A Note on Differential Geometry

Hifive Zheng

April 22, 2024

Chapter 1

Geometry in Euclidean Space

In this chapter, we are going to derive what a tangent vector or a cotangent vector is, give them a good definition and have a new view on differential and derivative. *(We will ignore many detials in order to grasp the very idea instead of being lost in tedious derivation.)*

We must wander what is a geometry in our space. Is it a curve surface or a field? There are two basic views on it. If we think of it as a curve surface, that is to say it's essence is beyond our given space. *(Later we'll call it as a submanifold embedded into our given space.)* As a curve surface is two-dimensional, a point on the surface can be identified within two coordinates $(u, v) \in C \subset \mathbb{R}^2$, where C denotes a collection of points on the curve surface and is a subset of $\mathbb{R}^2$. *(We shall clarify what "ponits" actually mean later.)* Now we only adopt an intrinsic perspective on the surface without observing it's curvature until we use a **map** $\phi : C \to \mathbb{R}^3, (u, v) \mapsto (x, y, z) = \phi(u, v)$.

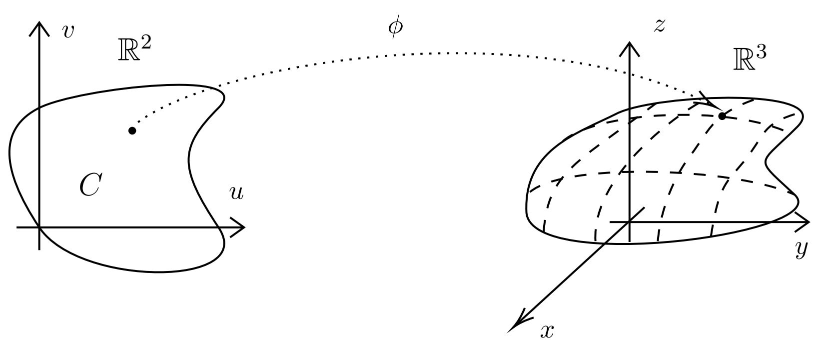

Figure 1.0.1: A map from C to $\mathbb{R}^3$

Another way to think of a geometry is **scalar field** $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto f(x, y)$ If we write this way: $z = f(x, y)$ where z represents each scalar value onto each point (x, y). There is a relationship between z and (x, y) that can be written as $((x, y), z)$ or just (x, y, z). All such relationships constitude a set $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$

However the relationship between x, y, z can be wider given by $F(x, y, z) = 0$. F is a strange function that can be considered as a constant scalar field. Before we use an simpler example to gain a deeper understanding of it, let's clarify some terms in the context.

Definition 1.0.1 (Relation)**.** *[1] A relation is a subset of a finite Cartes 'ian power $A^n = A \times \cdots \times A$ of a given set A, i.e. a set of tuples $(a_1, \ldots, a_n)$ of n elements of A.*

A subset $R \subseteq A^n$ is called an n-place, or an n-ary, relation on A. The number ***n*** *is called the* ***rank****, or* ***type****, of the relation R. The notation $R(a_1, \ldots, a_n)$ signifies that $(a_1, \ldots, a_n) \in R$.*

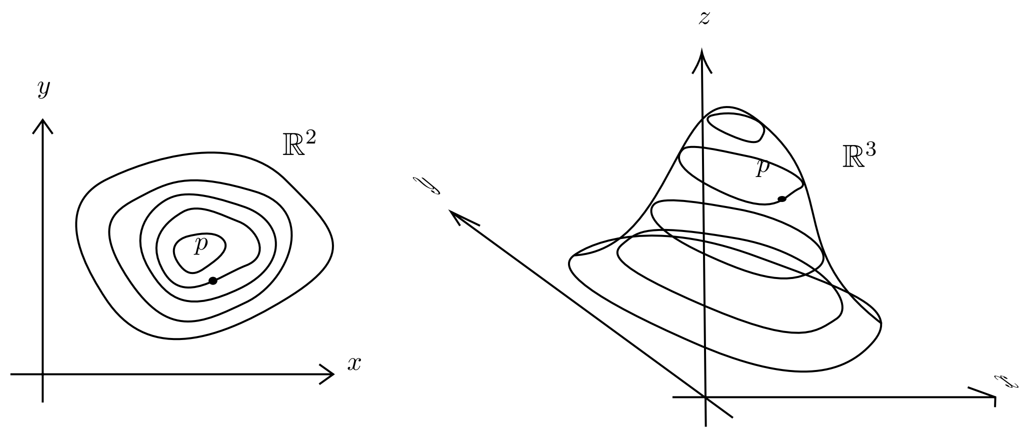

Figure 1.0.2: A scalar field over $\mathbb{R}^2$

The set A^n and the empty subset $\emptyset$ in R^n are called, respectively, the ***universal relation*** *and the* ***zero relation*** *of rank n on A. The diagonal of the set A^n, i.e. the set $\Delta = (a, a, \ldots, a) : a \in A$ is called the* ***equality relation*** *on A.*

The set of all n-ary relations on A is a **Boolean algebra** relative to the operations $\cup, \cap, '$. An $(n+1)$-place relation F on A is called functional if for any elements $a_1, \ldots, an, a, b$, from A it follows from $F(a_1, \ldots, a_n, a)$ and $F(a_1, \ldots, a_n, b)$ that $a = b$.

Definition 1.0.2 (Binary Relation)**.** *[2] A binary relation is a predicate on a given set. A binary relation is a special case of a relation. Let $R \subseteq A \times A$. If $(a, b) \in R$, then one says that the element a is in binary relation R to the element b. An alternative notation for $(a, b) \in R$ is aRb.*

The empty subset $\emptyset$ in $A \times A$ and the set $A \times A$ itself are called, respectively, the nil relation and the universal relation in the set A. The diagonal of the set $A \times A$, i.e. the set $\Delta = (a, a) : a \in A$, is the ***equality relation*** *or the* ***identity binary relation*** *in A.*

Let R, R_1, R_2 be binary relations in a set A. In addition to the set-theoretic operations of union $R_1 \cup R_2$, intersection $R_1 \cap R_2$, and negation or complementation $R' = (A \times A) \backslash R$, one has the inversion (also inverse, converse or transpose)

$$R^{-1} = (a, b) : (b, a) \in R,$$

as well as the operation of multiplication (or composition):

$$R_1 \circ R_2 = (a, c) \in A \times A : (\exists b \in A)(aR_1b \text{ and } bR_2c).$$

Definition 1.0.3 (Mapping)**.** *[3] In set theory mappings are special binary relations.A mapping f from a set A to a set B is an (ordered) triple $f = (A, B, G_f)$ where $Gf \subset A \times B$ such that*

(1) if (x, y) and $(x, y') \in G_f$ then $y = y'$, and

(2) the projection $\pi_1(G_f) = \{x|(x, y) \in G_f\} = A$.

Condition (1) expresses that f is single-valued. and condition (2) that it is defined on A. A is the ***domain****, B is the* **codomain***, and G_f is the* ***graph*** *of the mapping. Therefore, in this setting, mappings are equal if and only if all three corresponding components (domain, codomain, and graph) are equal.*

The mapping is usually denoted as $f : A \to B$, and $a \mapsto f(a)$ where $f(a) := b \Leftrightarrow (a, b) \in G_f$ is the value of f at a.

If two mappings $f_1 = (A_1, B_1, G_1)$ and $f_2 = (A_2, B_2, G_2)$ satisfy $A_1 \subset A_2$, $B_1 \subset B_2$ and $G_1 \subset G_2$ then f_2 is called an **extension** of f_1, and f_1 a **restriction** of f_2. In this case, f_1 is often denoted as $f_2|A_1$ and, clearly, $f_1(a) = f_2(a)$ holds for all $a \in A_1$.

> Notice that functions to be talked below are always C^∞ (smooth) without special illustration.

1.1 Explicit Function and Implicit Equation

It's time to clarify the difference between implicit function and explicit function! Consider a explicit function $F : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto z = F(x, y)$ which can be thought of a scalar field over $\mathbb{R}^2$ Esay, right? What about taking the value of z as 0? Is it a constant scalar field valued 0? Never! There is a huge diference between $F(x, y) = 0$ and $F(x, y) \equiv 0$. The first one is a function that tells you $z \equiv 0$ whatever values x and y take; and the second one tells that for the specific function $z = F(x, y)$, $F(x, y) = 0$ determines a set of (x, y) satisfying the equation $F(x, y) = 0$.

Actually, due to our laziness and ambiguity, when we say a function $y = f(x)$ or $f(x)$, we are not really say that $y = f(x)$ or $f(x)$ is a function - because $y = f(x)$ is a formula and $f(x)$ is a value! However, $y = f(x)$ is the definition formula of the funciton $f : x \mapsto y = f(x)$, implying a binary relation that for a given x in the defined domain, there is a unique y corresponding to it (or determined). Because a formula $y = f(x)$ defines a funciton f, no wonder we mixed them up! Same, y is the image of f, which is the same as $f(x)$, so it is always omitted.

Going back to the example we started with. **Given a formula that contains a function, if the function has been defined, then the formula is a equation; if the function has not been defined, then we can consider the formula as the definition formula of the function.** So, for a function $F : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto z = F(x, y)$, what's the formula $F(x, y) = 0$ means? It means the intersection curve between the x-y coordinate plane and the scalar filed surface in the $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$ space.

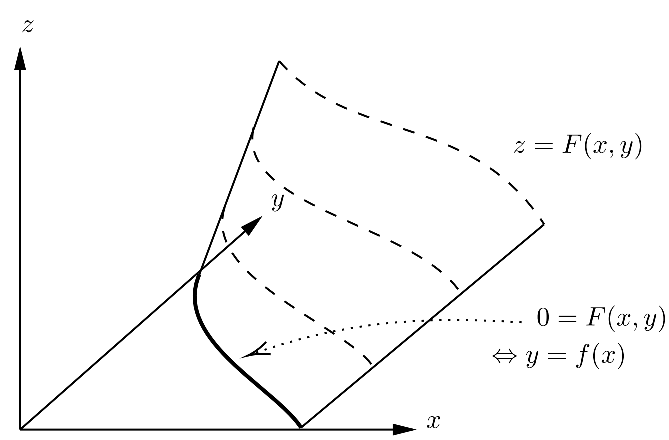

Figure 1.1.1: The intersection curve between the x-y coordinate plane and the scalar filed surface.

As figure 1.1.1 shows above, $F(x, y) = 0$ determines a function $f : \mathbb{R} \to \mathbb{R}, x \mapsto y = f(x)$ if we transform the formula to the form of $y = f(x)$.

1.2 Derivative and Tangent Vector

A function $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto z = f(x,y)$ is a explicit function and scalar filed over $\mathbb{R}^2$. f has good properties that we can define a derivative over it:

Definition 1.2.1. *For a real-valued function $f \in C^k : \mathbb{R} \to \mathbb{R}$, the **derivative** on $x_0 \in D(f)$ is defined by:*

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x_0} f := \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{t},$$

for $\Delta x \in \mathbb{R}$.

Definition 1.2.2. *Construct a function $\frac{\mathrm{d}}{\mathrm{d}x} f : C^k \to \mathbb{R}$ satisfying:*

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} f\right)(x_0) := \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{x_0} f.$$

*$\frac{\mathrm{d}}{\mathrm{d}x} f$ is called the **derivative function**. It is denoted as $\frac{\mathrm{d}}{\mathrm{d}x} f \equiv f'$ simultaneously when it is clear whom to take the derivative from (always x). It is also write as $\dot{f}$ when taking the derivative from t.*

There are abundant corollaries on the field of derivative and are needless to say more. Let's turn our attention to scalar field over $\mathbb{R}^n$, and we'll define the **derivative operator** and **directional derivative** as well as the **gradient**. And we'll further study their geometric significance in the next section.

Definition 1.2.3 (Parital Derivative)**.** *Given a function $f : \mathbb{R}^n \to \mathbb{R}, \boldsymbol{x} = (x^1, \dots, x^n) \mapsto y = f(\boldsymbol{x})$. The limitation is defined as the **partial derivative** at $\boldsymbol{x_0}$:*

$$\left.\frac{\partial}{\partial x^i}\right|_{\mathrm{x}_0} f := \lim_{\Delta x^i \to 0} \frac{f(\dots, x_0^i + \Delta x^i, \dots) - f(\boldsymbol{x_0})}{\Delta x^i} = \left.\frac{\mathrm{d}}{\mathrm{d}x^i}\right|_{x^i = x_0^i} f(x_0^1, \dots, x^i, \dots, x_0^n).$$

Definition 1.2.4.

$$\left(\frac{\partial}{\partial x^i} f\right)(\boldsymbol{x_0}) := \left.\frac{\partial}{\partial x^i}\right|_{\mathrm{x}_0} f$$

Theorem 1.2.1 (Chain Rule)**.** *Given a function $f : \mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{x} \mapsto \mathbf{y} = f(\boldsymbol{x})$ that eats n variables and spit out m variables where $\boldsymbol{x} = (x^1, \dots, x^n), \mathbf{y} = (y^1, \dots, y^m)$. It can be thought of a vector $\langle f_1, \dots, f_n \rangle$ in which $f_i : \mathbb{R} \to \mathbb{R}^m, \boldsymbol{x} \mapsto y^i = f(\boldsymbol{x})$. All partial derivatives can be written in a maritx, which we called a **partial derivative tensor**:*

$$\frac{\partial f_j}{\partial x^i} = \begin{bmatrix} \frac{\partial f_1}{\partial x^1} & \cdots & \frac{\partial f_1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x^1} & \cdots & \frac{\partial f_m}{\partial x^n} \end{bmatrix}.$$

Now we use vectors to represent a function chain: $\mathbf{z} = g(\mathbf{y}), \mathbf{y} = f(\boldsymbol{x})$. Then $\mathbf{z} = g(f(\boldsymbol{x})) = (g \circ f)(\boldsymbol{x})$. As can vertify, The partial derivative tensor of z over x is (here $z = g \circ f$ also):

$$\frac{\partial z^k}{\partial x^i} = \frac{\partial g_k}{\partial y^j} \frac{\partial f_j}{\partial x^i}, \tag{1.1}$$

*which is called the **Chain Rule**.*

The partial derivative tensor written above is also a **Jacobian Matrix** of a vector-valued function. But isn't it better to take tensor notation? When so, the gradient of each components can be represent more concise and intuitive!

$$D\vec{f} = \nabla f_j = \frac{\partial f_j}{\partial x^i}. \tag{1.2}$$

Example 1.2.1. *(1) Calculate the gradient of* $f(x,y) = \sin(xy) + \frac{x^2}{1+\ln y}$ *at* $(\pi, 1)$*.*

Solution. $\frac{\partial f}{\partial x} = y\cos(xy) + \frac{2x}{1+\ln y}$, $\quad \frac{\partial f}{\partial y} = x\cos(xy) - \frac{x^2}{y(1+\ln y)^2}$.

Then $\nabla f|_{(\pi,1)} = (\left.\frac{\partial f}{\partial x}\right|_{(\pi,1)}, \left.\frac{\partial f}{\partial y}\right|_{(\pi,1)}) = (-1+2\pi, -\pi-\pi^2)$. □

(2) Given a composite function $z = f(x+y^2, ye^x)$*, calculate* $\frac{\partial z}{\partial x}$*.*

Solution. Denote that $z = f(u,v) = f(u^1(x,y), u^2(x,y))$. According to the Chain Rule:
$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial u^i}\frac{\partial u^i}{\partial x} = \frac{\partial f}{\partial u^1} + \frac{\partial f}{\partial u^2} ye^x$. □

Definition 1.2.5 (Derivative Operator)**.** *For a scalar field* $f : \mathbb{R}^n \to \mathbb{R}$*, consider a column vector:* $\boldsymbol{v} = \langle v^1, \ldots, v^n \rangle$*. Define a coordinate function* $x : p \mapsto \boldsymbol{x} = x(p) \in \mathbb{R}^n$*, which can be determined by* $x^i : p \mapsto x^i = x^i(p) \in \mathbb{R}$*,* $\boldsymbol{x} = (x^1, \ldots, x^n)$*. Then the coordinate of a neighbor in* $\boldsymbol{v}$ *direction can be determined by a function* $c(t) = \boldsymbol{x} + t\boldsymbol{v} = (x^1 + tv^1, \ldots, x^n + tv^n)$*.*

The ***derivative operator*** $D_{v,p} : C^k \to \mathbb{R}$ *over* p *is a map satisfying:*

$$D_{v,p} f := \lim_{t\to 0} \frac{f(c(t)) - f(\boldsymbol{x})}{t} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ c)$$

By using the Chain Rule, we get $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f\circ c) = v^i \left.\frac{\partial}{\partial x^i}\right|_p f$. Since this is established for $\forall x_0 \in D(f)$, we obtain: $D_v f = v^i \cdot \left(\frac{\partial}{\partial x^i} f\right)$. All the derivative operators at point p constitude a set denoted by $\mathcal{D}_p(\mathbb{R}^n)$. As can be proven, $\mathcal{D}_p(\mathbb{R}^n)$ is a vector space. All the vectors at point p constitude a set denoted by $T_p(\mathbb{R}^n)$. As can be proven, $T_p(\mathbb{R}^n)$ is a vector space as well.

Definition 1.2.6 (Directional Derivative)**.** *Construct a function* $D_v f : D(f) \to \mathbb{R}$ *that satisfies:*

$$D_v f := v^i \cdot \left(\frac{\partial}{\partial x^i} f\right), \quad D_v f|_p := v^i \cdot \left.\frac{\partial}{\partial x^i}\right|_p f.$$

This means the ***directional derivative*** *in the direction of* $\boldsymbol{v}$*.*

At Figure 1.2.1 we can see that $\left\{ \left.\frac{\partial}{\partial x^i}\right|_p \right\}$ spans a vector space as is easy to prove. And if we construct a map $\phi : \mathcal{D}_p(\mathbb{R}^n) \to T_p(\mathbb{R}^n)$ at point p, we can prove that the map is an isomorphism! That means we can consider a vector as a directional derivative, so a vector v can act on a function f resulting in a real number(directional derivative) $v(f) = D_{v,p} f$ at point p. A directional derivative can be considered as a vector at p so we can draw it on $D(f) \in \mathbb{R}^n$.

In the sense of isomorphism of vector spaces $T_p(\mathbb{R}^n) \simeq \mathcal{D}_p(\mathbb{R}^n)$, the standard basis $\{e_i\}$ corresponds to the set of partial derivatives $\left\{ \left.\frac{\partial}{\partial x^i}\right|_p \right\}$. From then on, we treats them as the same thing: $\boldsymbol{v} = v^i = v^i e_i = v^i \left.\frac{\partial}{\partial x^i}\right|_p$. Interestingly, there are different interpretations of this formula. Using Einstein summation convention, $\boldsymbol{v}$ is a combination of the bases. Using tensor notation $e_i = \left.\frac{\partial}{\partial x^i}\right|_p = \mathbf{1}$ then $\boldsymbol{v} = v^i = \boldsymbol{v} \cdot \mathbf{1}$.

If we denote $\left.\frac{\partial}{\partial x^i}\right|_p = \left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right)_p$ as ∇, we obtain:

$$\boldsymbol{v} = \boldsymbol{v} \cdot \nabla \tag{1.3}$$

This is quit enlightening because if we know that $\left.\frac{\partial}{\partial x^1}\right|_p$ is a unit vector in $\mathcal{D}_p(\mathbb{R}^n)$, then $\nabla = \left.\frac{\partial}{\partial x^i}\right|_p$ forms a unit matrix, e.i. $\nabla = \mathbf{1}$.

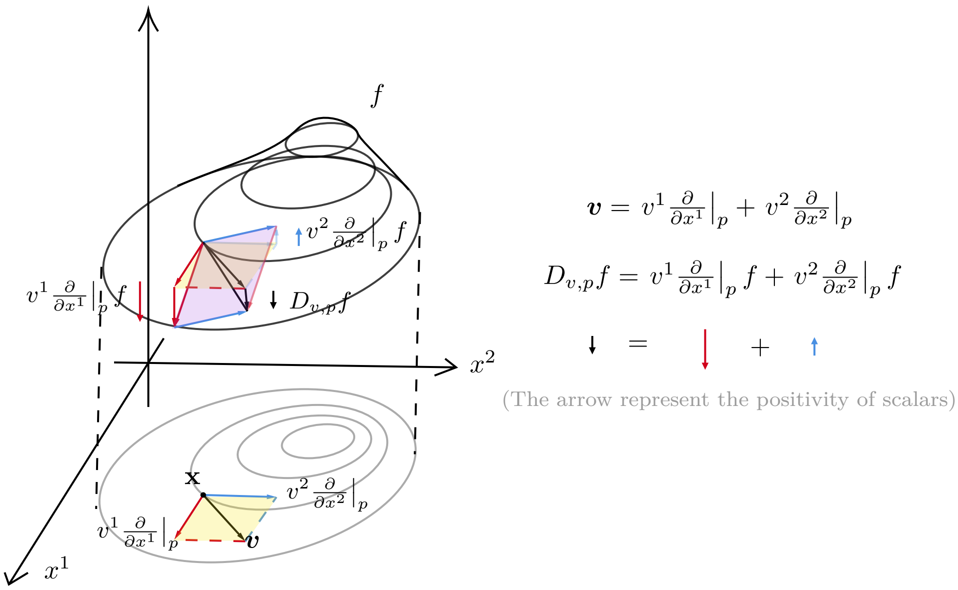

Figure 1.2.1: A vector and its derivatives

Definition 1.2.7 (Gradient)**.** ∇ *is a **differential operator** at point p. f is a function defined over $\mathbb{R}^n$. The gradient of f at p is defined by:*

$$\mathbf{grad} f := \nabla f$$

And ∇ acting on f is defined by:

$$\nabla f := \frac{\partial}{\partial x^i}\bigg|_p f$$

Proposition 1.2.1. *Direction derivative (unit) along the gradient is larger.*

Proof. Given a unit vector $\boldsymbol{v}$ at point p that $\|\boldsymbol{v}\| = 1$. The directional derivative

$$D_{v,p}f = \boldsymbol{v}(f) = (\boldsymbol{v} \cdot \nabla)(f) = \boldsymbol{v} \cdot (\nabla f).$$

Because

$$\boldsymbol{v} \cdot (\nabla f) = \|\boldsymbol{v}\|\|\nabla f\| \cos\theta \leq \|\nabla f\| = \|\mathbf{grad} f\|,$$

taken the equality when and only when $\boldsymbol{v} = \lambda(\mathbf{grad} f)$, where $\lambda \in \mathbb{R}$. $\square$

There is something to note (to be change):

(1) There are at least two notations of denoting the derivative over p: $\frac{\mathrm{d}}{\mathrm{d}x}\big|_p f$ and $\frac{\mathrm{d}}{\mathrm{d}x} f(p)$. We choose the first one as another one causes some ambiguity. The derivative at p is a map $C^\infty \to \mathbb{R}$ which eats a function and spit out a real number. And the second one seems that f eats p and spit out a real number first as if the derivative acts on $\mathbb{R}$! Right now $\frac{\mathrm{d}}{\mathrm{d}x}\big|_p$ is just the very function defined over p and its neighborhood, but if we pass through all the point in the domain we obtain a function from $\mathbb{R}$ to $\mathbb{R}$!

Definition 1.2.8 (Derivative Function $\frac{\mathrm{d}}{\mathrm{d}x}f$)**.** *Construct a function $\frac{\mathrm{d}}{\mathrm{d}x}f : \mathbb{R} \to \mathbb{R}$ satisfying:*

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}f\right)(p) := \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f,$$

*which shares the same domain with f. $\frac{\mathrm{d}}{\mathrm{d}x}f$ is called the **derivative function**.*

Example 1.2.2. *Calculate the derivative of the function $f(x) = x^2$*

Solution. Let $c(t) = (p + vt)$.

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f = \lim_{t\to 0}\frac{(p+vt)^2 - p^2}{t} = \lim_{t\to 0}(2vp + v^2t) = 2vp + o(t) = 2vp.$$

So $\frac{\mathrm{d}}{\mathrm{d}x}f$ satisfies:

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}f\right)(p) = \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f = 2vp, \quad \forall p \in D(f).$$

□

Definition 1.2.8 seems to be redundant because we see no difference in numerical values. But we shall keep in mind that derivative function is a constructed function. It assigns a derivative value to each point but is defined by $\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p$, that is, a map from the origin function to a derivative value over a certain point p.

So how shall we view derivatives? A value, a map, or a real-valued function? The difference between the first two point is the same that y between $y(x)$ in the function $y = y(x)$. So let's clarify it:

(i) $\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p$ is a map $C^\infty \to \mathbb{R}$ called a **derivative operator**.

(ii) $\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f$ is the image of the derivative operator, which called a **derivative**.

(iii) $\left(\frac{\mathrm{d}}{\mathrm{d}x}f\right)$ is a real-valued function called a **derivative function**.

(iv) $\left(\frac{\mathrm{d}}{\mathrm{d}x}f\right)(p)$ is the image of the derivative function, which is the same as (ii) on the same point.

(2) Notice that I use $c(t) = x_0 + vt$ instead of $x_0 + \Delta x$, because I want to introduce the parameter t into consideration. The reason is, if we talk about the process of taking the derivative, the picture in mind is that a point is approaching p gradually. Simply use $x \in D(f) \subseteq \mathbb{R}$ is a little vague because we think of x the axis function of point p ($x_0 = x(p)$), and we will see later that it fails to work when it comes to implicit function or multivariate function. So, we need another purer quantity to measure and represent the limit process.

$$\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f = \left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_{t=0}(f \circ c) = v\left.\frac{\mathrm{d}}{\mathrm{d}x}\right|_p f$$

In $y = f(x)$

Definition 1.2.9 (Name)**.** *Nothing.*

$$\begin{tikzcd}
 & & \mathbb{R} \\
\mathbb{R}|_t \arrow[r, "c(t)"] \arrow[urr, bend left, "\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(c(t))"] & U \arrow[ur, "D_v f|_p"] \arrow[dr, "f"'] \arrow[r, phantom, "\Uparrow D_v"] & \\
 & & \mathbb{R}
\end{tikzcd}$$

1.3 Differential and Cotangent Vector

1.4 Curve Theory and Surface Theory

1.5 Analytic Geometry

Chapter 2

Geometry in Oblique Coordinate System

2.1 Covariant and Contravariant

Chapter 3

Geometry as a Manifold

Bibliography

[1] J.L. Bell, M. Machover, ”A course in mathematical logic” , North-Holland (1977)

[2] Binary relation. Encyclopedia of Mathematics.

[3] Mapping. Encyclopedia of Mathematics.