# Mathematical Analysis II

Felix Chen

## Contents



**Definition 0.0.1** (Tangent map)**.** Let $f : M \to N$ be a map between manifolds, $v \in T_pM$. Let $\gamma : (-\varepsilon, \varepsilon) \to M$ be a parametrized curve with $\gamma(0) = p, \gamma'(0) = v$, then $f(\gamma(t))$ is a curve on $N$.

$$\mathrm{d}f(p)(v) = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\Big|_{t=0} \in T_{f(p)}N.$$

Thus $\mathrm{d}f(p) : T_pM \to T_{f(p)}N$ is a map between tangent spaces.

In fact, if $f = F\big|_M$, then $\mathrm{d}f(p)(v) = \mathrm{d}F(p) \cdot v$.

**Definition 0.0.2** (Tangent bundle)**.** Let $M$ be a manifold, $\forall p \in M$, there's a tangent space $T_pM$. Define the **tangent bundle** of $M$ to be

$$TM = \bigsqcup_{p \in M} T_pM.$$

If $X$ is a map $M \to TM$: $p \mapsto X(p)$, with $X(p) \in T_pM$, then it's called a **tangent vector field**.

In other words, a tangent vector field is just to assign a tangent vector to every point in $M$.

**Proposition 0.0.3**

Let $M \subset \mathbb{R}^n$ be a manifold, all its tangent vector field form a $C^\infty$ module $T(M, TM)$, i.e. $\forall f \in C^\infty(M)$, $X, Y$ are smooth vector fields, then $fX, X + Y$ are both smooth vector fields.

**Proposition 0.0.4**

Let $M \subset \mathbb{R}^n$ be a smooth manifold, we have

$$TM = \{(x, v) \mid x \in M, v \in T_xM\}$$

is a smooth manifold in $\mathbb{R}^{2n}$, and $\dim TM = 2 \dim M$.

*Proof.* There exists a local homeomorphism $\phi : V \to U \subset \mathbb{R}^n$ s.t. $V \subset \mathbb{R}^d$, $\phi(V) = M \cap U$.

Define map $T\phi : V \times \mathbb{R}^n \to U \times \mathbb{R}^n$, $(x, v) \mapsto (\phi(x), \mathrm{d}\phi(x) \cdot v)$. Since $T\phi$ is injective($\phi$ is homeomorphism), and

$$\mathrm{d}T\phi = \begin{pmatrix} \mathrm{d}\phi & 0 \\ \mathrm{d}(\mathrm{d}\phi)(v) & \mathrm{d}\phi \end{pmatrix}$$

is non-degenerate, so $T\phi$ is a bijection and hence differential homeomorphism.

Since the tangent space of $V$ is just $\mathbb{R}^d$, so $T(U \cap M)$ is the image of $T\phi$ restricted on $V \times \mathbb{R}^d$. (Note that $\mathrm{d}\phi(x) \cdot v \in T_{\phi(x)}M$) Thus $TM$ is a manifold in $\mathbb{R}^{2n}$ with dimension $2d$. □

**Definition 0.0.5** (Tangent maps)**.** Earlier we know that $\mathrm{d}f(p)$ is a map $T_pM \to T_{f(p)}N$, combined with tangent bundle we can write $\mathrm{d}f : TM \to TN$, this map is called the **tangent map** or the **differentiation** of $f$.

If we have a vector field $X$ and a smooth function $f : M \to \mathbb{R}^n$, consider

$$X(f)(p) = \mathrm{d}f(X)(p) := \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\Big|_{t=0}, \quad \gamma(0) = p, \gamma'(0) = X(p).$$

So $X$ induces a smooth map $C^\infty(M) \to C^\infty(M)$.

Now we can generalize a well known result to manifolds:

**Proposition 0.0.6**

Let $M \subset \mathbb{R}^n$ be a smooth manifold, $f \in C^\infty(M)$. If $f$ achieves a local extremum at $p \in M$, we must have $\mathrm{d}f(p) = 0$.

*Proof.* It suffices to prove $\mathrm{d}f(p)(v) = 0$, $\forall v \in T_pM$. Take $\gamma$ s.t. $\gamma(0) = p, \gamma'(0) = v$, then $f(\gamma(t))$ achieves its extremum at $t = 0$, so $\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\big|_{t=0} = 0 = \mathrm{d}f(p)(v)$. □

## §0.1 Conditional extremum problem

Consider a function $f(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$ and some constraint conditions

$$\begin{cases} g_1(x_1, \ldots, x_n) = 0 \\ \vdots \\ g_m(x_1, \ldots, x_n) = 0 \end{cases}$$

We want to compute the extremum of $f$ under these conditions.

Well, you probably heard of *Lagrange multipliers*, i.e. let

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x) - \sum_{j=1}^{m} \lambda_j g_j(x).$$

But here we'll provide a different point of view. Let $M$ be the manifold in $\mathbb{R}^n$ under those conditions, Suppose $p \in M$ is a local extremum of $f$, then $T_pM \subset \ker \mathrm{d}f(p)$.

Also recall that $T_pM = \ker \mathrm{d}g(p) = \bigcap_{j=1}^m \ker \mathrm{d}g_j(p)$. This means that, $\exists \lambda_1, \ldots, \lambda_m$ s.t.

$$\mathrm{d}f(p) = \sum_{j=1}^{m} \lambda_j \, \mathrm{d}g_j(p).$$

Surprisingly, we get the same result of Lagrange multipliers! Hence what we've done is to give a geometrical comprehension of Lagrange multipliers.

**Example 0.1.1**

Let $g : \mathbb{R}^n \to \mathbb{R}$ be the constraint function, then $f$ can achieve its extremum only if $\mathrm{d}f = \lambda\,\mathrm{d}g$.

For example, let $f(x) = d(x,z)^2$, $\mathrm{d}f(x) = 2(x_1 - z_1, \dots, x_n - z_n)$, so $\mathrm{d}f = \lambda\,\mathrm{d}g$ means the vector $\mathrm{d}f(p)$ is orthogonal to the tangent plane of $M = \{g = 0\}$.

**Proposition 0.1.2** (Hadamard's inequality)

Let $v_1, \dots, v_n \in \mathbb{R}^n$, then

$$|\det(v_1, \dots, v_n)| \le |v_1| \cdots |v_n|.$$

*Proof.* Let $f : \mathbb{R}^{n^2} \to \mathbb{R}$, $(v_1, \dots, v_n) \mapsto \det(v_1, \dots, v_n)$ with constraint $|v_i| = 1$. Let $v_{ij} \in \mathbb{R}$,

$$g_i(V) = -1 + \sum_{j=1}^{n} v_{ij}^2.$$

The manifold determined by $g_i$ is $M = (S^{n-1})^n$. The extremum point of $f$ in $M$ must satisfy:

$$\frac{\partial f}{\partial v_{i_0 j}} - \lambda_{i_0} \frac{\partial g_{i_0}}{\partial v_{i_0 j}} = 0.$$

This implies $v_{i_0 j}^* = 2\lambda_{i_0} v_{i_0 j}$, where $v_{i_0 j}^*$ is the *cofactors* of $v_{i_0 j}$.

This means that $\sum_{j=1}^n v_{i_0 j} v_{kj} = 0$, so $V$ must be an orthogonal matrix, so $|f| \le 1$. □

## §0.2 Convex functions

**Definition 0.2.1** (Hesse matrix)**.** Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a $C^2$ function, we call the Jacobi matrix of $\nabla f$ to be the **Hesse matrix** of $f$. (Also called Hessian matrix)

$$H_f(p) = \nabla^2 f(p) = \left( \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(p) \right)_{i,j}.$$

Since the partial derivatives commute, so $H_f$ is a symmetrical matrix, hence diagonalizable.

**Proposition 0.2.2**

Let $f \in C^2(\Omega)$, let $x_0$ be a minimum of $f$, then $\nabla f(x_0) = 0$, and $H_f(x_0)$ is semi positive definite.

*Proof.* By Taylor's expansion,

$$f(x) = f(x_0) + \mathrm{d}f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) + o(|x - x_0|^2).$$

If $H_f(x_0)$ has a negative eigenvalue $-\lambda$, with eigenvector $v$, then $f(x_0 + tv) = f(x_0) - \frac{1}{2}\lambda t^2 |v|^2 + o(|tv|^2)$, which contradicts with the minimality of $x_0$. □

**Proposition 0.2.3**
If $\nabla f(x_0) = 0$, $H_f(x_0)$ is positive definite, then $x_0$ is a local minimum of $f$.

*Proof.* Same as previous one. $\square$

**Definition 0.2.4** (Convex functions)**.** If $f$ and $\Omega$ satisfies:

$$\forall x, y \in \Omega, tx + (1-t)y \in \Omega, \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

we say $\Omega$ is a **convex set** and $f$ a **convex function**.

**Theorem 0.2.5** (Jensen's inequality)
Let $f$ be a convex function on $\Omega$. Real numbers $t_i \geq 0, \sum_{i=1}^N t_i = 1$, for $x_i \in \Omega$,

$$f\left(\sum_{i=1}^N t_i x_i\right) \leq \sum_{i=1}^N t_i f(x_i).$$

**Example 0.2.6** (Convex functions)
Linear functions $f(x) = Ax + b$ are convex.
The norm function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is convex. Also let $A$ be an $n \times n$ positive definite matrix, then $f(x) = x^T Ax$ is convex.

Just like the one dimensional case, convex functions have nice properties.

**Theorem 0.2.7**
Let $f$ be a convex function on an open convex set $\Omega$, then $f$ is continuous, and Lipschitz continuous in any compact set, i.e.

$$|f(x) - f(y)| \leq M|x - y|, \quad x, y \in U$$

where $U$ is a compact set.

*Proof.* WLOG $0 \in \Omega$, take an orthogonal basis $e_1, \dots, e_n$. Let

$$x = \sum_{i=1}^n \lambda_i \bar{e}_i, \quad \bar{e}_i = e_i \text{ or } -e_i, \lambda_i \geq 0.$$

When $|x|$ sufficiently small, $\sum_{i=1}^n \lambda_i < 1$, so by Jensen's inequality,

$$f(x) \leq \sum_{i=1}^n \lambda_i f(\bar{e}_i) + \lambda f(0),$$

$$f(x) - f(0) \leq \sum_{i=1}^n \lambda_i (f(\bar{e}_i) - f(0)) \leq \left(\sum_{i=1}^n \lambda_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n (f(\bar{e}_i) - f(0))^2\right)^{\frac{1}{2}} \leq |x|C,$$

since we can change the length of $e_i$, and $f$ is continuous on a straight line.

This means $f$ is continuous. For the second part, let $\lambda_0 = \frac{1}{1+\sum_{i=1}^n \lambda_i}$, since $0 = \lambda_0 x + \sum_{i=1}^n \lambda_0\lambda_i(-\bar{e}_i)$, by Jensen's inequality, we'll get the desired property. $\square$

**Proposition 0.2.8**

Let $f$ be a differentiable function on a covex set $\Omega$,
$f$ is convex $\iff f(x) \geq f(x_0) + \mathrm{d}f(x_0)(x - x_0)$.

*Proof.* If $f$ is convex, just use the definition and let $t \to 0$:

$$f(x_0) + f'(x_0)t(x - x_0) + o(t(x - x_0)) \leq tf(x) + (1-t)f(x_0).$$

Conversely, let $z = tx + (1-t)x_0$,

$$f(x) \geq f(z) + f'(z)(1-t)(x - x_0), f(x_0) \geq f(z) + f'(z)t(x_0 - x).$$

Thus adding these together we get

$$tf(x) + (1-t)f(x_0) \geq f(z).$$

$\square$

**Theorem 0.2.9**

Let $\Omega \subset \mathbb{R}^n$ be an open convex set, $f \in C^2(\Omega)$, $f$ convex $\iff H_f(x)$ semi positive definite.

*Proof.* One direction can be proved using Taylor's expansion.

On the other hand, let $H(t) = f(x_0 + t(x - x_0)) - f(x_0) - t\,\mathrm{d}f(x_0)(x - x_0)$, then $H'(t) = \mathrm{d}f(x_0 + t(x - x_0))(x - x_0) - \mathrm{d}f(x_0)(x - x_0)$,

$$H''(t) = (x - x_0)^T H_f(p)(x_0 + t(x - x_0))(x - x_0) \geq 0.$$

So $H(t)$ is a convex function, $H(0) = 0, H'(0) = 0$. $\square$

# §1 Integrals on surfaces

## §1.1 Measures on manifolds

To define integrals, we need to define a measure on it first.

For example, let $v_1, \ldots, v_d \in \mathbb{R}^n$ be linearly independent vectors, and unit vectors $v_{d+1}, \ldots, v_n$ complete them to a basis, satisfying $v_j \perp v_i, j > d, j > i$.

Let $A$ be a linear map s.t. $Ae_i = v_i$, then the volume of $A(E)$ is $|\det A| = \sqrt{\det(G \cdot G^T)}$, where $G = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix}$ is a $d \times n$ matrix.

Since $AA^T = \begin{pmatrix} GG^T & 0 \\ 0 & I_{n-d} \end{pmatrix}$, $|\det A| = \sqrt{\det GG^T}$, we say $GG^T$ is the **Gram matrix** of $G$.

Another example is the length of a curve. Recall that we have the formula

$$L(\gamma) = \int_0^1 |\gamma'(t)| \, \mathrm{d}t.$$

The length of a curve is essentially the "volume" of a 1-dimensional manifold, so the idea of higher dimensional manifold is the same.

**Definition 1.1.1.** Let $M$ be a manifold in $\mathbb{R}^n$. Let $\Phi : V \subset \mathbb{R}^d \to U \subset M$ be a smooth homeomorphism, $\operatorname{rank} \Phi = d$. We can split $U$ to many small regions and use the paraloids to approximate the volume of each regoin.

Thus we define:

$$m(U) = \int_V \sqrt{\det(\mathrm{d}\Phi(x)^T \, \mathrm{d}\Phi(x))} \, \mathrm{d}x_1 \, \mathrm{d}x_2 \cdots \mathrm{d}x_d.$$

The above differential form inside the integral is called the **volume form**:

$$\mathrm{d}\sigma(y) = \sqrt{\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi)} \, \mathrm{d}y.$$

Moreover, with this measure we can define the integral of general measurable function $f$ (measurable means locally measurable on $\mathbb{R}^d$):

$$\int_U f \, \mathrm{d}\sigma = \int_V f(\Phi(x)) \sqrt{\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi)} \, \mathrm{d}x.$$

Just like the length of a curve, the volume defined above is also a geometric quantity, i.e. independent of coordinates.

**Example 1.1.2**

Let $d = 1$, $\gamma : (-1, 1) \to \mathbb{R}^n$, $\gamma'(0) \neq 0$. For fixed $-1 < a < b < 1$ and a function $f$ on $\gamma$, let $C_a^b$ denote the curve between $\gamma(a), \gamma(b)$,

$$\int_{C_a^b} f \, \mathrm{d}\sigma = \int_a^b f(\gamma(t)) |\gamma'(t)| \, \mathrm{d}t$$

is called the **curve integral of the first type**.

**Example 1.1.3**

Let $d = n - 1$, $f : \mathbb{R}^{n-1} \to \mathbb{R}$, the graph of $f$ is a hyper-surface $\Gamma_f = \{(x, f(x)) \mid x \in \mathbb{R}^{n-1}\}$. It has a parametrization $\Phi(x) = (x, f(x))$, so

$$\mathrm{d}\Phi = \begin{pmatrix} I_{n-1} \\ \nabla f \end{pmatrix} \implies \mathrm{d}\Phi^T \, \mathrm{d}\Phi = I_{n-1} + \nabla f^T \cdot \nabla f.$$

Hence $\det(\mathrm{d}\Phi^T \, \mathrm{d}\Phi) = 1 + |\nabla f|^2$. (This can be obtained by looking at the eigenvectors)

Therefore for $\varphi$ on $\mathbb{R}^n$, we have

$$\int_{\Gamma_f} \varphi \, \mathrm{d}\sigma = \int_{\mathbb{R}^{n-1}} \varphi(x, f(x)) \sqrt{1 + |\nabla f|^2} \, \mathrm{d}x.$$

Next we'll compute the surface area of unit sphere $S^{n-1}$.

Let $c_n$ denote the volume of unit sphere in $\mathbb{R}^n$,

$$c_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}.$$

We claim in advance that the surface area of unit sphere $\omega_{n-1} = nc_n$. Here we use the sphere coordinates:

$$x_1 = r\prod_{k=1}^{n-1}\sin\theta_k, \quad x_i = r\sin\theta_{n-1}\cdots\sin\theta_i\cos\theta_{i-1}, 2 \le i \le n.$$

Let $F_n(r,\theta_1,\ldots,\theta_{n-1}) = (x_1,\ldots,x_n)$.

$$\mathrm{d}F_n = \begin{pmatrix} \sin\theta_{n-1}\,\mathrm{d}F_{n-1} & \cos\theta_{n-1}F_{n-1}^T \\ (\cos\theta_{n-1},0,\ldots,0) & -r\sin\theta_{n-1} \end{pmatrix}.$$

So we can compute its determinant (expand using the last row), note that the first column of $\mathrm{d}F_{n-1}$ is $r^{-1}F_{n-1}^T$,

$$\begin{aligned} \det\mathrm{d}F_n =& -r\sin\theta_{n-1}(\sin\theta_{n-1})^{n-1}\det(\mathrm{d}F_{n-1}) \\ &+ (-1)^{n-1}(\cos\theta_{n-1})^2(-1)^{n-2}r(\sin\theta_{n-1})^{n-2}\det(\mathrm{d}F_{n-1}) \\ =& -r(\sin\theta_{n-1})^{n-2}\det(\mathrm{d}F_{n-1}). \end{aligned}$$

Hence $\mathrm{d}x = r^{n-1}\sin^{n-2}\theta_{n-1}\cdots\sin\theta_2\,\mathrm{d}r\,\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1} = r^{n-1}\,\mathrm{d}r\,\mathrm{d}\omega$.

Denote $F_n^S$ to be the function $F_n$ restricted to $S^{n-1}$. Then $\mathrm{d}F_n = (r^{-1}F_n^T, \mathrm{d}F_n^S)$. We can compute that the Gram determinant of $\mathrm{d}F_n^S$ is just $\det\mathrm{d}F_n$ with $r=1$.

The rest is some integrals with gamma function and beta function, which is left out.

## §1.2 Stolkes' formula

Intuitively, Stolkes' formula states that: Let $D$ be a region, $\mathrm{d}\omega$ be a differential form, then

$$\int_D \mathrm{d}\omega = \int_{\partial D}\omega.$$

Here $\partial D$ means the "boundary" of $D$.

Of course we need some "regularity" requirements of $D$ and $\omega$, and it's the generalization of Newton-Lebniz formula into higher dimensions.

**Definition 1.2.1** (Bounded regions with boundary)**.** Let $\Omega \subset \mathbb{R}^n$ be a compact set, we say it's a **bounded region with boundary** if $\forall x \in \partial\Omega$, there exists open sets $U, V \subset \mathbb{R}^n$, $x \in U$ and a continuous homeomorphism $\Phi : U \to V$, such that

$$\Phi(U\cap\Omega) = V\cap\{x_n \ge 0\}, \quad \Phi(U\cap\partial\Omega) = V\cap\{x_n = 0\}.$$

If $\Phi$ is also $C^1$, we say $x \in \partial\Omega$ is a **regular point**, otherwise a **singular point**.

**Lemma 1.2.2**

Let $\Omega$ be a bounded region with boundary, for all regular $p \in \partial\Omega$, there exists a unique unit vector $\nu(p) \in \mathbb{R}^n$, and $\varepsilon > 0$, s.t.

$$\nu(p) \perp T_p\partial\Omega, |\nu(p)| = 1, p - t\nu(p) \in \Omega, \forall 0 < t < \varepsilon.$$

We call $\nu(p)$ the **outward unit normal vector** of $p$.

*Proof.* By the definition of regular points, we may assume that:

$$\Omega \cap V = \{x \in V \mid f(x) \geq 0\}, \quad \partial\Omega \cap V = \{x \in V \mid f(x) = 0\}.$$

Where $f$ is a $C^1$ function.

Since $\nabla f$ is nonzero, the tangent space $T_p \partial\Omega = \{v \mid v \cdot \nabla f = 0\}$.

Let $\nu(p) = -\frac{\nabla f}{|\nabla f|}$, then it's obvious $\nu(p)$ points outside of $\Omega$. □

Now for a cuboid $I$ and a $C^1$ function $\phi$,

$$\begin{aligned}\int_I \frac{\mathrm{d}\phi}{\mathrm{d}x_n} \,\mathrm{d}x &= \int_{I_{n-1}} \phi(x_1, \ldots, x_{n-1}, b_n) \,\mathrm{d}x_1 \cdots \mathrm{d}x_{n-1} - \int_{I_{n-1}} \phi(x_1, \ldots, x_{n-1}, a_n) \,\mathrm{d}x_1 \cdots \mathrm{d}x_{n-1} \\ &= \int_{\partial I} \phi \cdot \nu_n \,\mathrm{d}\sigma.\end{aligned}$$

Where $\sigma$ is the measure on the boundary, $\nu$ is the outward unit normal vector.

**Lemma 1.2.3**

Let $K$ be a compact set in $\mathbb{R}^n$, $U \supset K$ is open, there exists a smooth function $f$ such that $\operatorname{supp} f \subset U$, and $f\big|_K > 0$.

*Proof.* Let $\rho(x)$ be a smooth function s.t. $\rho(x) = 1$ for $|x| \leq 1$ and $\rho(x) = 0$ for $|x| \geq 2$. Let

$$g(x) = \int_{|y| \leq 2} f(x - \delta y)\rho(y) \,\mathrm{d}y.$$

Then $g$ is a smooth non-negative function. □

**Theorem 1.2.4** (Unit decomposition on compact sets)

Let $K$ be a compact set, $\{U_1, \ldots, U_k\}$ is an open covering of $K$. There exists smooth functions $f_1, \ldots, f_k$ s.t.

$$1 = f_1(x) + f_2(x) + \cdots + f_k(x), \quad \operatorname{supp} f_i(x) \subset U_i.$$

*Proof.* For $1 \leq i \leq k$, $\delta > 0$, define

$$K_i^\delta = \{x \in U_i \mid d(x, U_i^c) > \delta\}.$$

Note that $\{\bigcup_{i=1}^k K_i^{\frac{1}{m}}\}_{m=1}^\infty$ is also an open covering of $K$, thus there exists $N$ s.t.

$$K \subset \bigcup_{i=1}^k K_i^{\frac{1}{N}}.$$

Hence by lemma we have $g_i$ s.t. $\operatorname{supp} g_i \subset U_i$ and $g_i > 0$ on the closure of $K_i^{\frac{1}{N}}$. Similarly we have a smooth function $g$ s.t. $g(x) = 0$ on $K$, and $g > 0$ outside of the closure of $\bigcup_{i=1}^k K_i^{\frac{1}{N}}$.

Let $G(x) = g_1(x) + \cdots + g_k(x) + g(x) > 0$ on $\bigcup_{i=1}^k U_i$, then we can define $f_i(x) = \frac{g_i(x)}{G(x)}$ which satisfy the condition. □

**Theorem 1.2.5**

Let $\Phi$ be a $C^1$ homeomorphism from a cuboid $I$ to $\Omega$, then $\Omega$ satisfies Stolkes' formula:

$\forall \phi \in C^1(\mathring{D}) \cap C(\overline{D})$, we have

$$\int_D \nabla\phi \,\mathrm{d}x = \int_{\partial D} \phi\nu \,\mathrm{d}\sigma.$$