and $[a, a+h] \subseteq U$ then

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2} f''(a) + \cdots + \frac{h^k}{k!} f_{(k)} + R_k(a,h)$$

where

$$R_k(a,h) = \int_a^{a+h} \frac{(a+h-\tau)^k}{k!} f_{(k+1)}(\tau) d\tau$$

and satisfies

$$\lim_{h\to 0} \frac{R_k(a,h)}{h^k} = 0.$$

There is a general Taylor's theorem for multivariable functions but we will consider only the first two cases.

Taylors Theorem

Theorem 3.8 (Multivariable Taylor's theorem)

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be C^{k+1}. Assume that for some $R > 0$ we have $B(\boldsymbol{a}, R) \subseteq U$. Then if $||\boldsymbol{h}|| < R$ we have

1. *if $k = 1$ then*
$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\boldsymbol{a}) + R_1(\boldsymbol{a}, \boldsymbol{h}).$$
2. *if $k = 2$ then*
$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\boldsymbol{a}) + \frac{1}{2}\sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) + R_2(\boldsymbol{a}, \boldsymbol{h}).$$

and in both cases
$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{R_k(\boldsymbol{a}, \boldsymbol{h})}{||\boldsymbol{h}||^k} = 0.$$

Linear and second order approximation

If $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ we have the best linear approximation to f given by

$$P_1(f, \boldsymbol{a})(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{a})\ (\boldsymbol{x}_i - \boldsymbol{a}_i)$$

and the best second-order approximation given by

$$P_2(f, \boldsymbol{a})(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{a})\ (\boldsymbol{x}_i - \boldsymbol{a}_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\ (\boldsymbol{x}_i - \boldsymbol{a}_i)(\boldsymbol{x}_j - \boldsymbol{a}_j)$$

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be C^2. We define the *Hessian of f at $\boldsymbol{a}$* to be the second order term in the Taylor expansion:

$$H_f(\boldsymbol{a})(\boldsymbol{h}) = \frac{1}{2} \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})$$

and the *Hessian matrix* to be the matrix of second derivatives

$$\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right].$$

Note that because f is C^2 the Hessian matrix is *symmetric*.

Extrema of scalar functions

Let $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a scalar function. A point $\boldsymbol{a} \in U$ is called a *local minimum* if there is an open ball $B(\boldsymbol{a}, \delta) \subseteq U$ such that $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B(\boldsymbol{a}, \delta)$. Similarly a point $\boldsymbol{a} \in U$ is called a *local maximum* if there is an open ball $B(\boldsymbol{a}, \delta) \subseteq U$ such that $f(\boldsymbol{a}) \geq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B(\boldsymbol{a}, \delta)$. A point that is a local minimum or a local maximum is called a *local extremum*. If we have strict inequalities we call it a strict local minimum, strict local maximum etc.

A point $a \in U$ is called a *critical point of f* if either f is not differentiable at $\boldsymbol{a}$ or $Df(\boldsymbol{a}) = 0$.

A critical point which is not a local extremum is called a *saddle point*.

Theorem 3.9 (First derivative test for local extrema)

Let $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ with U open be a differentiable function. If $\boldsymbol{a} \in U$ is a local extremum of f then $Df(\boldsymbol{a}) = 0$.

Second derivative test for local extrema

We say a function $H\colon \mathbb{R}^n \to \mathbb{R}$ is a *quadratic* function if $H(\boldsymbol{h}) = (1/2)\sum_{i,j=1}^{n} h_i h_j H_{ij}$ for some symmetric matrix $[H_{ij}]$. Notice that if $\boldsymbol{h}$ is a column vector we can write this as $H(\boldsymbol{h}) = (1/2)\boldsymbol{h}[H_{ij}]\boldsymbol{h}^\top$. We call a quadratic function *positive definite* if $H(\boldsymbol{h}) \geq 0$ for all $\boldsymbol{h} \in \mathbb{R}^n$ and $H(\boldsymbol{h}) = 0$ only if $\boldsymbol{h} = \boldsymbol{0}$. We call a quadratic function *negative definite* if $-H$ is positive definite.

Lemma 3.10

If $H(\boldsymbol{h}) = (1/2)\boldsymbol{h}X\boldsymbol{h}^\top$ where X is a symmetric matrix then H is positive (negative) definite if and only if X has all of its eigenvalues positive (negative).

Lemma 3.11

If $H(\boldsymbol{h})$ is positive definite then there is an $M \geq 0$ such that $H(\boldsymbol{h}) \geq M\|\boldsymbol{h}\|^2$ for all $\boldsymbol{h} \in \mathbb{R}^n$.

Theorem 3.12 (Definiteness test for extrema)

Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be C^3 with a critical point at $\boldsymbol{a} \in \mathbb{R}^n$. If $H_f(\boldsymbol{a})$ is positive definite then $\boldsymbol{a}$ is a strict local minimum of f and if $H_f(\boldsymbol{a})$ is negative definite then $\boldsymbol{a}$ is a strict local maximum of f.

Positive definiteness of the Hessian

We define the *principal minors* of a quadratic function $H(\boldsymbol{h})$ to be the numbers

$$H_{11}, \det\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \det\begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix}, \cdots, \det(H).$$

Theorem 3.13 (Sylvester's criterion)

Let $H(\boldsymbol{h})$ be a quadratic function. Then $H(\boldsymbol{h})$ is positive definite if and only if all the principal minors are positive and negative definite if and only if $-H(\boldsymbol{h})$ is positive definite.

If the principal minors of $H_f(\boldsymbol{a})(\boldsymbol{h})$ are all non-zero but it is neither positive or negative definite then $\boldsymbol{a}$ is a saddle point of f. Otherwise we don't know.

Constrained extrema: Lagrange multipliers

Let $g_1, \dots, g_k \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be C^1 and assume that for all $\boldsymbol{a}$ in

$$S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) = g_2(\boldsymbol{x}) = \dots = g_k(\boldsymbol{x}) = 0\}.$$

the vectors $\nabla g_1(\boldsymbol{a}), \dots, \nabla g_k(\boldsymbol{a})$ are linearly independent. Then we call S a *C^1 submanifold of $\mathbb{R}^n$ defined by constraints $g_1, \dots, g_k$*.

Proposition 3.14

Let $g \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be C^1 and assume that for all $\boldsymbol{a}$ in $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) = 0\}$ we have that the $\nabla g(\boldsymbol{a}) \neq 0$ then S is a C^1 submanifold of $\mathbb{R}^n$ defined by the constraint g.

Theorem 3.15 (Constrained extrema)

Let S be a C^1 submanifold of $\mathbb{R}^n$ defined by k constraints $g_1, \dots, g_k \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$. If $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ is C^1 and has an extremum on S at $\boldsymbol{a}$ then there exist $\lambda_1, \dots, \lambda_k$ such that

$$\nabla f(\boldsymbol{a}) = \lambda_1 \nabla g_1(\boldsymbol{a}) + \dots + \lambda_k \nabla g_k(\boldsymbol{a}).$$

Corollary 1

Let S be a C^1 submanifold of $\mathbb{R}^n$ defined by the constraint $g \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$. If $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ is C^1 and has an extremum on S at $\boldsymbol{a}$ then there exist λ such that

$$\nabla f(\boldsymbol{a}) = \lambda \nabla g(\boldsymbol{a}).$$

Review

If $f\colon [a,b] \to \mathbb{R}$ we define the Riemann integral as follows. First we divide $[a,b]$ into n equal intervals $a = x_0 < x_1 < \cdots < x_n = b$, choose $c_i \in [x_{i-1}, x_i]$ and define the *Riemann sum* by

$$S_n = \sum_{i=1}^{n} f(c_i)|x_i - x_{i-1}|.$$

The *Riemann integral* is the limit

$$\int_a^b f(x)dx = \lim_{n\to\infty} S_n$$

if it exists and is independent of how we choose the c_i. In such a case we say that f is *Riemann integrable* on $[a,b]$.
Calculation is usually done using:

Theorem 4.1 (The fundamental theorem of calculus)

Let $U \subseteq \mathbb{R}$ be an an open set containing an interval $[a,b]$. If F is C^1 on U and $f(x) = F'(x)$ for all $x \in [a,b]$ then

$$\int_a^b f(x)dx = F(b) - F(a).$$

Double integrals over rectangles

Let $R = [a, b] \times [c, d] \subseteq \mathbb{R}^2$. Partition $[a, b]$ into n subintervals $[x_{i-1}, x_i]$ and $[c, d]$ into m subintervals $[y_{j-1}, y_j]$. Let $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ and pick $\boldsymbol{c}_{ij} \in R_{ij}$. Define

$$S_{n,m} = \sum_{i=1,j=1}^{n,m} f(\boldsymbol{c}_{ij})|x_i - x_{i-1}|\,|y_j - y_{j-1}|.$$

We define the *double integral* of f over R to be

$$\iint_R f(x, y)dA = \lim_{n,m\to\infty} S_{n,m}$$

if the limit exists and is independent of the choice of c_{ij}. In such a case we say that f is *integrable* on R. If $f\colon R \to \mathbb{R}$ is continuous then f is integrable but we shall see below that more general functions can be integrated.
To calculate we use

Theorem 4.2 (Fubini's theorem)

Let f be continuous on $R = [a, b] \times [c, d]$. Then

$$\iint_R f(x, y)dA = \int_c^d \left(\int_a^b f(x, y)dx\right) dy = \int_a^b \left(\int_c^d f(x, y)dy\right) dx.$$

The second two integrals here are called *iterated integrals*.

Properties of double integrals

Theorem 4.3 (Criteria for integrability)

If $f\colon R = [a, b] \times [c, d] \to \mathbb{R}$ is bounded and discontinuous only on a finite union of graphs of continuous functions then f is integrable on R.

Let $f, g\colon R \to \mathbb{R}$ be integrable and c a constant. Then we have

$$\iint_R f(x, y) + g(x, y)dA = \iint_R f(x, y)dA + \iint_R g(x, y)dA$$

and

$$\iint_R cf(x, y)dA = c \iint_R f(x, y)dA.$$

If $f(x, y) \leq g(x, y)$ for all $(x, y) \in R$ then

$$\iint_R f(x, y)dA \leq \iint_R g(x, y)dA.$$

Also

$$\left|\iint_R f(x, y)dA\right| \leq \iint_R |f(x, y)|dA.$$

Integration over more general regions I

Suppose $\phi_1, \phi_2 \colon [a, b] \to \mathbb{R}$ are continuous and satisfy $\phi_1(x) \leq \phi_2(x)$ for all $x \in [a, b]$. Let

$$D = \{(x, y) \mid x \in [a, b], \phi_1(x) \leq y \leq \phi_2(x)\}.$$

We call such a region *vertically simple*.

Similarly if $\psi_1, \psi_2 \colon [c, d] \to \mathbb{R}$ are continuous with $\psi_1(y) \leq \psi_2(y)$ for all $y \in [c, d]$ we let

$$D = \{(x, y) \mid y \in [c, d], \psi_1(y) \leq x \leq \psi_2(y)\}$$

and call D *horizontally simple*.

Call a region *simple* if it is horizontally and vertically simple and *elementary* if it is one or the other.

Let D be an elementary region inside a rectangle $R = [a, b] \times [c, d]$ and let $f \colon D \to \mathbb{R}$ be continuous and therefore bounded. Define $f_* \colon R \to \mathbb{R}$ by

$$f_*(x, y) = \begin{cases} 0 & \text{if } (x, y) \notin D \\ f(x, y) & \text{if } (x, y) \in D \end{cases}$$

Integration over more general regions II

As f_* is discontinuous on at most four continuous curves we have by Theorem 4.3 that it is integrable so we can define

$$\iint_D f(x,y)dA = \int_R f_*(x,y)dA.$$

We can evaluate this as an iterated integral. If D is vertically simple then

$$\iint_D f(x,y)dA = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)dy \right) dx$$

and if D is horizontally simple then

$$\iint_D f(x,y)dA = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx \right) dy$$

For a non-elementary region we divide it into elementary regions and add the corresponding integrals. Unbounded regions can be dealt with by a limiting process similar to the case of one-variable unbounded intervals.

Integration in $\mathbb{R}^3$ I

The same ideas can be applied to three variables to integrate over regions in $\mathbb{R}^3$. Let $R = [a, b] \times [c, d] \times [e, f] \subset \mathbb{R}^3$ and $f\colon R \to \mathbb{R}$. Partition $[a, b]$ into n subintervals $[x_{i-1}, x_i]$, $[c, d]$ into m subintervals $[y_{j-1}, y_j]$ and $[e, f]$ into l subintervals $[z_{k-1}, z_k]$. Let $R_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$ and pick $\boldsymbol{c}_{ijk} \in R_{ijk}$.

Define

$$S_{n,m,l} = \sum_{i=1,j=1,k=1}^{n,m,l} f(\boldsymbol{c}_{ijk})|x_i - x_{i-1}||y_j - y_{j-1}||z_k - z_{k-1}|.$$

We say that f is integrable over R and write

$$\iint_R f(x, y, z)dV = \lim_{n,m,l\to\infty} S_{n,m,l}$$

if the limit exists and is independent of the choice of $\boldsymbol{c}_{ijk}$.

Theorem 4.4 (Fubini's theorem again)

Let f be continuous on $R = [a, b] \times [c, d] \times [e, f]$. Then

$$\iint_R f(x, y, z)dV = \int_e^f \left(\int_c^d \left(\int_a^b f(x, y, z)dx\right) dy\right) dz$$

Integration in $\mathbb{R}^3$ II

We can also compute the triple integral by an iterated integral in any of the five other re-orderings of the variables x, y and z.

Theorem 4.5 (Criteria for integrability)

If $f\colon R = [a,b] \times [c,d] \times [e,f] \to \mathbb{R}$ is bounded and discontinuous only on a finite union of graphs of continuous functions then f is integrable on R.

We can extend the notion of an elementary region to volumes in $\mathbb{R}^3$ but it becomes more complicated and we will do some examples in lectures.

Change of variables for double integrals

Theorem 4.6 (Change of variables for double integrals)

Let $T\colon R \subseteq \mathbb{R}^2 \to T(R) \subseteq \mathbb{R}^2$ be a one to one C^1 map and $f\colon T(R) \to \mathbb{R}$ be integrable. Then $f \circ T\colon R \to \mathbb{R}$ is integrable and

$$\iint_{T(R)} f\, dA = \iint_D f \circ T |\det D(T)|\, dA$$

If $T(u,v) = (x(u,v), y(u,v))$ then

$$D(T) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

In more detail

$$\iint_{T(R)} f(x,y)\, dxdy = \iint_R f(x(u,v), y(u,v)) \left|\det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}\right| dudv$$

Change of variables for triple integrals

Theorem 4.7 (Change of variables for triple integrals)

Let $T\colon W \subseteq \mathbb{R}^3 \to T(W) \subseteq \mathbb{R}^3$ be a one to one C^1 map and $f\colon T(W) \to \mathbb{R}$ be integrable. Then $f \circ T\colon W \to \mathbb{R}$ is integrable and

$$\iiint_{T(W)} f\, dV = \iiint_W f \circ T |\det D(T)|\, dV$$

If $T(u, v, w) = (x, y, z)$ then

$$D(T) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix}.$$

Curves in $\mathbb{R}^n$

A subset C of $\mathbb{R}^n$ is a *curve* if it is the image of a C^1 function $\boldsymbol{\gamma}\colon [a,b] \to \mathbb{R}^n$ with $\gamma'(t) \neq 0$ for all $t \in [a,b]$. We call $\boldsymbol{\gamma}$ a parametrization of C. A curve is *closed* if $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$. A curve is *simple* if it $\boldsymbol{\gamma}$ is one to one except possibly $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$. The length of a curve is

$$L(C) = \int_a^b \|\boldsymbol{\gamma}'(t)\|\, dt.$$

We say that C is *parametrized by arc-length* if $\boldsymbol{\gamma}$ satisfies

$$\int_a^\tau \|\boldsymbol{\gamma}'(t)\|\, dt = \tau - a$$

for any $a < t < b$. Differentiating this condition we see that an arc-length parametrization is determined by the requirement that $\boldsymbol{\gamma}'(t) = 1$ for all $t \in [a,b]$.

If C is a simple curve parametrized by γ. If $t \in (a,b)$ we define $T_{\gamma(t)}C$, the tangent space to C at $\gamma(t)$, to be all multiples of $\gamma'(t)$.

Definition 4.8

Let $D \subseteq \mathbb{R}^n$ be closed and $f\colon D \to \mathbb{R}$. We say that g is C^k if there exists an open set U with $D \subset U$ and a C^k function $F\colon U \to \mathbb{R}$ such that $f(\boldsymbol{x}) = F(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$.

Integrating a function along a curve and line integrals I

Let f be a function defined on a curve C. If $\boldsymbol{\gamma}$ is a parametrization we define the integral of f along C by

$$\int_C f\, ds = \int_a^b f(\gamma(t))\|\boldsymbol{\gamma}'(t)\|\, dt$$

If $\gamma(s)$ is arc-length parametrized then

$$\int_C f\, ds = \int_a^b f(s)ds.$$

Notice that the length of C is $\int_C 1\, ds$. We can also integrate vector fields along oriented curves. This is called a *line integral*. Let $\boldsymbol{u}$ be a vector field defined on a curve C parametrized by $\boldsymbol{\gamma}$.

If we associate a direction to C we call it *oriented*. Usually we choose the parametrization to be such that the direction of orientation corresponds to increasing t. If C is an oriented curve and $\boldsymbol{c} \in C$ we define $\widehat{T}(\boldsymbol{c})$ to be the unique *unit tangent vector* to C at $\boldsymbol{c}$ pointing in the direction of orientation. We define the *line integral* by

$$\int_C \boldsymbol{u} \cdot d\boldsymbol{s} = \int_C \boldsymbol{u} \cdot \widehat{\boldsymbol{T}}\, ds.$$

Integrating a function along a curve and line integrals II

If C is an oriented curve with parametrization $\gamma(t)$ we have

$$\widehat{\boldsymbol{T}} = \frac{\gamma'(t)}{\|\gamma'(t)\|}$$

and

$$\int_C \boldsymbol{u} \cdot d\boldsymbol{s} = \int_C \boldsymbol{u} \cdot \widehat{\boldsymbol{T}}\, ds. = \int_a^b \boldsymbol{u}(\gamma(t)) \cdot \left(\frac{\gamma'(t)}{\|\gamma'(t)\|}\right) \|\gamma'(t)\| dt = \int_a^b \boldsymbol{u}(\gamma(t)) \cdot \gamma'(t) dt$$

We can extend these definitions to curves C which are unions of a finite number of curves $C_1, C_2, \ldots, C_n$ joined end to end.

Theorem 4.9 (Fundamental theorem of calculus for curves)

Let $\boldsymbol{u}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ be a conservative vector field on U open in $\mathbb{R}^3$ and assume that $\boldsymbol{u} = \nabla\phi$ where $\phi\colon U \to \mathbb{R}$. If $C \subset U$ is an oriented curve with endpoints $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ then

$$\int_C \boldsymbol{u} \cdot d\boldsymbol{s} = \phi(\boldsymbol{c}_2) - \phi(\boldsymbol{c}_1).$$

We can do the same thing in any $\mathbb{R}^n$.

Integration over surfaces I

Let $\sigma\colon R \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ be a parametrisation of a surface $\Sigma = \sigma(R)$. This means that σ is C^1, one to one and $D(\sigma)$ has no kernel, that is it has rank 2. The columns of $D(\sigma)$ are

$$\sigma_u = \left(\frac{\partial \sigma_1}{\partial u}, \frac{\partial \sigma_2}{\partial u}, \frac{\partial \sigma_3}{\partial u}\right) \qquad \text{and} \qquad \sigma_v = \left(\frac{\partial \sigma_1}{\partial v}, \frac{\partial \sigma_2}{\partial v}, \frac{\partial \sigma_3}{\partial v}\right).$$

If $(u, v) \in R$ then $\sigma_u(u, v)$ and $\sigma_v(u, v)$ span $T_{\sigma(u,v)}\Sigma$, the *tangent space* to the *surface* Σ at the point $\sigma(u, v)$. The vector $\boldsymbol{n} = \sigma_u \times \sigma_v$ is the normal to the tangent space and $\hat{\boldsymbol{n}}$ the unit normal. Hence

$$\hat{\boldsymbol{n}} \circ \sigma = \frac{\sigma_u \times \sigma_v}{\|\sigma_u \times \sigma_v\|}.$$

We call a surface Σ oriented if we have continuously chosen a unit normal to the tangent space everywhere on Σ. If Σ is oriented we always choose the parametrisation so that $\boldsymbol{n}$ points in the chosen direction.

Let $\sigma\colon R \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ be a parametrization of a surface Σ and $f\colon \Sigma \to \mathbb{R}$. We define the *surface integral* of f over Σ by

$$\iint_\Sigma f dS = \iint_R f \circ \sigma \|\sigma_u \times \sigma_v\| \, dudv$$

Integration over surfaces II

If $\boldsymbol{w}\colon U \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field we define the *flux integral* by

$$\iint_\Sigma \boldsymbol{w} \cdot d\boldsymbol{S} = \iint_\Sigma \boldsymbol{w} \cdot \hat{\boldsymbol{n}}\, dS.$$

We have

$$\begin{aligned}\iint_\Sigma \boldsymbol{w} \cdot d\boldsymbol{S} &= \iint_\Sigma \boldsymbol{w} \cdot \hat{\boldsymbol{n}}\, dS \\ &= \iint_R (\boldsymbol{w} \circ \sigma) \cdot (\hat{\boldsymbol{n}} \circ \sigma)\, \|\sigma_u \times \sigma_v\|\, du dv \\ &= \iint_R (\boldsymbol{w} \circ \sigma) \cdot (\sigma_u \times \sigma_v)\, du dv\end{aligned}$$

Green's and Stoke's Theorems

Theorem 4.10 (Green's Theorem)

Let $R \subset \mathbb{R}^2$ be a simple region and $\boldsymbol{w}\colon R \to \mathbb{R}^2$ a C^1 vector field $\boldsymbol{w}(x, y) = (u(x, y), v(x, y))$. Then

$$\iint_R \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dA = \oint_{\partial R} \boldsymbol{w} \cdot d\boldsymbol{s}.$$

We say a surface Σ is simple if it has a C^2 parametrization $\sigma\colon R \to \Sigma$ where R is simple.

Theorem 4.11 (Stokes' Theorem)

Let Σ be a simple oriented surface in $\mathbb{R}^3$. Let $\boldsymbol{w}$ be a C^1 vector field in an open set containing Σ. Then

$$\iint_\Sigma (\nabla \times \boldsymbol{w}) \cdot d\boldsymbol{S} = \oint_{\partial \Sigma} \boldsymbol{w} \cdot d\boldsymbol{s}.$$

Conservatve vector fields and Gauss's divergence Theorem

Theorem 4.12 (Conservative equals irrotational)

Let $\boldsymbol{u}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be a C^1 vector field. Then $\boldsymbol{u}$ is conservative if and only if it is irrotational. That is $\boldsymbol{u} = \nabla\phi$ for some $\phi\colon \mathbb{R}^3 \to \mathbb{R}$ if and only if $\nabla \times \boldsymbol{u} = 0$.

Theorem 4.13 (Gauss's (divergence) theorem)

Let $W \subset \mathbb{R}^3$ be a simple volume with closed boundary ∂W oriented by the outward normal. Let $\boldsymbol{u}\colon W \to \mathbb{R}^3$ be a C^1 vector field. Then

$$\iiint_W \nabla \cdot \boldsymbol{u}\, dV = \iint_{\partial W} \boldsymbol{u} \cdot d\boldsymbol{S}.$$

Review of complex numbers I

Recall that complex numbers are written $z = x + iy$ and added and multiplied bearing in mind the rule that $i^2 = -1$. So that $(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + y_1) + i(x_2 + y_2)$ and $(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$. We denote the set of all complex numbers by $\mathbb{C}$.

We can identify $z = x + iy$ with an element of $\mathbb{R}^2$ by mapping z to (x, y). The addition of complex numbers becomes vector addition. Under this identification $1 = (1, 0)$ and $i = (0, 1)$

If $z = x + iy$ we define the *real part* of z to be $\mathrm{Re}(z) = x$ and the imaginary part of z to be $\mathrm{Im}(z) = y$. We call z *real* if $z = x$ and *imaginary* if $z = iy$. We define $\bar{z} = x - iy$ and call it the *complex conjugate of z* and we have

$$\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z}) \quad \text{and} \quad \mathrm{Im}(z) = \frac{1}{2i}(z - (\bar{z})).$$

Complex numbers form a *field* and satisfy the following for all $u, v, w \in \mathbb{C}$:

$$\begin{array}{rclrcl}
u + (v + w) & = & (u + v) + w, & u(vw) & = & (uv)w \\
u + v & = & v + u, & uv & = & vu \\
u + 0 & = & 0 + u, & u1 & = & u \\
u + (-u) & = & 0, & uu_{-1} & = & 1 \quad \text{where} \quad u_{-1} = \frac{\bar{u}}{|u|^2} \quad \text{and} \quad u \neq 0 \\
u(v + w) & = & uv + uw. & & &
\end{array}$$

Review of complex numbers II

If $w \neq 0$ we define

$$\frac{u}{w} = uw^{-1} = \frac{u\bar{w}}{|w|^2}.$$

Theorem 5.1 (Fundamental theorem of algebra)

Let $a_0, a_1, \ldots, a_n \in \mathbb{C}$ with $n \geq 1$ and $a_n \neq 0$. Then

$$a_0 + a_1 z + \cdots + a_n z^n = 0$$

has n solutions in $\mathbb{C}$ (counting multiplicity).

De Moivre's formula

We can write complex numbers in *polar form* as

$$z = r(\cos(\theta) + i\sin(\theta))$$

where $r \in \mathbb{R}, r \geq 0$ and $\theta = \arg(z)$ is called the *argument* of z. We call $z = x + iy$ the *cartesian form* of z. If we require the argument to be in $[0, 2\pi)$ we call this the *principal argument* of z and denote it by $\operatorname{Arg}(z)$.

Trigonometric formulae show that if $z_1 = r_1(\cos(\theta_1) + i\sin(\theta_1))$ and $z_2 = r_2(\cos(\theta_2) + i\sin(\theta_2))$ then

$$z_1 z_2 = r_1 r_2(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2))$$

De Moivres' formula follows from this and says that if $z = r(\cos(\theta) + i\sin(\theta))$

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta))$$

Functions of one complex variable and complex differentiability

A *complex function of one variable* is a function

$$f\colon U \subseteq \mathbb{C} \to \mathbb{C}$$

The same definitions of open ball, open set, boundary, closed set, limits and continuity all apply as we regard $\mathbb{C} = \mathbb{R}^2$. Note that if $z = x + iy$ then $|z| = \|(x, y)\|$. Often we change between writing a complex function as $f(z)$ or $f(x, y)$ where $z = x + iy$ and we also write the value of the function as $f(x + iy) = u(x, y) + iv(x, y)$. We show in lectures that complex polynomials $p(z)$ are continuous and ratios of complex polynomials $p(z)/q(z)$ are continuous away from points where $q(z) = 0$.

Let $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ where U is open. We say that f is *complex differentiable* at $z_0 \in U$ if the limit

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. If the limit exists we write it as $f'(z_0)$ or $(df/dz)(z_0)$ and call it the *derivative* of f at z_0. If f is complex differentiable at all points of U we say that f is *holomorphic* or *analytic* on U. If $f\colon \mathbb{C} \to \mathbb{C}$ is holomorphic on all of $\mathbb{C}$ we say that f is *entire*.

A criterion for complex differentiability

Theorem 5.2 (Cauchy-Riemann Theorem)

Let $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ be written as $f(x+iy) = u(x,y) + iv(x,y)$ and let $z_0 = x_0 + iy_0 \in U$. Then $f'(z_0)$ exists if and only if f is differentiable at (x_0, y_0) in the real two-variable sense and satisfies the Cauchy Riemann equations:

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \qquad \text{and} \qquad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0).$$

If $f'(z_0)$ exists then it is given by

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i\frac{\partial v}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) - i\frac{\partial u}{\partial y}(x_0, y_0).$$

This gives us an easy test for complex differentiability. If f is C^1 and satisfies the Cauchy-Riemann equations then it is complex-differentiable.

Product, quotient and chain rules for complex functions

Proposition 5.3

If $f\colon U \subseteq \mathbb{C} \to \mathbb{C}$ with U is complex differentiable at $z_0 \in U$ then f is continuous at z_0.

Proposition 5.4

If $f, g\colon U \subseteq \mathbb{C} \to \mathbb{C}$, where U is open, is complex differentiable at $z_0 \in U$ and $\alpha \in \mathbb{C}$ then $f + g$, αf and fg are complex differentiable at z_0 and

(a) $(f+g)'(z_0) = f'(z_0) + g'(z_0)$ and $(\alpha f)'(z_0) = \alpha f'(z_0)$

(b) $(fg)'(z_0) = f(z_0)g'(z_0) + f'(z_0)g(z_0)$

If $g(z_0) \neq 0$ then f/g is complex differentiable at z_0 and $\left(\frac{f}{g}\right)'(z_0) = \frac{g(z_0)f'(z_0) - f(z_0)g'(z_0)}{g(z_0)^2}$

Proposition 5.5 (Chain rule)

Let $g\colon U \subseteq \mathbb{C} \to \mathbb{C}$ and $f\colon V \subseteq \mathbb{C} \to \mathbb{C}$ where U and V are open and $g(V) \subseteq U$. If g is complex differentiable at z_0 and f is complex differentiable at $g(z_0)$ then $f \circ g$ is complex differentiable at z_0 and

$$(f \circ g)'(z_0) = f'(g(z_0))g'(z_0).$$

Hence, polynomials are entire and $\frac{p(z)}{q(z)}$ is holomorphic on the open set $\{z \in \mathbb{C} \mid q(z) \neq 0\}$.

Harmonic functions

A function $h\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ is called *harmonic* if it satisfies Laplace's equation

$$\nabla^2 h = \frac{\partial^2 h}{\partial x_2} + \frac{\partial^2 h}{\partial y_2} = 0,$$

where ∇^2 is called the *Laplacian*.

If $f(x+iy) = u(x,y) + iv(x,y)$ is C^2 and holomorphic then u and v are harmonic. A pair of harmonic functions u and v for which

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

are called *harmonic conjugate*. In this case $f(z) = f(x+iy) = u(x,y) + iv(x,y)$ is holomorphic.

Elementary functions I

If $z \in \mathbb{C}$ we define the *complex exponential function* $\exp\colon \mathbb{C} \to \mathbb{C}$ by

$$\exp(z) = e^z = e^x(\cos(y) + i\sin(y))$$

where $z = x + iy$.

Proposition 5.6

The complex exponential function satisfies

(a) $r(\cos(\theta) + i\sin(\theta)) = re^{i\theta}$

(b) *For all* $u, v \in \mathbb{C}$ $e^u e^v = e^{u+v}$.

(c) $|e^z| = e^x > 0$.

(d) *For any* $k \in \mathbb{Z}$ *we have* $e^{z+2\pi ik} = e^z$.

(e)
$$\frac{d}{dz}e^z = e^z$$

(f) $e^{\pi i} = -1$

We can use the complex exponential to define other complex functions:

Elementary functions II

$$\cos(z) = \frac{1}{2}\left(e^{iz} + e^{-iz}\right) \quad \text{and} \quad \sin(z) = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right)$$

$$\cosh(z) = \frac{1}{2}\left(e^{z} + e^{-z}\right) \quad \text{and} \quad \sinh(z) = \frac{1}{2}\left(e^{z} - e^{-z}\right).$$

We define the *complex logarithm* with branch $[\Theta, \Theta + 2\pi)$ to be

$$\log(z) = \log|z| + i\arg(z)$$

if $\Theta \leq \arg(z) < \theta + 2\pi$. We write $\mathrm{Log}(z)$ for the branch $[0, 2\pi)$. The function $\log(z)$ is holomorphic on $\{z \in \mathbb{C} \mid z \neq 0, \arg(z) \neq \Theta\}$.

Contour integration and Cauchy's Theorem

For $z\colon [a,b] \to \mathbb{C}$ an oriented parametrisation of a simple, oriented curve C and f a complex function on an open set containing C, the *contour integral of f along C* is defined by

$$\int_C f(z)dz = \int_a^b f(z(t))z'(t)dt$$

Theorem 5.7 (Cauchy's Theorem)

Let C be a simple oriented closed curve in $\mathbb{C}$ and assume that f is holomorphic in an open set containing C and the region bounded by C. Then

$$\oint_C f(z)dz = 0.$$

Contour deformation

Let $U \subseteq \mathbb{C}$. We say that two curves C_1 and C_2 are *homotopic with endpoints fixed* if they have the same endpoints and be continuously deformed one into the other inside U without moving the endpoints. We say they are *homotopic* if one can be deformed into the other.

Theorem 5.8 (Contour deformation theorem)

Let f be holomorphic on $U \subseteq \mathbb{C}$. Let C_1 and C_2 be two curves in U which are either closed and homotopic to each other or homotopic to each other endpoints fixed. Then

$$\int_{C_1} f(z)dz = \int_{C_2} f(z)dz.$$

Cauchy's Integral Formula

Theorem 5.9 (Cauchy's integral formula)

Let C be a simple closed curve in $\mathbb{C}$ oriented anti-clockwise. Assume that f is holomorphic in an open set containing C and the region bounded by C. Let a be inside C. Then

$$\oint_C \frac{f(z)}{(z-a)}\, dz = 2\pi i f(a).$$

We can show from this, by repeated differentiation under the integral sign, that

$$f_{(k)}(a) = \frac{k!}{2\pi i} \oint_C \frac{f(z)}{(z-a)^{k+1}}\, dz,$$

so that all derivatives at a exist and hence f is C^∞. Moreover we can use this integral formula to calculate bounds on the derivatives and show that $f(z)$ has a convergent complex Taylor series

$$f(z) = f(a) + f'(a)(z-a) + \frac{1}{2!} f_{(2)}(a)(z-a)^2 + \dots$$

in an open ball around a.

Calculus of residues

Let $U \subseteq \mathbb{C}$ be open and $a \in U$. If $f\colon U - \{a\} \to \mathbb{C}$ is holomorphic we say it has a *singularity* at a.

In such a case we define the *residue* of f at a by

$$\operatorname{res}(f, a) = \frac{1}{2\pi i} \oint_C f(z)dz$$

where C is a simple closed curve oriented anti-clockwise, a is inside C and U contains C and the region bounded by C.

Theorem 5.10 (Residue Theorem)

Let C be a simple closed curve in $\mathbb{C}$ oriented anti-clockwise. Assume that f is holomorphic in an open set which contains C and the region bounded by C except points $a_1, \ldots, a_n$ which are inside C. Then

$$\oint_C f(z)\, dz = 2\pi i \sum_{j=1}^{n} \operatorname{res}(f, a_j).$$

Poles and the residue formula

Assume f has a singularity at a and $(z-a)^k f(z)$ is holomorphic in an open set containing a. If k is the smallest natural number for which this is true we say that f has a *pole of order k at a*. If f has a pole of order 1 at a we say that f has a *simple pole* at a. If there is no such k we say that f has an *essential singularity at a*. If f is not defined at a but we can make it holomorphic near a by defining it at a we say that f has a *removable singularity at a*.

If f has a pole of order k at a it can be shown that it has a unique Laurent expansion at a given by

$$f(z) = \frac{a_{-k}}{(z-a)^k} + \cdots + \frac{a_{-1}}{(z-a)} + a_0 + a_1(z-z_0) + \ldots$$

which converges in some open ball about a. In such a case it is easy to show that

$$\operatorname{res}(f,a) = a_{-1}.$$

Theorem 5.11 (Residue formula)

If f has a pole at a of order k then

$$\operatorname{res}(f,a) = \frac{1}{(k-1)!}\left[\frac{d^{k-1}}{dz^{k-1}}\left((z-a)^k f(z)\right)\Big|_{z=a}\right].$$