Chapter 4

Differential calculus

Remark 4.1. Recall that for a function $f : \mathbb{R} \supset \Omega \to \mathbb{R}$

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

This can also be formulated with sequences:

$$\forall (x_n) \subset \Omega \backslash \{x_0\},\ \lim_{n \to \infty} x_n = x_0 : \quad \lim_{n \to \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}$$

Definition 4.2 (Partial derivative)**.**
Let $n \in \mathbb{N}$, $\Omega \subset \mathbb{R}^n$ open. For $x \in \Omega$ and $j \in \{1, \ldots, n\}$ a function $f : \Omega \to \mathbb{R}^m$ is called *partially differentiable* in the j^{th} coordinate direction at x, if the limit:

$$\lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h} = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

exists. One writes:

$$\frac{\partial f}{\partial x_j}(x) = \partial_j f(x) := \lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h}$$

and calls the vector $\partial_j f(x) \in \mathbb{R}^m$ the j^{th} partial derivative at x.
If f is partially differentiable in all directions, at all Ω and all the partial derivatives $\partial_j f : \Omega \to \mathbb{R}^m$ are continuous functions, then f is called *continuously partially differentiable*. The vector space of the continuously partially differentiable functions on $\Omega \subset \mathbb{R}^n$ is denoted by $C^1(\Omega, \mathbb{R}^m)$.

Remark 4.3. Observe that the definition can immediately be extended to maps between any finite dimensional vector spaces. Indeed, the definitions also work if the target space is a general normed vector space. Also, we may analogously define derivatives in arbitrary directions.

Definition 4.4 (Vector field)**.**
A continuous map $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^n$ is called a *vector field* on Ω.

Definition 4.5 (Higher order partial derivatives)**.** A function $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$ is called *r-times continuously partially differentiable*, if for all $j = (j_1, \ldots, j_r)$, $j_i \in \{1, \ldots, n\}$

1. f is c.p.d.
2. $\partial_{j_1} f$ is c.p.d.
3. $\partial_{j_2}\partial_{j_1}$ is c.p.d.

   $\vdots$

4. $\partial_{j_r} \ldots \partial_{j_1} f$ is continuous

The real vector space of r-times c.p.d. functions is denoted by $C^r(\Omega, \mathbb{R}^m)$.

Definition 4.6. Let $\Omega \subset \mathbb{R}^n$, $g \in C^1(\Omega, \mathbb{R}^n)$, $f \in C^2(\Omega, \mathbb{R})$. Then:

$$\nabla f : \Omega \to \mathbb{R}^n \quad x \mapsto \nabla f(x) = (\partial_1 f(x), \ldots, \partial_n f(x))$$

is called the *gradient* of f,

$$\operatorname{div} g : \Omega \to \mathbb{R}, \qquad x \mapsto \operatorname{div} g(x) = \sum_{j=1}^{n} \frac{\partial g_j}{\partial x_j}(x)$$

is called the *divergence* of g,

$$\operatorname{curl} g : \Omega \to \mathbb{R}^n, \qquad x \mapsto \operatorname{curl} g(x) = \begin{pmatrix} \partial_2 g_3(x) - \partial_3 g_2(x) \\ \partial_3 g_1(x) - \partial_1 g_3(x) \\ \partial_1 g_2(x) - \partial_2 g_1(x) \end{pmatrix}$$

for $n = 3$ is called the *curl* of g, and:

$$\Delta f : \Omega \to \mathbb{R}, \qquad x \mapsto \Delta f(x) = \operatorname{div}(\nabla f)(x) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}(x)$$

is called the *Laplace* of f.

Exercise 4.1. Let $g : \mathbb{R}^n \to \mathbb{R}^n$, $\quad x \mapsto g(x) = x$ and $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$, $\quad x \mapsto \|x\|_2$. Compute $\operatorname{div} g$ and Δf.

Theorem 4.7 (Schwarz)**.**
Let $f \in C^2(\Omega, \mathbb{R}^m)$. Then $\forall x \in \Omega,\ i, j \in \{1, \ldots, n\}$

$$\partial_j \partial_i f(x) = \partial_i \partial_j f(x)\,.$$

Corollary 4.8. *Let $\Omega \subset \mathbb{R}^3$, $f \in C^2(\Omega, \mathbb{R})$ and $g \in C^2(\Omega, \mathbb{R}^3)$. Then* $\operatorname{curl}(\nabla f) = 0$ *and* $\operatorname{div}(\operatorname{curl} g) = 0$.

The derivative as linear approximation

For $f : \mathbb{R} \to \mathbb{R}$ differentiability at $x_0 \in \mathbb{R}$ means

$$\lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right) = \lim_{x \to x_0} \frac{\varphi(x, x_0)}{x - x_0} = 0\,,$$

where $\varphi(x, x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0)$ or, after reshuffling

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varphi(x, x_0)\,,$$

where

$$\varphi(x, x_0) = o(|x - x_0|) \quad :\Leftrightarrow \quad \lim_{x \to x_0} \frac{|\varphi(x, x_0)|}{|x - x_0|} = 0\,.$$

The map $\mathbb{R} \to \mathbb{R}$, $x \mapsto f'(x_0) \cdot x$ is $\mathbb{R}$-linear and the map $\mathbb{R} \to \mathbb{R}$, $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ is affine-$\mathbb{R}$-linear. Hence, we think of $f(x_0) + f'(x_0)(x - x_0)$ as the (affine) linear approximation to f near x_0.

Definition 4.9 (Total derivative)**.**
Let $\Omega \subset \mathbb{R}^n$ open and $f : \Omega \to \mathbb{R}^m$. We call f differentiable at $x_0 \in \Omega$, if there exists an linear-map $A : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - A(x - x_0)\|_2}{\|x - x_0\|_2} = 0$$

Then A is uniquely determined by the above equation, is denoted by $Df|_{x_0}$, and called the *total derivative* or the differential of f at x_0.
If $f : \Omega \to \mathbb{R}^m$ is differentiable at all $x \in \Omega$, then f is called *differentiable on Ω* and

$$Df : \Omega \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), \qquad x \mapsto Df|_x$$

is a $m \times n$ matrix-valued function on Ω.

Remark 4.10. $f : \Omega \to \mathbb{R}^m$ is differentiable at $x_0 \Rightarrow f(x) = f(x_0) + Df|_{x_0}(x - x_0) + o(\|x - x_0\|_2)$

Theorem 4.11 (Jacobi matrix)**.**
Let $f : \Omega \to \mathbb{R}^m$ be differentiable at a point $x_0 \in \Omega$. Then

$$(Df|_{x_0})_{ij} = \frac{\partial f_i}{\partial x_j}(x_0)$$

or, more explicitly

$$Df|_{x_0} = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \partial_n f_1(x_0) \\ \vdots & & \vdots \\ \partial_1 f_m(x_0) & \dots & \partial_n f_m(x_0) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x_0) \\ \vdots \\ \nabla f_m(x_0) \end{pmatrix}$$

is called the Jacobian matrix.

Proposition 4.12. *Let $\Omega \subset \mathbb{R}^n$ open and $f \in C^1(\Omega, \mathbb{R}^m)$. Then f is differentiable.*

$$\begin{array}{ccccc} \text{cont. part. diff.} & \Rightarrow & \text{differentiable} & \Rightarrow & \text{part. diff.} \\ & & \Downarrow & & \\ & & \text{continuous} & & \end{array}$$

None of the implications holds in the reversed direction! But cont. part. diff. $\Leftrightarrow$ differentiable with continuous derivative.

Proposition 4.13 (Chain rule)**.** *Let $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^m$ be open. Consider maps*

$$\mathbb{R}^n \supset \Omega \xrightarrow{g} \Omega' \subset \mathbb{R}^m \xrightarrow{f} \mathbb{R}^l \ .$$

If g is differentiable at $x \in \Omega$ and f is differentiable at $g(x) \in \Omega'$, then $f \circ g : \Omega \to \mathbb{R}^l$ is differentiable at x and

$$D(f \circ g)|_x = Df|_{g(x)} \circ Dg|_x \,.$$

Remark 4.14. Often times the chain rule is written with a multiplication instead of a composition, i.e. $D(f \circ g)|_x = Df|_{g(x)} \cdot Dg|_x$. This is because if we think of linear maps in terms of matrices, composition is just matrix multiplication. Let, for instance $v \in \mathbb{R}^n$. Then, the components of $D(f \circ g)|_x$ are

$$(D(f \circ g)|_x)_i = \sum_{k=1}^{n} \sum_{j=1}^{m} (Df|_{g(x)})_{ij} (Dg|_x)_{jk} v_k \,.$$

Corollary 4.15. *For a function $f \in C^1(\Omega, \mathbb{R}^m)$, $x_0 \in \Omega$, and $v \in \mathbb{R}^n$ we have*

$$\partial_v f(x_0) = Df|_{x_0} v \,.$$

Fundamental theorems

First want to state the Taylor theorem. To do so, however, we have to understand higher-order differentials. For $f : \Omega \to W$, with $\Omega \subset V$ an open subset of a finite dimensional vector space and W a normed space, the differential Df is a map

$$Df : \Omega \to \mathcal{L}(V, W).$$

Thus the second differential $D(Df)$ is a map

$$D(Df) : \Omega \to \mathcal{L}(V, \mathcal{L}(V, W)) \cong \mathcal{L}_2(V \times V, W)$$

($=$ bilinear maps $V \times V \to W$) and the k^{th} derivative:

$$D^k f : G \to \mathcal{L}_k(\underbrace{V \times \ldots \times V}_{k\text{-times}}, W) \,.$$

Theorem 4.16 (Taylor). *Let $\Omega \subset V$ open, $x_0 \in \Omega$, and $\delta > 0$ such that $B_\delta(x_0) \subset \Omega$. Then for any function $f \in C^k(\Omega, W)$ and $x \in B_\delta(x_0)$*

$$f(x) = f(x_0) + Df|_{x_0}(x - x_0) + \frac{1}{2}D^2 f|_{x_0}(x - x_0, x - x_0) + \ldots + \frac{1}{k!}D^k f|_{x_0}(x - x_0, \ldots, x - x_0) + o(\|x - x_0\|_V^k)\,.$$

Definition 4.17 (Local extremum). Let X be a topological space and $f : X \to \mathbb{R}$. A point $x_0 \in X$ is called a *local maximum* of f, if

$$\exists U \subset \mathcal{U}(x_0) : \forall x \in U \backslash \{x_0\} : \quad f(x) \leq f(x_0)$$

If the strict inequality $<$ holds, we call it a *strict local maximum*. The definitions of *local minimum* and *strict local minimum* are analogous, by reversing the direction of the inequalities.

Proposition 4.18. *Let $\Omega \subset V$ and $f \in C^1(\Omega, \mathbb{R})$ have a local extremum at $x_0 \in \Omega$. Then $Df|_{x_0} = 0$.*

Proposition 4.19. *Let $\Omega \subset V$ and $f \in C^2(\Omega, \mathbb{R})$ and $x_0 \in \Omega$ such that $Df|_{x_0} = 0$.*

1. *If $D^2 f|_{x_0}(h, h) > 0$ $\forall h \in V \backslash \{0\}$, then f has a strict local minimum at x_0.*
2. *$D^2 f|_{x_0}(h, h) < 0$ $\forall h \in V \backslash \{0\}$, then f has a stric local maximum at x_0.*
3. *If $D^2 f|_{x_0}$ is indefinite, then f has no local extremum at x_0.*

Theorem 4.20 (Fundamental theorem of calculus). *Let $f : [a, b] \to \mathbb{R}$ be continuous in $[a, b]$ and differentiable in (a, b). Then ,*

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

Theorem 4.21 (Mean value theorem). *For a function $f : [a, b] \to \mathbb{R}$ continuous and differentiable on (a, b). Then $\exists x_0 \in (a, b)$:*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}\,.$$

Theorem 4.22 (FTC in higher dimension). *Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{K}^n$ be continuously differentiable. Let $\gamma : [a, b] \to \Omega$ continuously differentiable. Then*

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b \underbrace{Df|_{\gamma(t)}}_{\mathbb{K}^n \leftarrow \mathbb{R}^n} \circ \underbrace{\gamma'(t)}_{\in \mathbb{R}^n}\, dt\,.$$

Here the integral of a vector should be understood as the integral of each component, i.e.

$$\begin{aligned} f(\gamma(b))_i - f(\gamma(a))_i &= \int_a^b \sum_{j=1}^n (Df|_{\gamma(t)})_{ij} (\gamma'(t))_j \, dt \\ &= \sum_{j=1}^n \int_a^b \frac{\partial f_i}{\partial x_j}(\gamma(t)) \frac{d\gamma_j}{dt}(t)\, dt \end{aligned}$$

for every $i = 1, \dots m$.

Theorem 4.23 (MVT in higher dimension)**.** $\Omega \subset \mathbb{R}^n$, $f \in C^1(\Omega, \mathbb{K}^m)$. *Let* $x, h \in \mathbb{R}^n$ *such that* $\{x + th \mid t \in [0,1]\} \subset \Omega$. *Then*

$$f(x+h) - f(x) = \left(\int_0^1 Df|_{x+th}\, dt \right) \cdot h \, .$$

Corollary 4.24. *The setup is as in the Theorem 4.23. Then*

$$\|f(x) - f(y)\| \leq \underbrace{\left\| \int_0^1 Df|_{x+th}\, dt \right\|_{op}}_{\sup_{z \in \overline{xy}} \|Df|_z\|} \cdot \|x - y\|$$

For $f : \mathbb{R}^n \to \mathbb{R}$ *we obtain again*

$$f(y) - f(x) = Df|_z \cdot (y - x)$$

Definition 4.25 (Equivalence of norms)**.**
Two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on a vector space V are equivalent, if $\exists c, C > 0 \forall x \in V$:

$$c\|x\|_a \leq \|x\|_b \leq C\|x\|_a \, .$$

Theorem 4.26. *On finite dimensional vector spaces, all norms are equivalent.*

Theorem 4.27. *All finite dimensional normed spaces are complete (Banach spaces).*

Definition 4.28 (Frechet derivative)**.**
Let X and Y be Banach spaces and $\Omega \subset X$ open. A map $f : \Omega \to Y$ is *Frechet differentiable* at $x \in \Omega$, if there exists a **continuous** linear map $A : X \to Y$ such that

$$f(x+h) = f(x) + Ah + o(\|h\|_X)$$

for h in a neighbourhood of $0 \in X$. The notation $A = Df|_x$ remains.

Example 4.29. $X = C^2([0,T],\mathbb{R}^n)$. An element $x \in X$ is a map $x : [0,T] \to \mathbb{R}^n$. We can equip this space with a norm

$$\|x\|_X = \|x\|_\infty + \|\dot{x}\|_\infty + \|\ddot{x}\|_\infty$$

an turn it into a Banach space with respect to that norm. The action is given by

$$S : X \to \mathbb{R} \quad x \mapsto S(x) = \int_0^T L(x(t),\dot{x}(t))\,dt\,,$$

where $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (q,v) \mapsto L(q,v) \in C^2$ (e.g. $L(q,v) = \frac{1}{2}m\|v\|^2 - V(q)$). We compute the derivative $DS|_x$ of S: $x,h \in X$

$$\begin{aligned}
S(x+h) &= \int_0^T L\left(x(t)+h(t),\dot{x}(t)+\dot{h}(t)\right)\,dt \\
&= \int_0^T \left(L(x(t),\dot{x}(t)) + \left\{D_qL|_{(x(t),\dot{x}(t))}\cdot h(t) + D_vL|_{(x(t),\dot{x}(t))}\cdot \dot{h}(t)\right\}\right) + o(\|h\|_X^2) \\
&= S(x) + \underbrace{\int_0^T \left(D_qL|_{(x(t),\dot{x}(t))}\cdot h(t) + D_vL|_{(x(t),\dot{x}(t))}\cdot \dot{h}(t)\right)}_{DS|_x\cdot h} + o(\|h\|_X^2) \\
&= S(x) + D_vL|_{(x(T),\dot{x}(T))}\cdot h(T) - D_vL|_{(x(0),\dot{x}(0))} \\
&+ \int_0^T \left(D_qL|_{(x(t),\dot{x}(t))} - \left(\frac{d}{dt}D_vL|_{(x(t),\dot{x}(t))}\right)\right) h(t)\,dt + o(\|h\|_X^2)
\end{aligned}$$

for $h \in X$ such that $h(0) = h(T) = 0$

$$DS|_x h = 0 \quad \Leftrightarrow \quad D_qL|_{(x(t),\dot{x}(t))} - \frac{d}{dt}D_vL|_{(x(t),\dot{x}(t))} = 0$$

the Euler-Lagrange equation.

Exercises

1. Let $f : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ and $F : \mathbb{R}^3 \to R^3$ be defined by

$$f(x) = \|x\|_2 \,, \quad F(x) = \lambda x - x_0$$

for some $\lambda \in \mathbb{R}$ and $x_0 \in \mathbb{R}^3$. Compute ∇f, $\nabla^2 f$, Δf, $\operatorname{div} F$ and $\operatorname{curl} F$.

2. Let $f : V \to W$ be a map between finite dimensional normed spaces and fix $v_0 \in V$. Show that there exists at most one linear map $L \in \mathcal{L}(V, W)$ such that

$$\lim_{\|v\|_V \to 0} \frac{\|f(v_0 + v) - f(v_0) - L(v)\|_W}{\|v\|_V} = 0\,.$$

3. Fix $\alpha \in (0, 1)$. We say that $f \in C^{0,\alpha}([a, b])$ if

$$\exists C > 0 \; \forall x, y \in [a, b] \; : \quad |f(x) - f(y)| \leq C|x - y|^\alpha \,.$$

Show the following inclusions:

$$C^1([a, b]) \subset C^{0,1}([a, b]) \subset C^{0,\alpha}([a, b]) \subset C([a, b])\,.$$

4. Let $f : \mathbb{R}^n \supset B_r(0) \to \mathbb{R}$ be a C^1 function such that

$$\sup_{x \in B_r(0)} \|\nabla f(x)\|_2 \leq \frac{1}{r}\,.$$

Show that if there is some $x \in B_r(0)$ such that $f(x) = 0$, then $\|f\|_\infty \leq 1$.

Chapter 5

Implicit functions and ordinary differential equations

Implicit function theorem

Say we have a system of m algebraic equations on n variables

$$
\begin{aligned}
F_1(x_1, \dots, x_n) &= 0 \\
&\vdots \\
F_m(x_1, \dots, x_n) &= 0
\end{aligned}
$$

In the case of linear equations, if $n = m$, basic linear algebra tells us that the solvability depends on the degeneracy of the coefficient matrix, whereas if $n < m$, the degeneracy of a coefficient sub-matrix determines the parametrizability of the space solutions.
In the nonlinear case, one simply "linearizes" the problem around a point and obtains a similar statement locally. Consider a function

$$F : \underbrace{\mathbb{R}^n \times \mathbb{R}^m}_{\mathbb{R}^{n+m}} \to \mathbb{R}^m, \qquad (x, y) \mapsto F(x, y)$$

and think of level sets as solutions to a system of algebraic equations, i.e.

$$F(x, y) = 0 \quad \iff \begin{cases} F_1(x_1, \dots, x_n, y_1, \dots, y_m) & = 0 \\ \vdots & \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) & = 0 \end{cases}$$

where we want to solve for the $(y_1, \dots, y_m)$ variables in terms of the extra $(x_1, \dots, x_n)$ parameters.

Theorem 5.1 (Implicite function theorem)**.** *Let $\Omega \subset \mathbb{R}^{n+m}$ be open, $F \in C^1(\Omega, \mathbb{R}^m)$, and*

$$N \doteq \{(x, y) \in \Omega \mid F(x, y) = 0\}.$$