

Chapter 3

Line and Surface Integrals

3.1 Parametric Curves

We will use bold letter $\mathbf{x} = (x_1, x_2, \dots, x_n)$ to denote a point (or a vector) in \mathbb{R}^n . When $n = 2$ and 3 , we also write $\mathbf{x} = (x, y) = x\mathbf{i} + y\mathbf{j}$ where $\mathbf{i} = (1, 0)$, $\mathbf{j} = (0, 1)$, and $\mathbf{x} = (x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, where $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$, respectively. The norm or length of a vector \mathbf{x} is given by

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

A *parametric curve* is a continuous map from an interval to $\mathbb{R}^n, n \geq 1$. Although the interval could be arbitrary, we will take it to be $[a, b]$ for definiteness. Let $\mathbf{r}(t) = (x_1(t), \dots, x_n(t)), t \in [a, b]$, be a parametric curve. We also let $\mathbf{x}(t), \gamma(t)$, or $\mathbf{c}(t)$ to denote a parametric curve. For a space curve, the vector notation: $\mathbf{x}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ is often used. When the context is clear, we drop the adjective parametric. The independent variable is usually taken to be t, z , etc.

It is best to imagine a parametric curve as the trajectory of a particle, where $|\mathbf{r}'(t)|$ stands for its speed at time t . Indeed, the average speed of the particle in the time duration $[t, t + \Delta t]$ is given by

$$\frac{|\mathbf{r}(t + \Delta t) - \mathbf{r}(t)|}{\Delta t}.$$

By the mean value theorem, $x_j(t + \Delta t) - x_j(t) = x'_j(t^*)\Delta t$ for some mean value $t^* \in [t, t + \Delta t]$. Hence

$$\mathbf{r}(t + \Delta t) - \mathbf{r}(t) = (x'_1(t_1^*), x'_2(t_2^*), \dots, x'_n(t_n^*))\Delta t, \quad t_j^* \in [t, t + \Delta t],$$

and

$$\frac{|\mathbf{r}(t + \Delta t) - \mathbf{r}(t)|}{\Delta t} = \sqrt{x_1'^2(t_1^*) + x_2'^2(t_2^*) + \dots + x_n'^2(t_n^*)} \rightarrow |\mathbf{r}'(t)|,$$

as $\Delta t \rightarrow 0$.

There are some common definitions for a parametric curves.

- It is a C^1 -curve if every $x_i = x_i(t), i = 1, \dots, n$, is continuously differentiable on $[a, b]$;
- It is *regular* if it is a C^1 -curve and $|\mathbf{r}'| > 0$ on $[a, b]$. Here $|\mathbf{r}'| = \sqrt{x_1'^2 + \dots + x_n'^2}$;
- It is a *closed curve* if $\mathbf{r}(a) = \mathbf{r}(b)$;
- It is a *simple curve* if it is one-to-one on $[a, b)$. (The point b is excluded because $\mathbf{r}(b) = \mathbf{r}(a)$ for a closed curve.)
- A *piecewise C^1 -* (resp. *regular*) *curve* consists of finitely many C^1 (resp. regular) regular curves \mathbf{r}_j on $[a_j, b_j], j = 1, \dots, n$, that satisfies $\mathbf{r}_j(b_j) = \mathbf{r}_{j+1}(a_{j+1}), j = 1, \dots, n - 1$. It allows \mathbf{r} to be non-differentiable at the endpoints of the intervals (although one-sided derivatives exist there).

The constant map $\mathbf{r}(t) = \mathbf{x}_0$ is a parametric curve which degenerates into a point. It satisfies $\mathbf{r}' = \mathbf{0}$. The condition $|\mathbf{r}'| > 0$ ensures \mathbf{r} really looks like a curve. The parametric curves studied in this course are all piecewise regular curves.

Example 3.1. Straight lines. A general parametric form of a straight line is given by

$$\mathbf{r}(t) = \mathbf{a} + \mathbf{d} t, \mathbf{d} \neq \mathbf{0}, \quad t \in (-\infty, \infty),$$

where \mathbf{a} and \mathbf{d} are given points in \mathbb{R}^n . The straight line passing two different points \mathbf{a} and \mathbf{b} is given by

$$\mathbf{r}(t) = \mathbf{a} + (\mathbf{b} - \mathbf{a})t, \quad t \in (-\infty, \infty),$$

so that $\mathbf{r}(0) = \mathbf{a}$ and $\mathbf{r}(1) = \mathbf{b}$. From $|\mathbf{r}'(t)| = |\mathbf{b} - \mathbf{a}|$ we see that this parametrization is regular.

Example 3.2. The circle of radius r centered at the origin is given by

$$\gamma_1(\theta) = (r \cos \theta, r \sin \theta) = r \cos \theta \mathbf{i} + r \sin \theta \mathbf{j}, \quad \theta \in [0, 2\pi].$$

When θ increases from 0 to 2π , a particle starting at $(1, 0)$ travels around the circle once in the anticlockwise way.

On the other hand, the parametric curve

$$\gamma_2(\theta) = (r \sin \theta, -r \cos \theta), \quad \theta \in [0, 2\pi],$$

goes around the circle in the clockwise way.

One may also consider $\gamma_3(\theta) = (r \cos 2\theta, r \sin 2\theta)$, $\theta \in [0, 2\pi]$, which, though the image is the same circle, goes around it twice as θ runs from 0 to 2π .

All these curves are regular.

Example 3.3. We know that the graph of the equation $y = ax^2 + bx + c$ represents a parabola. A convenient parametrization is given by

$$\mathbf{r}(x) = (x, ax^2 + bx + c) = x\mathbf{i} + (ax^2 + bx + c)\mathbf{j}, \quad x \in (-\infty, \infty).$$

In general, whenever a function $y = f(x)$, $x \in I$, where I is some interval is given, the expression

$$\mathbf{r}(x) = (x, f(x)) = x\mathbf{i} + f(x)\mathbf{j}, \quad x \in I,$$

gives a parametrization of the graph of f as a parametric curve in the plane.

Example 3.4. The helix is a space curve given by

$$\mathbf{h}(t) = r \cos t \mathbf{i} + r \sin t \mathbf{j} + t \mathbf{k}, \quad t \in (-\infty, \infty).$$

From $|\mathbf{h}'| = \sqrt{r^2 + 1}$ we see that the helix is always regular. Whether the curve is regular or not depends on the choice of the parametrization. For instance, if now we consider

$$\gamma(t) = r \cos t^2 \mathbf{i} + r \sin t^2 \mathbf{j} + t^2 \mathbf{k}, \quad t \in (-\infty, \infty),$$

whose image is identical to the image of \mathbf{h} , from $|\gamma'| = 2t\sqrt{r^2 + 1}$ we see that it is not regular at $t = 0$. One may say this is not a good parametrization.

3.2 Line Integrals of Functions

Now we define the integral of a function on a parametric curve.

Let \mathbf{r} be a C^1 parametric curve in \mathbb{R}^n , $n \geq 1$ defined on $[a, b]$ and f a continuous function defined on C , the image of $[a, b]$ under \mathbf{r} . We would like to introduce a concept of integration of f over \mathbf{r} . The guiding principle is that it should give the total mass of the curve when regarded as a very thin object in space with density f .

A partition $P, a = t_0 < t_1 < \dots < t_n = b$ on $[a, b]$ introduces points $\mathbf{p}_j = \mathbf{r}(t_j)$ on the curve. We may use the polygonal lines connecting \mathbf{p}_j and \mathbf{p}_{j+1} to approximate the curve. In particular, the mass of C is given approximately by

$$\sum_j f(\mathbf{q}_j) |\mathbf{p}_{j+1} - \mathbf{p}_j| = \sum_j f(\mathbf{r}(t_j^*)) |\mathbf{r}(t_{j+1}) - \mathbf{r}(t_j)|,$$

where $\mathbf{q}_j = \mathbf{r}(t_j^*)$ is a tag point. By applying the mean value theorem, $x_i(t_{j+1}) - x_i(t_j) = x_i'(t_j^*)\Delta t_j$ for some $t_j^* \in [t_j, t_{j+1}]$. Hence,

$$\sum_j f(\mathbf{q}_j) |\mathbf{r}(t_{j+1}) - \mathbf{r}(t_j)| = \sum_j f(\mathbf{q}_j) \sqrt{x_1'^2(t_{j1}^*) + x_2'^2(t_{j2}^*) + \cdots + x_n'^2(t_{jn}^*)} \Delta t_j, \quad t_{ji}^* \in [t_j, t_{j+1}].$$

Letting $\|P\| \rightarrow 0$, we see that

$$\sum_j f(\mathbf{q}_j) |\mathbf{p}_{j+1} - \mathbf{p}_j| \rightarrow \int_a^b f(\mathbf{r}(t)) \sqrt{x_1'^2 + \cdots + x_n'^2(t)} dt .$$

This consideration leads us to the following definition. Let C be a C^1 parametric curve in \mathbb{R}^n . For a continuous function f defined on C , its line integral along C is defined to be

$$\int_C f ds \equiv \int_a^b f(\mathbf{r}(t)) |\mathbf{r}'(t)| dt . \quad (3.1)$$

When \mathbf{r} maps bijectively onto its image and f is positive and, this integral gives the total mass of the image of this parametric curve with density f . Taking $f \equiv 1$,

$$|C| \equiv \int_a^b |\mathbf{r}'(t)| dt , \quad (3.2)$$

yields the length of the image of this parametric curve.

We have seen that one may alter the way of parametrization while keeping the image unchanged. Let γ_1 be a regular parametric curve on $[a, b]$. Another regular parametric curve γ_2 on $[c, d]$ is called a *reparametrization* of γ_1 if there is a C^1 bijection φ from $[a, b]$ to $[c, d]$ such that $\gamma_1(t) = \gamma_2(\varphi(t))$. Applying the chain rule to this relation, $\gamma_1'(t) = \gamma_2'(\varphi(t))\varphi'(t)$, which implies $|\gamma_1'(t)| = |\gamma_2'(\varphi(t))||\varphi'(t)|$. Since both curves are regular, we see that φ' never vanishes, hence either $\varphi' > 0$ or $\varphi' < 0$. For instance, let $\gamma_1(t) = (\cos t, \sin t)$ $t \in [0, \pi]$ and $\gamma_2(z) = (z, \sqrt{1-z^2})$, $z \in [-1, 1]$. γ_2 is a reparametrization of γ_1 and they are related by $z = \varphi(t) = \cos t$. Here $\varphi'(t) = -\sin t < 0$ on $[0, \pi]$.

Given two regular curves $\gamma_i, i = 1, 2$, which map bijectively onto the same image $C \subset \mathbb{R}^n$. Clearly we can find a bijective map φ satisfying the relation $\gamma_1(t) = \gamma_2(\varphi(t))$. One can show that φ is also C^1 so that indeed γ_2 is a reparametrization of γ_1 .

When we talk about the length or the mass of a curve, certainly we mean it to be a quantity which depends only on its form in the ambient space. In other words, no matter which parametrization you choose to perform the calculation, the end result should be the same. This is the content of the following theorem.

Theorem 3.1. *The line integral (3.1) is independent of reparametrization.*

More precisely, let \mathbf{r}_1 and \mathbf{r}_2 be two regular curves on $[a, b]$ and $[c, d]$ respectively sharing the same image. Then

$$\int_a^b f(\mathbf{r}_1(t))|\mathbf{r}'_1(t)| dt = \int_c^d f(\mathbf{r}_2(t))|\mathbf{r}'_2(t)| dt .$$

Proof. Let φ be the C^1 -map that relates the two curves. We have $\mathbf{r}_1(t) = \mathbf{r}_2(\varphi(t))$, $t \in [a, b]$. Since φ maps $[a, b]$ onto $[c, d]$ bijectively and φ' never vanishes, either $\varphi' > 0$ or $\varphi' < 0$ on $[a, b]$. In the first case, $\varphi(a) = c$ and $\varphi(b) = d$. From $\mathbf{r}_1(t) = \mathbf{r}_2(\varphi(t))$, we have $\mathbf{r}_1(t) = \mathbf{r}_2(\varphi(t))\varphi'(t)$. Therefore, by change of variables,

$$\begin{aligned} \int_c^d f(\mathbf{r}_2(z))|\mathbf{r}'_2(z)| dz &= \int_a^b f(\mathbf{r}_2(\varphi(t))) \frac{|\mathbf{r}'_1(t)|}{|\varphi'(t)|} \varphi'(t) dt \\ &= \int_a^b f(\mathbf{r}_1(t))|\mathbf{r}'_1(t)| dt . \end{aligned}$$

In case $\varphi(a) = d$ and $\varphi(b) = c$, $\varphi' < 0$. We have

$$\begin{aligned} \int_c^d f(\mathbf{r}_2(z))|\mathbf{r}'_2(z)| dz &= \int_b^a f(\mathbf{r}_2(\varphi(t))) \frac{|\mathbf{r}'_1(t)|}{|\varphi'(t)|} \varphi'(t) dt \\ &= \int_a^b f(\mathbf{r}_1(t))|\mathbf{r}'_1(t)| dt . \end{aligned}$$

□

The definition is readily extended to any piecewise regular curve C using the decomposition $C = C_1 + \cdots + C_n$ where C_j 's are regular:

$$\int_C f ds = \sum_j \int_{C_j} f ds . \quad (3.3)$$

Formula (3.1) takes a special form when the curve is the graph of a function. Consider the plane curve $C = \{(x, \varphi(x)) : x \in [a, b]\}$ where φ is continuous. In the trivial parametrization $\mathbf{r} : x \mapsto (x, \varphi(x))$. From $\mathbf{r}' = (1, \varphi'(x))$ and $|\mathbf{r}'| = \sqrt{1 + \varphi'^2}$, (3.1) becomes

$$\int_C f ds = \int_a^b f(x, \varphi(x))\sqrt{1 + \varphi'^2(x)} dx . \quad (3.4)$$

Example 3.5. Evaluate

$$\int_C (x - 3y^2 + z) ds ,$$

where C is the line segment connecting the origin to $(1, 1, 1)$ given by $t(1, 1, 1)$, $t \in [0, 1]$.

Clearly, $\mathbf{r}(t) = (t, t, t)$ implies $|\mathbf{r}'(t)| = \sqrt{3}$. Therefore,

$$\int_C (x - 3y^2 + z) ds = \int_0^1 (t - 3t^2 + t)\sqrt{3} dt = 0.$$

Example 3.6. Consider the following two line segments C_1 and C_2 given respectively by $\mathbf{r}_1 = t\mathbf{i} + t\mathbf{j}$, $t \in [0, 1]$ and $\mathbf{r}_2(t) = \mathbf{i} + \mathbf{j} + t\mathbf{k}$, $t \in [0, 1]$. Here $\mathbf{r}_1(1) = \mathbf{r}_2(0)$, so they form a piecewise regular curve C . Evaluate

$$\int_C (x - 3y^2 + z) ds.$$

Well, it just has to do things separately.

$$\int_{C_1} (x - 3y^2 + z) ds = \int_0^1 (t - 3t^2 + 0)\sqrt{2} dt = -\frac{\sqrt{2}}{2}.$$

$$\int_{C_2} (x - 3y^2 + z) ds = \int_0^1 (1 - 3 + t) \times 1 dt = -\frac{3}{2}.$$

So

$$\int_C (x - 3y^2 + z) ds = -\frac{\sqrt{2}}{2} - \frac{3}{2}.$$

There are numerous ways to parametrize a curve. One may ask: Among all these possible parametrizations, is there an optimal one? The following theorem suggests an answer to this question.

Theorem 3.2. *Each regular curve in \mathbb{R}^n admits a parametrization \mathbf{r} satisfying $|\mathbf{r}'| \equiv 1$.*

This special parametrization is called the parametrization by arc-length. It represents a particle running along the curve in unit speed. An arc-length parametrization is unique among all parametrizations of the same orientation.

Proof. Pick a regular parametrization \mathbf{r} of the given curve C . That is, \mathbf{r} is a regular parametrization which maps some $[a, b]$ bijectively onto C . Define

$$S(t) = \int_a^t |\mathbf{r}'(z)| dz, \quad t \in [a, b].$$

Since $|\mathbf{r}'| > 0$, as t runs from a to b , S runs from 0 to L , the length of C . Let ψ be the inverse map of $S : [0, L] \rightarrow [a, b]$. Let

$$\gamma(s) = \mathbf{r}(\psi(s)), \quad s \in [0, L].$$

Then γ maps $[0, L]$ bijectively onto C , and

$$\begin{aligned}\gamma'(s) &= \mathbf{r}'(\psi(s))\psi'(s) \\ &= \mathbf{r}'(t) \frac{1}{|S'(\psi(s))|} \\ &= \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}, \quad t = \psi(s).\end{aligned}$$

Hence $|\gamma'(s)| = 1$ on $[0, L]$. □

3.3 Geometric Curves

What is a (geometric) curve? A set C in \mathbb{R}^n is called a C^1 - (resp. *regular*) *curve* if it is the bijective image of a C^1 - (resp. *regular*) parametric curve. (When the curve is closed, it is allowed that the endpoints of the interval being mapped to the same.) Such parametric curve is called a C^1 - (resp. *regular*) *parametrization* of the curve C . A curve admits many different parametrization. Let \mathbf{r}_1 be such a regular parametrization on $[a, b]$ and \mathbf{r}_2 be another on $[c, d]$. From the definition of regular parametrization, there is a bijective C^1 -map φ so that $\mathbf{r}_2(z) = \mathbf{r}_1(\varphi(z))$ for $z \in [c, d]$, and the following alternative holds: Either

- (a) $\varphi(a) = \varphi(c), \varphi(b) = \varphi(d)$ and $\varphi' > 0$ on $[a, b]$, or
- (b) $\varphi(a) = \varphi(d), \varphi(b) = \varphi(c)$ and $\varphi' < 0$ on $[a, b]$.

Proposition 3.1 which asserts the line integral (3.1) is independent of reparametrization shows that (3.1) is a property of the curve, that is, a subset in the ambient space rather than its parametrization.

Example 3.7. Find the mass of the arc in the shape of a half disk $y^2 + z^2 = 1, z \geq 0$ in \mathbb{R}^3 with density $2 - z$.

We have freedom to pick a parametrization to perform the line integral. Let us take $\gamma(t) = (\cos t, \sin t, 0)$, $t \in [0, \pi]$. Then $\gamma'(t) = (-\sin t, \cos t, 0)$, $|\gamma'(t)| = 1$, so

$$\int_C (2 - z) ds = \int_0^\pi (2 - \sin t) dt = 2\pi - 2.$$

If we choose another parametrization, say, let $\mathbf{r}(y) = y\mathbf{j} + \sqrt{1 - y^2}\mathbf{k}$, $y \in [-1, 1]$. We have $\mathbf{r}'(y) = \sqrt{1/(1 - y^2)}$, so

$$\int_C (2 - z) ds = \int_{-1}^1 (2 - \sqrt{1 - y^2}) \frac{1}{\sqrt{1 - y^2}} dy = 2\pi - 2,$$

same as above.

From the above discussion, all regular reparametrizations of a curve can be put into two groups according to their “orientation”. An oriented curve is a curve with a chosen orientation, that is, a choice of preferred parametrization. Let \mathbf{r}_1 and \mathbf{r}_2 be two parametrizations of an oriented curve C and let φ be the C^1 -map such that $\mathbf{r}_2(z) = \mathbf{r}_1(\varphi(z))$. By the chain rule, $\mathbf{r}'_2(z) = \mathbf{r}'_1(t)\varphi'(z)$ where $t = \varphi(z)$. As φ' is positive, the unit vector $\mathbf{r}'_1/|\mathbf{r}'_1| = \mathbf{r}'_2/|\mathbf{r}'_2|$ which is tangential to the curve C is independent of parametrization of the same orientation. It is called the *unit tangent* of the curve at the point $\mathbf{r}_1(t)$ and will be denoted by \mathbf{t} . When the orientation of the parametrization is reversed, the unit tangent changes to $-\mathbf{t}$ since now φ' is negative.

Example 3.8. Determine the unit tangent of the unit circle at the point $(\sqrt{3}/2, 1/2)$.

First, we choose a parametrization of the circle. The handiest one is

$$\mathbf{r}(\theta) = \cos \theta \mathbf{i} + \sin \theta \mathbf{j}, \quad \theta \in [0, \pi],$$

where $\mathbf{r}(\pi/6) = (\sqrt{3}/2, 1/2)$. We have

$$\begin{aligned} \mathbf{t}(\sqrt{3}/2, 1/2) &= \frac{\mathbf{r}'}{|\mathbf{r}'|} \\ &= \frac{-\sin \theta \mathbf{i} + \cos \theta \mathbf{j}}{1} \\ &= (-1/2, \sqrt{3}/2). \end{aligned}$$

Alternatively, one may use the parametrization

$$\gamma(t) = (-t, \sqrt{1-t^2}), \quad t \in [-1, 1],$$

where $\gamma(-\sqrt{3}/2) = (\sqrt{3}/2, 1/2)$. We have $\gamma'(t) = (-1, -t/\sqrt{1-t^2})$ and $|\gamma'(t)| = 1/\sqrt{1-t^2}$. Therefore,

$$\begin{aligned} \mathbf{t}(\sqrt{3}/2, 1/2) &= \frac{\gamma'(-\sqrt{3}/2)}{|\gamma'(-\sqrt{3}/2)|} \\ &= (-1/2, \sqrt{3}/2), \end{aligned}$$

which is the same as above.

The same consideration can be extended to all piecewise regular curves.

Oriented curves are relevant when we integrate a vector field along a curve in the next section.

Finally, we point out there are two operations on curves we will encounter. First, for an oriented curve C we use $-C$ to denote the oriented curve obtained by reversing the

orientation of C . More precisely, let $\mathbf{r} : [a, b] \rightarrow C$ be a regular parametrization of the curve C . The parametric curve $\mathbf{r}_1(t) = \mathbf{r}(a + b - t)$ becomes a regular parametrization of C in reverse orientation.

Second, let C_1 be a oriented curve parametrized by \mathbf{r}_1 on $[a, b]$ and C_2 be another oriented curve parametrized by \mathbf{r}_2 on $[c, d]$. In case $\mathbf{r}_1(b) = \mathbf{r}_2(c)$, we can put these two curves together to form a new curve $C_1 + C_2$. It is again a piecewise regular curve when both C_1 and C_2 are piecewise regular. In fact, let \mathbf{r}_1 be on $[a, b]$ and \mathbf{r}_2 on $[c, d]$. The curve $\mathbf{r}(t) = \mathbf{r}_1(t)$, $t \in [a, b]$ and $\mathbf{r}(t) = \mathbf{r}_2(t + c - b)$, $t \in [b, b + d - c]$ is a parametrization of \mathbf{r} on $[a, b + d - c]$.

3.4 Vector Fields

Vector fields are usually defined in open sets in \mathbb{R}^n . It is necessary to define an open set first.

An open set is a set consisting of interior points or a set without boundary points. A formal definition is that S is *open* if for every $\mathbf{x} \in S$, there is a ball containing \mathbf{x} which is completely contained in S . The ball $\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < 1\}$ is an open set. In general, the regions we performed integration in the previous two chapters consist of interior points and boundary points. Without counting the boundary points, that is, the interior of a region, forms an open set.

Let S be a subset of \mathbb{R}^n . A *vector field* $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))$ in S is an assignment of an n -tuple of function to each point $\mathbf{x} \in S$. It is a continuous vector field if all components F_i 's are continuous in S . A vector field \mathbf{F} defined in an open set S is called a *C^1 -vector field* if all components F_j 's are C^1 -functions in S . Note that since F_j is well-defined in all nearby points around a specific point in S , one is able to form difference quotients and hence partial derivatives, and that is the reason we need the set to be open.

The definition of a vector field which is just putting some functions together looks rather trivial. In fact, it becomes meaningful when one considers it on a surface or a manifold (a generalized surface). It is required further that the vector field lies on the tangent space of the manifold. Many laws of physics are described by differential equations and differential equations are completely determined by its associated vector field. Hence the study of vector fields on manifolds is essentially equal to the study of differential equations on the manifolds.

Example 3.8. Consider the vector field

$$\mathbf{F} = \frac{1}{|\mathbf{x}|}(x_1, x_2, \dots, x_n).$$

This expression is well-defined, and in fact, infinitely many times differentiable away from

the origin. Hence its natural domain of definition is the open set $\mathbb{R}^2 \setminus \{\mathbf{0}\}$. The direction of this vector field at the point \mathbf{x} is $\mathbf{x}/|\mathbf{x}|$ and its magnitude is always equal to 1. Any vector field of the form $f(|\mathbf{x}|)\mathbf{x}$ is called a radial vector field. \mathbf{F} is a radial vector field.

Example 3.9. Let

$$\mathbf{R}(x, y) = -y\mathbf{i} + x\mathbf{j} .$$

This is a smooth vector field in the entire plane. A vector field is smooth if all its components are infinitely differentiable functions. Since $\mathbf{R}(x, y) \cdot (x, y) = 0$, this vector field is perpendicular to its position at every point. It is a rotating vector field.

Example 3.10. The notion of a vector field is originated from physics. The presentation of a point mass (say the Sun) generates a force field in space. At the center the force becomes infinity, hence the vector field is only defined in the entire space minus the center (the location of the point mass). Any point mass (say the Earth) would be attracted to the center by a gravitational force. According to Newton's law, the gravitational force is a field given by

$$\mathbf{G}(\mathbf{x}) = -\frac{GMm}{|\mathbf{x}|^3} \mathbf{x} ,$$

where M, m are respectively the masses of the Sun and the Earth, and G is the gravitational constant. The magnitude of the force is

$$|\mathbf{G}(\mathbf{x})| = \frac{GMm}{r^2} ,$$

where $r = |\mathbf{x}|$ is the distance between the Earth and the Sun. The direction of force is $-\mathbf{x}/r$, that is, it is attracting toward the Sun.

Example 3.11. The gravitational force of the Earth exerting on objects on its surface may be approximated by

$$\mathbf{V} = (0, 0, -g) = -g\mathbf{k} ,$$

where g is the gravitational constant. This is a constant vector field.

3.5 Line Integral Of Vector Fields

In this section we show how to integrate a vector field along a curve.

The background comes from physics. Suppose an object is moved from Point \mathbf{a} to Point \mathbf{b} under the influence of a constant force field \mathbf{F} . Its work done by the force field is given by $\mathbf{F} \cdot (\mathbf{b} - \mathbf{a})$. Now we would like to calculate the work done of a non-constant force field along a parametric curve $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^n$. Let P be a partition $t_0 = a < t_1 < \dots < t_n = b$ on

$[a, b]$ and $\mathbf{p}_j = \mathbf{r}(t_j)$. We approximate the curve by polygonal lines formed by connecting the points \mathbf{p}_j 's. The work done moving from \mathbf{p}_j to \mathbf{p}_{j+1} being $\mathbf{F}(\mathbf{p}_j) \cdot (\mathbf{p}_{j+1} - \mathbf{p}_j)$ where \mathbf{p}_j is taken to be a tag point. When the partition is very fine, the force field is almost a constant near \mathbf{p}_j . Therefore, a good approximation to the work done along the curve is given by

$$\sum_j \mathbf{F}(\mathbf{p}_j) \cdot (\mathbf{p}_{j+1} - \mathbf{p}_j) = \sum_j \mathbf{F}(\mathbf{p}_j) \cdot \frac{\mathbf{p}_{j+1} - \mathbf{p}_j}{\Delta t_j} \Delta t_j.$$

As the term

$$\begin{aligned} & \frac{\mathbf{p}_{j+1} - \mathbf{p}_j}{\Delta t_j} \\ &= \frac{\mathbf{r}(t_{j+1}) - \mathbf{r}(t_j)}{\Delta t_j} \\ &\rightarrow \mathbf{r}' \\ &= \mathbf{t}|\mathbf{r}'| \end{aligned}$$

as $\|P\| \rightarrow 0$ where \mathbf{t} is the unit tangent of the parametric curve, the work done by \mathbf{F} along the curve \mathbf{r} is and should be defined to be

$$\int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) dt ,$$

or

$$\int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{t}|\mathbf{r}'(t)| dt .$$

Let G be a nonempty open set in \mathbb{R}^n and C an oriented C^1 -parametric curve in G . For a continuous vector field \mathbf{F} in G , one defines the line integral of the vector field \mathbf{F} along C to be

$$\int_C \mathbf{F} \cdot d\mathbf{r} \equiv \int_C \mathbf{F} \cdot \mathbf{t} ds , \quad (3.5)$$

where \mathbf{t} is the unit tangent of C . From this definition one readily sees that the line integral of a vector field is independent of reparametrization of the same orientation, but changes to a negative sign when the orientation is reversed.

Using $\mathbf{F} = (F_1, \dots, F_n)$ and $\mathbf{t} = \mathbf{r}'/|\mathbf{r}'|$, we have

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b (F_1(\mathbf{r}(t))x'_1(t) + \dots + F_n(\mathbf{r}(t))x'_n(t)) dt , \quad (3.6)$$

where \mathbf{r} is a parametrization of C on $[a, b]$.

When $n = 2$, a vector field is usually written as $\mathbf{F} = (M, N)$ or $M\mathbf{i} + N\mathbf{j}$. Then (3.6) is the same as

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b (M(\mathbf{r}(t))x'(t) + N(\mathbf{r}(t))y'(t)) dt ,$$

where $\mathbf{r} : [a, b] \rightarrow C$ is a parametrization of C and $\mathbf{r}(t) = (x(t), y(t))$. In view of this formula, we also use

$$\int_C M dx + N dy$$

to denote the line integral of the vector field \mathbf{F} .

Similarly, when $n = 3$, we let $\mathbf{F} = (M, N, P)$ or $M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, and express the line integral of \mathbf{F} as

$$\int_C M dx + N dy + P dz .$$

When a parametrization \mathbf{r} is given,

$$\int_C M dx + N dy + P dz = \int_a^b (M(\mathbf{r}(t))x'(t) + N(\mathbf{r}(t))y'(t) + P(\mathbf{r}(t))z'(t)) dt .$$

Example 3.12. Find the work done of the vector field $\mathbf{F} = (y - x^2)\mathbf{i} + (z - y^2)\mathbf{j} + (x - z^2)\mathbf{k}$ along the curve $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$, $t \in [0, 1]$, from $(0, 0, 0)$ to $(1, 1, 1)$.

We have

$$\frac{d\mathbf{r}}{dt} = \mathbf{i} + 2t\mathbf{j} + 3t^2\mathbf{k} ,$$

$$\mathbf{F}(\mathbf{r}(t)) = (t^2 - t^2)\mathbf{i} + (t^3 - t^4)\mathbf{j} + (t - t^6)\mathbf{k} ,$$

and

$$\mathbf{F}(\mathbf{r}(t)) \cdot \frac{d\mathbf{r}}{dt} = 2t^4 - 2t^5 + 3t^3 - 3t^8 .$$

Therefore,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (2t^4 - 2t^5 + 3t^3 - 3t^8) dt = \frac{29}{60} .$$

There are other physical interpretation of the line integral (3.5) other than work done. Let C be an oriented curve which admits a regular parametrization \mathbf{r} on $[a, b]$. The term $\mathbf{F}(\mathbf{r}) \cdot \mathbf{t}$ is the projection of \mathbf{F} onto the tangential direction of C at the point $\mathbf{r}(t)$. Image now instead of a force field, \mathbf{F} is the velocity of some flow (fluid or gas whatsoever). The integral on the right hand side of (3.5) becomes the amount of the flow along the curve. When the curve is a closed one, it is called the *circulation* of the flow around the curve. And the notation is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} .$$

Example 3.13. Find the flow of the velocity field $\mathbf{V} = x\mathbf{i} + z\mathbf{j} + y\mathbf{k}$ along the curve

$$\gamma(\theta) = \cos \theta \mathbf{i} + \sin \theta \mathbf{j} + \theta \mathbf{k}, \quad \theta \in [0, \pi/2] .$$

We have

$$\frac{d\gamma}{d\theta} = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j} + \mathbf{k},$$

and

$$\mathbf{V}(\gamma(t)) = \cos\theta\mathbf{i} + \theta\mathbf{j} + \sin\theta\mathbf{k}.$$

Therefore, the flow through the curve is

$$\begin{aligned} \int_0^{\pi/2} \mathbf{V}(\gamma(\theta)) \cdot \frac{d\gamma}{d\theta} d\theta &= \int_0^{\pi/2} (-\sin\theta\cos\theta + t\cos\theta + \sin\theta) d\theta \\ &= \frac{\pi}{2} - \frac{1}{2} \end{aligned}$$

There is another interpretation of the line integral (3.5) when $n = 2$. Let $\mathbf{t} = (x'(t), y'(t))/|\mathbf{r}'|$ be the unit tangent of the oriented curve at the point $(x(t), y(t))$. The unit vector $\mathbf{n} = (y'(t), -x'(t))/|\mathbf{r}'|$ is a vector field along the curve. From $\mathbf{n} \cdot \mathbf{t} = y'(t)x'(t) - x'(t)y'(t) = 0$ we see that \mathbf{n} is perpendicular to \mathbf{t} . It is called the (unit) *normal* vector field along the curve. When the curve is a closed one and \mathbf{t} is in the anticlockwise direction, it is easy to see that \mathbf{n} points outward, so it is the unit outward normal. For a vector field \mathbf{F} defined in some open set containing the closed curve C oriented in the anticlockwise direction, the amount of \mathbf{F} flowing across the boundary at \mathbf{r} in unit time is given by $\mathbf{F} \cdot \mathbf{n}$. Consequently, the integral

$$\int_C \mathbf{F} \cdot \mathbf{n} ds$$

gives the flux of \mathbf{F} across the curve C .

When $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$, the flux is given by

$$\int_C \mathbf{F} \cdot \mathbf{n} ds = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{n}(\mathbf{r}(t)) dt = \int_a^b (M(\mathbf{r}(t))y'(t) - N(\mathbf{r}(t))x'(t)) dt,$$

which is the line integral of $-N\mathbf{i} + M\mathbf{j}$ along C . Hence, the flux is also given by the formula

$$\int_C \mathbf{F} \cdot \mathbf{n} ds = \int_C M dy - N dx. \quad (3.7)$$

Example 3.14. Find the flux of $\mathbf{F} = (x - y)\mathbf{i} + x\mathbf{j}$ across the circle $x^2 + y^2 = 1$.

We choose the standard parametrization for the unit circle: $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j}$, $t \in [0, 2\pi]$. Then $\mathbf{r}' = -\sin t\mathbf{i} + \cos t\mathbf{j}$ and $\mathbf{n} = \cos t\mathbf{i} + \sin t\mathbf{j}$. The flux is given by

$$\begin{aligned} \int_C M dy - N dx &= \int_0^{2\pi} ((\cos t - \sin t)\cos t + \cos t \sin t) dt \\ &= \int_0^{2\pi} \cos^2 t dt \\ &= \pi. \end{aligned}$$

Summarizing, the line integral of a vector field is

- When \mathbf{F} is a force field, the line integral of \mathbf{F} along a curve C beginning at point A and ending at B gives the work done of the force on a particle (or person) moving from A to B .
- When \mathbf{F} is the velocity of some fluid, the line integral of \mathbf{F} along a closed curve C is called the circulation of the fluid around C . It measures the amount of fluid going around the curve in unit time.
- When $n = 2$ and \mathbf{F} is the velocity of some fluid, the line integral

$$\int_C Q dx - P dy = \int_C \mathbf{F} \cdot \mathbf{n} ds ,$$

gives the flux of the fluid across the curve C , where \mathbf{n} is the preferred unit normal of C . When C is a closed curve oriented in the anticlockwise way, the unit normal points outward.

3.6 Independence of Path

A main theme on the line integral of a vector field is independence of path. To formulate it let us consider a vector field \mathbf{F} in an open, connected subset G in \mathbb{R}^n . A set is *connected* if two points inside can be connected by a continuous curve lying inside this set. When the set is open, one may always approximate a curve inside this set by a regular curve inside the set. The vector field \mathbf{F} is called *independent of path* if for any two points $A, B \in G$, the line integral

$$\int_C \mathbf{F} \cdot d\mathbf{x} ,$$

gives the same value as long as C is a piecewise regular parametric curve running from A to B inside G . Again it suffices to take C to be a regular parametric curve, since we can approximate a piecewise regular parametric curve by regular parametric curves which are formed by smoothing out the junctions of the given curve.

A vector field \mathbf{F} is called a *conservative vector field* or a *gradient vector field* if it is the gradient of some function. In other words, it is conservative if there exists some differentiable Φ in G such that $\mathbf{F} = \nabla\Phi$. The function Φ is called a *potential* of \mathbf{F} . The potential (function) is essentially unique; in a connected set, two potential functions differ by a constant.

The following vector fields are conservative.

- The constant vector field $\mathbf{F} = (a_1, \dots, a_n)$ whose potential is $\Phi(x) = a_1x_1 + \dots + a_nx_n$. For instance, the gravity on the surface of the earth is $(0, 0, -g)$ where g is the gravitational constant.

- The gravitational force between two objects is given by

$$\mathbf{G}(x, y, z) = -\frac{GmM}{r^3} \mathbf{r}, \quad r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}.$$

Its potential is given by

$$\Phi(x, y, z) = \frac{GmM}{r}.$$

Theorem 3.3. Let \mathbf{F} be a continuous vector field in a connected, open set G in \mathbb{R}^n . The following statements are equivalent:

- (a) \mathbf{F} is independent of path;
 (b) For any closed curve C in G ,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0,$$

- (c) \mathbf{F} is conservative.

Proof. (a) \Rightarrow (b). Let C be a closed piecewise regular curve in G . Mark any two points A and B on the curve and express the curve as $C = C_1 - C_2$ where both C_1 and C_2 run from A to B . By (a)

$$\begin{aligned} \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_{C_1 - C_2} \mathbf{F} \cdot d\mathbf{r} \\ &= \int_{C_1} \mathbf{F} \cdot d\mathbf{r} - \int_{C_2} \mathbf{F} \cdot d\mathbf{r} \\ &= 0. \end{aligned}$$

(b) \Rightarrow (a). Reverse the above argument.

(a) \Rightarrow (c). Fix a point O in G . Let $A(\mathbf{x})$ be any point in G and connect O to A by a piecewise regular parametric curve C inside G . Note that here we have used the assumption that G is connected, otherwise we cannot always connect two points by a path. Define a function Φ by

$$\Phi(x) = \int_C \mathbf{F} \cdot d\mathbf{r}.$$

Since the vector field is independent of path, $\Phi(\mathbf{x})$ is independent of the choice of C . We claim that $\nabla\Phi = \mathbf{F}$ in G . For, we connect \mathbf{x} to the point $(x_1 + h, x_2, \dots, x_n)$ (h is small) by the curve $C_h : t \mapsto \mathbf{x} + ht\mathbf{e}_1$, $t \in [0, 1]$. Then $C + C_h$ becomes a curve from O to

$(x_1 + h, x_2, \dots, x_n)$. We have

$$\begin{aligned}
 & \frac{\Phi(\mathbf{x} + h\mathbf{e}_1) - \Phi(\mathbf{x})}{h} \\
 &= \frac{1}{h} \left(\int_{C+C_h} \mathbf{F} \cdot d\mathbf{r} - \int_C \mathbf{F} \cdot d\mathbf{r} \right) \\
 &= \frac{1}{h} \left(\int_C \mathbf{F} \cdot d\mathbf{r} + \int_{C_h} \mathbf{F} \cdot d\mathbf{r} - \int_C \mathbf{F} \cdot d\mathbf{r} \right) \\
 &= \frac{1}{h} \int_{C_h} \mathbf{F} \cdot d\mathbf{r} \\
 &= \frac{1}{h} \int_0^1 \mathbf{F}(\mathbf{x} + ht\mathbf{e}_1) \cdot (h, 0, \dots, 0) dt \\
 &= \int_0^1 F_1(x_1 + ht, x_2, \dots, x_n) dt \\
 &\rightarrow F_1(\mathbf{x}),
 \end{aligned}$$

as $h \rightarrow 0$. Hence $\partial\Phi/\partial x_1 = F_1$. Similarly, one can show $\partial\Phi/\partial x_j = F_j$ for $j = 2, \dots, n$.

(c) \Rightarrow (a). Let Φ be the potential for \mathbf{F} . Let C be a regular curve running from points A to B whose parametrization is given by \mathbf{r} on $[a, b]$ such that $\mathbf{r}(a) = A$ and $\mathbf{r}(b) = B$. We have

$$\begin{aligned}
 \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \sum_j F_j(\mathbf{r}(t)) x'_j(t) dt \\
 &= \int_a^b \sum_j \frac{\partial\Phi}{\partial x_j}(x_1(t), \dots, x_n(t)) x'_j(t) dt \\
 &= \int_a^b \frac{d}{dt} \Phi(x_1(t), \dots, x_n(t)) dt \\
 &= \Phi(x_1(t), \dots, x_n(t)) \Big|_a^b \\
 &= \Phi(B) - \Phi(A),
 \end{aligned}$$

which shows that the path integral is independent of the choice of path. Since we can approximate each piecewise regular curve by regular curves (simply by smoothing out the junctions), the vector field is also independent of the path formed by any piecewise regular curve. □

Remark. By the definition of connectedness, it is possible to connect two points in the set under consideration by a path (that is, a continuous curve). However, it is rather obvious that this path can be chosen as a simple, regular curve. Using this fact, one can construct the potential function for the vector field only involving simple, regular curves. With a

potential function available, the line integral of the vector field along any closed curve (not necessarily simple) vanishes. Therefore, in order a vector field to be conservative, it suffices to show it is independent of all simple regular curves.

From a practical point of view, one would like to find ways to determine whether a given vector field is conservative, and if yes, how to recover its potential.

The answer for this question starts with a necessary condition. Assume that \mathbf{F} is conservative and $\nabla\Phi = \mathbf{F}$. We have

$$F_j = \frac{\partial\Phi}{\partial x_j}, \quad i = 1, \dots, n.$$

Taking partial derivative in x_k ,

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial^2\Phi}{\partial x_j \partial x_k}.$$

By switching j and k ,

$$\frac{\partial F_k}{\partial x_j} = \frac{\partial^2\Phi}{\partial x_k \partial x_j}.$$

However, $\Phi_{jk} = \Phi_{kj}$. We arrive at

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial F_k}{\partial x_j}. \quad (3.8)$$

This condition does not involves the potential function. (3.8), called a compatibility condition, must be satisfied when the vector field \mathbf{F} is conservative. In the following we show that it is also sufficient as long as the vector field is defined everywhere.

Theorem 3.4. *Let \mathbf{F} be a C^1 -vector field in \mathbb{R}^n . It is conservative if and only if (3.8) holds.*

Proof. We have shown that (3.8) is necessary for \mathbf{F} to be a conservative vector field. To establish the converse, we claim the function Φ given by

$$\Phi(\mathbf{x}) = \int_0^1 \mathbf{F}(t\mathbf{x}) \cdot \mathbf{x} dt,$$

is a potential for \mathbf{F} . For,

$$\begin{aligned} \frac{\partial \Phi}{\partial x_j}(\mathbf{x}) &= \frac{\partial}{\partial x_j} \int_0^1 \sum_k F_k(t\mathbf{x}) x_k dt \\ &= \int_0^1 \sum_k \frac{\partial F_k}{\partial x_j}(\mathbf{x}) t x_k + F_j(t\mathbf{x}) dt \\ &= \int_0^1 \sum_k \frac{\partial F_j}{\partial x_k}(\mathbf{x}) t x_k + F_j(t\mathbf{x}) dt \\ &= \int_0^1 \frac{d}{dt} t F_j(t x_1, \dots, t x_n) dt \\ &= F_j(\mathbf{x}) , \end{aligned}$$

after using (3.8) in the third line, done. □

When $n = 2$ and $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$, the compatibility condition (3.8) becomes

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} . \quad (3.9)$$

When $n = 3$ and $\mathbf{F} = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, the compatibility condition (3.8) becomes

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} , \quad \frac{\partial M}{\partial z} = \frac{\partial P}{\partial x} , \quad \frac{\partial N}{\partial z} = \frac{\partial P}{\partial y} . \quad (3.10)$$

Example 3.15. Find the potential function (if exists) for the vector field

$$\mathbf{F} = (e^x \cos y + yz)\mathbf{i} + (xz - e^x \sin y)\mathbf{j} + (xy + z)\mathbf{k} .$$

We verify the existence of potential using Theorem 3.4,

$$\frac{\partial N}{\partial x} = -e^x \sin y + z = \frac{\partial M}{\partial y} ,$$

$$\frac{\partial P}{\partial y} = x = \frac{\partial N}{\partial z} ,$$

and

$$\frac{\partial M}{\partial z} = y = \frac{\partial P}{\partial x} .$$

It follows that the vector field is conservative.

To find the potential, we solve the equations

$$\frac{\partial \Phi}{\partial x} = e^x \cos y + yz ,$$

to get $\Phi(x, y, z) = e^x \cos y + xyz + g(y, z)$, where g is to be determined. Next, we use

$$\frac{\partial \Phi}{\partial y} = xz - e^x \sin y = -e^x \sin y + xz + \frac{\partial g}{\partial y},$$

which implies $\partial g / \partial y = 0$, that is, $g(y, z) = h(z)$ for some function h only depends on z . Finally, from

$$\frac{\partial \Phi}{\partial z} = xy + z = xy + \frac{dh}{dz},$$

we conclude $h(z) = z^2/2 + C$ where C is an arbitrary constant. The potential for the vector field is given by

$$\Phi(x, y, z) = e^x \cos y + xyz + \frac{z^2}{2} + C.$$

It is remarkable that the compatibility condition (3.8) may not be sufficient when the vector field is not defined in the entire space. Let us look at a remarkable example.

Consider the vector field

$$\mathbf{R}(x, y) = \frac{-y}{x^2 + y^2} \mathbf{i} + \frac{x}{x^2 + y^2} \mathbf{j}$$

which is C^1 in the plane except at the origin. Hence the vector field is defined only in the set $\mathbb{R}^2 \setminus \{(0, 0)\}$ which is open and connected. We compute

$$\frac{\partial M}{\partial y} = \frac{-1}{x^2 + y^2} + \frac{2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2},$$

$$\frac{\partial N}{\partial x} = \frac{1}{x^2 + y^2} + \frac{-2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

We see that (3.9) holds. However, consider the circle $C : [0, 2\pi] \rightarrow \mathbb{R}^2$ given by $\theta \mapsto (\cos \theta, \sin \theta)$. We have

$$\begin{aligned} \oint_C M dx + N dy &= \int_0^{2\pi} M(\mathbf{r}(\theta))x'(\theta) + N(\mathbf{r}(\theta))y'(\theta) d\theta \\ &= \int_0^{2\pi} (-\sin \theta)(-\sin \theta) + \cos \theta \cos \theta d\theta \\ &= 2\pi. \end{aligned}$$

According to Theorem 3.3, \mathbf{R} is not conservative.

3.7 Surfaces in Space

The treatment of surfaces is essentially parallel to that of curves.

A parametric surface is a continuous map from a region in \mathbb{R}^2 to \mathbb{R}^3 . Although it is okay to study surfaces in \mathbb{R}^n , our main concern is surfaces in space. To this end we only consider $n = 3$. For definiteness we take the region to be $R = [a, b] \times [c, d]$. In fact, most results continue to hold in a region in the plane. A *parametric surface* $\mathbf{r} : R \rightarrow \mathbb{R}^3, \mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v))$, is C^1 - (resp. *regular*) if all components of S are C^1 (resp. C^1 - and \mathbf{r}_u and \mathbf{r}_v are linearly independent everywhere). Using the property of the cross product of 3-vectors, the last condition is equivalent to $\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$. Geometrically the two vectors \mathbf{r}_u and \mathbf{r}_v span the tangent plane of the surface at \mathbf{r} . It is most suitable to visualize them as vectors pointing out from \mathbf{r} , the position vector of the surface. The linear independence of the two tangent vectors ensures that the image of \mathbf{r} cannot degenerate into curves or even points, hence looks like a piece of surface. For instance, the non-regular C^1 -parametric surface $\mathbf{r}(u, v) = (u, u, 0)$ has image being the straight line given by $t \rightarrow (t, t, 0)$ in space.

A (*geometric*) *surface* is the image of a C^1 -parametric surface \mathbf{r} from a region D to \mathbb{R}^3 which is one-to-one and regular when restricted to the interior of D . We continue to call such parametrization a regular parametrization of the surface.

Very often a surface is given in the form of the graph of a function with two variables. Let φ be a C^1 -function in some region $D \in \mathbb{R}^2$. Then

$$S = \{(x, y, \varphi(x, y)) : (x, y) \in D\} ,$$

is a parametric surface. Here $\mathbf{r}_x = (1, 0, \varphi_x)$ and $\mathbf{r}_y = (0, 1, \varphi_y)$, so $\mathbf{r}_x \times \mathbf{r}_y = (-\varphi_x, -\varphi_y, 1)$ which is always non-zero. We conclude that \mathbf{r} is a regular surface.

Example 3.16. The sphere centered at the origin and with radius a . First of all, a parametrization is given by (as inspired by the polar coordinates)

$$\mathbf{r} : (\varphi, \theta) \rightarrow (a \sin \varphi \cos \theta, a \sin \varphi \sin \theta, a \cos \varphi) , \quad \varphi \in [0, \pi], \theta \in [0, 2\pi] .$$

The parametrization is one-to-one on the interior of its region, that is, $(0, \pi) \times (0, 2\pi)$. It maps $\{0\} \times [0, 2\pi]$ to the north point $(0, 0, a)$ and $\{\pi\} \times [0, 2\pi]$ to the south point $(0, 0, -a)$. Also it maps the points $(\varphi, 0)$ and $(\varphi, 2\pi)$ to the same point. Then

$$\mathbf{r}_\varphi \times \mathbf{r}_\theta = (a^2 \sin^2 \varphi \cos \theta, a^2 \sin^2 \varphi \sin \theta, a^2 \cos \varphi \sin \varphi) ,$$

which is equal to $\mathbf{0}$ if and only if $\varphi = 0, \pi$. It follows that \mathbf{r} is a regular parametrization of the sphere in $(0, \pi) \times [0, 2\pi)$ after removing the north pole $(0, 0, a)$ and the south pole $(0, 0, -a)$.

This sphere can also be described as the union of two graphs of functions, namely, S_+ and S_- given respectively by

$$S_+ = \{(x, y, \sqrt{a^2 - x^2 - y^2}) : x^2 + y^2 < a^2\},$$

and

$$S_- = \{(x, y, -\sqrt{a^2 - x^2 - y^2}) : x^2 + y^2 < a^2\}.$$

Example 3.17. The cylinder whose axis is the z -axis is given by the parametrization

$$\mathbf{r}(\theta, z) = (r \cos \theta, r \sin \theta, z), \quad (\theta, z) \in [0, 2\pi) \times (-\infty, \infty).$$

As

$$\begin{aligned} \mathbf{r}_\theta &= (r \cos \theta, r \sin \theta, 0), \quad \mathbf{r}_z = (0, 0, 1), \\ \mathbf{r}_\theta \times \mathbf{r}_z &= (r \cos \theta, -r \sin \theta, 0), \end{aligned}$$

which is never equal to $\mathbf{0}$, the cylinder is a regular surface. Since the cylinder is perpendicular to the xy -plane, its projection on this plane is not a region. It cannot be represented as graphs over the xy -plane. On the other hand, one may project it to the yz -plane to express it as the union of two graphs. We will not do it here.

Example 3.18. Equation for a circular cone is given by $z = (x^2 + y^2)^{1/2}$. It is already in graphic form. We have the representation

$$\mathbf{r}(x, y) = (x, y, \sqrt{x^2 + y^2}), \quad (x, y) \in \mathbb{R}^2,$$

which is regular away from the origin.

One may use polar coordinates to get another parametrization of the circular cone

$$\mathbf{s}(r, \theta) = (r \cos \theta, r \sin \theta, r), \quad (r, \theta) \in \mathbb{R} \times [0, 2\pi),$$

which is regular except at $r = 0$.

A class of surfaces commonly encountered is surfaces of revolution.

Let $(x(t), z(t)), x(t) > 0$, be a curve in the xz -plane parametrized by some t and $x > 0$. Rotate it around the z -axis to get a surface S . The surface can be described by adding a parameter θ :

$$S : (\alpha, t) \mapsto (r(t) \cos \alpha, r(t) \sin \alpha, z(t)),$$

after changing the notation x to r . Now

$$\frac{\partial \mathbf{r}}{\partial \alpha} = (-r \sin \alpha, r \cos \alpha, 0), \quad \frac{\partial \mathbf{r}}{\partial t} = (r' \cos \alpha, r' \sin \alpha, z').$$

When $z' \neq 0$, these two vectors are linearly independent. When $z' = 0$ but $r' \neq 0$, the determinant of $(-r \sin \alpha, r \cos \alpha)$ and $(r' \cos \alpha, r' \sin \alpha)$ is equal to rr' which is non-zero.

It shows that as long as $(x(t), z(t))$ is a regular curve, the surface obtained by rotation this curve is a regular surface.

When a plane curve is given in the explicit form $y = f(x)$, the surface of revolution it produces is given in the form

$$z = f(r), \quad r = \sqrt{x^2 + y^2}.$$

Example 3.19. Let us examine the torus, which is obtained by rotating a circle staying away from the origin. Let

$$(x(\theta), z(\theta)) = (b + a \cos \theta, a \sin \theta), \quad 0 < a < b.$$

This describes the circle centers at $(b, 0)$ with radius a . Since $a < b$, the circle does not touch the z -axis. The torus is given by the parametrization

$$(\alpha, \theta) \mapsto ((b + a \cos \theta) \cos \alpha, (b + a \cos \theta) \sin \alpha, a \sin \theta).$$

We have

$$\frac{\partial \mathbf{r}}{\partial \alpha} = (-(b + a \cos \theta) \sin \alpha, (b + a \cos \theta) \cos \alpha, 0),$$

$$\frac{\partial \mathbf{r}}{\partial \theta} = (-a \sin \theta \cos \alpha, -a \sin \theta \sin \alpha, a \cos \theta),$$

and

$$\frac{\partial \mathbf{r}}{\partial \alpha} \times \frac{\partial \mathbf{r}}{\partial \theta} = (a(b + \cos \theta) \cos \theta \cos \alpha, a(b + a \cos \theta) \cos \theta \sin \alpha, a(b + a \cos \theta) \sin \theta).$$

So

$$\left| \frac{\partial \mathbf{r}}{\partial \alpha} \times \frac{\partial \mathbf{r}}{\partial \theta} \right| = a(b + a \cos \theta) > 0.$$

We confirm again that the torus is a regular surface.

The equation of the torus can be found as follows. Starting with $(x - a)^2 + y^2 - a^2 = 0$, we replace x and y by r and z respectively to get $(r - b)^2 + z^2 = a^2$. In other words,

$$\sqrt{x^2 + y^2} = \sqrt{a^2 - z^2} + b,$$

$$x^2 + y^2 = a^2 - z^2 + 2b\sqrt{a^2 - z^2} + b^2,$$

and finally we arrive at the implicit form of the torus

$$(x^2 + y^2 + z^2 - a^2 - b^2)^2 = 4b(a^2 - z^2).$$

3.8 Surface Area

In this section we derive the surface area formula for a surface. Let $\mathbf{r} : D \rightarrow \mathbb{R}^3$ be the regular parametrization of a surface S from a region D in the plane. Let $P = \{D_{ij}\}$ be a partition on D which is taken to be a rectangle for simplicity. The parametrization introduces a “generalized partition” on S whose “subrectangles” are denoted by S_{ij} . Fix a typical subrectangle with vertices at (u, v) , $(u + h, v)$, $(u, v + k)$, $(u + h, v + k)$. Its image under the parametrization has vertices at $\mathbf{r}(u, v)$, $\mathbf{r}(u + h, v)$, $\mathbf{r}(u, v + k)$, $\mathbf{r}(u + h, v + k)$. We approximate the surface area of S_{ij} by the area of the parallelogram P formed by the vectors from $\mathbf{r}(u, v)$ to $\mathbf{r}(u + h, v)$ and from $\mathbf{r}(u, v)$ to $\mathbf{r}(u, v + k)$. By Taylor’s expansion theorem,

$$\mathbf{r}(u + h, v) = \mathbf{r}(u, v) + \mathbf{r}_u(u, v)h + T_1 ,$$

$$\mathbf{r}(u, v + k) = \mathbf{r}(u, v) + \mathbf{r}_v(u, v)k + T_2 ,$$

$$\mathbf{r}(u + h, v + k) = \mathbf{r}(u, v) + \mathbf{r}_u(u, v)h + \mathbf{r}_v(u, v)k + T_3 ,$$

where $T_j, j = 1, 2, 3$ are higher order terms. Assuming that \mathbf{r} is of C^2 -class, $|T_1| \leq Ch^2, |T_2| \leq Ck^2$, for some constant C . (When \mathbf{r} is only in C^1 -class, it is a bit messy but the arguments are essential the same.) By the property of the cross product, the area of P is equal to the absolute value of

$$\begin{aligned} & (\mathbf{r}_u(u, v)h + T_1) \times (\mathbf{r}_v(u, v)k + T_2) \\ &= \mathbf{r}_u \times \mathbf{r}_v h k + \mathbf{r}_u h T_2 + \mathbf{r}_v k T_1 + T_1 T_2 \\ &\equiv \mathbf{r}_u \times \mathbf{r}_v h k + T_4 . \end{aligned}$$

Observing that $\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$ and $|T_4|/hk \rightarrow 0$ as $\sqrt{h^2 + k^2} \rightarrow 0$, the area of P can be further approximated by $|\mathbf{r}_u \times \mathbf{r}_v| h k$.

Replacing (u, v) by (u_i, v_j) and (h, k) by $(\Delta u_i, \Delta v_j)$, the surface area of S is approximately given by

$$\sum_{i,j} |\mathbf{r}_u \times \mathbf{r}_v|(u_i, v_j) \Delta u_i \Delta v_j ,$$

which is a Riemann sum on D . Letting $\|P\| \rightarrow 0$, the Riemann sum tends to the limit

$$\iint_D |\mathbf{r}_u \times \mathbf{r}_v|(u, v) dA(u, v) .$$

In view of this, for an regular parametrization $\mathbf{r} : D \rightarrow \mathbb{R}^3$ of the surface S , we define its surface area to be

$$|S| = \iint_D |\mathbf{r}_u \times \mathbf{r}_v|(u, v) dA(u, v) . \quad (3.11)$$

From the derivation above we see that the surface area is independent of the choice of the parametrization.

When the surface is given by the graph of a function $\varphi(x, y)$ where (x, y) belongs to a region D , we have $|\mathbf{r}_x \times \mathbf{r}_y| = (1 + \varphi_x^2 + \varphi_y^2)^{1/2}$, so the surface area becomes

$$|S| = \int_D \sqrt{1 + \varphi_x^2 + \varphi_y^2} dA(x, y) . \quad (3.12)$$

Introducing the notation

$$|\nabla\varphi| = \sqrt{\varphi_x^2 + \varphi_y^2} ,$$

we have

$$|S| = \int_D \sqrt{1 + |\nabla\varphi|^2} dA(x, y) .$$

For a surface of revolution obtained from revolving a curve $(x(t), z(t))$, $t \in [a, b]$, the formula for its surface area is

$$|S| = \int_0^{2\pi} \int_a^b x(t) \sqrt{x'^2(t) + z'^2(t)} dt . \quad (3.13)$$

To see this, let $(\alpha, t) \mapsto (x(t) \cos \alpha, x(t) \sin \alpha, z(t))$, be the parametrization of the surface. Then

$$\mathbf{r}_\alpha = (-x \sin \alpha, x \cos \alpha, 0) , \quad \mathbf{r}_t = (x' \cos \alpha, x' \sin \alpha, z') ,$$

and

$$\mathbf{r}_\alpha \times \mathbf{r}_t = x \sqrt{x'^2 + z'^2} ,$$

and the formula follows.

Let us examine the surface area of the sphere. We will do it in three different ways. First, the polar coordinates gives a parametrization of the sphere of radius a ;

$$\mathbf{r} : (\varphi, \theta) \mapsto (a \cos \varphi \cos \theta, a \cos \varphi \sin \theta, a \cos \varphi) .$$

This map is bijective from $(0, \pi) \times [0, 2\pi)$ to the sphere but missing the north and south pole. As missing two points does not affect the surface area, the area formula (3.6) above still works. We have

$$\mathbf{r}_\varphi \times \mathbf{r}_\theta = a \sin \varphi .$$

Therefore, the surface area of the sphere of radius a is equal to

$$\int_0^{2\pi} \int_0^\pi a^2 \sin \varphi d\varphi d\theta = 4\pi a^2 .$$

Next, the sphere can be described as the union of two graphs one of which is $f(x, y) = \sqrt{a^2 - x^2 - y^2}$ and the other is $-\sqrt{a^2 - x^2 - y^2}$ over the disk D_a . By symmetry, it suffices

to calculate the surface area of the upper half sphere. As $1 + |\nabla f|^2 = a^2/(a^2 - x^2 - y^2)$, by (3.7) the surface area of the sphere is

$$\begin{aligned} 2 \iint_{D_a} \frac{a}{\sqrt{a^2 - x^2 - y^2}} dA &= 2a \int_0^{2\pi} \int_0^a \frac{1}{\sqrt{a^2 - x^2 - y^2}} r dr d\theta \\ &= 4\pi a^2 . \end{aligned}$$

Finally, the sphere can be viewed as the surface of revolution by rotation the curve $\theta \mapsto (a \cos \theta, a \sin \theta)$, $\theta \in (-\pi/2, \pi/2)$ around the y -axis. By (3.8), since $x_\theta^2 + y_\theta^2 = a^2$, the surface area is given by

$$\int_0^{2\pi} \int_{-\pi/2}^{\pi/2} a \cos \theta \times a dt = 4\pi a^2 .$$

Example 3.20. Let S be the surface obtained by rotating the curve $x = \cos z$, $z \in [-\pi/2, \pi/2]$. Find its surface area.

The curve in the xz -plane is given by $(\cos z, z)$, $z \in [-\pi/2, \pi/2]$. According our formula, a parametrization of S is given by

$$\mathbf{r}(\alpha, z) = (\cos z \cos \alpha, \cos z \sin \alpha, z) .$$

We have

$$\mathbf{r}_\alpha = (-\cos z \sin \alpha, \cos z \cos \alpha, 0) , \quad \mathbf{r}_z = (-\sin z \cos \alpha, -\sin z \sin \alpha, 1) .$$

It follows that

$$\mathbf{r}_\alpha \times \mathbf{r}_z = (-\cos z \cos \alpha, -\cos z \sin \alpha, -(\sin z \cos z \cos^2 \alpha + \cos z \sin z \sin^2 \alpha)) ,$$

and

$$|\mathbf{r}_\alpha \times \mathbf{r}_z| = \cos z \sqrt{1 + \sin^2 z} .$$

Therefore, the surface area of S is given by

$$\begin{aligned} |S| &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos z \sqrt{1 + \sin^2 z} dz d\alpha \\ &= 2 \int_0^{2\pi} \int_0^1 \sqrt{1 + w^2} dw d\alpha \\ &= 2\pi(\sqrt{2} + \log(1 + \sqrt{2})) . \end{aligned}$$