# 4 Differentials in normed vector spaces

## 4.1 Differential and partial derivatives

In the first year calculus, we have seen the notion of derivative of a function $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval. In particular, Taylor's formula allows us to develop $f$ around $x \in I$ in the following way,

$$f(x+h) = f(x) + hf'(x) + o(h),$$

where the term $h \mapsto hf'(x)$ is a linearisation of $f$ around $x$. If the function takes values in a higher dimensional Euclidean space such as $\mathbb{R}^n$, similar theories can also be developped. Below, we are going to see how to generalize these notions to functions from an open subset of a normed vector space with values in another normed vector space.

### 4.1.1 Differential

Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector spaces. Let us consider an open set $A \subseteq V$ and $f : A \to W$.

**Definition 4.1.1 :** Let $a \in A$. We say that $f$ is *differentiable*[1] at $a$ if there exists $\varphi \in \mathcal{L}_c(V, W)$ such that

$$f(a+h) = f(a) + \varphi(h) + o(\|h\|_V), \quad \text{when} \quad h \to 0. \tag{4.1}$$

If such a map $\varphi$ exists, it is unique, and is called the differential (微分) of $f$ at $a$, denoted by $Df(a)$ or $\mathrm{d}f_a$.

**Remark 4.1.2 :** Since $A$ is an open set and $a$ is an interior point, for $h$ close enough to $0$, we know that $a+h$ is also in $A$. Therefore, the condition "when $h \to 0$ is important in Eq. (4.1), since the relation only makes sense when $h$ is close enough to $0$.

**Definition 4.1.3 :** If $f$ is differentiable at every $a \in A$, we say that $f$ is differentiable on $A$, and the map

$$\begin{array}{rccc} Df : & A & \to & \mathcal{L}_c(V, W) \\ & a & \mapsto & \mathrm{d}f_a \end{array}$$

is called the *differential map* of $f$. If $Df$ is continuous, we say that $f$ is of class $\mathcal{C}^1$.

[1] Also known as Fréchet differentiable. In Exercise 4.10 we will see a more general notion of differentiability, called Gâteaux differentiability.

**Remark 4.1.4 :**

(1) If $V = \mathbb{R}$, then the notion corresponds to the classical notion of derivative, that is the continuous linear map $Df(a)$ writes $Df(a)(h) = \mathrm{d}f_a(h) = f'(a)h$. So we may also just write $Df(a) = \mathrm{d}f_a = f'(a)$.

(2) In general, the definition of $\mathrm{d}f_a$ may depend on the norms $\|\cdot\|_V$ and $\|\cdot\|_W$. However, if $V$ and $W$ are finite dimensional vector spaces, we have seen in Theorem 3.2.22 that all the norms are equivalent, so the existence of $\mathrm{d}f_a$ does not depend on the norms that we equip on the spaces.

(3) It is important to require the differential $\mathrm{d}f_a$ to be a continuous map. In finite dimensional spaces, all the linear maps are continuous (Corollary 3.2.24), so in such spaces, we only need to check the linearity, then the continuity follows automatically.

**Example 4.1.5 :**

(1) If $f \in \mathcal{L}_c(V, W)$, then the relation $f(a+h) = f(a) + f(h)$ implies that $f$ is differentiable on $V$ with $\mathrm{d}f_a = f$ for every $a \in V$.

(2) Consider the product on $\mathbb{R}^2$,

$$\psi : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & xy \end{array}.$$

Then,

$$\psi(x + h_x, y + h_y) - \psi(x, y) = xh_y + h_x y + h_x h_y.$$

Since the map $(h_x, h_y) \mapsto xh_y + yh_x$ is linear, and $h_x h_y = o(\|(h_x, h_y)\|)$, we deduce that $\mathrm{d}\psi_{x,y}(h) = xh_y + yh_x$ for $h = (h_x, h_y) \in \mathbb{R}^2$.

(3) Consider the matrix product on $\mathcal{M}_n(\mathbb{R})$,

$$\psi : \begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) & \to & \mathcal{M}_n(\mathbb{R}) \\ (M,N) & \mapsto & MN \end{array}.$$

We equip the vector space $\mathcal{M}_n(\mathbb{R})$ with the norm $|||\cdot|||$ defined in Remark 3.2.16. Let $M, N \in \mathcal{M}_n(\mathbb{R})$ be fixed. Then, for $H, K \in \mathcal{M}_n(\mathbb{R})$, we have

$$\psi(M + H, N + K) - \psi(M, N) = MK + HN + HK.$$

The map $(H, K) \mapsto MK + HN$ is linear, and $|||HK||| \leqslant |||H||| |||K||| \leqslant \|(H, K)\|^2$. Therefore, we find $\mathrm{d}\psi_{M,N}(H, K) = MK + HN$.

**Example 4.1.6 :** Let $V$ be a normed vector space, and

$$\mathcal{GL}_c(V) = \{u \in \mathcal{L}(V, V) : u \text{ and } u^{-1} \text{ are continuous}\}.$$

Define the map $\mathrm{Inv} : \mathcal{GL}_c(V) \to \mathcal{GL}_c(V), u \mapsto u^{-1}$. For $h \in \mathcal{GL}_c(V)$ such that $|||h||| < 1$, we know that $\mathrm{id} + h$ is invertible with inverse

$$(\mathrm{id} + h)^{-1} = \mathrm{id} - h + \sum_{n \geqslant 2} (-1)^n h^n.$$

We have

$$\left\|\left|\sum_{n\geqslant 2}(-1)^n h^n\right|\right\| \leqslant \sum_{n\geqslant 2} |||h|||^n = \frac{|||h|||^2}{1-|||h|||}.$$

Thus, when $h \to 0$, we have

$$(\mathrm{id}+h)^{-1} = \mathrm{id} - h + o(|||h|||).$$

This means that Inv is differentiable at id with differential $\mathrm{d}\,\mathrm{Inv}_{\mathrm{id}} : h \mapsto -h$.

**Proposition 4.1.7 :** *If $f$ is differentiable at $a \in A$, then $f$ is also continuous at $a$.*

**Proof :** Suppose that $f$ is differentiable at $a \in A$. Then, we can find a continuous linear function $\varphi : V \to W$ and $r > 0$ such that

$$\forall h \in B_V(0,r), \quad f(a+h) = f(a) + \varphi(h) + \|h\|_V\, \varepsilon(h),$$

where $\lim_{h\to 0}\varepsilon(h) = 0$. Fix $\delta > 0$ and $0 < r' \leqslant r$ such that $\|\varepsilon(h)\|_W < \delta$ for $h \in B_V(0,r')$. Then,

$$\forall h \in B_V(0,r'), \quad \|f(a+h)-f(a)\|_W \leqslant \|\varphi(h)\|_W + \|h\|_V\,\delta \leqslant (M+\delta)\,\|h\|_V\,,$$

where $M = |||\varphi|||$. This implies the continuity of $f$ at $a$. □

**Proposition 4.1.8 :** *Let $V, W$ be two normed vector spaces, $A \subseteq V$ be an open subset, and $f, g : A \to W$ be two differentiable functions at $a \in A$. Then,*

*(1) $f+g$ is differentiable at $a$, and $\mathrm{d}(f+g)_a = \mathrm{d}f_a + \mathrm{d}g_a$,*

*(2) for every $\lambda \in \mathbb{K}$, $\lambda f$ is differentiable at $a$, and $\mathrm{d}(\lambda f)_a = \lambda\,\mathrm{d}f_a$.*

**Proof :** Complete the proof by yourself using directly the definition in Definition 4.1.1. □

**Proposition 4.1.9** (Chain rule) **:** *Let $V, W, X$ be normed $\mathbb{K}$-vector spaces, $A \subseteq V$ and $B \subseteq W$ be two open subsets. Consider two functions $f : A \subseteq V \to W$ and $g : B \subseteq W \to X$ satisfying $f(A) \subseteq B$. Suppose that $f$ is differentiable at $a \in A$ and $g$ is differentiable at $f(a)$. Then, $g \circ f : A \subseteq V \to X$ is differentiable at $a$, and we have*

$$\mathrm{d}(g\circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a. \tag{4.2}$$

**Remark 4.1.10 :** If $V = W = X = \mathbb{R}$, Eq. (4.2) becomes $(g\circ f)'(a) = g'(f(a)) \cdot f'(a)$, which is the chain rule we have seen in the first-year calculus.

**Proof :** By the differentiability of $f$ at $a$, we can write

$$f(a+h) = f(a) + \mathrm{d}f_a(h) + o(\|h\|_V), \quad \text{when } h \to 0.$$

When we compose with $g$ and by the differentiability of $g$ at $b = f(a)$, we get

$$\begin{aligned}(g \circ f)(a+h) &= g(\underbrace{f(a)}_{b} + \underbrace{\mathrm{d}f_a(h) + o(\|h\|_V)}_{h'}) \\ &= g(f(a)) + \mathrm{d}g_b(h') + o(\|h'\|_V).\end{aligned}$$

Since $\mathrm{d}f_a \in \mathcal{L}_c(V, W)$, by Theorem 3.2.12, we know that $h' = O(\|h\|_V)$. Similarly, due to the fact that $\mathrm{d}g_b \in \mathcal{L}_c(W, X)$, we have

$$\mathrm{d}g_b(h') = \mathrm{d}g_b \circ \mathrm{d}f_a(h) + \mathrm{d}g_b(o(\|h\|_V)) = \mathrm{d}g_b \circ \mathrm{d}f_a(h) + o(\|h\|_V),$$

and the map $\mathrm{d}g_b \circ \mathrm{d}f_a$ is linear and continuous being composition of such functions. In consequence,

$$(g \circ f)(a+h) = (g \circ f)(a) + \mathrm{d}g_b \circ \mathrm{d}f_a(h) + o(\|h\|_V), \quad \text{when } h \to 0.$$

implying that $\mathrm{d}(g \circ f)_a = \mathrm{d}g_b \circ \mathrm{d}f_a$. □

**Corollary 4.1.11 :** *Let $f, g : A \subseteq V \to \mathbb{R}$ be differentiable at $a \in A$, then the product $fg$ is also differentiable at $a$, and*

$$\mathrm{d}(fg)_a = g(a) \cdot \mathrm{d}f_a + f(a) \cdot \mathrm{d}g_a.$$

**Proof :** It is a direct application of Proposition 4.1.9. Actually, let us consider the functions

$$\varphi : \begin{array}{ccc} A & \to & \mathbb{R}^2 \\ x & \mapsto & (f(x), g(x)) \end{array}, \quad \text{and} \quad \psi : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & xy \end{array}.$$

Then, the product $fg$ is the composition $x \mapsto (\psi \circ \varphi)(x)$, and we have

$$\begin{aligned}\mathrm{d}\varphi_x(h) &= (\mathrm{d}f_x(h), \mathrm{d}g_x(h)) \\ \mathrm{d}\psi_{x,y}(h_x, h_y) &= h_x y + h_y x.\end{aligned}$$

Therefore, by composition, we find, for $h \to 0$,

$$\mathrm{d}(fg)_a(h) = \mathrm{d}\psi_{\varphi(a)} \circ \mathrm{d}\varphi_a(h) = g(a)\,\mathrm{d}f_a(h) + f(a)\,\mathrm{d}g_a(h)$$

□

### 4.1.2 Mean-value theorem

We recall from the first-year calculus that for a continuous and differentiable function $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an open interval, we have the *mean-value theorem* stated as below. For $a, b \in I$ with $a < b$, there exists $c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a). \tag{4.3}$$

In particular, if we know that $\sup_{t\in[a,b]} |f'(c)| \leqslant M$, then $|f(b)-f(a)| \leqslant M(b-a)$, which is known as the *mean-value inequality*. Below, we are going to generalize the mean-value theorem and the mean-value inequality to functions defined on an open subset of a normed vector space, with values in another normed vector space.

**Lemma 4.1.12 :** *Let $a<b$ be real numbers, and $W$ be a normed vector space. Let $f:[a,b]\to W$ and $g:[a,b]\to\mathbb{R}$ be two continuous functions on $[a,b]$ and differentiable on $(a,b)$. If $\|f'(t)\|_W \leqslant g'(t)$ for all $t\in(a,b)$, then $\|f(b)-f(a)\|_W \leqslant g(b)-g(a)$.*

**Proof :** First, let us assume that $\|f'(t)\|_W < g'(t)$ for all $t\in(a,b)$. This means that,

$$\begin{aligned} & \forall t\in(a,b), \quad \lim_{\substack{x\to t\\ x>t}} \left\|\frac{f(x)-f(t)}{x-t}\right\|_W - \frac{g(x)-g(t)}{x-t} < 0 \\ \Rightarrow\quad & \forall t\in(a,b), \exists y>t, \forall x\in(t,y), \quad \left\|\frac{f(x)-f(t)}{x-t}\right\|_W < \frac{g(x)-g(t)}{x-t} \\ \Rightarrow\quad & \forall t\in(a,b), \exists y>t, \forall x\in[t,y], \quad \|f(x)-f(t)\|_W \leqslant g(x)-g(t). \end{aligned} \tag{4.4}$$

Let $[\alpha,\beta]\subseteq(a,b)$, and we want to show that

$$\|f(\beta)-f(\alpha)\|_W \leqslant g(\beta)-g(\alpha). \tag{4.5}$$

Let

$$\Gamma = \{\theta\in(\alpha,\beta] : \forall x\in[\alpha,\theta], \|f(x)-f(\alpha)\|_W \leqslant g(x)-g(\alpha)\}.$$

It follows from Eq. (4.4) that $\Gamma$ is nonempty. Let $\gamma=\sup\Gamma$, and we want to show that $\gamma=\beta$, which will imply Eq. (4.5).

We prove by contradiction. Suppose that $\gamma<\beta$. Since $f$ and $g$ are continuous, we also have

$$\|f(\gamma)-f(\alpha)\|_W \leqslant g(\gamma)-g(\alpha). \tag{4.6}$$

But from Eq. (4.4), we know that

$$\exists\delta\in(\gamma,\beta], \forall x\in[\gamma,\delta], \quad \|f(x)-f(\gamma)\|_W \leqslant g(x)-g(\gamma). \tag{4.7}$$

Then, it follows from Eq. (4.6) and Eq. (4.7) that there exists $\delta\in(\gamma,\beta]$ such that

$$\forall x\in[\gamma,\delta], \quad \|f(x)-f(\alpha)\|_W \leqslant g(x)-g(\alpha).$$

This shows that $\delta\in\Gamma$, which is not possible because we assumed that $\delta>\gamma=\sup\Gamma$. Therefore, Eq. (4.5) is true. Then, we may take $\alpha\to a$ and $\beta\to b$ in Eq. (4.5), and by continuity of $f$ and $g$, we also have $\|f(b)-f(a)\|_W \leqslant g(b)-g(a)$.

To conclude, we need to deal with the case with the original hypothesis $\|f'(t)\|_W \leqslant g'(t)$ for all $t\in(a,b)$. Fix $\varepsilon>0$, we may consider $g_\varepsilon(t)=g(t)+\varepsilon t$ for $t\in[a,b]$. Then, $\|f'(t)\|_W < g'_\varepsilon(t)$ for $t\in(a,b)$. We may apply the above arguments to obtain $\|f(b)-f(a)\|_W \leqslant g_\varepsilon(b)-g_\varepsilon(a)$. By taking $\varepsilon\to 0$, we find the desired result. □

**Theorem 4.1.13** (Mean-value inequality) **:** *Let $V$ and $W$ be two normed vector spaces, and $A \subseteq V$ be an open subset. Let $f : A \subseteq V \to W$ be a function. Consider $a, b \in A$ such that the line segment $[a, b] \subseteq A$. Suppose that*

*(a) $f$ is continuous on $[a, b]$,*

*(b) $f$ is differentiable on $(a, b)$,*

*(c) there exists $M > 0$ such that $|||\mathrm{d}f_c||| \leqslant M$ for $c \in (a, b)$.*

*Then,*

$$\|f(b) - f(a)\|_W \leqslant M \|b - a\|_V . \tag{4.8}$$

**Proof :** Let $g : [0, 1] \to W$ be defined by $g(t) = f(a + t(b - a))$ for $t \in [0, 1]$. Then, $g$ is continuous on $[0, 1]$ and differentiable on $(0, 1)$, with derivative

$$g'(t) = \mathrm{d}f_{a+t(b-a)}(b - a), \quad \forall t \in (a, b).$$

Therefore, $\|g'(t)\|_W \leqslant M \|b - a\|_V$ for $t \in (0, 1)$. By Lemma 4.1.12, we find the desired result. □

**Remark 4.1.14 :** We note that here in general normed vector spaces (dimension larger or equal to 2), the best result we can get is only an inequality, even when the operator norm of the differential is always equal to $M$ in the condition (c) of Theorem 4.1.13. We may consider for example the map

$$\begin{array}{rccc} f : & \mathbb{R} & \to & \mathbb{R}^2 \\ & t & \mapsto & (\cos t, \sin t). \end{array}$$

It is not hard to check that for every $t \in \mathbb{R}$, we have $\mathrm{d}f_t = (-\sin t, \cos t)$ which satisfies $|||\mathrm{d}f_t||| = 1$. However, we have $\|f(0) - f(2\pi)\| = 0 \neq 2\pi \cdot 1$.

**Theorem 4.1.15** (Mean-value theorem) **:** *Let $V$ be a normed vector space and $W = \mathbb{R}^n$ be an Euclidean space, and $A \subseteq V$ be an open subset. Consider a function $f : A \subseteq V \to \mathbb{R}^n$ that is differentiable on $A$. Let $a, b \in A$ such that $[a, b] \subseteq A$. Then, for any vector $v \in \mathbb{R}^n$, there exists $c \in (a, b)$ such that*

$$v \cdot [f(b) - f(a)] = v \cdot \mathrm{d}f_c(b - a). \tag{4.9}$$

**Proof :** Let $h = b - a$. Since $A$ is open and $[a, a + h] \subseteq A$, there exists $\delta > 0$ such that $a + th \in A$ for $t \in (-\delta, 1 + \delta)$. Fix a vector $v \in \mathbb{R}^n$ and let $g : (-\delta, 1 + \delta) \to \mathbb{R}$ be defined by

$$g(t) = v \cdot f(a + th), \quad \forall t \in (-\delta, 1 + \delta).$$

Then, $f$ is differentiable on $(-\delta, 1 + \delta)$ and its derivative writes

$$g'(t) = v \cdot \mathrm{d}f_{a+th}(h).$$

By the classical one-dimensional mean-value theorem (Eq. (4.3)), we have

$$g(1) - g(0) = g'(t), \quad \text{for some } t \in (0,1),$$

which is exactly Eq. (4.9). □

### 4.1.3 Directional derivative

**Definition 4.1.16 :** Let $a \in A$. We say that the *directional derivative* of $f$ at $a$ in the direction $u \in V$ exists, denoted by $f'_u(a)$, if the following limit exists

$$f'_u(a) = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}. \tag{4.10}$$

**Proposition 4.1.17 :** *If $f$ is differentiable at $a$, then its directional derivative at $a$ in any direction $u \in V$ is well defined, and we have $f'_u(a) = \mathrm{d}f_a(u) = Df(a)(u)$.*

**Remark 4.1.18 :** We note that if the directional derivative of $f$ at $a$ in any direction exists, it does not necessarily imply that $f$ is differentiable at $a$. Actually, even the continuity at $a$ does not hold in general. We may consider $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = \begin{cases} \frac{y^2}{x}, & \text{if } x \neq 0, \\ y, & \text{if } x = 0. \end{cases}$$

Then, $f$ is not continuous at $(0,0)$ because for example,

$$\lim_{x \to 0} f(x, \sqrt{x}) = 1 \neq 0 = f(0,0).$$

However, for any $u = (a,b) \in \mathbb{R}^2$, the direction derivative of $f$ at $(0,0)$ in the direction $u$ exists,

$$f'_u(0,0) = \lim_{h \to 0} \frac{f(h(a,b)) - f(0,0)}{h} = \begin{cases} \frac{b^2}{a}, & \text{if } a \neq 0, \\ b, & \text{if } a = 0. \end{cases}$$

Below, let us take $V = \mathbb{R}^n$ to be the $n$-dimensional Euclidean space, with the canonical basis given by $(e_1, \ldots, e_n)$. Let $A$ be an open subset of $V$, and $f : A \to W$.

**Definition 4.1.19 :** For $1 \leqslant i \leqslant n$, if the directional derivative of $f$ at $a$ in the direction $e_i$ exists, we say that its partial derivative at $a$ with rerspect to the $i$-th coordinate exists and define

$$\frac{\partial f}{\partial x_i}(a) = f'_{e_i}(a) \tag{4.11}$$

**Remark 4.1.20 :**

(1) Following Remark 4.1.18, it is possible that all the partial derivatives of $f$ at $a$ exist without $f$ being differentiable or continuous at $a$.

(2) If $A \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $a \in A$, then all the partial derivatives at $a$ exist, and

$$Df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\,\mathrm{d}x_i, \quad \overrightarrow{\mathrm{grad}}_a\, f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) e_i,$$

where $(\mathrm{d}x_i = e_i^*)_{1 \leqslant i \leqslant n}$ is the dual basis in $(\mathbb{R}^n)^* = \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ of the canonical basis $(e_i)_{1 \leqslant i \leqslant n}$ of $\mathbb{R}^n$, that is

$$\mathrm{d}x_i(e_j) = e_i^*(e_j) = \delta_{i,j}, \quad \forall 1 \leqslant i, j \leqslant n.$$

In particular, we have

$$Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i = \big(\overrightarrow{\mathrm{grad}}_a\, f\big) \cdot h. \tag{4.12}$$

**Theorem 4.1.21 :** *Let $f : A \subseteq \mathbb{R}^n \to W$. Suppose that*

*(a) all the partial derivatives of $f$ exist on $A$,*

*(b) the partial derivatives are continuous at $a$.*

*Then, $f$ is differentiable at $a$ with*

$$Df(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\,\mathrm{d}x_i. \tag{4.13}$$

**Remark 4.1.22 :** We recall that $Df(a)$ is a linear map from $\mathbb{R}^n$ to $W$. For each $1 \leqslant i \leqslant n$, the partial derivative $\frac{\partial f}{\partial x_i}(a)$ is a vector in $W$, $\mathrm{d}x_i$ is a linear form on $\mathbb{R}^n$, that is a linear (continuous) function from $\mathbb{R}^n$ to $\mathbb{R}$. If we evaluate Eq. (4.13) at $u \in \mathbb{R}^n$, the left-hand side gives us $Df(a)(u) \in W$, and each term on the right-hand side gives us a scalar $\mathrm{d}x_i(u) = u_i \in \mathbb{R}$, multiplied by the vector $\frac{\partial f}{\partial x_i}(a) \in W$.

**Proof :** We equip $\mathbb{R}^n$ with the norm $\|x\| = \sum_{i=1}^n |x_i|$. Let

$$\begin{array}{rccc} g: & A & \to & W \\ & x & \mapsto & f(x) - \displaystyle\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(a). \end{array}$$

We want to show that when $x \to a$, we have $g(x) - g(a) = o(\|x - a\|)$.

Let $\varepsilon > 0$. The continuity in assumption (b) guarantees that there exists $r > 0$ such that for $1 \leqslant i \leqslant n$, we have

$$\forall x \in A \cap B(a,r), \quad \left\| \frac{\partial g}{\partial x_i}(x) \right\|_W = \left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a) \right\|_W < \varepsilon. \tag{4.14}$$

Since $A$ is an open set, by choosing a smaller $r > 0$, we may assume that $B(a,r) \subseteq A$.

For $x \in B(a,r)$, we consider the following points

$$\begin{aligned} y_0 &= (a_1, \dots, a_n) = a, \\ y_k &= (x_1, \dots, x_k, a_{k+1}, \dots, a_n), \quad \forall k = 1, \dots, n. \end{aligned}$$

We note that $y_0 = a$, $y_n = x$, and the intermediate $y_k$'s are obtained by replacing coordinates of $a$ by those of $x$ one by one. For $1 \leqslant k \leqslant n$, define

$$\begin{array}{rccc} g_k : & [a_k, x_k] & \to & W \\ & t & \mapsto & g(x_1, \dots, x_{k-1}, t, a_{k+1}, \dots, a_n). \end{array}$$

The derivative of $g_k$ writes

$$g_k'(t) = \frac{\partial g}{\partial x_k}(x_1, \dots, x_{k-1}, t, a_{k+1}, \dots, a_n),$$

and it follows from Eq. (4.14) that $\|g_k'(t)\|_W < \varepsilon$ on $[a_k, x_k]$. Therefore, it follows from Lemma 4.1.12 that

$$\|g_k(x_k) - g_k(a_k)\|_W \leqslant \varepsilon |x_k - a_k|.$$

Since $g_k(a_k) = g(y_{k-1})$ and $g_k(x_k) = g(y_k)$, we get

$$\begin{aligned} \|g(x) - g(a)\|_W = \left\| \sum_{k=1}^{n} [g(y_k) - g(y_{k-1})] \right\|_W &\leqslant \sum_{k=1}^{n} \|g(y_k) - g(y_{k-1})\|_W \\ &\leqslant \varepsilon \sum_{k=1}^{n} |x_k - a_k| = \varepsilon \|x - a\|. \end{aligned}$$

Thus, we have obtained

$$\forall x \in B(a,r), \quad \|g(x) - g(a)\|_W \leqslant \varepsilon \|x - a\|.$$

Or equivalently, $g(x) - g(a) = o(\|x - a\|)$. □

**Remark 4.1.23 :** Note that the converse of Theorem 4.1.21 is false. We have functions which are differentiable whose partial derivatives need not to be continuous. For example, consider the classical example $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 \sin(\frac{1}{x}), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

We can compute the derivative of $f$ at $0$ as below,

$$f'(0) = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} h \sin(\tfrac{1}{h}) = 0.$$

However, the derivative of $f$ at $x \neq 0$ writes

$$f'(x) = 2x \sin(\tfrac{1}{x}) - \cos(\tfrac{1}{x}).$$

And clearly, the $f'$ is not continuous at $0$.

### 4.1.4 Jacobian matrix

We look at the special case where our normed vector spaces are taken to be Euclidean spaces, that is $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ for some $n, m \geqslant 1$. Let $(v_1, \ldots, v_n)$ be the canonical basis of $V = \mathbb{R}^n$ and $(w_1, \ldots, w_m)$ be the canonical basis of $W = \mathbb{R}^m$. Let $A \subseteq \mathbb{R}^n$ be an open subset and $f : A \to \mathbb{R}^m$ be a differentiable function at $a \in A$. Since $\mathrm{d}f_a \in \mathcal{L}_c(\mathbb{R}^n, \mathbb{R}^m)$, it can also be represented by an $m \times n$ real-valued matrix using the canonical bases, that is, with matrix coefficients given by

$$\mathrm{d}f_a(v_j) \cdot w_i, \quad 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n.$$

**Definition 4.1.24 :** The *Jacobian matrix* of $f$ at $a$ is the matrix $J_f(a) \in \mathcal{M}_{m,n}(\mathbb{R})$, given by

$$J_f(a) = \left[\frac{\partial f_i}{\partial x_j}(a)\right]_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}},$$

where $f_i = \mathrm{Proj}_i \circ f$ for $1 \leqslant i \leqslant m$ and $f = \sum_{i=1}^m f_i w_i$. When $m = n$, the Jacobian matrix is a square matrix, and we call its determinant $\det(J_f(a))$ the *Jacobian determinant* or simply the *Jacobian*.

**Remark 4.1.25 :** We note that the $i$-th row of the Jacobian matrix $J_f(a)$ is the gradient of $f_i$, that is

$$\overrightarrow{\mathrm{grad}}_a\, f_i = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a) v_j, \quad \text{or} \quad \left(\overrightarrow{\mathrm{grad}}_a\, f_i\right)_{v_1, \ldots, v_n} = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1 \leqslant j \leqslant n}.$$

We may also write the differential of $f$ at $a$ as follows, using Eq. (4.12), we find, for all $h \in \mathbb{R}^n$, that

$$Df(a)(h) = \sum_{i=1}^m Df_i(a)(h) w_i = \sum_{i=1}^m \left[\left(\overrightarrow{\mathrm{grad}}_a\, f_i\right) \cdot h\right] w_i.$$

This is exactly the matrix multiplication between $J_f(a)$ and $h$, where the vector $h$ is represented in the canonical basis $(v_1, \ldots, v_n)$ as an $n \times 1$ column matrix, and the resulting matrix is an $m \times 1$ matrix, which is $Df(a)(h)$ represented in the canonical basis $(w_1, \ldots, w_m)$ of $\mathbb{R}^m$.

**Proposition 4.1.26** (Composition and Jacobian matrices) **:** *Let $m, n, k \geqslant 1$ and $A \subseteq \mathbb{R}^m$, $B \subseteq \mathbb{R}^n$ be two open subsets. Let $f : A \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ be such that $f(A) \subseteq B$. Suppose that $f$ is differentiable at $a$ and $g$ is differentiable at $f(a)$. For $1 \leqslant i \leqslant n$, we also write $f_i = \mathrm{Proj}_i \circ f$ to be the $i$-th coordinate of the function $f$. Then, the function $h = g \circ f : A \to \mathbb{R}^k$ is differentiable at $a$ and its Jacobian matrix writes*

$$J_h(a) = J_g(f(a)) \cdot J_f(a).$$

*Alternatively, we may also write, for $1 \leqslant j \leqslant m$,*

$$\frac{\partial h}{\partial x_j}(a) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_j}(a) \frac{\partial g}{\partial y_i}(f(a)).$$

**Proof :** It is a direct consequence of Proposition 4.1.9 written in terms of the Jacobian matrices defined in Definition 4.1.24. □

**Example 4.1.27 :** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a $\mathcal{C}^1$ function. Consider the map

$$\begin{array}{rccc} \varphi: & \mathbb{R}_{>0} \times \mathbb{R} & \to & \mathbb{R}^2 \\ & (r,\theta) & \mapsto & (r\cos\theta, r\sin\theta). \end{array}$$

Then, the composition $F = f \circ \varphi$ is a $\mathcal{C}^1$ function, and can be seen as the function $f$ written in the polor coordinates. We have

$$\begin{aligned} & J_F(r,\theta) = J_f(r\cos\theta, r\sin\theta) J_\varphi(r,\theta) \\ \Leftrightarrow \quad & \begin{pmatrix} \frac{\partial F}{\partial r} & \frac{\partial F}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial \varphi_1}{\partial r} & \frac{\partial \varphi_1}{\partial \theta} \\ \frac{\partial \varphi_2}{\partial r} & \frac{\partial \varphi_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}. \end{aligned}$$

In other words,

$$\frac{\partial f}{\partial x} = \cos\theta \frac{\partial F}{\partial r} - \frac{\sin\theta}{r}\frac{\partial F}{\partial \theta} \quad \text{and} \quad \frac{\partial f}{\partial y} = \sin\theta \frac{\partial F}{\partial r} + \frac{\cos\theta}{r}\frac{\partial F}{\partial \theta}.$$

## 4.2 Higher-order derivatives

In this subsection, we will focus on the case of finite dimensional vector spaces. However, we will still mention a generalization of higher-order differentials to general normed vector spaces in Section 4.2.3.

### 4.2.1 Schwarz theorem

Let $A$ be an open subset of $\mathbb{R}^n$ and $f : A \to W$ be a function. Let $p \geqslant 1$ be an integer, and $1 \leqslant i_1, \dots, i_p \leqslant n$. We may define the partial derivative of order $p$ by induction, under the assumption of existence,

$$\frac{\partial^p f}{\partial x_{i_p} \dots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_p}} \Big( \frac{\partial^{p-1} f}{\partial x_{i_{p-1}} \dots \partial x_{i_1}} \Big).$$

We say that $f$ is of class $\mathcal{C}^p$ if all its partial derivatives up to order $p$ exist and are continuous on $A$.

**Theorem 4.2.1** (Schwarz theorem) **:** *Let $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a function, where $A \subseteq \mathbb{R}^2$ is an open subset. Suppose that the partial derivative*

$$\frac{\partial^2 f}{\partial x \partial y} \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial x}$$

*exist on $A$, and are continuous at $a \in A$. Then,*

$$\frac{\partial^2 f}{\partial x \partial y}(a) = \frac{\partial^2 f}{\partial y \partial x}(a). \tag{4.15}$$

**Remark 4.2.2 :** It follows from the above theorem that, under the assumption of existence and continuity, the order of partial derivative does not count.

**Example 4.2.3 :** This example is due to Peano. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

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

On one hand,

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Therefore, $\frac{\partial f}{\partial x}(0,y) = -y$ for $y \in \mathbb{R}$, giving us

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1.$$

On the other hand,

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Therefore, $\frac{\partial f}{\partial y}(x,0) = x$ for $x \in \mathbb{R}$, giving us

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$$

Actually, one can easily check that, the second partial derivatives are not continuous. In fact, for $(x,y) \neq (0,0)$, we have

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2+y^2)^3},$$

which gives

$$\lim_{x \to 0} \frac{\partial^2 f}{\partial y \partial x}(x,0) = 1, \quad \text{and} \quad \lim_{y \to 0} \frac{\partial^2 f}{\partial y \partial x}(0,y) = -1.$$

You may also see this discontinuity using an antisymmetry argument, without doing computations.

**Proof :** Without loss of generality, we may assume that $a = (0,0) \in A$. Let $h, k > 0$ such that $[0,h] \times [0,k] \subseteq A$ and

$$\delta(h,k) = f(h,k) - f(h,0) - f(0,k) + f(0,0).$$

Consider the function $\varphi$ defined by

$$\begin{array}{rccc} \varphi : & [0,h] & \to & \mathbb{R} \\ & x & \mapsto & f(x,k) - f(x,0). \end{array}$$

Then, $\delta(h,k) = \varphi(h) - \varphi(0)$. Since $\varphi$ is continuous on $[0,h]$ and differentiable on $(0,h)$, it follows from the mean-value inequality on $\mathbb{R}$ (Eq. (4.3)) that there exists $t_1 \in (0,1)$ such that

$$\delta(h,k) = h\varphi'(t_1 h) = h\Big[\frac{\partial f}{\partial x}(t_1 h, k) - \frac{\partial f}{\partial x}(t_1 h, 0)\Big].$$

The function $y \mapsto \frac{\partial f}{\partial x}(t_1 h, y)$ is continuous on $[0,1]$ and differentiable on $(0,1)$, it follows again from the mean-value inequality that there exists $t_2 \in (0,1)$ such that

$$\delta(h,k) = hk\frac{\partial^2 f}{\partial y \partial x}(t_1 h, t_2 k). \tag{4.16}$$

If we consider the function $\psi$ defined by

$$\begin{array}{rccc} \psi : & [0,k] & \to & \mathbb{R} \\ & y & \mapsto & f(h,y) - f(0,y) \end{array}$$

and follow the same steps as above, we may find $t_3, t_4 \in (0,1)$ such that

$$\delta(h,k) = hk\frac{\partial^2 f}{\partial x \partial y}(t_3 h, t_4 k). \tag{4.17}$$

By putting Eq. (4.16) and Eq. (4.17) together and taking $h, k \to 0$, the continuity of the partial derivatives at $(0,0)$ implies that they are equal at $(0,0)$. ☐

**Corollary 4.2.4 :** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a function, where $A \subseteq \mathbb{R}^n$ is an open subset. Suppose that $f$ is of class $\mathcal{C}^p$, then the partial derivatives up to order $p$ do not depend on the order in which we take the derivative. Therefore, we may simply write these derivatives in the following form,*

$$\frac{\partial^k f}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} \quad \textit{where} \quad i_1 + \cdots + i_n = k \leqslant p.$$

### 4.2.2 Hessian matrix

**Definition 4.2.5 :** Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function, where $A \subseteq \mathbb{R}^n$ is an open subset. Suppose that all the second order partial derivatives of $f$ exist at $a \in A$. Then, the *Hessian matrix* of $f$ at $a$ is defined by

$$H_f(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right]_{1 \leqslant i,j \leqslant n}. \tag{4.18}$$

If the second derivatives are continuous at $a$, then Schwarz theorem (Theorem 4.2.1) implies that the Hessian matrix is symmetric at $a$.

Below, we will always consider a function $f$ whose second order derivatives are continuous, so that its Hessian matrix is symmetric.

**Proposition 4.2.6 :** *Under the same assumption as in Definition 4.2.5, we have*

$$H_f(a) = J_f(\overrightarrow{\text{grad}}\, f(a))^T.$$

**Proof :** It is a direct consequence by applying the definition of the Jacobian matrix to the gradient vector. □

When we study the local behavior of a function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ with some good assumptions (continuity of all the second derivatives), the Hessian matrix is symmetric and defines a quadratic form (二次型). The property of this quadratic form at a critical point can tell us whether this critical point is a local maximum, a local minimum, or a saddle point (鞍點). See Section 4.3.2 and Section 4.3.3 for more details.

### 4.2.3 Higher-order differentials

Given a function $f : A \subseteq V \to W$ betweeen an open subset $A$ of a normed vector space $V$ and another normed vector space $W$, we defined its differential at a point $a \in A$ in Definition 4.1.1, and its differential map $Df$ in Definition 4.1.3, under the condition that these notions exist. We may define its higher-order differentials by differentiating the differential map $Df : A \to \mathcal{L}_c(V, W)$.

From Definition 4.1.1, we know that the differential of $Df$ should take its values in $\mathcal{L}_c(V, \mathcal{L}_c(V, W))$, which may be identified as the space $\mathcal{L}_c^2(V \times V, W)$, the space of continuous bilinear maps from $V \times V$ to $W$, via the following map

$$\begin{array}{ccc} \mathcal{L}_c(V, \mathcal{L}_c(V, W)) & \to & \mathcal{L}_c^2(V \times V, W) \\ \Phi & \mapsto & \left\{ \begin{array}{ccc} V \times V & \to & W \\ (x, y) & \mapsto & \Phi(x)(y) \end{array} \right. \end{array}.$$

Similarly, the differential of order $p \geqslant 1$ takes values in the space $\mathcal{L}_c^p(V^p, W)$, which is the space of continuous $p$-linear maps.

**Definition 4.2.7 :** We define the higher-order differentials of $f$ recursively.

- For $p \geqslant 1$, we say that $f$ is differentiable $p+1$ times at $a \in A$ if its $p$-th differential $D^p f : A \to \mathcal{L}_c^p(V^p, W)$ is well defined, and writes

$$D^p f(a + h_{p+1})(h_1, \ldots, h_p) = D^p f(a)(h_1, \ldots, h_p) + \varphi_{p+1}(h_1, \ldots, h_p, h_{p+1}) + o(\|h_{p+1}\|_V)$$

  when $h_{p+1} \to 0$, uniformly for $(h_1, \ldots, h_p)$ in a bounded set of $V^p$, for some $\varphi_{p+1} \in \mathcal{L}_c^{p+1}(V^{p+1}, W)$. If such a map $\varphi_{p+1}$ exists, it is unique, and is called the $(p+1)$-th differential of $f$ at $a$, denoted by $D^{p+1} f(a)$.
- For $p \geqslant 1$, we say that $f$ is of class $\mathcal{C}^p$ if $D^p f$ is well defined on $A$ and is continuous on $A$.

**Remark 4.2.8 :** Let us take $V = \mathbb{R}^n$, $W = \mathbb{R}$, and $A \subseteq V$ be an open subset. Consider a $\mathcal{C}^1$ function $f : A \to W$ and suppose that its second partial derivatives exist. Fix $a \in A$, and take $\varepsilon > 0$ such that

$B(a,\varepsilon) \subseteq A$. Then, for $h_2 \in B(0,\varepsilon)$, we have

$$\begin{aligned} Df(a+h_2)(h_1) &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+h_2)\,\mathrm{d}x_i(h_1) \\ &= \sum_{i=1}^{n} \Big[\frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{2,j} + o(\|h_2\|_V)\Big]\,\mathrm{d}x_i(h_1) \\ &= Df(a)(h_1) + \sum_{i=1}^{n}\sum_{j=1}^{n} h_{2,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{1,i} + o(\|h_2\|_V)\, O(\|h_1\|_V). \end{aligned}$$

This implies that $D^2 f(a)$ is the (continuous) bilinear form given by the Hessian $H_f(a)$, written by

$$(h_1, h_2) \mapsto \sum_{i=1}^{n}\sum_{j=1}^{n} h_{2,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_{1,i} = h_2^T H_f(a) h_1.$$

Similar relations between higher-order differentials and higher-order derivatives exist as well. We do not discuss more here since this is not the main goal of this class.

In the following section, we will keep the same setting, that is $V = \mathbb{R}^n$ and $W = \mathbb{R}$, and look at the Taylor formulas of a function $f : A \to W$, where $A \subseteq V$ is an open subset. In this case, we will only need the higher-order differentials $D^p f$ evaluated at $(\underbrace{h, \dots, h}_{p \text{ times}})$.

## 4.3 Local behavior of real-valued functions

In this section, we are interested in real-valued functions and their local behaviors.

### 4.3.1 Taylor formulas

Let $p \geqslant 1$ be an integer. We recall that for a $\mathcal{C}^p$ function $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an open interval, we have the following Taylor formulas. Let $x \in I$ and $h \in \mathbb{R}$ be such that $x + h \in I$.

$$\begin{aligned} &\text{Taylor–Lagrange} && f(x+h) = f(x) + \sum_{m=1}^{p-1} f^{(m)}(x)\frac{h^m}{m!} + f^{(p)}(c)\frac{h^p}{p!} \text{ for some } c \in (x, x+h). \\ &\text{Taylor integral} && f(x+h) = f(x) + \sum_{m=1}^{p-1} f^{(m)}(x)\frac{h^m}{m!} + h^p \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} f^{(p)}(x+th)\,\mathrm{d}t. \\ &\text{Taylor–Young} && f(x+h) = f(x) + \sum_{m=1}^{p} f^{(m)}(x)\frac{h^m}{m!} + o(|h|^p) \text{ when } h \to 0. \end{aligned}$$

Below, we are going to generalize these formulas to real-valued functions defined on a subset of a higher dimensional Euclidean space $\mathbb{R}^n$.

Let $A$ be an open subset of $\mathbb{R}^n$, $f : A \to \mathbb{R}$ be a function of class $\mathcal{C}^p$ with $p \geqslant 1$, and $a \in A$. We have already defined the differential $\mathrm{d}f_a$ of $f$ at $a$ in Definition 4.1.1, and we gave the relation between the differential and the directional derivative $f'_u(a)$ in Proposition 4.1.17. Moreover, it follows from Theorem 4.1.21 that this can also be expressed using partial derivatives of $f$ at $a$. Below, we are going to define higher-order directional derivatives of $f$.

We will see that the Taylor formulas in higher dimensions are not too much different from their one-dimensional counterparts, due to the fact that when we restrict the function $f$ on a segment $[x, x+h]$, we are actually studying a function defined on a one-dimensional subspace.

**Definition 4.3.1 :** For $1 \leqslant m \leqslant p$, we may define the $m$th derivative of $f$ at $a$ in the direction $u \in \mathbb{R}^n$ as follows,

$$f_u^{(m)}(a) = \sum_{i_m=1}^{n} \cdots \sum_{i_1=1}^{n} \frac{\partial^m f}{\partial x_{i_m} \ldots \partial x_{i_1}}(a) u_{i_1} \ldots u_{i_m}, \tag{4.19}$$

$$= \sum_{j_1+\cdots+j_n=m} \frac{m!}{j_1! \ldots j_n!} \frac{\partial^m f}{\partial x_1^{j_1} \ldots \partial x_n^{j_n}}(a) u_1^{j_1} \ldots u_n^{j_n}, \tag{4.20}$$

where the equality is a direct consequence of Theorem 4.2.1.

**Theorem 4.3.2** (Taylor–Lagrange formula) **:** *Let $x \in A$ and $h \in \mathbb{R}^n$ such that $[x, x+h] \subseteq A$. Then, there exists $t \in (0,1)$ such that*

$$f(x+h) = f(x) + \sum_{m=1}^{p-1} \frac{f_h^{(m)}(x)}{m!} + \frac{f_h^{(p)}(x+th)}{p!}. \tag{4.21}$$

**Proof :** Since $A$ is open and $[x, x+h] \subseteq A$, there exists $\delta > 0$ such that $x+th \in A$ for all $t \in (-\delta, 1+\delta)$. Let $g : (-\delta, 1+\delta) \to \mathbb{R}$ be defined by

$$g(t) = f(x+th), \quad \forall t \in (-\delta, 1+\delta). \tag{4.22}$$

We note that $g$ is still a function of class $\mathcal{C}^p$ being composition of such functions. We also have $f(x+h) - f(x) = g(1) - g(0)$. Let us apply the classical Taylor formula to $g$, that is

$$g(1) - g(0) = \sum_{m=1}^{p-1} \frac{g^{(m)}(0)}{m!} + \frac{g^{(p)}(t)}{p!} \quad \text{for some } t \in (0,1).$$

We may explicit the derivatives of $g$ as below using the chain rule (Proposition 4.1.9). For $t \in (-\delta, 1+\delta)$, we have

$$g'(t) = \mathrm{d}f_{x+th}(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+th) h_i = f_h'(x+th),$$

$$g''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x+th) h_i h_j = f_h^{(2)}(x+th).$$

And by induction, we easily find that

$$g^{(m)}(t) = f_h^{(m)}(x+th), \quad \text{and} \quad g^{(m)}(0) = f_h^{(m)}(x), \quad \forall m \geqslant 1.$$

This allows us to conclude. □

Using the same technique by setting the function $g$ as in Eq. (4.22) and the other one-dimensional Taylor formulas, we easily deduce the following results.

**Theorem 4.3.3** (Taylor formula with integral remainder) **:** *Let $x \in A$ and $h \in \mathbb{R}^n$ such that $[x, x+h] \subseteq A$. Then,*

$$f(x+h) = f(x) + \sum_{m=1}^{p-1} \frac{f_h^{(m)}(x)}{m!} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} f_h^{(p)}(x+th)\,\mathrm{d}t. \tag{4.23}$$

**Proof :** See Exercise 4.19. □

**Theorem 4.3.4** (Taylor-Young formula) **:** *Let $x \in A$. Then, for $h \to 0$, we have*

$$f(x+h) = f(x) + \sum_{m=1}^{p} \frac{f_h^{(m)}(x)}{m!} + o(|h|^p). \tag{4.24}$$

**Proof :** See Exercise 4.19. □

### 4.3.2 Quadratic form

**Definition 4.3.5 :** Given a symmetric matrix $A \in \mathcal{M}_n(\mathbb{R})$, we can define a quadratic form (二次型) on $\mathbb{R}^n$ by

$$q_A(x) = q_A(x_1, \ldots, x_n) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j = x^T A x, \quad \forall x \in \mathbb{R}^n, \tag{4.25}$$

where a vector in $\mathbb{R}^n$ can be seen as a column vector.

**Definition 4.3.6 :** Given a quadratic form $Q$, we say that it is

- *positive* if $Q(x) \geqslant 0$ for all $x \in \mathbb{R}^n$,
- *positive-definite* if $Q(x) > 0$ for all $x \in \mathbb{R}^n \backslash \{0\}$,
- *negative* if $Q(x) \leqslant 0$ for all $x \in \mathbb{R}^n$,
- *negative-definite* if $Q(x) < 0$ for all $x \in \mathbb{R}^n \backslash \{0\}$.

**Remark 4.3.7 :** Under the condition that all the second partial derivatives are continuous at $a \in A$, we may rewrite $f_u^{(2)}(a)$ as follows,

$$f_u^{(2)}(a) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} u_i u_j = u^T H_f(a) u,$$

where $H_f(a)$ is a symmetric matrix, the vector $u$ can be seen as a column vector, and $u^T$ is its transposition. This defines a quadratic form in the sense of Definition 4.3.5.

**Remark 4.3.8 :** From the class of linear algebra, we know that a symmetric matrix $A$ is diagonalizable, that is, we may find a diagonal matrix $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_1 \geqslant \dots \geqslant \lambda_n$ and an orthogonal matrix $P$ (that is, $PP^T = P^T P = I_n$) such that $A = P^T DP$. This means that, after a proper change of basis given by $P$, the quadratic form is diagonal. More precisely, let $v = Pu$, then

$$u^T Au = (Pu)^T D(Pu) = v^T Dv,$$

meaning that

$$q_A(u) = q_D(v) = \sum_{i=1}^{n} \lambda_i |v_i|^2.$$

Therefore, we may conclude that if $\lambda_n > 0$, then the quadratic form is positive-definite; if $\lambda_1 < 0$, then the quadratic form is negative-definite.

### 4.3.3 Local extrema

Below, let $A$ be a subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ be a function. We want to study the local extrema of $f$. To do so, we are going to use the Taylor formula that we got in Section 4.3.1.

**Definition 4.3.9 :** If $f$ is differentiable at an interior point $a \in \mathring{A}$ with $\mathrm{d}f_a = 0$, then we call $a$ a *critical point* of $f$.

**Proposition 4.3.10 :** *Suppose that $f$ attains a local extremum at an interior point $a \in \mathring{A}$ and $f$ is differentiable at $a$. Then, $a$ is a critical point of $f$.*

**Proof :** Without loss of generality, we may assume that $f$ attains its local maximum at $a$. Let $h \in \mathbb{R}^n$, and we want to show that $\mathrm{d}f_a(h) = 0$. Since $a \in \mathring{A}$, there exists $\eta > 0$ such that $[a - \eta h, a + \eta h] \subseteq A$. We define the map $\varphi : [-\eta, \eta] \to \mathbb{R}, t \mapsto f(a + th)$, which has a local maximum at $t = 0$. Since $f$ is differentiable at $a$, we know that $\varphi$ is differentiable at $0$, and we have $\varphi'(0) = \mathrm{d}f_a(h)$. Additionally, we have

$$\varphi'(0) = \lim_{\substack{t \to 0 \\ t > 0}} \frac{\varphi(t) - \varphi(0)}{t} \leqslant 0, \quad \text{and} \quad \varphi'(0) = \lim_{\substack{t \to 0 \\ t < 0}} \frac{\varphi(t) - \varphi(0)}{t} \geqslant 0,$$

which gives us $\varphi'(0) = 0$. □

**Remark 4.3.11 :** Proposition 4.3.10 tells us that, to look for local extrema of a function $f : A \to \mathbb{R}$, we need to look at the following types of points,

(i) $a \in \mathring{A}$ which is a critical point of $f$;

(ii) $a \in \mathring{A}$ where $f$ is not differentiable;

(iii) $a \in A \backslash \mathring{A}$.

**Theorem 4.3.12 :** *Suppose that $f$ is of class $\mathcal{C}^2$ and there exists $a \in A$ such that $\mathrm{d}f_a = 0$. Taylor-Young formula (Eq. (4.24)) gives us*

$$f(a+h) = f(a) + \frac{1}{2}Q(h) + o(\|h\|^2), \quad \text{when } h \to 0.$$

*(1) If $f$ attains a local minimum (resp. maximum) at $a$, then $Q$ is a positive (resp. negative) quadratic form.*

*(2) If $Q$ is a positive-definite (resp. negative-definite) quadratic form, then $f$ attains a local minimum (resp. maximum) at $a$.*

**Example 4.3.13 :** In Theorem 4.3.12 (2), it is not enough for the quadratic form to be only positive to have a local minimum. Indeed, we may consider the function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ at $a = 0$, then the quadratic form is $Q \equiv 0$ but $f$ does not attain a local extremum.

**Proof :**

(1) Suppose that $f$ attains a local minimum at $a$. Let $h \in \mathbb{R}^n$ and $t \in \mathbb{R}$. When $t$ is sufficiently close to 0, we have

$$f(a+th) = f(a) + \frac{1}{2}Q(th) + o(\|th\|^2) \geqslant f(a).$$

This implies that

$$0 \leqslant Q(th) + o(\|th\|^2) = t^2(Q(h) + o(1)),$$

that is $Q(h) \geqslant 0$ when we take $t \to 0$.

(2) Suppose that $Q$ is a positive-definite quadratic form, then for $h \in \mathbb{R}^n$, $h \neq 0$, we have $Q(h) > 0$. Since the unit sphere $S(0,1)$ of $\mathbb{R}^n$ is compact, we deduce that $m = \inf_{h \in S(0,1)} Q(h) > 0$. Therefore, for $h \to 0$, we have

$$f(a+h) - f(h) = \frac{1}{2}[Q(h) + o(\|h\|^2)] = \frac{\|h\|^2}{2}\Big[Q\Big(\frac{h}{\|h\|}\Big) + o(1)\Big] \geqslant \frac{\|h\|^2}{2}(m + o(1)).$$

For $h$ close enough to 0, we have $m + o(1) \geqslant 0$, leading to $f(a+h) \geqslant f(a)$.

□

**Example 4.3.14 :** Let us consider the case $n = 2$ as an example. A quadratic form on $\mathbb{R}^2$ may be represented by a symmetric a matrix

$$A = \begin{pmatrix} r & s \\ s & t \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

Following Remark 4.3.8, we know that $A = P^T DP$, where $P$ is an orthogonal matrix and $D = \mathrm{Diag}(\lambda_1, \lambda_2)$ is a diagonal matrix with $\lambda_1 \geqslant \lambda_2$. We obtain the following relations for the eigenvalues

$\lambda_1 \geqslant \lambda_2$,

$$\begin{cases} \lambda_1 + \lambda_2 = \operatorname{tr}(D) = \operatorname{tr}(A) = r + t, \\ \lambda_1 \lambda_2 = \det(D) = \det(A) = rt - s^2. \end{cases}$$

Therefore, we have the following cases,

(i) When $rt - s^2 > 0$, and $r + t > 0$, the quadratic form associated with $A$ is positive-definite.

(ii) When $rt - s^2 > 0$, and $r + t < 0$, the quadratic form associated with $A$ is negative-definite.

When we apply this to a $\mathcal{C}^2$ function $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$, and $a \in \mathring{A}$ is a critical point of $f$. Write

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a), \quad t = \frac{\partial^2 f}{\partial y^2}(a).$$

Then, from the discussion above, we know that

(i) When $rt - s^2 > 0$, and $r + t > 0$, $f$ attains a local minimum at $a$.

(ii) When $rt - s^2 > 0$, and $r + t < 0$, $f$ attains a local maximum at $a$.

(iii) When $rt - s^2 < 0$, $f$ does not have an extremum at $a$, and we call it a saddle point (鞍點).

(iv) When $rt - s^2 = 0$, we cannot say anything.

## 4.4 Implicit function theorem

### 4.4.1 Inversion theorems

For a $\mathcal{C}^1$ function $f : \mathbb{R} \to \mathbb{R}$, we know that if $f'(x) \neq 0$ for all $x \in \mathbb{R}$, then $f$ is a bijection and its inverse $f^{-1}$ is also a $\mathcal{C}^1$ function satisfying $(f^{-1})'[f(x)] = [f'(x)]^{-1}$ for all $x \in \mathbb{R}$.

Let $V$ and $W$ be two Banach spaces, and $A \subseteq V$ be an open subset of $V$.

**Theorem 4.4.1** (Local inversion theorem) **:** *Let $f : A \to W$ be a function of class $\mathcal{C}^1$. Suppose that there exists $a \in A$ such that $(\mathrm{d}f_a)^{-1}$ exists and that $\mathrm{d}f_a$ and $(\mathrm{d}f_a)^{-1}$ are continuous (we say that $\mathrm{d}f_a$ is a bicontinuous isomorphism). Then, there exists an open set $X$ containing $a$ and an open set $Y$ containing $f(a)$ such that*

*(i) the function $f_{|X}$ is a bijection between $X$ and $Y$;*

*(ii) the inverse function $g := (f_{|X})^{-1} : Y \to X$ is continuous;*

*(iii) $g$ is of class $\mathcal{C}^1$ and $\mathrm{d}g_{f(x)} = (\mathrm{d}f_x)^{-1}$ for all $x \in X$.*

*In this case, we also say that $f_{|X} : X \to Y$ is a $\mathcal{C}^1$*-diffeomorphism *between $X$ and $Y$, or $f : A \to W$ is a local $\mathcal{C}^1$*-diffeomorphism *around $a$.*

**Remark 4.4.2 :**

(1) This theorem is called local inversion because it only describes the local behavior around $a \in X$ and $f(a) \in Y$. Later in Corollary 4.4.5, we will see how to upgrade this local inversion theorem into a global inversion theorem.

(2) If we consider $V = W = \mathbb{R}^n$ for some $n \geqslant 1$, since $\mathcal{L}(V,W) = \mathcal{L}_c(V,W)$, the condition for the local inversion at $a \in A \subseteq V$ reduces to the condition $\mathrm{d}f_a$ is invertible, that is $\det J_f(a) \neq 0$.

**Example 4.4.3 :**

(1) If we consider $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$, which is a $\mathcal{C}^1$ function on $\mathbb{R}$. For $a \in \mathbb{R}\backslash\{0\}$, the derivative $f'(a) = 2a \neq 0$, and it follows from the local inversion theorem that when $f$ is restricted to an open set $X$ containing $a$, its inverse is well defined. Actually, when $a > 0$, we may take $X = Y = (0,\infty)$, and define $g(y) = \sqrt{y}$ for $y \in Y$; and when $a < 0$, we may take $X = (-\infty, 0), Y = (0,\infty)$, and define $g(y) = -\sqrt{y}$ for $y \in Y$.

(2) If we define the transformation between the polar coordinates and the Cartesian coordinates,

$$\begin{array}{rccc} \varphi: & (0,\infty)\times\mathbb{R}\subseteq\mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (r,t) & \mapsto & (r\cos t, r\sin t), \end{array}$$

then its differential at $(r,t) \in (0,\infty)\times\mathbb{R}$ writes

$$\mathrm{d}\varphi_{r,t}(r',t') = (r'\cos t - t'r\sin t, r'\sin t + t'r\cos t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix}\begin{pmatrix} r' \\ t' \end{pmatrix}.$$

This gives us

$$\det J_\varphi(r,t) = r \neq 0, \quad \text{where } J_\varphi(r,t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix}.$$

From the local inversion theorem, at all $(r,t) \in (0,\infty)\times\mathbb{R}$, we may find an open set $X$ containing $(r,t)$ such that $f$ is invertible on $X$. However, $f$ does not have a global inverse, because it is clearly not injective.

**Proof :** Without loss of generality, we may consider $x \mapsto (\mathrm{d}f_a)^{-1}[f(a+x) - f(a)]$ instead of $f$, so that we can assume $a = 0$, $f(a) = 0$, and $\mathrm{d}f_0 = \mathrm{d}f_a = \mathrm{id}_V$, so $V = W$. Using the assumption that $f$ is of class $\mathcal{C}^1$, there exists $r > 0$ such that

$$\overline{B}(0,r) \subseteq A \quad \text{and} \quad |||\mathrm{d}f_x - \mathrm{d}f_0||| = |||\mathrm{d}f_x - \mathrm{id}_V||| \leqslant \tfrac{1}{2}, \quad \forall x \in B(0,r).$$

Then, for $x \in B(0,r)$, we have $\mathrm{d}f_x = \mathrm{id}_V - u$, where $u = \mathrm{id}_V - \mathrm{d}f_x$ with $|||u||| \leqslant \frac{1}{2}$, and it follows from Proposition 3.2.20 that

$$\begin{aligned} (\mathrm{d}f_x)^{-1} &= \mathrm{id}_V + \sum_{n\geqslant 1} u^n, \\ \left|\left|\left|(\mathrm{d}f_x)^{-1}\right|\right|\right| &\leqslant \sum_{n\geqslant 0} |||u|||^n \leqslant 2. \end{aligned} \tag{4.26}$$

(i) First, let us show that $f$ has a local inverse. More precisely, we want to prove that for every $y \in B(0, \frac{r}{2})$, there exists a unique $x \in B(0, r)$ satisfying $f(x) = y$. We are going to construct a function and apply the fixed point theorem (Theorem 3.2.7) to show this.

Let $y \in B(0, \frac{r}{2})$ and consider the function

$$\begin{array}{rccc} h: & B(0,r) & \to & V \\ & x & \mapsto & y + x - f(x). \end{array}$$

The function $h$ is of class $\mathcal{C}^1$, and for every $x \in B(0, r)$, we have $|||\mathrm{d}h_x||| = |||\mathrm{id}_V - \mathrm{d}f_x||| \leqslant \frac{1}{2}$. Thus, by the mean-value inequality (Theorem 4.1.13), we find

$$\forall x, x' \in \overline{B}(0,r), \quad \|h(x) - h(x')\| \leqslant \frac{1}{2}\|x - x'\|. \tag{4.27}$$

Therefore, for $x \in \overline{B}(0, r)$, we have

$$\|h(x)\| \leqslant \|y\| + \|x - f(x)\| = \|y\| + \|h(x) - h(0)\| \leqslant \|y\| + \frac{1}{2}\|x\| < r.$$

It means that $h$ is a contraction from $\overline{B}(0, r)$ to $B(0, r) \subseteq \overline{B}(0, r)$, so the fixed point theorem (Theorem 3.2.7) implies the existence and uniqueness of $x \in \overline{B}(0, r)$ such that $h(x) = x$. But since $h$ takes values in $B(0, r)$, it follows that the fixed point $x$ belongs to $B(0, r)$, and we have $f(x) = y$.

To conclude, let $Y = B(0, \frac{r}{2})$ and $X = f^{-1}(Y) \cap B(0, r)$. Due to the continuity of $f$ and $f(0) = 0$, the open set $X$ also contains 0. Then from what we have shown above, the restriction $f_{|X} : X \to Y$ is a bijection.

(ii) Let $g : Y \to X$ be the inverse $f_{|X}$, i.e. $g = (f_{|X})^{-1}$. Consider the function $h : X \to V, x \mapsto x - f(x)$, so we have $x = f(x) + h(x)$ for $x \in X$. Then, for $x, x' \in B(0, r)$, we have

$$\begin{aligned} & \|x - x'\| \leqslant \|h(x) - h(x')\| + \|f(x) - f(x')\| \leqslant \frac{1}{2}\|x - x'\| + \|f(x) - f(x')\| \\ \Leftrightarrow \quad & \|x - x'\| \leqslant 2\|f(x) - f(x')\|. \end{aligned}$$

Therefore, for $y, y' \in Y$, we have

$$\|g(y) - g(y')\| \leqslant 2\|f(g(y)) - f(g(y'))\| = 2\|y - y'\|. \tag{4.28}$$

This implies that $g$ is a Lipschitz function, so continuous.

(iii) Let $x \in X$ and $y = f(x) \in Y$. Let us first check that $\mathrm{d}g_y = (\mathrm{d}f_x)^{-1}$. Let $w \in W$ such that $y + w \in Y$, and $v = g(y + w) - g(y)$, which is equivalent to $w = f(x + v) - f(x)$. By Eq. (4.28), we have $\|v\| \leqslant 2\|w\|$. Let

$$\begin{aligned} \Delta(w) &= g(y + w) - g(y) - (\mathrm{d}f_x)^{-1}(w) \\ &= (\mathrm{d}f_x)^{-1} \circ \mathrm{d}f_x(v) - (\mathrm{d}f_x)^{-1}[f(x + v) - f(x)] \\ &= -(\mathrm{d}f_x)^{-1}[f(x + v) - f(x) - \mathrm{d}f_x(v)]. \end{aligned}$$

It follows from Eq. (4.26) that

$$\|\Delta(w)\| \leqslant 2\,\|f(x+v) - f(x) - \mathrm{d}f_x(v)\| = 2\,\|v\|\,\varepsilon(v),$$

for some function $\varepsilon$ satisfying $\lim_{v\to 0}\varepsilon(v) = 0$. Let $\widetilde{\varepsilon}(w) = \varepsilon\big(g(y+w) - g(y)\big)$. Since $g$ is continuous, we also have $\lim_{w\to 0}\widetilde{\varepsilon}(w) = 0$. Thus,

$$\frac{\|\Delta(w)\|}{\|w\|} \leqslant \frac{2\,\|v\|}{\|w\|}\widetilde{\varepsilon}(w) \xrightarrow[w\to 0]{} 0.$$

This means that $g$ is differentiable at $y$ with $\mathrm{d}g_y = (\mathrm{d}f_x)^{-1}$.

To conclude, since $u \mapsto u^{-1}$ on the space of invertible endomorphisms is continuous (Example 4.1.6 and Proposition 4.1.7), and $g$ is continuous, we deduce that the map $y \mapsto \mathrm{d}g_y = (\mathrm{d}f_{g(y)})^{-1}$ is also continuous, that is $g$ is of class $\mathcal{C}^1$.

□

**Corollary 4.4.4 :** *Let $f : A \to W$ be a function of class $\mathcal{C}^1$. Suppose that $\mathrm{d}f_x$ is invertible and bicontinuous for all $x \in A$. Then, $f$ is an open map, that is for any open subset $X \subseteq A$, the image $f(X)$ is open in $W$.*

**Proof :** It is enough to prove for the case that $X = A$. For each $a \in A$, the local inversion theorem (Theorem 4.4.1) gives an open subset $X_a$ containing $a$ and an open subset $Y_a$ containing $f(a)$ such that $f_{|X_a}$ is a bijection between $X_a$ and $Y_a$, i.e., $f(X_a) = Y_a$. Therefore,

$$f(A) = f\Big(\bigcup_{a\in A} X_a\Big) = \bigcup_{a\in A} f(X_a) = \bigcup_{a\in A} Y_a,$$

which is still an open subset of $W$. □

**Corollary 4.4.5** (Global inversion theorem) **:** *Let $f : A \to W$ be an injective function of class $\mathcal{C}^1$. Then, the following properties are equivalent.*

*(a) The differential $\mathrm{d}f_a$ is invertible and bicontinuous for all $a \in A$.*

*(b) $B = f(A)$ is open in $W$ and $f^{-1} : B \to A$ is of class $\mathcal{C}^1$.*

*If one of the above properties is satisfied, we say that $f : A \to B$ is a $\mathcal{C}^1$-*diffeomorphism *between $A$ and $B$.*

**Proof :**

- (a) ⇒ (b). It follows from Corollary 4.4.4 that $B = f(A)$ is open. Since $f$ is injective, we deduce that $f$ is bijective from the open set $A$ to the open set $B$. Now, we need to check that $f^{-1}$ is of class $\mathcal{C}^1$. Let $x \in A$ and $y = f(x) \in B$. The local inversion theorem (Theorem 4.4.1), we can find an open set $A_x$ containing $x$ and an open set $B_x$ containing $f(x)$ such that $f_{|A_x} : A_x \to B_x$ is bijective and $(f_{|A_x})^{-1}$ is of class $\mathcal{C}^1$. Since $(f^{-1})_{|B} = (f_{|A})^{-1}$, $(f^{-1})_{|B_x} = (f_{|A_x})^{-1}$, and being

$\mathcal{C}^1$ is a local property, we know that $f^{-1}$ is of class $\mathcal{C}^1$ around $f(x)$. This holds for all $x \in A$, so $f^{-1}$ is of class $\mathcal{C}^1$ on $B$.

- (b) $\Rightarrow$ (a). Write $g = f^{-1}$. Since both $f$ and $g$ are of class $\mathcal{C}^1$, the relation $g \circ f = \mathrm{Id}_A$ and the chain rule gives us $\mathrm{d}g_{f(x)} \circ \mathrm{d}f_x = \mathrm{Id}_V$ for all $x \in A$. Similarly, the relation $f \circ g = \mathrm{Id}_B$ gives $\mathrm{d}f_x \circ \mathrm{d}g_{f(x)} = \mathrm{Id}_W$ for all $x \in A$. Therefore, for all $x \in A$, the differential $\mathrm{d}f_x$ is invertible with inverse $\mathrm{d}g_{f(x)}$, which is continuous.

□

**Remark 4.4.6 :** We make a similar remark as in Remark 4.4.2 (2). If we consider an Euclidean space $V = W = \mathbb{R}^n$ for some $n \geqslant 1$, since $\mathcal{L}(V, W) = \mathcal{L}_c(V, W)$, we may replace the property (a) by

(a') $\mathrm{d}f_a$ is invertible, or $\det J_f(a) \neq 0$,

without requiring the bicontinuity.

### 4.4.2 Diffeomorphisms

**Definition 4.4.7 :** Let $V, W$ be two normed vector spaces, and $A \subseteq V$ and $B \subseteq W$ be open subsets. For $k \geqslant 1$, a function $f : A \to B$ is said to be a $\mathcal{C}^k$-diffeomorphism if $f$ is bijective, of class $\mathcal{C}^k$ and $f^{-1}$ is also of class $\mathcal{C}^k$.

The following two corollaries give the conditions under which a map is a local diffeomorphism and a global diffeomorphism in the setting of Euclidean spaces. Their proofs are based on the local inversion theorem and the global inversion theorem. We note that they can also be generalized to Banach spaces by adding the bicontinuity in the condition. Since we did not really discussed $\mathcal{C}^k$ functions in general Banach spaces or normed vector spaces when $k \geqslant 2$ (see Section 4.2.3), we keep our statements to Euclidean spaces for which we had a thorough discussion about regularity in Section 4.2.1.

**Corollary 4.4.8 :** *Let $A \subseteq \mathbb{R}^n$ be an open subset. Let $f : A \to \mathbb{R}^n$ be a $\mathcal{C}^k$ function. Suppose that there exists $a \in A$ such that $\mathrm{d}f_a$ is invertible (or equivalently, $\det J_f(a) \neq 0$), then there exists an open set $X_a$ containing $a$ and an open set $Y_a$ containing $f(a)$ such that $f_{|X_a}$ is a $\mathcal{C}^k$-diffeomorphism from $X_a$ to $Y_a$. We also have $\mathrm{d}(f_{|X_a}^{-1})_{f(x)} = (\mathrm{d}f_x)^{-1}$ for all $x \in X_a$.*

**Proof :** Since $\mathrm{d}f_a$ is invertible, and we work with finite dimensional vector spaces, $\mathrm{d}f_a$ is automatically bicontinuous. Then, we may apply the local inversion theorem (Theorem 4.4.1) to find $X_a$ and $Y_a$ as stated, such that $f_{|X_a}$ is a $\mathcal{C}^1$-diffeomorphism. It remains to show that $g = f_{|X_a}^{-1}$ is a $\mathcal{C}^k$ function.

We recall the notation $J_f(a)$ for the Jacobian matrix of $f$ at $a$, and $J_g(f(a))$ for the Jacobian matrix of $g$ at $f(a)$. For all $x \in X_a$, since $(\mathrm{d}g)_{f(x)} = (\mathrm{d}f_x)^{-1}$, we deduce that $J_g(f(x)) = J_f(x)^{-1} = (\det J_f(x))^{-1}\widetilde{J}(x)$, where $\widetilde{J}(x)$ is the transpose of the comatrix of $J_f(x)$ (also called the adjugate matrix), whose coefficients are linear combinations of products of coefficients of $J_f(x)$. Therefore, the first-order partial derivatives of $g$ are rational fractions of first-order partial derivatives of $f$, which are of class $\mathcal{C}^{k-1}$, implying that the first-order partial derivatives of $g$ are also of class $\mathcal{C}^{k-1}$. We can conclude that $g$ is of class $\mathcal{C}^k$. □

**Corollary 4.4.9 :** *Let $A \subseteq \mathbb{R}^n$ be an open subset and $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be an injective function of class $\mathcal{C}^k$ with $k \geqslant 1$. Then, the following properties are equivalent.*

*(a) The differential $\mathrm{d}f_a$ is invertible for all $a \in A$.*

*(b) $B = f(A)$ is open in $W$ and $f$ is a $\mathcal{C}^k$-diffeomorphism from $A$ to $B$.*

**Proof :** The proof is similar to Corollary 4.4.5 and Corollary 4.4.8. □

### 4.4.3 Implicit function theorem

We describe some motivation behind the implicit function theorem. We are given a function $f : A \subseteq \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and want to look at its level lines, that is we look for $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ such that $f(x, y) = c$ for some given $c \in \mathbb{R}^n$. The implicit function theorem provides *local* sufficient conditions such that $y$ can be written as a function $\varphi$ of $x$, that is $f(x, \varphi(x)) = c$. In other words, in a neighborhood of such $x$, the solutions of $f(x, y) = c$ can be represented by a graph. More generally speaking, we may take $c$ to be a variable, and we obtain a function $\varphi$ in $x$ and $c$. See the theorem below for a more precise statement.

Let

$$f = (f_1, \ldots, f_n) : \begin{array}{ccc} A \subseteq \mathbb{R}^m \times \mathbb{R}^n & \to & \mathbb{R}^n \\ (x, y) = (x_1, \ldots, x_m; y_1, \ldots, y_n) & \mapsto & f(x, y). \end{array} \tag{4.29}$$

We may define the partial Jacobian matrices and their determinants (called partial Jacobian determinants) with respect to the variables $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_n)$ at $(a, b) \in A$ as below,

$$J_{f,x}(a, b) = \left[\frac{\partial f_i}{\partial x_j}(a, b)\right]_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant m}} \qquad \text{and} \qquad J_{f,y}(a, b) = \left[\frac{\partial f_i}{\partial y_j}(a, b)\right]_{1 \leqslant i,j \leqslant n}.$$

**Theorem 4.4.10** (Implicit function theorem) **:** *Let $m, n \geqslant 1$ be integers and $A \subseteq \mathbb{R}^m \times \mathbb{R}^n$ be an open subset. Suppose that we are given a $\mathcal{C}^k$ function $f$ with $k \geqslant 1$ as in Eq. (4.29). Let us fix $(a, b) \in A$. If the partial Jacobian determinant $\det J_{f,y}(a, b)$ is nonzero, then there exist*

- *an open subset $X$ containing $a$, an open subset $W$ containing $f(a, b)$ and an open subset $Z$ containing $(a, b)$,*
- *a $\mathcal{C}^k$ function $\varphi : X \times W \to \mathbb{R}^n$*

*such that for all $x \in X$ and $w \in W$, $y = \varphi(x, w)$ is the unique solution to $f(x, y) = w$ with condition $(x, y) \in Z$. In particular, we have $f(x, \varphi(x, w)) = w$ for all $x \in X$ and $w \in W$.*

*Moreover, for $(a, c) \in X \times W$, write $b = \varphi(a, c)$, then we also have the following relations between the partial Jacobian matrices,*

$$J_{\varphi,x}(a, c) = -\left[J_{f,y}(a, b)\right]^{-1} J_{f,x}(a, b) \qquad \textit{and} \qquad J_{\varphi,y}(a, c) = \left[J_{f,y}(a, b)\right]^{-1}.$$

**Remark 4.4.11 :** The inversion theorems were stated and proven for general Banach spaces, where we require the differential to be invertible and bicontinuous. We also noted in Remark 4.4.2 and Remark 4.4.6 that when we take Euclidean spaces (or finite dimensional normed vector spaces), the bicontinuity property automatically holds, so need not be checked. Here, for simplicity, we state the implicit function theorem for Euclidean spaces, but you need to bear in mind that when we work with general Banach spaces, the only additional condition you need to add to the assumption is the bicontinuity.

**Proof :** Let $F = (F_1, \dots, F_m; F_{m+1}, \dots, F_{m+n})$ be a function defined on $A$ with values in $\mathbb{R}^{m+n}$ whose components are defined by $F_i(x, y) = x_i$ if $1 \leqslant i \leqslant m$ and $F_{m+i} = f_i(x, y)$ if $1 \leqslant i \leqslant n$. The Jacobian matrix of $F$ is a block matrix, given by

$$\left(\begin{array}{c|c} I_m & \mathbf{0} \\ \hline \star & \left(\dfrac{\partial f_i}{\partial y_j}(a,b)\right)_{1\leqslant i,j\leqslant n} \end{array}\right).$$

Its determinant is the same as the partial Jacobian determinant given by $\det J_{f,y}(a, b)$, which is nonzero by assumption. Then, it follows from Corollary 4.4.8 that there exists an open set $Z$ containing $(a, b)$ and an open set $Y$ containing $F(a, b) = (a, f(a, b))$ such that $F_{|Z}$ is a $\mathcal{C}^k$-diffeomorphism from $Z$ to $Y$. We may restrict $Y$ to $X \times W \subseteq Y$, where $X$ is an open set containing $a$ and $W$ an open set containing $f(a, b)$. Then, we may write $F^{-1} : X \times W \subseteq Y \to Z$ as $F^{-1}(x, w) = (x, \varphi(x, w))$, where $\varphi$ is a $\mathcal{C}^k$ function. Therefore, we deduce that for any $(x, w) \in X \times W$, there exists a unique $y$ such that $(x, y) \in Z$ with $f(x, y) = w$; and additionally, $y = \varphi(x, z)$.

To get the identities between the partial Jacobian matrices, we just need to apply the relation between the composition of functions and multiplication of Jacobian matrices as mentioned in Proposition 4.1.26. □

In the above theorem, we may take $w$ to be a constant, leading to the following corollary.

**Corollary 4.4.12 :** *Under the same assumption as Theorem 4.4.10, we may find*

- *an open subset $X$ containing $a$ and an open subset $Y$ containing $b$,*
- *a $\mathcal{C}^k$ function $\varphi : X \to \mathbb{R}^n$*

*such that for all $x \in X$, $y = \varphi(x)$ is the unique solution to $f(x, y) = c$ with condition $y \in Y$. This allows us to write $f(x, \varphi(x)) = c$ for all $x \in X$.*

*Moreover, for $a \in X$, write $b = \varphi(x)$, then we also have the following relation between the partial Jacobian matrices,*

$$J_{f,x}(a, b) + J_{f,y}(a, b)J_{\varphi,x}(a) = 0 \qquad \text{or} \qquad J_{\varphi,x}(a) = -\left[J_{f,y}(a, b)\right]^{-1} J_{f,x}(a, b). \tag{4.30}$$

**Corollary 4.4.13 :** *Let $A \subseteq \mathbb{R}^2$ be an open set and $f : \mathbb{R}^2 \to \mathbb{R}$ be a $\mathcal{C}^k$ function with $k \geqslant 1$. Let $(a, b) \in A$ and suppose that*

$$f(a, b) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(a, b) \neq 0.$$

*Then, there exists $\alpha, \beta > 0$ such that for all $x \in (a - \alpha, a + \alpha)$, the equation $f(x, y) = 0$ has a unique* *solution $y = \varphi(x)$ in $(b-\beta, b+\beta)$. Moreover, the function $\varphi$ is of class $\mathcal{C}^k$ on $(a-\alpha, a+\alpha)$ and we have*

$$\varphi'(x) = -\frac{\partial f}{\partial x}(x,\varphi(x)) \Big/ \frac{\partial f}{\partial y}(x,\varphi(x)), \quad \forall x \in (a-\alpha, a+\alpha).$$

**Proof :** The existence of $\alpha, \beta > 0$ and regularity of $\varphi$ follows from Corollary 4.4.12. To compute $\varphi'$, we differentiate the relation $f(x,\varphi(x)) = 0$, giving us

$$\frac{\partial f}{\partial x}(x,\varphi(x)) + \varphi'(x)\frac{\partial f}{\partial y}(x,\varphi(x)) = 0.$$

This can also be obtained directly from Eq. (4.30). □

The following corollary can be shown in a similar way.

**Corollary 4.4.14 :** *Let $A \subseteq \mathbb{R}^3$ be an open set and $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^k$ function with $k \geqslant 1$. Let $(a,b,c) \in A$ and suppose that*

$$f(a,b,c) = 0 \quad \textit{and} \quad \frac{\partial f}{\partial z}(a,b,c) \neq 0.$$

*Then, there exists $\alpha, \beta, \gamma > 0$ such that for all $(x,y) \in (a-\alpha, a+\alpha) \times (b-\beta, b+\beta)$, the equation $f(x,y,z) = 0$ has a unique solution $z = \varphi(x,y)$ in $(c-\gamma, c+\gamma)$. Moreover, the function $\varphi$ is of class $\mathcal{C}^k$ on $(a-\alpha, a+\alpha) \times (b-\beta, b+\beta)$, and we have*

$$\frac{\partial \varphi}{\partial x}(x,y) = -\frac{\partial f}{\partial x}(x,y,\varphi(x,y)) \Big/ \frac{\partial f}{\partial z}(x,y,\varphi(x,y)),$$
$$\frac{\partial \varphi}{\partial y}(x,y) = -\frac{\partial f}{\partial y}(x,y,\varphi(x,y)) \Big/ \frac{\partial f}{\partial z}(x,y,\varphi(x,y)).$$

**Example 4.4.15 :** Let us consider a $\mathcal{C}^\infty$ function $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto \sin(y) + xy^4 + x^2$. We want to look at the graph of $f(x,y) = 0$ and its asymptotic behavior around $(x,y) = (0,0)$.

- It is not hard to check that $f(0,0) = 0$. The partial derivatives of $f$ write

$$\frac{\partial f}{\partial x}(x,y) = y^4 + 2x \quad \text{and} \quad \frac{\partial f}{\partial y}(x,y) = \cos(y) + 4xy^3.$$

  We have $\frac{\partial f}{\partial x}(0,0) = 0$ and $\frac{\partial f}{\partial y}(0,0) = 1$. Therefore, it follows from Corollary 4.4.13 that there exist $\alpha, \beta > 0$ and a $\mathcal{C}^\infty$ function $\varphi : (-\alpha, \alpha) \to \mathbb{R}$ such that for every $x \in (-\alpha, \alpha)$, $y = \varphi(x)$ is the unique solution to $f(x,y) = 0$ in $(-\beta, \beta)$.

- Let us find the Taylor expansion of $\varphi$ around 0. First, from the above computations, we have $\varphi(0) = 0$ and $\varphi'(0) = 0$, so we may write $\varphi(x) = \mathcal{O}(x^2)$ when $x \to 0$. To get a higher-order expansion, we will substitute this expression into $f(x,\varphi(x)) = 0$ and expand $\sin(y) = y + \mathcal{O}(y^3)$

when $y = \varphi(x) \to 0$. We find

$$\begin{aligned}\varphi(x) &= \varphi(x) - \sin(\varphi(x)) - x\varphi(x)^4 - x^2 \\ &= \mathcal{O}(\varphi(x)^3) - x^2 \\ &= -x^2 + \mathcal{O}(x^6) = -x^2(1 + \mathcal{O}(x^4)).\end{aligned}$$

If we want to get a even higher-order expansion of $\varphi$, we expand the $\sin$ function to a higher order, that is $\sin(y) = y - \frac{y^3}{6} + \mathcal{O}(y^5)$ when $y = \varphi(x) \to 0$. We find

$$\begin{aligned}\varphi(x) &= \varphi(x) - \sin(\varphi(x)) - x\varphi(x)^4 - x^2 \\ &= -x^2 - \frac{\varphi(x)^3}{6} + \mathcal{O}(\varphi(x)^5) - x\varphi(x)^4 \\ &= -x^2 + \frac{x^6}{6}(1 + \mathcal{O}(x^4)) + \mathcal{O}(x^{10}) - x^9(1 + \mathcal{O}(x^4)) \\ &= -x^2 + \frac{x^6}{6} - x^9 + \mathcal{O}(x^{10}).\end{aligned}$$

One may also proceed further by expanding the $\sin$ function to higher orders of $\varphi$.

### 4.4.4 Conditional extrema

Let $A \subseteq \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}$ be a function. Let $g_1, \ldots, g_r : A \to \mathbb{R}$ be functions and

$$\Gamma = \{x \in A : g_1(x) = \cdots = g_r(x) = 0\}.$$

We want to look for the extrema of $f$ on $\Gamma$. Such a problem is called *conditional extrema.*

**Theorem 4.4.16 :** *Suppose that $f, g_1, \ldots, g_r$ are $\mathcal{C}^1$ functions. Suppose that $f_{|\Gamma}$ attains a local extremum at $a \in \Gamma$ and that $\mathrm{d}g_{1,a}, \ldots, \mathrm{d}g_{r,a}$ are linearly independent, then there exist $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ such that*

$$\mathrm{d}f_a = \lambda_1\,\mathrm{d}g_{1,a} + \cdots + \lambda_r\,\mathrm{d}g_{r,a}. \tag{4.31}$$

**Remark 4.4.17 :** The coefficients $\lambda_i$'s are called Lagrange multipliers. They are unique because the linear forms $\mathrm{d}g_{1,a}, \ldots, \mathrm{d}g_{r,a}$ are linearly independent.

**Proof :** Let $s = n - r$ and write $\mathbb{R}^n = \mathbb{R}^s \times \mathbb{R}^r$. An element of $\mathbb{R}^n$ can be written in the form $(x, y) = (x_1, \ldots, x_s; y_1, \ldots, y_r)$. Let $a = (x_a, y_a) \in \mathbb{R}^n$ with $x_a \in \mathbb{R}^s$ and $y_a \in \mathbb{R}^r$.

First, we note that we necessarily have $r \leqslant n$ because $(\mathrm{d}g_{i,a})_{1 \leqslant i \leqslant r}$ are linear independent, and the dimension of $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$, the space of lineaer forms, is equal to $n$. If $n = r$, then the theorem is trivial because $(\mathrm{d}g_{i,a})_{1 \leqslant i \leqslant r}$ forms a basis. Thus, let us assume that $r \leqslant n - 1$ in what follows, that is $s \geqslant 1$.

Due to the linear independence of $(\mathrm{d}g_{i,a})_{1 \leqslant i \leqslant r}$, the Jacobian matrix of $g = (g_1, \ldots, g_r)$ at $a$ has rank $r$. Without loss of generality, we may assume that the following $r \times r$ submatrix has nonzero determinant,

$$\det\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{1\leqslant i,j\leqslant r} \neq 0.$$

Therefore, it follows from Corollary 4.4.12 that there exists an open set $X$ of $\mathbb{R}^s$ containing $x_a$ and an open set $W$ of $\mathbb{R}^n$ containing $a = (x_a, y_a)$ and a $\mathcal{C}^1$ function $\varphi = (\varphi_1, \ldots, \varphi_r) : X \to \mathbb{R}^r$ such that

$$g(x,y) = 0 \text{ with } x \in X \text{ and } (x,y) \in W \quad \Leftrightarrow \quad y = \varphi(x).$$

In other words, for $x \in X$, the elements of $\Gamma = \{z : g(z) = 0\}$ can be written as $(x, \varphi(x))$. Let $h(x) = f(x, \varphi(x))$, which has a local extremum at $x = a$ by assumption. This leads to

$$0 = \frac{\partial h}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(x_a)\frac{\partial f}{\partial y_j}(a), \quad \forall i = 1, \ldots, s. \tag{4.32}$$

Additionally, by differentiating the relation $g(x, \varphi(x)) = 0$, we find

$$0 = \frac{\partial g_k}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(x_a)\frac{\partial g_k}{\partial y_j}(a), \quad \forall k = 1, \ldots, r, \forall i = 1, \ldots, s. \tag{4.33}$$

Putting Eq. (4.32) and Eq. (4.33) in the matrix form, we find the matrix

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \ldots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \ldots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \ldots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \ldots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \ldots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \ldots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}.$$

whose first $s$ columns are linear combination of its last $r$ columns, which implies that $\operatorname{rank} M \leqslant r$. Since $\operatorname{rank} M^t = \operatorname{rank} M$, it means that the $r+1$ rows of $M$ are linearly dependent, that is, there exist $\mu_0, \ldots, \mu_r$ that are not identically zero such that

$$\mu_0 \,\mathrm{d}f_a + \mu_1 \,\mathrm{d}g_{1,a} + \cdots + \mu_r \,\mathrm{d}g_{r,a} = 0. \tag{4.34}$$

From the assumption that $(\mathrm{d}g_{i,a})_{1\leqslant i\leqslant r}$ is linearly independent, it follows that $\mu_0 \neq 0$, therefore, by dividing Eq. (4.34) by $\mu_0$, we prove the theorem. □

**Example 4.4.18 :** Find the minimum and the maximum values of the function $f(x,y) = x^2 - 2x + 4y^2 + 8y$ in the first quadrant, that is $(x,y) \in A := \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0}$, under the condition $g(x,y) = 0$ with $g(x,y) = x + 2y - 7$. Since the domain $A$ is not closed in $\mathbb{R}^2$, we need to distinguish between interior points and other points. The extrema of $f$ are attained either at $a \in \mathring{A} = \mathbb{R}_{>0} \times \mathbb{R}_{>0}$ satsifying Eq. (4.31), or at some point in $A \backslash \mathring{A} = (\{0\} \times \mathbb{R}_{\geqslant 0}) \cup (\mathbb{R}_{\geqslant 0} \times \{0\})$, that are not covered by Theorem 4.4.16.

- Let us look for an interior point $(x,y) \in \mathring{A}$ satisfying $g(x,y) = 0$ and such that $\mathrm{d}f_{(x,y)} =$

$\lambda \, \mathrm{d}g_{(x,y)}$ has a nonzero solution $\lambda \in \mathbb{R}$. First, we want to solve

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) &= \lambda \frac{\partial g}{\partial x}(x,y), \\ \frac{\partial f}{\partial y}(x,y) &= \lambda \frac{\partial g}{\partial y}(x,y), \end{cases} \quad \Leftrightarrow \quad \begin{cases} 2x-2 &= \lambda, \\ 8y+8 &= 2\lambda. \end{cases}$$

Thus, we find $x = 2y+3$. We put this back to the condition $g(x,y) = 0$ and find $(x,y) = (5,1)$. We compute the value $f(5,1) = 27$.

- The points $(x,y)$ in $A \backslash \mathring{A}$ satisfying $g(x,y) = 0$ are exactly $(x,y) = (7,0)$ or $(0, \frac{7}{2})$. We compute the values $f(7,0) = 35$ and $f(0, \frac{7}{2}) = 77$.

From above, we conclude that in the first quadrant with condition $g(x,y) = 0$, the maximum value of $f$ is attained at $(0, \frac{7}{2})$ with value 77, and the minimum value of $f$ is attained at $(5,1)$ with value 27.