Multivariable Calculus

Yuxin Gong

13/12/2024

Contents

1 Conventions in vectors **3**

1.1 Einstein summation convention 3

1.2 Kronecker delta 3

1.3 Permutation symbol 4

1.4 Vector Product 5

2 Gradient **6**

2.1 Formula for Gradient 7

2.1.1 Formula in cylindrical coordinate 7

2.2 Tangent plane to a surface 8

3 Divergence and Curl **8**

3.1 Divergence of the gradient 10

3.2 Curl of the gradient 10

3.3 Curl of the curl 11

3.4 Divergence of the curl 11

4 Path Integrals **11**

4.1 Path element 13

4.2 Conservative forces 13

5 Surface Integral **15**

5.1 Integral in $\mathbb{R}^2$ 15

5.1.1 example of a circle 17

5.2 General Surface Integral 18

5.2.1 Types of Surfaces 18

5.2.2 Projection Theorem 19

6 Volume Integrals **21**

7 Green's Theorem **21**

1 Conventions in vectors

1.1 Einstein summation convention

We only deal with vectors in $\mathbb{R}^3$ in this course. Although this notation is originated from physics, we have to admit that it is really helpful when you are dealing with vectors in $\mathbb{R}^3$.

Definition 1.1 (Einstein Summation Convention)**.** The sum of the product

$$\sum_{i=1}^{3} a_i x_i \xrightarrow{\text{shorthand}} a_i x_i$$

where i in the right hand side can be changed to arbitrary subscript in this convention, that is

$$a_i x_i = a_j x_j$$

Notice that the Einstein Summation Convention can only work in the case where it gives something like $a_i b_i c$, where c is a product or sum that have irrelevant subscript of i. Otherwise the expression is said to be invalid in Einstein Summation Convention. I will give some invalid expressions in Einstein summation convention

1. $a_i x_i c_i$, invalid because it has i repeated three times.
2. $a_i b_j a_i a_i$, invalid for same reason.

1.2 Kronecker delta

Definition 1.2 (Kronecker delta)**.** Kronecker delta δ_{ij} is defined to be

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Let's consider the following example. Simplify $\delta_{ji} a_i$ for $1 \leq j \leq 3$ if it is in Einstein Summation Convention

$$\delta_{ji} a_i \xrightarrow{actual} \sum_{i=1}^{3} \delta_{ji} a_i = \delta_{jj} a_j = a_j$$

From now on, we will ignore the arrow which convert between this convention and real expression. This means we will directly use $=$ for summation convention.

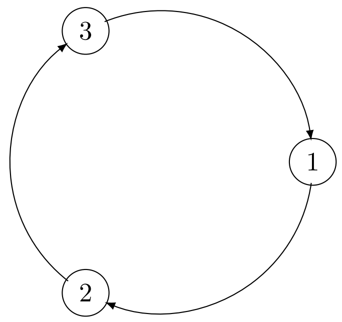


Figure 1: Cyclic Permutation

1.3 Permutation symbol

Definition 1.3 (Permutation symbol)**.** Permutation symbol ε_{ijk} is defined to be

$$\varepsilon = \begin{cases} 0, & \text{any of two subscripts are same.} \\ 1, & \text{if } (ijk) \text{ is a cyclic permutation.} \\ -1, & \text{if } (ijk) \text{ is an acyclic permutation.} \end{cases}$$

So what is a cyclic permutation and what is an acyclic permutation. Consider the following diagram If we are going clockwise, starting at any node on this cycle, it is a cyclic permutation. Otherwise, they are acyclic permutation.

Example: $\varepsilon_{231} = 1$, $\varepsilon_{132} = -1$. There will be 6 permutations in total, you can run over them to see all possibilities.

Theorem 1.1. $\varepsilon_{ijk}\varepsilon_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}$

This theorem will be useful later, let's just prove it now.

Proof. Focus on the pair $(p, q)$and (j, k). If (p, q) is not a permutation of (j, k). Then for $(ijk), (ipq)$, there will be at least one equal 0. LHS hence is always 0. For the RHS, it will certainly be zero, since the non-zero things happens iff (p, q) is a permutation of (j, k). Now it remains to deal with the case where (p, q) is a permutation of (j, k). Let's do case work to prove it.

Case I $(p, q) = (j, k)$, i.e. $p = j,\ q = k$. LHS is $\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{ipq} = \varepsilon_{rjk}\varepsilon_{rpq}$ where $r \in \{1, 2, 3\} \setminus \{j, k\}$. Now (rpq), (rjk) must be same type of permutation, that is either both cyclic or acyclic. So LHS must always be 1. Consider RHS, it can be computed easily that first term is 1 while second term is 0. Finally LHS = RHS = $1 - 0 = 1$.

Case II $(p, q) = (k, j)$. Now LHS must be different permutation type and RHS has the second term 1 and first term 0. Thus LHS = RHS = -1.

$\square$

This is an very important equality! But before using this equality, there are some extra equalities can be deduced from it.

Example: Since change of the start number of a cycle will not change the value of permutation symbol of that cycle. $\varepsilon_{ijk} = \varepsilon_{jki}$. This implies that $\varepsilon_{jki}\varepsilon_{ipq} = \delta_{jp}\delta_{kq} - \delta_{jq}\delta_{kp}$

1.4 Vector Product

Recall that for vectors in $\mathbb{R}^3$, we can define cross product of two vectors. Cross product can be defined because of important property in $\mathbb{R}^3$. Actually we can not define cross product generally in $\mathbb{R}^n$, but this is away from the purpose of this course. We will only focus on cross product in $\mathbb{R}^3$.

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Why we want to recall this concept here. Stop to think for a while, is there any connection between the tools we have developed before and cross product of vectors?

$(\mathbf{a} \times \mathbf{b})_1 = a_2b_3 - a_3b_2$, $(\mathbf{a} \times \mathbf{b})_2 = a_3b_1 - a_1b_3$, $(\mathbf{a} \times \mathbf{b})_3 = a_1b_2 - a_2b_1$. Check with this $(\mathbf{a} \times \mathbf{b})_i = \varepsilon_{ijk}a_jb_k$! We have extremely a beautiful way to write the cross product of two vectors.

Example: For dot product $\mathbf{a} \cdot \mathbf{b} = a_ib_i$.

Proposition 1.4.1.

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

Proof.

$$\begin{aligned} \text{LHS} &= a_i(\mathbf{b} \times \mathbf{c})_i = a_i\varepsilon_{ijk}b_jc_k \\ &= \varepsilon_{kij}a_ib_jc_k = (\mathbf{a} \times \mathbf{b})_kc_k \\ &= (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \text{RHS} \end{aligned}$$

$\square$

Proposition 1.4.2.

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\, \mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\, \mathbf{c}$$

Proof. We will do this basically using the same way as what we did above.

$$\begin{aligned} \mathrm{LHS}_i &= \varepsilon_{ijk} a_j (\mathbf{b} \times \mathbf{c})_k = \varepsilon_{ijk} a_j \varepsilon_{kpq} b_p c_q \\ &= (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}) a_j b_p c_q \\ &= a_j c_j b_i - a_j b_j c_i = \mathrm{RHS}_i \end{aligned}$$

□

2 Gradient

Let ϕ be a differentiable scalar function. By differentiable, we mean if ϕ depends on three random variables $x,\ y,\ z$. Then ϕ is differentiable with respect to $x,\ y,\ z$.
Consider the family $\phi =$ constant. This will define a surface in the $\mathbb{R}^3$. Using the following diagram first.

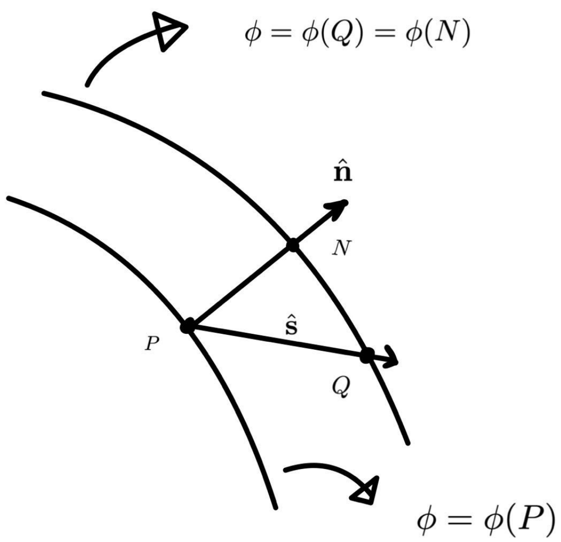


Figure 2: Graph for gradient

For the plane $\phi = \phi(P)$, suppose Q is some other point in the plane. $\phi = \phi(Q)$ fixes another plane. Drawing the normal line to $\phi = \phi(P)$ at point P, intersecting the plane $\phi = \phi(Q)$ at point N. Suppose P and Q can fixes a unit vector $\hat{\mathbf{s}}$ where it is marked on the diagram. Define

$$\frac{\partial \phi}{\partial s} := \lim_{PQ \to 0} \frac{\phi(P) - \phi(Q)}{PQ}$$

Using the following argument(not very rigorous), we have

$$\frac{\partial \phi}{\partial s} = \lim_{PQ \to 0} \frac{\phi(P) - \phi(Q)}{PN} \cdot \frac{PN}{PQ} = \lim_{PQ \to 0} \frac{\phi(P) - \phi(Q)}{PN} (\hat{\mathbf{n}} \cdot \hat{\mathbf{s}})$$

Because $\phi(Q) = \phi(N)$, we have

$$\frac{\partial \phi}{\partial s} = \lim_{PN \to 0} \frac{\phi(P) - \phi(N)}{PN} (\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}) = \frac{\partial \phi}{\partial n} (\hat{\mathbf{n}} \cdot \hat{\mathbf{s}})$$

Interested in $\frac{\partial \phi}{\partial n} \hat{\mathbf{n}}$. So we have a special name for it, **gradient** of scalar function ϕ. Denote the gradient of a scalar function by $\nabla \phi$.

2.1 Formula for Gradient

Gradient is a vector, $\nabla \phi = A_1 \mathbf{i} + A_2 \mathbf{j} + A_3 \mathbf{k}$. Able to show that

$$\frac{\partial \phi}{\partial x} = \lim_{h \to 0} \frac{\phi(x+h, y, z) - \phi(x, y, z)}{h} = \nabla \phi \cdot \mathbf{i} = A_1$$

Therefore,

$$\nabla \phi = \frac{\partial \phi}{\partial x} \mathbf{i} + \frac{\partial \phi}{\partial y} \mathbf{j} + \frac{\partial \phi}{\partial z} \mathbf{k}$$

2.1.1 Formula in cylindrical coordinate

However, you should notice that this is only the formula in Cartesian coordinate, so it has such beautiful form. If you put it in a cylindrical coordinate, everything changes. We will do this using a simple diagram.

As you can see, we use $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{k}}$ as the new basis. To work out the gradient in that coordiante, write $\nabla \phi = A_1 \hat{\mathbf{r}} + A_2 \hat{\boldsymbol{\theta}} + A_3 \hat{\mathbf{k}}$. Now A_1, A_2, A_3 can be worked out by

$$A_1 = \nabla \phi \cdot \hat{\mathbf{r}},\ A_2 = \nabla \phi \cdot \hat{\boldsymbol{\theta}},\ A_3 = \nabla \phi \cdot \hat{\mathbf{k}}$$

This means that we are only required to find the dot product between the $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and this basis. The result is

$$\nabla \phi = \hat{\mathbf{r}} \frac{\partial \phi}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial \phi}{\partial \theta} + \hat{\mathbf{k}} \frac{\partial \phi}{\partial z}$$

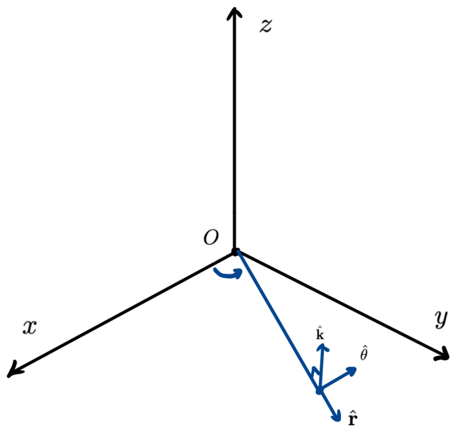


Figure 3: Cylindrical coordinate

2.2 Tangent plane to a surface

Using gradient, the tangent plane at one point to a plane curve is rather trivial to calculate. Since $\nabla\phi|_P$ is normal to the tangent plane at point P. The tangent plane is just basically

$$(\mathbf{r} - \mathbf{r}_P) \cdot \nabla\phi|_P = 0$$

3 Divergence and Curl

View ∇ as the operator sends a scalar function to a vector field, rigorously

$$\nabla : \mathrm{Func}(\mathbb{R}^3, \mathbb{R}) \to \mathrm{Func}(\mathbb{R}^3, \mathbb{R}^3)$$

$$\phi \mapsto \nabla\phi$$

Making convenient symbol, writing $\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}$ is a good choice. Similarly, define $\nabla\cdot$ to be another vector operator which sending $\mathbf{A} \mapsto \nabla \cdot \mathbf{A}$. By using the formula we write for ∇, divergence and curl can be established naturally. Define

$$\mathrm{Div}(\mathbf{A}) = \nabla \cdot \mathbf{A}, \ \mathrm{curl}(\mathbf{A}) = \nabla \times \mathbf{A}$$

It would be better to write down the Cartesian form of them.

$$\begin{aligned}
\operatorname{div} \mathbf{A} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \right) \cdot (A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k}) \\
&= \frac{\partial A_1}{\partial x} + \frac{\partial A_2}{\partial y} + \frac{\partial A_3}{\partial z}
\end{aligned}$$

$$\begin{aligned}
\operatorname{curl} \mathbf{A} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_1 & A_2 & A_3 \end{vmatrix} \\
&= \mathbf{i}\left(\frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} \right) - \mathbf{j}\left(\frac{\partial A_3}{\partial x} - \frac{\partial A_1}{\partial z} \right) + \mathbf{k}\left(\frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y} \right)
\end{aligned}$$

Combine those vector operators together, there will be some formulas from them. I will list some of them,

$$\begin{aligned}
&\text{(i)} \quad \nabla(\varphi_1 + \varphi_2) = \nabla\varphi_1 + \nabla\varphi_2, \\
&\text{(ii)} \quad \operatorname{div}(\mathbf{A} + \mathbf{B}) = \operatorname{div} \mathbf{A} + \operatorname{div} \mathbf{B}, \\
&\text{(iii)} \quad \operatorname{curl}(\mathbf{A} + \mathbf{B}) = \operatorname{curl} \mathbf{A} + \operatorname{curl} \mathbf{B}, \\
&\text{(iv)} \quad \nabla(\varphi\psi) = \varphi\nabla\psi + \psi\nabla\varphi, \\
&\text{(v)} \quad \operatorname{div}(\varphi\mathbf{A}) = \varphi \operatorname{div} \mathbf{A} + \nabla\varphi \cdot \mathbf{A}.
\end{aligned}$$

First three are obvious, (iv) is quite straightforward. Let's prove (v) using our summation convention.

Proof.

$$\operatorname{div}(\varphi\mathbf{A}) = \frac{\partial}{\partial x_i}(\varphi\mathbf{A})_i = \frac{\partial}{\partial x_i}(\varphi A_i) = \frac{\partial \varphi}{\partial x_i} A_i + \frac{\partial A_i}{\partial x_i}\varphi = \nabla\varphi \cdot \mathbf{A} + \varphi \operatorname{div} \mathbf{A}.$$

Using one word that my high school Chemistry teacher told me: doing summation is just simple and sweet. □

Now you can try the following thing yourself, it's just the same.

$$\begin{aligned}
&\text{(vi)} \quad \operatorname{curl}(\varphi\mathbf{A}) = \varphi \operatorname{curl} \mathbf{A} + \nabla\varphi \times \mathbf{A}, \\
&\text{(vii)} \quad \operatorname{div}(\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \operatorname{curl} \mathbf{A} - \mathbf{A} \cdot \operatorname{curl} \mathbf{B}, \\
&\text{(viii)} \quad \operatorname{curl}(\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A} \operatorname{div} \mathbf{B} - \mathbf{B} \operatorname{div} \mathbf{A}, \\
&\text{(ix)} \quad \nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} + (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{B} \times \operatorname{curl} \mathbf{A} + \mathbf{A} \times \operatorname{curl} \mathbf{B}.
\end{aligned}$$

One thing to clarify here is that if you notice the last two identities, $\mathbf{B} \cdot \nabla$ means the operator $B_i(\partial/\partial x_i)$, when it is applied to a vector field, it produce a vector field by operating all its component.

3.1 Divergence of the gradient

Let's imagine what will happen if we take the divergence of gradient of a scalar function. First, it will end up with a scalar.

$$\operatorname{div}(\nabla\varphi) = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot \left(\frac{\partial \varphi}{\partial x}\mathbf{i} + \frac{\partial \varphi}{\partial y}\mathbf{j} + \frac{\partial \varphi}{\partial z}\mathbf{k}\right)$$

final result would be $\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}$. We often write div as $\nabla\cdot$, so usually, we call this operator laplacian operator and denote it by $\nabla \cdot \nabla = \nabla^2$, that is $\operatorname{div}(\nabla\varphi) \coloneqq \nabla^2\varphi$.
Extend this idea a little bit on the vector field $\mathbf{A}$.

$$\nabla^2\mathbf{A} = \frac{\partial^2\mathbf{A}}{\partial x^2} + \frac{\partial^2\mathbf{A}}{\partial y^2} + \frac{\partial^2\mathbf{A}}{\partial z^2}$$

3.2 Curl of the gradient

Expand $\operatorname{curl}(\nabla\varphi)$ using Einstein summation convention,

$$[\operatorname{curl}(\nabla\varphi)]_i = \varepsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_k}\right)$$

Rearrange the position of j and k since they are just a symbol. Assuming that those partial derivative exists and continuous, we have

$$\varepsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_k}\right) = \varepsilon_{ikj}\frac{\partial}{\partial x_k}\left(\frac{\partial\varphi}{\partial x_j}\right)$$

$$\implies \varepsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_k}\right) = -\varepsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_k}\right) = 0$$

Schwartz's inequality ensures that we can change the i and j-th derivative for the scalar function which ends up with the second one.

$$[\operatorname{curl}(\nabla\varphi)]_i = 0 \implies \operatorname{curl}(\nabla\varphi) = \mathbf{0}$$

3.3 Curl of the curl

Given a vector field $\mathbf{A}$, let's see what's the curl of the curl of $\mathbf{A}$. Result is $\text{curl}(\text{curl}\mathbf{A}) = \nabla(\text{div}\mathbf{A}) - \nabla^2\mathbf{A}$

$$\begin{aligned}
[\text{curl}(\text{curl}\mathbf{A})]_i &= \epsilon_{ijk}\frac{\partial}{\partial x_j}(\text{curl}\mathbf{A})_k \\
&= \epsilon_{ijk}\frac{\partial}{\partial x_j}\left(\epsilon_{klm}\frac{\partial A_m}{\partial x_l}\right) \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\frac{\partial^2 A_m}{\partial x_j \partial x_l} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\frac{\partial^2 A_m}{\partial x_j \partial x_l} \\
&= \frac{\partial^2 A_j}{\partial x_j \partial x_i} - \frac{\partial^2 A_i}{\partial x_j^2} \\
&= \frac{\partial}{\partial x_i}\left(\frac{\partial A_j}{\partial x_j}\right) - \frac{\partial^2 A_i}{\partial x_j^2} \\
&= [\nabla(\text{div}\mathbf{A})]_i - [\nabla^2\mathbf{A}]_i
\end{aligned}$$

3.4 Divergence of the curl

The divergence of a curl is always zero, as shown by the following argument:

$$\text{div}(\text{curl } \mathbf{A}) = \frac{\partial}{\partial x_i}(\text{curl } \mathbf{A})_i = \epsilon_{ijk}\frac{\partial}{\partial x_i}\frac{\partial A_k}{\partial x_j} = \frac{1}{2}\epsilon_{ijk}\frac{\partial}{\partial x_i}\frac{\partial A_k}{\partial x_j} + \frac{1}{2}\epsilon_{jik}\frac{\partial}{\partial x_j}\frac{\partial A_k}{\partial x_i}$$

Since $\epsilon_{ijk} = -\epsilon_{jik}$,

$$= \frac{1}{2}\epsilon_{ijk}\left(\frac{\partial^2 A_k}{\partial x_i \partial x_j} - \frac{\partial^2 A_k}{\partial x_j \partial x_i}\right)$$

Since partial derivatives commute,

$$= \frac{1}{2}\epsilon_{ijk} \times 0 = 0$$

Thus,

$$\text{div}(\text{curl } \mathbf{A}) = 0$$

4 Path Integrals

Let's do a quick review of what is

$$\int_0^1 f(x)\,\mathrm{d}x$$

in sense of non-rigorous intuition.

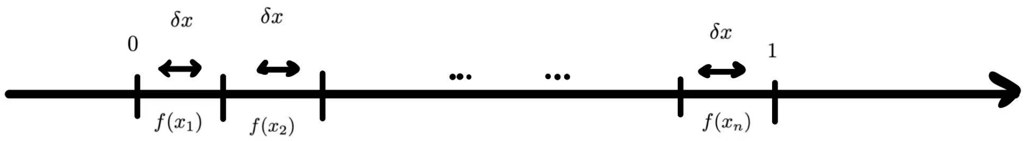


Figure 4: Integral in $\mathbb{R}$

At each interval, we assign a value $f(x_i)$ to each interval. Basically, the intergal $\int_0^1 f(x)\ \mathrm{d}x$ is evaluating

$$\sum_{i=1}^{n} f(x_i)\delta x$$

If we let the interval going to 0, it will just be the value of $\int_0^1 f(x)\ \mathrm{d}x$. Using this idea, we are now trying to define what is meant by path integral. Let's assume A and B are two points in $\mathbb{R}^3$ plane. There is a curve γ between them, illustrating in the diagram below. Suppose there is also a scalar function $\varphi : \mathbb{R}^3 \to \mathbb{R}$. For each point P_i in the diagram, $\varphi(P_i)$ is a constant in $\mathbb{R}$. Suppose we want to estimate the sum

$$\sum_{i=1}^{n} \varphi(P_i)\delta s_i$$

since we want the same intuition inherited from normal integral on $\mathbb{R}$. Taking the $\max(\delta s_i) \to 0$, this summation becomes a series, if it is convergent, then we define it to be the **line integral** of φ along the path γ.

Usually we denote it by

$$\int_\gamma \varphi\ ds$$

This is just the intuition of the line integral, actually there are some more questions we need to deal with to make the definition well-defined. Firstly, when will the value of the line integral be independent of the parametrization. Secondly, when is the integral convergent. However, in this course, we assume that all the line integral is convergent and as long as the parametrization is bijective from $[a, b] \to \gamma$ the value of line integral is independent from parametrization. Therefore, if γ is parametrized by $\mathbf{c} : [a, b] \to \gamma$ with $\mathbf{c}(a) = A,\ \mathbf{c}(b) = B$, then

$$\int_\gamma \varphi\ \mathrm{d}s = \int_a^b \varphi(\mathbf{c}(t))|\mathbf{r}'(t)|\ \mathrm{d}t$$

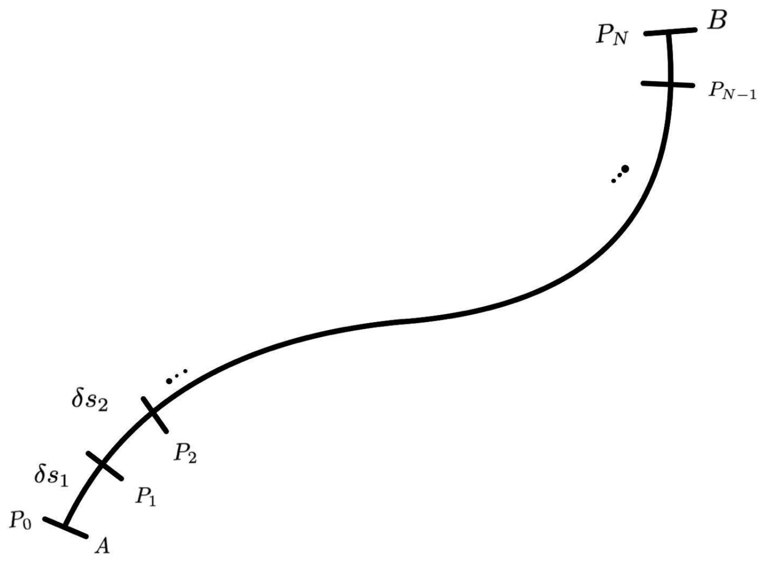


Figure 5: A path in $\mathbb{R}^3$

4.1 Path element

Let $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, suppose we are on the path γ. $\mathbf{r}$ is going to change when s changes, where s is the length we have already passed. Now define

$$\hat{\mathbf{t}} = \lim_{\delta s \to 0} \frac{\delta \mathbf{r}}{\delta s} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$$

$\hat{\mathbf{t}}$ is a unit vector because $|\delta \mathbf{r}| \to \delta s$ as $\delta s \to 0$. Let's write $\mathrm{d}\mathbf{r} = \hat{\mathbf{t}}\ \mathrm{d}s$ in order to be convenient so that we can convert between two element. When this is established, we can define what it means by line intergal of a vector field. If $\mathbf{c}$ is the same parametrization as we defined in line integral of a scalar function, then

$$\int_\gamma \mathbf{F} \cdot \mathrm{d}\mathbf{r} \coloneqq \int_\gamma \mathbf{F} \cdot \hat{\mathbf{t}}\ \mathrm{d}s = \int_a^b \mathbf{F}(\mathbf{c}(t)) \cdot \mathbf{r}'(t)\ \mathrm{d}t$$

4.2 Conservative forces

Now we introduce an important concept in line intergal, this happens when $\mathbf{F} = \nabla\varphi$. Consider the line integral of a vector field,

$$\int_\gamma \mathbf{F} \cdot \mathrm{d}\mathbf{r}$$

Using what we derive before,

$$\begin{aligned}
\int (\nabla\varphi \cdot \hat{\mathbf{t}})\, ds &= \int \frac{\partial \varphi}{\partial x_i}\hat{e}_i \cdot \frac{d\mathbf{r}}{ds}\, ds \\
&= \int \frac{\partial \varphi}{\partial x_i}\hat{e}_i \cdot \frac{dx_j}{ds}\hat{e}_j\, ds \\
&= \int \frac{\partial \varphi}{\partial x_i}\frac{dx_i}{ds}\, ds \\
&= \int \left(\frac{d\varphi}{ds}\right)\, ds \\
&= [\varphi]_A^B \\
&= \varphi(B) - \varphi(A)
\end{aligned}$$

Several points are needed to be mentioned here. s is the length we already passed along the curve γ, this definition allowed us to view $x,\ y,\ z$ as function of s and therefore view φ as a function of s. That is $\varphi : \mathbb{R} \to \mathbb{R}$, sending $s \mapsto \varphi(x(s), y(s), z(s))$.

But let's keep an eye on what we derive, $\int_\gamma \mathbf{F}\cdot \mathrm{d}\mathbf{r}$ is independent of the path between! Those vector fields are really nice, so we give them a name. The term "conservative" was introduced in the context of physics in the mid-19th century by scientists studying mechanical energy conservation. Question is, how do we define a conservative force. If γ is a closed curve, that is a curve with start point equals end point. In this case, we define a new symbol to denote the line integral,

$$\oint_\gamma \mathbf{F}\cdot \mathrm{d}\mathbf{r} := \int_\gamma \mathbf{F}\cdot \mathrm{d}\mathbf{r}$$

Nothing has changed except for adding a circle. A conservative force is defined to be a vector field $\mathbf{F}$ such that $\oint_\gamma \mathbf{F} = 0$ for all closed curve γ. Let's now ask a question, will conservative forces be the gradient of some scalar function φ. Answer is yes!

Proof. Let $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$. Since we know that $\mathbf{F}$ is conservative, it must be the case that $\int_A^P \mathbf{F}\cdot d\mathbf{r}$ is independent of the path from A to P and hence:

$$\int_{C_1} \mathbf{F}\cdot d\mathbf{r} = \int_{C_2} \mathbf{F}\cdot d\mathbf{r} = -\int_{C_3} \mathbf{F}\cdot \mathrm{d}\mathbf{r}$$

where C_1 and C_2 are any two curves drawn from A to P and C_3 is the line joining P and A. Suppose that the point A is fixed. Then,

$$\int_A^P \mathbf{F}\cdot d\mathbf{r} = \varphi(P) = \varphi(x, y, z).$$

Our claim now becomes φ is the scalar we search for. Let Q be the point $(x+\delta x, y, z)$ and let P be the point (x, y, z). Consider the quantity:

$$\varphi(x+\delta x, y, z) - \varphi(x, y, z) \equiv \int_A^Q \mathbf{F} \cdot d\mathbf{r} - \int_A^P \mathbf{F} \cdot d\mathbf{r}.$$

This simplifies to:

$$= \int_P^Q \mathbf{F} \cdot d\mathbf{r}.$$

But we can choose the path from P to Q so that only x varies, in which case $d\mathbf{r} = \mathbf{i}\, dx$. Thus:

$$\varphi(x+\delta x, y, z) - \varphi(x, y, z) = \int_x^{x+\delta x} F_1\, dx.$$

The derivative of φ with respect to x will just be F_1 here by fundamental theorem of calculus given we have good properties of $\mathbf{F}$. Because of symmetry,

$$\frac{\partial \varphi}{\partial y} = F_2, \quad \frac{\partial \varphi}{\partial z} = F_3$$

□

5 Surface Integral

To actually do the surface integral, it would be better to deal with the area integral in $\mathbb{R}^2$.

5.1 Integral in $\mathbb{R}^2$

Suppose there is some area in the plane $\mathbb{R}^2$ and we want to calculate the area. What's the way we can do this?

Let $A \subset \mathbb{R}^2$, and $f : \mathbb{R}^2 \to \mathbb{R}$. Let's assume the straight lines we draw in this picture divides the A into N pieces of area which are $A_1, A_2, \ldots, A_N$. Inside A_i, let's define an arbitrary point $\mathbf{r}_i$, but usually we might want to define this point to be the middle point of that small square so it would be convenient. Suppose we try to evaluate

$$\sum_{i=1}^{N} f(\mathbf{r}_i)\delta A_i$$

Suppose in each A_i, we can control the 'length' for each of them. Basically, to control the δx and δy presented in the diagram. Let $m = \max(\delta x, \delta y)$ We define the surface integral in $\mathbb{R}^2$ as

$$\int_A f \;\mathrm{d}A = \lim_{m \to 0} \sum_{i=1}^{N} f(\mathbf{r}_i)\delta A_i$$

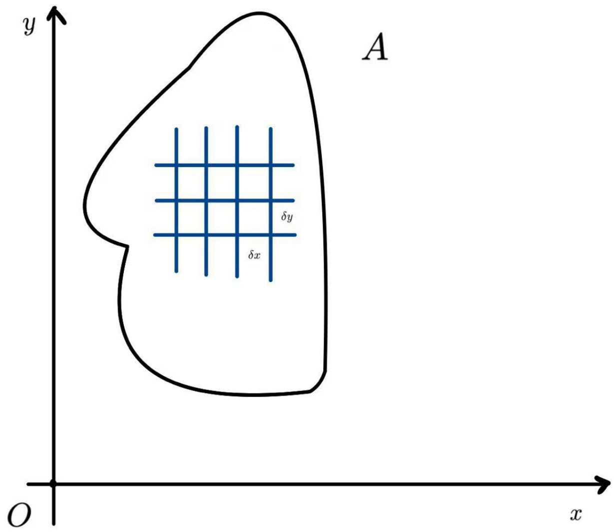


Figure 6: Double Integral

As $m \to 0$, we can approximate δA_i by $\delta x \delta y$. We have now

$$\int_A f \, \mathrm{d}A = \lim_{\delta y \to 0} \delta y \left(\lim_{\delta x \to 0} \sum_{i=1}^{N} f(\mathbf{r}_i) \delta x \right)$$

Think about the definition in one random variable and denote X_y the set of all values of x at fixed y. Denote the Y as the range of y, we have

$$\int_A f \, \mathrm{d}A = \lim_{\delta y \to 0} \sum_Y \delta y \int_{X_y} f(x, y) \, \mathrm{d}x = \int_Y \int_{X_y} f(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

Notice that all the implications are not rigorous, at least when comparing to language in analysis. However this is still an intuitive way to establish those result, giving a better understand of the formula why the double integral makes sense when we integrate separately to two variables x and y.

There are few theorems tell us about the double integral we have developed so far.

Theorem 5.1 (Fubini's theorem)**.** If f is a continuous function and D is a bounded and closed subset of $\mathbb{R}^2$, then

$$\iint f \, \mathrm{d}x \, \mathrm{d}y = \iint f \, \mathrm{d}y \, \mathrm{d}x.$$

This theorem means that under mild conditions, we can change the order of integral x and y. This sometimes helps us a lot.

Now set the $f = 1$ for all x and y, the thing we are evaluating is just the area in $\mathbb{R}^2$. An example will be the area of the circle. Let's do this example to check whether our technology (double intergal) is well-developed.

5.1.1 example of a circle

We want to find the area enclosed by the circle $x^2+y^2 = a^2$. Using a double integral in Cartesian coordinates:

$$\text{Area} = \int_{x=-a}^{a} \int_{y=-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} 1 \, \mathrm{d}y \, \mathrm{d}x.$$

First, integrate with respect to y:

$$\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} 1 \, \mathrm{d}y = y|_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} = \sqrt{a^2-x^2} - \big(-\sqrt{a^2-x^2}\big) = 2\sqrt{a^2-x^2}.$$

Hence,

$$\text{Area} = \int_{-a}^{a} 2\sqrt{a^2-x^2} \, \mathrm{d}x = 2\int_{-a}^{a} \sqrt{a^2-x^2} \, \mathrm{d}x.$$

A common technique is to use the trigonometric substitution $x = a\sin\theta$. Then $\mathrm{d}x = a\cos\theta \, \mathrm{d}\theta$ and $\sqrt{a^2-x^2} = a\cos\theta$. When $x = -a$, we have $\sin\theta = -1$, so $\theta = -\frac{\pi}{2}$. When $x = a$, we have $\sin\theta = 1$, so $\theta = \frac{\pi}{2}$. Therefore,

$$2\int_{-a}^{a} \sqrt{a^2-x^2} \, \mathrm{d}x = 2\int_{\theta=-\frac{\pi}{2}}^{\theta=\frac{\pi}{2}} \underbrace{\sqrt{a^2-(a\sin\theta)^2}}_{a\cos\theta} \underbrace{(a\cos\theta \, \mathrm{d}\theta)}_{\mathrm{d}x} = 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \big(a\cos\theta\big)\big(a\cos\theta\big) \, \mathrm{d}\theta.$$

Simplify:

$$= 2a^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta \, \mathrm{d}\theta = 2a^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1+\cos(2\theta)}{2} \, \mathrm{d}\theta$$

$$= a^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \big(1+\cos(2\theta)\big) \, \mathrm{d}\theta = a^2 \left[\theta + \tfrac{1}{2}\sin(2\theta)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = a^2\,(\pi + 0) = \pi a^2.$$

Hence, the area of the circle $x^2+y^2 = a^2$ is:

$$\boxed{\pi a^2}$$

This exactly coincides with our knowledge of the area of the circle!

5.2 General Surface Integral

In this subsection, we explain the intuition of the surface integral. To define a surface integral of $f = f(P)$ over a surface S, we divide S into elements of area $\delta S_1, \delta S_2, \ldots, \delta S_N$. Let $f_1, f_2, \ldots, f_N$ be the values of f at typical points $P_1, P_2, \ldots, P_N$ of $\delta S_1, \delta S_2, \ldots, \delta S_N$, respectively. We calculate the quantity

$$\sum_{n=1}^{N} f_n \delta S_n.$$

We now let $N \to \infty$, $\max \delta S_n \to 0$. The resulting limit, if it exists, is called the **surface integral** of f over S, and we write it as

$$\int_S f\, dS = \sum_{n=1}^{N} f_n \delta S_n, \quad N \to \infty\,, \max(\delta S_n) \to 0$$

Which is very similar to the double integral we have introduced before. But at this time the surface might be fancy and it could be any surface you can imagine in $\mathbb{R}^3$. However we do not have any method currently that we can calculate the surface integral directly.

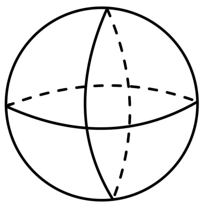
5.2.1 Types of Surfaces

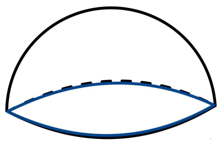
Before we actually introduce the technique to compute the surface integral, let's classify the type of surfaces

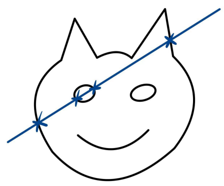
Closed surface: this divides three-dimensional space into two non-connected regions – an interior region and an exterior region.

Convex surface: this is a surface which is crossed by a straight line at most twice.

Open surface: this does not divide space into two non-connected regions – it has a rim which can be represented by a closed curve. (A closed surface can be thought of as the sum of two open surfaces).







Let's describe those surfaces, the first surface is a sphere. It is closed because it divides $\mathbb{R}^3$ into two connected regions, one is the region inside, another is the region outside. The second surface is a hemispherical shell. The circle marked blue is the rim of that surface and it's open. The third surface is a cute cat in $\mathbb{R}^2$! But do not be captivated by its cute face, it is not a convex surface!

5.2.2 Projection Theorem

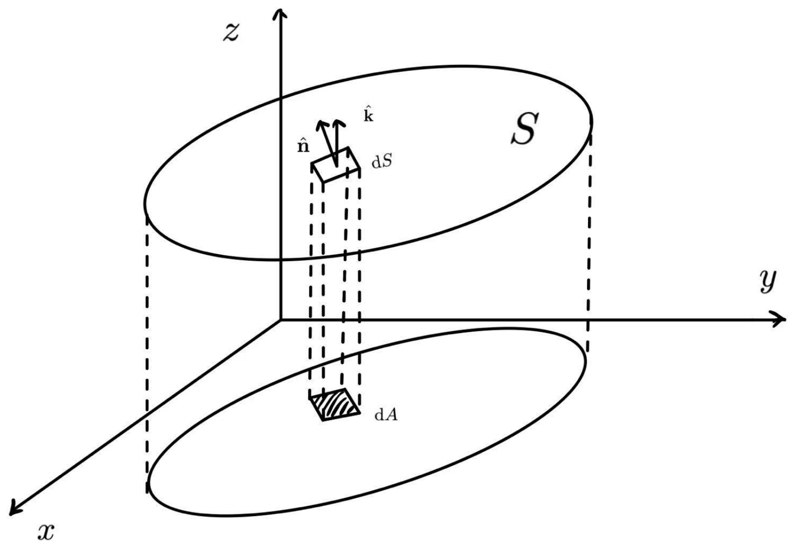


Figure 7: Projection

Imagine the case illustrated in the figure. $\mathrm{d}S$ can be approximated by $\cos(\theta)\mathrm{d}A$, where θ is angle between $\mathrm{d}S$ and $\mathrm{d}A$. You can do this by geometric argument easily. Now for appropriate surface, we can apply projection theorem. I will list some conditions that are necessary. All $\mathrm{d}S$ in that plane should not be vertical. That is suppose we have the surface $\varphi =$ constant, it has normal $\nabla\varphi|_P$ at point P. This condition tells us that $\nabla\varphi|_P$ should not perpendicular to $\hat{\mathbf{k}}$ if we are projecting the surface onto $x - y$ plane.

This condition limits many conditions where we can apply the projection theorem, since for a smooth closed surface like sphere. There are some points where normal is perpendicular to $\hat{\mathbf{k}}$. But we will later give a more generally way of doing all the surface integrals. Following would be a brief proof of this theorem.

Let P denote a general point of a surface S which at no point is orthogonal to the direction $\mathbf{k}$. Then:

$$\int_S \varphi(P)\, dS = \int_A \varphi(P) \frac{dx\, dy}{|\hat{\mathbf{n}} \cdot \mathbf{k}|},$$

where A is the projection of S onto the plane $z = 0$, and $\hat{\mathbf{n}}$ is the unit normal to S.

$$\int_S \varphi(P)\, dS = \lim_{N\to\infty,\, \delta S_r \to 0} \sum_{r=1}^{N} \varphi(P_r)\delta S_r = \lim_{N\to\infty,\, \delta S_r \to 0} \sum_{r=1}^{N} \varphi(P_r) \left(\frac{\delta A_r}{|\hat{\mathbf{n}}_r \cdot \mathbf{k}|} + \epsilon_r \right),$$

where $\epsilon_r \to 0$ as $\delta S_r \to 0$. (Here $\hat{\mathbf{n}}_r$ is the unit vector normal to S at P_r, and δA_r is the projection of δS_r onto the plane $z = 0$). It therefore follows that

$$\int_S \varphi(P)\, dS = \int_A \varphi(P) \frac{dA}{|\hat{\mathbf{n}} \cdot \mathbf{k}|}$$

as required. Note that $\varphi(P)$ is evaluated at $P(x, y, z)$ on S in both integrals.

If, for example, the equation of S is $z = \phi(x, y)$, then the theorem gives

$$\int_S \varphi(x, y, z)\, dS = \int_A \varphi(x, y, \phi(x, y)) \frac{dx\, dy}{|\hat{\mathbf{n}} \cdot \mathbf{k}|}.$$

Alternatively, we may choose to project the surface onto $x = 0$ or $y = 0$ to give

$$\int_S \varphi(P)\, dS = \int_{A_x} \varphi(P) \frac{dy\, dz}{|\hat{\mathbf{n}} \cdot \mathbf{i}|} = \int_{A_y} \varphi(P) \frac{dx\, dz}{|\hat{\mathbf{n}} \cdot \mathbf{j}|},$$

where A_x is the projection of S onto $x = 0$ and A_y is the projection of S onto $y = 0$. Let's apply the technique we developed above to solve an example. Consider $z = \sin(x)$, we want to find the surface area of it in the region $0 \le x \le \pi$ and $0 \le y \le 1$.

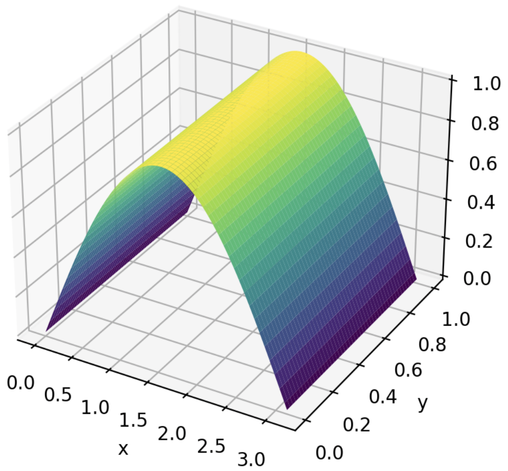


Figure 8: Diagram of $z = \sin(x)$

Let's do this surface integral by first let $\varphi(x, y, z) = z - \sin(x)$. At any given point $P = (x, y, z)$, the normal can be calculate by

$$\nabla\varphi = \frac{\partial\varphi}{\partial x}\mathbf{i} + \frac{\partial\varphi}{\partial y}\mathbf{j} + \frac{\partial\varphi}{\partial z}\mathbf{k} = -\cos(x)\mathbf{i} + \mathbf{k}$$

The $\hat{\mathbf{n}}$ in our theorem will be $(-\cos(x)\mathbf{i}+\mathbf{k})/\sqrt{1+\cos(x)^2}$. Hence

$$S(\text{Surface area}) = \int_S 1\,\mathrm{d}S = \int\int_A 1\,\frac{\mathrm{d}x\mathrm{d}y}{|\hat{\mathbf{n}}\cdot\mathbf{k}|} = \int\int_A \sqrt{1+\cos(x)^2}\,\mathrm{d}x\mathrm{d}y$$

However, it's unfortunate that we can't work out explicitly the value of $L = \int_0^\pi \sqrt{1+\cos(x)^2}\mathrm{d}x$. But we know that this is the length of $\sin(x)$ over the interval $[0,\pi]$! Hence

$$S = \int_0^1 \mathrm{d}y \cdot \int_0^\pi \sqrt{1+\cos(x)^2}\,\mathrm{d}x = L\cdot 1$$

which coincides our knowledge of the area of this surface since it is just length times width.

6 Volume Integrals

Consider a volume τ and split it up into N subregions $\delta\tau_1, \delta\tau_2, \ldots, \delta\tau_N$. Let $P_1, P_2, \ldots, P_N$ be typical points of $\delta\tau_1, \delta\tau_2, \ldots, \delta\tau_N$.
Consider the sum:

$$\sum_{i=1}^N f(P_i)\delta\tau_i$$

Now let $N\to\infty$, $\max\delta\tau_i\to 0$. If this sum tends to a limit, we call it the volume integral of f over τ and write this as:

$$\int_\tau f\,\mathrm{d}\tau$$

The function f may be a vector or a scalar. The volumne element for Cartesian coordinates is

$$\mathrm{d}\tau = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

7 Green's Theorem

Suppose R is a closed plane region bounded by a simple plane closed convex curve in the x–y plane. Let L, M be continuous functions of x, y having continuous derivatives throughout R. Then:

$$\oint_C (L\,\mathrm{d}x + M\,\mathrm{d}y) = \int_R \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y}\right)\mathrm{d}x\,\mathrm{d}y,$$

where C is the boundary of R described in the counter-clockwise (positive) sense. Before giving the proof, we will use a diagram to illustrate the theorem. Let A be the point with highest y value along the curve C. Other points in the diagram have similar definitions as A which you

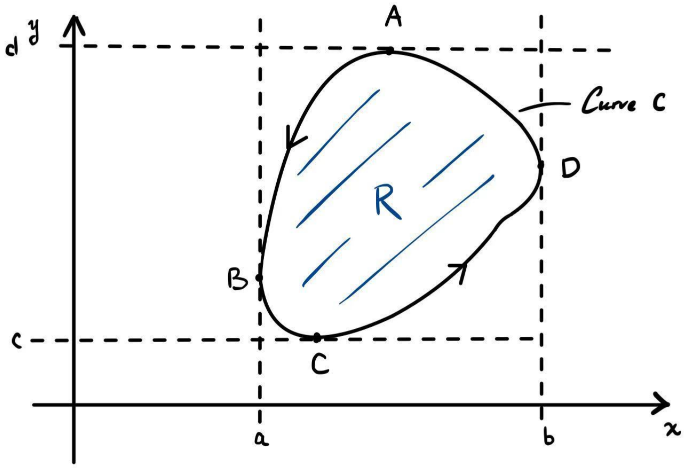


Figure 9: Green's Theorem

can see obviously from the diagram given, for example, the point B is the point with lowest x value. One thing worth noticing here is what is meant by

$$\oint_C L \, \mathrm{d}x$$

Assume $(x(t), y(t))$ is a good parametrization of curve C where t ranges from a to b

$$\oint_C L \, \mathrm{d}x = \int_a^b L(x(t), y(t)) \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t$$

To understand why it is this, we can view

$$\oint_C L \, \mathrm{d}x = \oint_C \begin{pmatrix} L \\ 0 \end{pmatrix} \cdot \mathrm{d}\mathbf{r} = \oint_C \begin{pmatrix} L \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix} = \oint L \cdot \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t$$

define After doing those settings, let's start the proof.

Proof. Starting from $\int_R \frac{\partial M}{\partial x} \, \mathrm{d}A$ first. Assume the curve ABC and curve ADC can be parametrized as $x = X_1(y)$ and $x = X_2(y)$. Then

$$\int_R \frac{\partial M}{\partial x} \, \mathrm{d}A = \int_c^d \int_{X_1(y)}^{X_2(y)} \frac{\partial M}{\partial x} \, \mathrm{d}x \, \mathrm{d}y = \int_c^d M(X_2(y), y) - M(X_1(y), y) \, \mathrm{d}y$$

Just analyse on $\int_c^d M(X_2(y), y) \, \mathrm{d}y$. View it as a parametrization by y, so

$$\int_c^d M(X_2(y), y) \, \mathrm{d}y = \int_c^d M(X_2(y), y) \cdot \frac{\mathrm{d}y}{\mathrm{d}y} \, \mathrm{d}y = \int_{C_1} M \, \mathrm{d}y$$

where C_1 denotes the curve ADC going from C to A. By taking the minus sign into the integral $\int_c^d M(X_1(y), y)\,\mathrm{d}y$, it can be viewed as $\int_{C_2} M\,\mathrm{d}y$. But this time, it means that we travel from A to C. Notice here the sign matters, this means that the direction we are travelling at matters. $\square$