

# Vector Calculus

Willoughby Seago

January 13, 2020

These are my notes for the *vector calculus* part of the *dynamics and vector calculus* course from the University of Edinburgh as part of the second year of the theoretical physics degree. When I took this course in the 2019/20 academic year it was taught by Professor Roman Zwicky<sup>1</sup>. These notes are based on the lectures delivered as part of this course, and the notes provided as part of this course. The content within is correct to the best of my knowledge but if you find a mistake or just disagree with something or think it could be improved please let me know.

These notes were produced using L<sup>A</sup>T<sub>E</sub>X<sup>2</sup>. Graphs were plotted using Python<sup>3</sup>, Matplotlib<sup>4</sup>, NumPy<sup>5</sup>, and SciPy<sup>6</sup>. Diagrams were drawn with tikz<sup>7</sup>.

This is version 1.0 of these notes, which is up to date as of 04/01/2021.

Willoughby Seago  
s1824487@ed.ac.uk

---

<sup>1</sup><https://www.ph.ed.ac.uk/people/roman-zwicky>

<sup>2</sup><https://www.latex-project.org/>

<sup>3</sup><https://www.python.org/>

<sup>4</sup><https://matplotlib.org/>

<sup>5</sup><https://numpy.org/>

<sup>6</sup><https://scipy.org/scipylib/>

<sup>7</sup><https://www.ctan.org/pkg/pgf>

# Contents

<b>I</b>	<b>Vector Differential Calculus</b>	<b>5</b>
<b>1</b>	<b>Recap</b>	<b>5</b>
1.1	Vectors and Scalars	5
1.2	Fields	5
1.3	Gravitation and Electrostatics	5
1.4	Need for Vector Calculus	6
1.5	Revision of Vector Algebra	7
<b>2</b>	<b>Index Notation</b>	<b>7</b>
2.1	Scalar Triple Product	7
2.2	Tensors	7
2.3	Kronecker and Levi-Civita Tensors	8
2.4	Vector Algebra with Index Notation	8
2.5	Rules and Properties	8
2.6	Symmetric Tensors	9
2.7	Vector Triple Product	9
<b>3</b>	<b>Gradient</b>	<b>10</b>
3.1	Tensor Addendum	10
3.2	Level Surfaces	11
3.3	Gradient	11
<b>4</b>	<b>More on the Gradient</b>	<b>12</b>
4.1	Geometric Interpretation	12
4.2	Directional Derivative	12
4.3	Physics Examples	13
4.4	Specific Gradients	13
4.5	Gradient Identities	14
<b>5</b>	<b>Divergence and Curl</b>	<b>14</b>
5.1	More Gradient Examples	14
5.2	The Operator Del	15
5.3	Equations of Points, Lines and Planes	15
5.3.1	Equation of a Point	15
5.3.2	Equation of a Line	15
5.3.3	Equation of a Plane	15
5.4	Divergence	16
5.5	Curl	16
<b>6</b>	<b>Geometric Interpretation and the Laplacian</b>	<b>17</b>
6.1	Specific Examples	17
6.2	Geometric Interpretation	18
6.3	Laplacian	18
6.4	Vector Operator Identities	19
6.4.1	Distributive Laws	19
6.4.2	Product Laws (One Vector Field)	19
<b>7</b>	<b>More Identities</b>	<b>19</b>
7.1	Product Laws (Two Vector Fields)	19
7.2	Identities with two $\nabla$	20
7.3	Summary of Elementary Results	21

<b>II</b>	<b>Vector Integral Calculus</b>	<b>21</b>
<b>8</b>	<b>One Dimensional Integrals</b>	<b>21</b>
8.1	Ordinary Integrals	21
8.2	Motivation	21
8.3	Parametrising A Curve	21
8.4	Different Types of Line Integral	23
<b>9</b>	<b>Surface Integrals and Polar Coordinates</b>	<b>24</b>
9.1	Surface Elements	24
9.2	Parametric Form	25
9.3	Plane Polar Coordinates	25
9.4	Cylindrical Coordinates	26
9.5	Spherical Polar Coordinates	26
<b>10</b>	<b>More Surface Integrals and Flux</b>	<b>26</b>
10.1	Flux	26
10.2	Other Surface Integrals	27
<b>11</b>	<b>Volume Integrals</b>	<b>28</b>
11.1	Integrals Over Vector Fields	29
<b>12</b>	<b>Divergence Theorem</b>	<b>29</b>
12.1	Integral Definition of Divergence	29
12.2	Showing Equivalence	29
12.3	Intuitive Understanding of Divergence	30
12.4	The Divergence Theorem	30
12.5	Volume of a Body Using the Divergence Theorem	31
<b>13</b>	<b>Continuity Equation and Gauss's Law</b>	<b>31</b>
13.1	Continuity Equation	31
13.2	Sources and Sinks	32
13.3	Electrostatics	32
13.4	Corollaries of the Divergence Theorem	33
<b>14</b>	<b>Stokes' Theorem</b>	<b>34</b>
14.1	Integral Definition of Curl	34
14.2	Showing Equivalence	34
14.3	Intuitive Understanding of Curl	35
14.4	Stokes' Theorem	35
<b>15</b>	<b>Stokes' Theorem In Use</b>	<b>36</b>
15.1	Planar Areas as a Special Case	37
15.2	Corollaries of Stokes' Theorem	37
15.3	Divergence and Stokes' Theorems Synopsis	38
<b>16</b>	<b>Potentials</b>	<b>38</b>
16.1	Maxwell's Equations	38
16.2	Potentials Introduction	39
16.3	Scalar Potential	39
<b>17</b>	<b>Calculating Scalar Potentials</b>	<b>40</b>
17.1	The First Method	40
17.2	The Second Method	42
17.3	The Third Method	42
17.4	Conservative Forces and Conservation of Energy	43

17.5 Gravitation and Electrostatics . . . . .	44
17.6 Poisson and Laplace Equations . . . . .	44
<b>18 Vector Potentials . . . . .</b>	<b>45</b>
18.1 Upcoming Result . . . . .	45
18.2 Helmholtz's Theorem . . . . .	45
18.3 Gauge Invariance . . . . .	46
18.4 Physics Examples . . . . .	46
 <b>III Orthogonal Curvilinear Coordinates . . . . .</b>	 <b>47</b>
<b>19 OCC line, area and volume elements, grad and divergence . . . . .</b>	<b>47</b>
19.1 What Are Orthogonal Curvilinear Coordinates? . . . . .	47
19.2 Length, Area and Volume Elements in OCC . . . . .	47
19.2.1 Length . . . . .	47
19.2.2 Area . . . . .	48
19.2.3 Volume . . . . .	48
19.3 Components of a Vector Field . . . . .	48
19.4 Gradient in OCC . . . . .	48
19.5 Divergence in OCC . . . . .	48
 <b>20 Curl and the Laplacian in OCC . . . . .</b>	 <b>49</b>
20.1 Curl in OCC . . . . .	49
20.2 Laplacian in OCC . . . . .	50
20.2.1 Laplacian of a Scalar Field . . . . .	50
20.3 Laplacian of a Vector Field . . . . .	50
20.4 Alternative Derivation of Divergence . . . . .	51

## List of Figures

6.1 Vector fields $\mathbf{a} = \mathbf{r}$ and $\mathbf{b} = \boldsymbol{\omega} \times \mathbf{r}$ in the $\mathbf{e}_1 - \mathbf{e}_2$ plane . . . . .	18
---	----

## Part I

# Vector Differential Calculus

## 1 Recap

### 1.1 Vectors and Scalars

A scalar is a quantity that is invariant under coordinate transformation. In Newtonian mechanics we are interested in rotations as the coordinate transform since euclidean space is rotationally symmetric. This means that a scalar is any quantity that remains constant after rotation. Examples are mass and energy (in non-relativistic physics).

A vector in  $n$ -dimensional space is  $n$  scalars that transform in a specific way under rotation. In this course we are only concerned with 3-dimensional vectors. Examples of vectors are force and position.

We can use an orthonormal basis set  $\mathbf{e}_i$  to describe a vector. Orthonormal means orthogonal and unit length. A vector  $\mathbf{a}$  in this basis is represented as

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 = \sum_i a_i\mathbf{e}_i \stackrel{\text{ESC}}{=} a_i\mathbf{e}_i$$

where ESC indicates the application of the Einstein summation convention which is that a repeated index (in this case  $i$ ) implies summation over that index. Such an index is called a dummy index. It is possible to find the value of  $a_i$  by doing  $\mathbf{a} \cdot \mathbf{e}_i$ .

The norm or magnitude of a vector can be thought of as its length. It is a scalar. A vector  $\mathbf{a}$  has norm  $|\mathbf{a}| = a$ , this can be calculated as

$$a = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{\mathbf{a} \cdot \mathbf{a}}$$

For historical reasons a position vector  $\mathbf{r}$  often has components  $x_i$  instead of  $r_i$  as one might expect.

### 1.2 Fields

A scalar field  $\varphi$  is a scalar quantity that depends on position, that is  $\varphi = \varphi(\mathbf{r}) = \varphi(x_1, x_2, x_3)$ . An example of a scalar field is temperature distribution  $T(\mathbf{r})$  or pressure  $P(\mathbf{r})$ .

A vector field  $\mathbf{v}$  is a vector quantity that depends on position, that is  $\mathbf{v} = \mathbf{v}(\mathbf{r}) = \mathbf{v}(x_1, x_2, x_3)$ . An example of a vector field is the electric field  $\mathbf{E}(\mathbf{r})$  or gravitational field  $\mathbf{G}(\mathbf{r})$ .

### 1.3 Gravitation and Electrostatics

Gravity is a force  $\mathbf{F}$  on a particle of mass  $m$  at position  $\mathbf{r}$  due to a particle of mass  $M$  at the origin. If  $M \gg m$  then the force is given classically by Newton's law of gravitation:

$$\mathbf{F}(\mathbf{r}) = -\frac{GM}{r^3}\mathbf{r}m = \mathbf{G}(\mathbf{r})m$$

where  $\mathbf{G}$  is the gravitational field given by

$$\mathbf{G}(\mathbf{r}) = -\frac{GM}{r^3}\mathbf{r}$$

note that this follows an inverse square law since

$$|\mathbf{G}| = \left| \frac{\mathbf{r}}{r^3} \right| = \frac{|\mathbf{r}|}{r^3} = \frac{1}{r^2}$$

The gravitational potential is

$$\varphi(\mathbf{r}) = -\frac{GM}{r}$$

which gives

$$-\nabla\varphi = \mathbf{G}$$

such that the potential energy of the small mass in the field is

$$V(\mathbf{r}) = m\varphi(\mathbf{r})$$

The force  $\mathbf{F}(\mathbf{r})$  on a particle of charge  $q$  due to a particle of charge  $Q$  at the origin is given by Coulomb's law:

$$\mathbf{F}(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0} \frac{\mathbf{r}}{r^3} q = \mathbf{E}(\mathbf{r})q$$

where  $\mathbf{E}$  is the electric field given by

$$\mathbf{E}(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0} \frac{\mathbf{r}}{r^3}$$

this follows an inverse square law. The electric potential is

$$\varphi(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0 r}$$

which gives

$$-\nabla\varphi = \mathbf{E}$$

such that the potential energy of the small charge in the field is

$$V(\mathbf{r}) = q\varphi(\mathbf{r})$$

Both of these laws clearly have similarities:

- Both follow an inverse square law
- Both have a constant:  $-G$  or  $1/4\pi\epsilon_0$
- The magnitude is proportional to the product of the masses or charges. The fact that charge can be negative but mass can't is one important difference.

For both laws if instead of being at the origin the second particle is at  $\mathbf{r}_2$  and the first is at  $\mathbf{r}_1$  then setting  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$  is equivalent to performing a coordinate transform such that particle 2 is at the origin.

## 1.4 Need for Vector Calculus

The questions we will mostly be interested in answering with vector calculus are:

1. How does a field change when we move from  $\mathbf{r} \rightarrow \mathbf{r} + d\mathbf{r}$  where  $d\mathbf{r}$  is an infinitesimal.

Answer: we use differential vector calculus and Taylor series.

2. What if we have more than two particles or a continuous distribution of mass or charge?

Answer: we use integral vector calculus

## 1.5 Revision of Vector Algebra

The scalar product of two vectors  $\mathbf{a}$  and  $\mathbf{b}$  gives the magnitude of  $\mathbf{a}$  projected onto  $\mathbf{b}$  times the magnitude of  $\mathbf{b}$ . If  $\vartheta_{ab}$  is the angle between  $\mathbf{a}$  and  $\mathbf{b}$  then the scalar product is

$$\begin{aligned}\mathbf{a} \cdot \mathbf{b} &= ab \cos \vartheta_{ab} \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3 \\ &= \sum_i a_i b_i \\ &\stackrel{\text{ESC}}{=} a_i b_i\end{aligned}$$

The vector product of two vectors  $\mathbf{a}$  and  $\mathbf{b}$  is the vector perpendicular to both that follows the right hand rule and has the magnitude of the parallelogram with vectors  $\mathbf{a}$  and  $\mathbf{b}$  as sides. If  $\hat{\mathbf{n}}$  is a vector normal to  $\mathbf{a}$  and  $\mathbf{b}$  satisfying the right hand rule and  $\vartheta_{ab}$  is the angle between  $\mathbf{a}$  and  $\mathbf{b}$  then

$$\mathbf{a} \times \mathbf{b} = ab \sin \vartheta_{ab} \hat{\mathbf{n}}$$

A useful (although mathematically meaningless) way of remembering this is

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

we know from the properties of determinants that interchanging two rows results in the determinant being multiplied by  $-1$ . This means that  $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ . In a right handed system

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$$

## 2 Index Notation

### 2.1 Scalar Triple Product

The scalar triple product  $(\cdot, \cdot, \cdot)$  takes three vectors and maps to a scalar:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

It gives the volume of the parallelepiped defined by the three vectors.

### 2.2 Tensors

An  $n$ -tensor in  $d$  dimensions is (for the sake of this course at least) an object  $t_{i_1 \dots i_n}$  where  $1 \leq i_1, \dots, i_n \leq d$ . We have come across some tensors already:

0. A 0-tensor is a scalar
1. A 1-tensor is a vector
2. A 2-tensor is a matrix
3. The Levi-Civita symbol is a 3-tensor

The components of tensors are basis dependent.

### 2.3 Kronecker and Levi-Civita Tensors

The Kronecker tensor  $\delta_{ij}$  is a 2-tensor defined as

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}$$

$\delta_{ij}$  is invariant under rotation. This is why length is invariant in Euclidean space as length can be found by a dot product which in turn uses  $\delta_{ij}$  as a metric in Euclidean space.  $\delta_{ij} = \mathbb{1}_{ij}$  where  $\mathbb{1}$  is the  $3 \times 3$  identity matrix.

The Levi-Civita tensor  $\varepsilon_{ijk}$  is a 3-tensor defined as

$$\varepsilon_{ijk} = \begin{cases} 1 & (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2) \\ -1 & (i, j, k) = (3, 2, 1), (2, 1, 3), (1, 3, 2) \\ 0 & \text{otherwise} \end{cases}$$

$\varepsilon_{ijk}$  is invariant under rotation which explains the covariance of the vector product (covariant for this purpose means transforms like a vector).

### 2.4 Vector Algebra with Index Notation

We can use these to calculate vector and scalar products with the Einstein summation convention:

$$\mathbf{a} \cdot \mathbf{b} = a_i \delta_{ij} b_j = a_i b_i$$

where we have used the tensor contraction and orthonormality properties defined in section 2.5.

$$\mathbf{a} \times \mathbf{b} = \varepsilon_{ijk} a_i b_j \mathbf{e}_k$$

or just one element can be written as

$$(\mathbf{a} \times \mathbf{b})_l = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{e}_l = [\varepsilon_{ijk} a_i b_j \mathbf{e}_k] \cdot \mathbf{e}_l = \varepsilon_{ijk} a_i b_j \delta_{kl} = \varepsilon_{ijl} a_i b_j$$

Care must be taken not to use the same dummy variable twice in the same product as this can lead to confusion, for example:

$$(\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \cdot \mathbf{d}) = a_i b_i c_j d_j \neq a_i b_i c_i d_i$$

these are clearly different since the first includes the term  $a_1 b_1 c_2 d_2$  but the second doesn't. We can also write the scalar triple product in index notation:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \varepsilon_{ijk} a_i b_j c_k$$

### 2.5 Rules and Properties

Contraction with  $\delta_{ij}$ :

$$\delta_{ij} t_{\dots i \dots} = t_{\dots j \dots}$$

This works since all terms where  $i \neq j$  are 0 so after summing over  $i$  only terms where  $i = j$  remain so we can replace  $i$  with  $j$ .

Orthonormality. It is clear that for an orthonormal basis set  $\{\mathbf{e}_i\}$   $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ .

The Lagrange identity is

$$\varepsilon_{ijk} \varepsilon_{klm} = \delta_{li} \delta_{mk} - \delta_{lj} \delta_{mi}$$



## 2.6 Symmetric Tensors

A tensor  $t_{\dots i_n \dots i_{n+m} \dots}$  is symmetric in  $i_n$  and  $i_{n+m}$  if

$$t_{\dots i_n \dots i_{n+m} \dots} = t_{\dots i_{n+m} \dots i_n \dots}$$

that is two indices can be swapped and leave the tensor unchanged. The same tensor is antisymmetric in  $i_n$  and  $i_{n+m}$  if

$$t_{\dots i_n \dots i_{n+m} \dots} = -t_{\dots i_{n+m} \dots i_n \dots}$$

$\delta_{ij}$  is symmetric and  $\varepsilon_{ijk}$  is antisymmetric. This allows us to write the most compact definition of  $\varepsilon_{ijk}$ : The Levi-Civita symbol is the antisymmetric 3-tensor with  $\varepsilon_{123} = 1$ .

Contraction of a symmetric and antisymmetric tensor:

Let  $s_{ij}$  be a symmetric tensor and  $a_{ij}$  an antisymmetric tensor. By this we know that

$$s_{ij} = +s_{ji} \quad a_{ij} = -a_{ji}$$

Let  $X = s_{ij}a_{ij}$ :

$$\begin{aligned} X &= s_{ij}a_{ij} = (+s_{ji})(-a_{ji}) = -s_{ij}a_{ij} = -X \\ X &= -X \implies X = 0 \end{aligned}$$

The second equality simply renames the dummy indices such that the names match the definition of  $X$ . We can illustrate this fact in 2 dimensions:

$$s_{ij}a_{ij} = s_{11}a_{11} + s_{22}a_{22} + s_{12}a_{12} + s_{21}a_{21}$$

The first two terms are 0 since  $a_{11} = a_{22} = 0$  as exchanging indices gives  $a_{11} = -a_{11} = 0$  and  $a_{22} = -a_{22} = 0$ .

$$\begin{aligned} s_{ij}a_{ij} &= s_{12}a_{12} + s_{21}a_{21} \\ &= s_{12}a_{12} + s_{12}a_{21} && \text{since } s_{12} = s_{21} \\ &= s_{12}a_{12} - s_{12}a_{12} && \text{since } a_{12} = -a_{21} \\ &= 0 \end{aligned}$$

This fact is useful since it means  $\delta_{ij}\varepsilon_{ijk} = 0$ .

## 2.7 Vector Triple Product

The vector triple product is

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

A proof of this is simple by considering one element at a time:

$$\begin{aligned} [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \varepsilon_{ijk}a_j(\mathbf{b} \times \mathbf{c})_k \\ &= \varepsilon_{ijk}a_j(\varepsilon_{klm}b_lc_m) \\ &= \varepsilon_{ijk}\varepsilon_{klm}a_jb_lc_m \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})a_jb_lc_m \\ &= \delta_{il}\delta_{jm}a_jb_lc_m - \delta_{im}\delta_{jl}a_jb_lc_m \\ &= (\delta_{il}b_l)(\delta_{jm}c_m)a_j - (\delta_{im}c_m)(\delta_{jl}b_l)a_j \\ &= b_ia_jc_j - c_ia_jb_j \\ &= b_i(\mathbf{a} \cdot \mathbf{c}) - c_i(\mathbf{a} \cdot \mathbf{b}) \end{aligned}$$

From here it is trivial to recover the full vector form of the vector triple product.

## 3 Gradient

### 3.1 Tensor Addendum

We define a  $n$  dimensional rotation matrix  $R$  to be a fundamental representation of  $SO(n)$ :

$$R \in SO(n) = \{O \in \mathbb{R}^{n \times n} | O^T O = \mathbb{1} \wedge \det O = 1\}$$

That is  $SO(3)$  is the set of all elements  $O$  which are real  $n \times n$  matrices ( $\mathbb{R}^{n \times n}$ ) that are orthogonal ( $O^T O = \mathbb{1}$ ) and have determinant 1 ( $\det O = 1$ ). It is easy to check that the rotation matrices fulfil all of these properties. In this course we are working in 3 dimensions so we are interested in the  $3 \times 3$  rotation matrices.

A constant tensor is any tensor  $a$  such that its elements are the same after an arbitrary allowed transformation:

$$M_{ij} = [M']_{ij}$$

In general this isn't true but there are specific tensors where it is. Every constant tensor can be used to construct an invariant product. For example:

- $\mathbf{a} \cdot \mathbf{b}$  is invariant  $\iff \delta_{ij}$  is a constant tensor  $\iff O^T O = \mathbb{1}$
- $(\mathbf{a}, \mathbf{b}, )$  is invariant  $\iff \varepsilon_{ijk}$  is a constant tensor  $\iff \det O = 1$

To show  $\delta_{ij}$  is a constant tensor consider the transformation  $\mathbf{a} \rightarrow O\mathbf{a}'$ :

$$\mathbf{a}' \cdot \mathbf{b}' = \mathbf{a}'^T \mathbf{b}' = (O\mathbf{a})^T O\mathbf{b} = \mathbf{a}^T O^T O\mathbf{b} = \mathbf{a}^T \mathbb{1} \mathbf{b} = \mathbf{a}^T \mathbf{b} = \mathbf{a} \cdot \mathbf{b}$$

so the fact that the dot product is invariant means that  $\delta_{ij}$  is a constant tensor.

How does a tensor transform?

Up to now we have been considering any object with indices to be a tensor. This isn't true. A tensor can be defined as an object that transforms like a tensor. What does this mean? First we will look at how a matrix transforms.

Define  $\mathbf{a} = M\mathbf{b}$  for some matrix  $M$ . Then  $\mathbf{a}' = O\mathbf{a} = OM\mathbf{b}$ . We also want to be able to transform  $M$  and  $\mathbf{b}$  and get the same result:

$$\mathbf{a}' = M'\mathbf{b}' = M'O\mathbf{b} = \underbrace{(OMO^T)}_{=M'} O\mathbf{b} = OM O^T O\mathbf{b} = OM \mathbb{1} \mathbf{b} = OM\mathbf{b}$$

Note that  $\mathbb{1}' = O\mathbb{1}O^T = OO^T = \mathbb{1}$  so  $\mathbb{1}$  is invariant. This is equivalent to saying that  $\delta_{ij}$  is invariant since  $\mathbb{1}_{ij} = \delta_{ij}$ . So a general matrix  $M$  transforms as

$$M'_{ij} = O_{ik} M_{kl} (O^T)_{lj}$$

we can use the fact that  $(O^T)_{lj} = O_{jl}$  and that once we look at individual components we are just looking at real numbers so they commute. This allows us to write

$$M'_{ij} = O_{ik} O_{jl} M_{kl}$$

This defines how a 2-tensor transforms and hence defines a 2 tensor. We can do the same for 1 and 3-tensors:

$$v'_i = O_{ij} v_j$$

$$t_{ijk} = O_{ia} O_{jb} O_{kc} t_{abc}$$

In general an  $n$ -tensor transforms by contraction with  $n$  components of  $O$  across all of its indices. Using this we can show that the Levi-Civita tensor is invariant:

$$\varepsilon'_{ijk} = O_{ia} O_{jb} O_{kc} \varepsilon_{abc}$$

$$\begin{aligned}
&= \varepsilon_{ijk} \det O \\
&= \varepsilon_{ijk}
\end{aligned}$$

where we have used the definition of the  $3 \times 3$  determinant using the Levi-Civita symbol ( $\det M = \varepsilon_{ijk} M_{1i} M_{2j} M_{3k}$ ) and the fact that  $\det O = 1$  by definition.

In a non-Euclidean geometry instead of the Kronecker delta tensor we use the metric  $g_{ij}$  in the dot product, which in general is not diagonal and dependent on the space.

$$\mathbf{a} \cdot \mathbf{b} = a_i g_{ij} b_j$$

### 3.2 Level Surfaces

The level surfaces or equipotentials of a scalar field  $\varphi$  are the surfaces where  $\varphi = c$  for some constant  $c$ . For example isobars on a weather map are surfaces of constant pressure or contour lines are lines of constant height.

#### Example 3.1

$\varphi(\mathbf{r}) = x_1^2 + x_2^2 + x_3^2 = r^2 = c$  The level surfaces are spheres of radius  $\sqrt{c}$  centred at the origin.

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r} - \mathbf{a}|} = c$$

The level surfaces are spheres centred on  $\mathbf{a}$  radius  $1/c$ .

$\varphi(\mathbf{r}) = \mathbf{k} \cdot \mathbf{r} = c$  for some constant vector  $\mathbf{k}$ . The level surfaces are planes normal to  $\mathbf{k}$ . This is most obvious in the case  $c = 0$  where any  $\mathbf{r}$  perpendicular to  $\mathbf{k}$  will suffice so the plane passes through the origin.

$\varphi = \exp(i\mathbf{k} \cdot \mathbf{r}) = c$ . The level surfaces are again a plane since we can take logarithms and divide through by  $i$  to get  $\mathbf{k} \cdot \mathbf{r} = -i \ln c = c'$  which is the same as the previous example.

In general for a function  $f$  if  $f^{-1}$  exists then if  $\varphi(\mathbf{r}) = c$  and  $f(\varphi(\mathbf{r})) = d$  for constants  $c$  and  $d$  then  $\varphi(\mathbf{r}) = f^{-1}(d) = c$  so the level surfaces have the same form.

We can also exploit the symmetry of the situation if  $\varphi(\mathbf{r}) = f(r)$  (ie no dependence on direction) then the function must be spherically symmetrical so the level surfaces must be spheres.

### 3.3 Gradient

The gradient gives the variation of a scalar function as a vector function. For this derivation we need a few concepts. We need partial derivatives for which we will introduce the notation

$$\frac{\partial}{\partial x_i} = \partial_{x_i} = \partial_i$$

The later is common but not in use in this course. We also need the Taylor series of a function. For a single variable function this is

$$f(x + \delta x) = f(x) + \delta x f'(x) + \mathcal{O}(\delta x^2)$$

For three variables it is

$$\varphi(\mathbf{r} + \delta \mathbf{r}) = \varphi(x_1 + \delta x_1, x_2 + \delta x_2, x_3 + \delta x_3) = \varphi(\mathbf{r}) + \delta x_i \partial_{x_i} \varphi(\mathbf{r}) + \mathcal{O}(\delta x_i \delta x_j)$$

with Einstein's summation convention. We define  $\delta \varphi = \varphi(\mathbf{r} + \delta \mathbf{r}) - \varphi(\mathbf{r})$  where as above  $\delta \mathbf{r} = (\delta x_1 \ \delta x_2 \ \delta x_3)^\top$ . The Taylor expansion of this is then

$$\delta \varphi = \varphi(\mathbf{r} + \delta \mathbf{r}) = \varphi(\mathbf{r}) + \delta x_i \partial_{x_i} \varphi(\mathbf{r}) + \mathcal{O}(\delta x_i \delta x_j) - \varphi(\mathbf{r})$$

$$\delta \varphi = \delta x_i \partial_{x_i} \varphi(\mathbf{r}) + \mathcal{O}(\delta x_i \delta x_j)$$

We now take the limit as  $\delta x_i \rightarrow 0 \forall i$ . This allows us to disregard the terms  $\delta x_i \delta x_j$  and above. We then define

$$\delta x_i \partial_{x_i} \varphi(\mathbf{r}) \stackrel{\text{def}}{=} \delta \mathbf{r} \cdot \nabla \varphi$$

Which leaves us with

$$\nabla \varphi = \partial_{x_i} \varphi(\mathbf{r}) \mathbf{e}_i$$

In Cartesian coordinates since  $\partial_{x_i} \mathbf{e}_i = 0$  we get that  $[\nabla \varphi]_i = \partial_{x_i} \varphi$ . This doesn't hold in general if the basis vectors have any dependence on position.

$\nabla$  is known as the nabla. It is a differential vector operator.  $\nabla \varphi$  is perpendicular to level surfaces since it points in the direction of fastest increase which is perpendicular to the level surfaces.

### Example 3.2

$\varphi(\mathbf{r}) = x_1^1 + x_2^2 + x_3^2 = r^2$ . What is the gradient of  $\varphi$ ?

$$\partial_{x_i} \varphi = 2x_i \implies \nabla \varphi = 2x_i \mathbf{e}_i = r \mathbf{r}$$

$\varphi(\mathbf{r}) = \mathbf{a} \cdot \mathbf{r} = a_i x_i$ . What is the gradient of  $\varphi$ ?

$$\partial_{x_i} \varphi = \partial_{x_i} a_j x_j = a_j \partial_{x_i} x_j = a_i \delta_{ij} = a_j \implies \nabla \varphi = a_i \mathbf{e}_i = \mathbf{a}$$

A useful shorthand applied here is that

$$\frac{\partial x_j}{\partial x_i} = \partial_{x_i} x_j = \delta_{ij}$$

## 4 More on the Gradient

### 4.1 Geometric Interpretation

Consider a small displacement  $\delta \mathbf{r}$  along a level surface of  $\varphi(\mathbf{r})$  from a point  $\mathbf{r}$ . The change in  $\varphi$  is zero as by definition  $\varphi$  has the same value everywhere on a level surface. We can also write this change in terms of the gradient like in the original derivation:

$$\delta \varphi = 0 = \varphi(\mathbf{r} + \delta \mathbf{r}) - \varphi(\mathbf{r}) = \nabla \varphi \cdot \delta \mathbf{r}$$

For this to be true for a step along the level surface in any direction we require  $\nabla \varphi$  to be normal to the level surface. This allows us to find a normal vector  $\hat{\mathbf{n}}$  for the level surface and its tangent plane at a point by evaluating:

$$\hat{\mathbf{n}} = \widehat{\nabla \varphi} = \frac{\nabla \varphi}{|\nabla \varphi|}$$

### 4.2 Directional Derivative

We want to know the gradient of  $\varphi$  in the direction  $\hat{\mathbf{s}}$ . We choose  $\delta \mathbf{r} = \delta s \hat{\mathbf{s}}$ . This means that a small change in  $\varphi$  is given by

$$\delta \varphi = \nabla \varphi \cdot \delta \mathbf{r} = (\nabla \varphi \cdot \hat{\mathbf{s}}) \delta s$$

$$\frac{\delta \varphi}{\delta s} = \nabla \varphi \cdot \hat{\mathbf{s}}$$

In a limiting process this then becomes

$$\frac{d\varphi}{ds} = \nabla \varphi \cdot \hat{\mathbf{s}}$$

This is the directional derivative.

We now ask when this has its maximal value. This occurs when  $\hat{\mathbf{s}}$  is parallel to  $\nabla \varphi$ . This means that  $\nabla \varphi$  points in the direction of maximum increase of  $\varphi$ .

### 4.3 Physics Examples

#### Example 4.1

If we have a temperature field  $T$  how does it change with time as we move through it? We consider a small change in the temperature field:

$$\delta T = \nabla T \cdot \delta \mathbf{r}$$

Dividing through by a small time change  $\delta t$

$$\frac{\delta T}{\delta t} = \nabla T \cdot \frac{\delta \mathbf{r}}{\delta t}$$

In a limiting process we get

$$\frac{dT}{dt} = \nabla T \cdot \frac{d\mathbf{r}}{dt} = \nabla T \cdot \mathbf{v}$$

#### Example 4.2

The gravitational potential of a mass  $m$  in the gravitational field of Earth is, close to the surface, given by

$$V(\mathbf{r}) = mgz$$

The force  $\mathbf{F}$  due to gravity is then

$$\mathbf{F} = -\nabla V = -mge_z$$

### 4.4 Specific Gradients

What is the gradient of  $\varphi(\mathbf{r}) = r$ ?

$$\begin{aligned}\varphi(\mathbf{r}) &= r = (x_1^2 + x_2^2 + x_3^2)^{1/2} \\ [\nabla \varphi]_i &= \partial_{x_i} r = x_i (x_1^2 + x_2^2 + x_3^2)^{-1/2} = \frac{x_i}{r} \\ \nabla \varphi &= \frac{x_i \mathbf{e}_i}{r} = \frac{\mathbf{r}}{r} = \hat{\mathbf{r}}\end{aligned}$$

What is the gradient of  $\varphi(\mathbf{r}) = 1/r$ ?

$$\begin{aligned}\varphi &= \frac{1}{r} = (x_1^2 + x_2^2 + x_3^2)^{-1/2} \\ [\nabla \varphi]_i &= \partial_{x_i} \frac{1}{r} = -x_i (x_1^2 + x_2^2 + x_3^2)^{-3/2} = -\frac{x_i}{r^3} \\ \nabla \varphi &= -\frac{x_i \mathbf{e}_i}{r^3} = -\frac{\mathbf{r}}{r^3} = -\frac{\hat{\mathbf{r}}}{r^2}\end{aligned}$$

#### Example 4.3

The gravitational potential is given by

$$V(\mathbf{r}) = -\frac{GMm}{r}$$

what is the force due to gravity?

$$\mathbf{F} = -\nabla V = -\nabla \left[ -\frac{GMm}{r} \right] = GMm \nabla \frac{1}{r} = -\frac{GMm}{r^2} \hat{\mathbf{r}}$$

One question that we will ask on this course is for what vector fields  $\mathbf{F}$  can we find an associated potential  $V$  such that  $\mathbf{F} = -\nabla V$ ? (The answer, as we will see, is this can be done only for conservative forces where  $\mathbf{F}$  is an irrotational field, that is  $\nabla \times \mathbf{F} = \mathbf{0}$ )

## 4.5 Gradient Identities

The distributive law is

$$\nabla(\varphi + \psi) = \nabla\varphi + \nabla\psi$$

This follows from the linearity of partial derivatives:

$$[\nabla(\varphi + \psi)]_i = \partial_{x_i}(\varphi + \psi) = \partial_{x_i}\varphi + \partial_{x_i}\psi = [\nabla\varphi]_i + [\nabla\psi]_i$$

The product rule for gradients is

$$\nabla(\varphi\psi) = (\nabla\varphi)\psi + \varphi(\nabla\psi)$$

This follows from the product rule for partial derivatives:

$$[\nabla(\varphi\psi)]_i = \partial_{x_i}(\varphi\psi) = (\partial_{x_i}\varphi)\psi + \varphi(\partial_{x_i}\psi) = [(\nabla\varphi)\psi]_i + [\varphi(\nabla\psi)]_i$$

The chain rule for the gradient allows us to find the gradient of functions  $F$  of other functions  $\varphi$  such as  $F(\varphi(\mathbf{r}))$ :

$$\nabla F(\varphi(\mathbf{r})) = \partial_\varphi F(\varphi) \nabla\varphi$$

This follows from the chain rule for partial derivatives:

$$[\nabla F(\varphi)]_i = \partial_{x_i} F(\varphi) = \partial_\varphi F(\varphi) \partial_{x_i} \varphi = \partial_\varphi F(\varphi) [\nabla\varphi]_i = \frac{dF}{d\varphi} \nabla\varphi$$

where we have used the normal chain rule in the second equality. Written out with Leibniz notation this is:

$$\frac{\partial F(\varphi)}{\partial x_i} = \frac{\partial F(\varphi)}{\partial \varphi} \frac{\partial \varphi}{\partial x_i} = \partial_\varphi F(\varphi) \partial_{x_i} \varphi$$

### Example 4.4

This gives us a much faster method for calculating the gradient of  $1/r$ . We take  $F(\varphi) = 1/\varphi$  and  $\varphi(\mathbf{r}) = r$ . We know that  $\nabla\varphi = \nabla r = \hat{\mathbf{r}}$  and that  $\partial_\varphi F(\varphi) = \partial_\varphi \varphi^{-1} = -\varphi^{-2}$ . Hence the gradient is given by

$$\nabla \frac{1}{r} = -\frac{1}{\varphi^2} \hat{\mathbf{r}} = -\frac{\hat{\mathbf{r}}}{r^2}$$

which is the same result as previously.

## 5 Divergence and Curl

### 5.1 More Gradient Examples

From before we have that  $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$  and  $\nabla r = \hat{\mathbf{r}}$ . Using the chain rule we can see that any function  $F$  of  $r$  has a gradient

$$\nabla F(r) = \frac{dF(r)}{dr} \nabla r = \frac{dF}{dr} \hat{\mathbf{r}}$$

For example  $F(r) = r^n$  then

$$\nabla r^n = \frac{dr^n}{dr} \hat{\mathbf{r}} = nr^{n-1} \hat{\mathbf{r}} = nr^{n-2} \mathbf{r}$$

Before we had that  $\nabla r^{-1} = -r^{-3} \hat{\mathbf{r}}$ . This is easily verified with the above formula for  $\nabla r^n$  with  $n = -1$ . If instead  $\varphi(\mathbf{r}) = r^n (\mathbf{a} \cdot \mathbf{r})^m$  for some constant vector  $\mathbf{a}$  then the gradient is

$$\begin{aligned} \nabla r^n (\mathbf{a} \cdot \mathbf{r})^m &= (\nabla r^n) (\mathbf{a} \cdot \mathbf{r})^m + r^n (\nabla [(\mathbf{a} \cdot \mathbf{r})^m]) \\ &= nr^{n-1} \hat{\mathbf{r}} (\mathbf{a} \cdot \mathbf{r})^m + m (\mathbf{a} \cdot \mathbf{r})^{m-1} r^n \nabla (\mathbf{a} \cdot \mathbf{r}) \\ &= nr^{n-1} \hat{\mathbf{r}} (\mathbf{a} \cdot \mathbf{r})^m + mr^n (\mathbf{a} \cdot \mathbf{r})^{m-1} \mathbf{a} \end{aligned}$$

## 5.2 The Operator Del

As previously discussed  $\nabla$  is a differential operator defined as  $\mathbf{e}_i \partial_{x_i}$ . It transforms like a vector since  $(x_1 \ x_2 \ x_3)^\top$  transforms like a vector. We can see this from the fact that  $\partial_{x_i} x_j = \delta_{ij} = \partial_{x'_i} x'_j$  so for this to be true it must transform like a vector. The directional derivative operator  $\hat{\mathbf{s}} \cdot \nabla$  shows that we can combine  $\nabla$  in multiple ways with other vectors, for a scalar field  $\varphi$  and vectors  $\mathbf{a}$  and  $\mathbf{b}$ :

- $\nabla \varphi$  is the gradient and is a vector.
- $(\mathbf{a} \cdot \nabla) \varphi$  is  $a(\hat{\mathbf{a}} \cdot \nabla) \varphi$  and is the directional derivative scaled by a constant  $a$ .
- $(\mathbf{a} \cdot \nabla) \mathbf{b}$  is a vector.
- $\mathbf{a} \cdot \nabla \mathbf{b}$  is the same as above but more ambiguous.
- $\mathbf{a} \cdot (\nabla \mathbf{b})$  is ambiguous and therefore wrong.

Note that in the definition of  $\nabla$  the unit vector comes first. This isn't important in Cartesian coordinates since  $\partial_{x_i} \mathbf{e}_j = \mathbf{0}$  but in coordinate systems where this isn't true it is important so it is best to be consistent and always write the unit vector first then the operator. In general operators aren't commutative. We can see that here:

$$\nabla \cdot \mathbf{a} \neq \mathbf{a} \cdot \nabla$$

the first is the divergence of  $\mathbf{a}$  and the second is an operator.

For small  $\mathbf{a}$  the following approximations can be made in analogy to a Taylor series

$$\varphi(\mathbf{r} + \mathbf{a}) = \varphi(\mathbf{r}) + \mathbf{a} \cdot \nabla \varphi + \mathcal{O}(a^2)$$

$$\mathbf{V}(\mathbf{r} + \mathbf{a}) = \mathbf{V}(\mathbf{r}) + (\mathbf{a} \cdot \nabla) \mathbf{V} + \mathcal{O}(a^2)$$

## 5.3 Equations of Points, Lines and Planes

### 5.3.1 Equation of a Point

A point is defined by a position vector  $\mathbf{r}$ .

### 5.3.2 Equation of a Line

A line has the point  $\mathbf{a}$  on it, the line is parallel to  $\mathbf{u}$ , the explicit equation for this line is

$$\mathbf{r} = \mathbf{a} + \lambda \mathbf{u}$$

where  $\lambda \in (-\infty, \infty)$ . This can be thought of as the first term  $\mathbf{a}$  taking you to the line and the second term  $\lambda \mathbf{u}$  moving you along the line to the desired point on the line. We can use the fact that  $\mathbf{u} \times \mathbf{u} = \mathbf{0}$  to get

$$\mathbf{r} \times \mathbf{u} = \mathbf{a} \times \mathbf{u} + \lambda \mathbf{u} \times \mathbf{u}$$

$$\mathbf{r} \times \mathbf{u} = \mathbf{a} \times \mathbf{u}$$

This is the implicit equation of the line. Note that the right hand side is constant. The specific case  $\mathbf{r} \times \mathbf{u} = \mathbf{0}$  is a line through the origin.

### 5.3.3 Equation of a Plane

A plane contains the point  $\mathbf{a}$  and the vectors  $\mathbf{u}$  and  $\mathbf{v}$  lie in the plane and are linearly independent of each other. Then the point  $\mathbf{r}$  on the plane is given by:

$$\mathbf{r} = \mathbf{a} + \lambda \mathbf{u} + \mu \mathbf{v}$$

where  $\lambda, \mu \in (-\infty, \infty)$ . This is the explicit equation of a plane. Like the line the first term  $\mathbf{a}$  takes you to the plane and the other two terms  $\lambda\mathbf{u} + \mu\mathbf{v}$  move you to the desired point on the plane. We define the normal to the plane as

$$\hat{\mathbf{n}} = \frac{\mathbf{u} \times \mathbf{v}}{|\mathbf{u} \times \mathbf{v}|}$$

and use the fact that this is perpendicular to  $\mathbf{u}$  and  $\mathbf{v}$  and hence  $\mathbf{u} \cdot \hat{\mathbf{n}} = \mathbf{v} \cdot \hat{\mathbf{n}} = 0$  to get

$$\begin{aligned}\mathbf{r} \cdot \hat{\mathbf{n}} &= \mathbf{a} \cdot \hat{\mathbf{n}} + \lambda\mathbf{u} \cdot \hat{\mathbf{n}} + \mu\mathbf{v} \cdot \hat{\mathbf{n}} \\ \mathbf{r} \cdot \hat{\mathbf{n}} &= \mathbf{a} \cdot \hat{\mathbf{n}}\end{aligned}$$

$$\begin{aligned}0 &= \mathbf{r} \cdot \hat{\mathbf{n}} - \mathbf{a} \cdot \hat{\mathbf{n}} \\ 0 &= (\mathbf{r} - \mathbf{a}) \cdot \hat{\mathbf{n}}\end{aligned}$$

This is the implicit equation of a plane. Note that the order  $\mathbf{r} - \mathbf{a}$  or  $\mathbf{a} - \mathbf{r}$  doesn't matter. The specific case  $\mathbf{r} \cdot \hat{\mathbf{n}} = 0$  (ie  $\mathbf{a} = \mathbf{0}$ ) is a plane through the origin.

## 5.4 Divergence

The divergence of a vector field  $\mathbf{V}$  is written as

$$\operatorname{div} \mathbf{V} = \nabla \cdot \mathbf{V}$$

As the dot product notation suggests in Cartesian coordinates

$$\nabla \cdot \mathbf{V} = \partial_{x_i} V_i$$

So in  $n$  dimensions the divergence maps a vector field  $\mathbf{V} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , which is differentiable over a region, to a scalar field  $\nabla \cdot \mathbf{V} : \mathbb{R}^n \rightarrow \mathbb{R}$  over that same region such that  $\mathbf{V} \mapsto \partial_{x_i} V_i$ .<sup>8</sup>

### Example 5.1

$$\nabla \cdot \mathbf{r} = \partial_{x_i} x_i = \delta_{ii} = 3$$

$$\begin{aligned}\mathbf{V}(\mathbf{r}) &= x_1^2 x_3 \mathbf{e}_1 - 2x_2^3 x_3^2 \mathbf{e}_2 + x_1 x_2^2 x_3 \mathbf{e}_3 \\ \nabla \cdot \mathbf{V} &= 2x_1 x_3 - 6x_2^2 x_3^2 + x_1 x_2^2\end{aligned}$$

## 5.5 Curl

The curl of a vector field  $\mathbf{V}$  is written as

$$\operatorname{curl} \mathbf{V} = \nabla \times \mathbf{V}$$

As the cross product notation suggests in Cartesian coordinates

$$\nabla \times \mathbf{V} = \varepsilon_{ijk} \mathbf{e}_i \partial_{x_j} V_k = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_{x_1} & \partial_{x_2} & \partial_{x_3} \\ V_1 & V_2 & V_3 \end{vmatrix}$$

Which version used depends on what is easier. Typically the index form is better for proofs and the determinant form is better for calculations. Note that like with the cross product the determinant has no physical meaning as  $\mathbf{e}_i$  and  $\partial_{x_i}$  are not valid entries in a matrix. Just like  $\nabla$  isn't really a vector. The curl

<sup>8</sup>Strictly speaking divergence is actually defined by an integral that gives the flux over a closed surface as the volume of that surface tends to zero. See the section on the divergence theorem for more.



maps a vector field  $\mathbf{V} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , which is differentiable over a region, to a vector field  $\nabla \times \mathbf{V} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  over the same region such that  $\mathbf{V} \mapsto \varepsilon_{ijk} \mathbf{e}_i \partial_{x_j} V_k$ .<sup>9</sup>

### Example 5.2

$$\nabla \times \mathbf{r} = \varepsilon_{ijk} \mathbf{e}_i \partial_{x_j} x_k = \mathbf{e}_i \varepsilon_{ijk} \delta_{jk} = \mathbf{0}$$

$$\begin{aligned} \mathbf{V}(\mathbf{r}) &= x_1^2 x_2 \mathbf{e}_1 + x_1 x_2^2 \mathbf{e}_2 + x_1 x_2 x_3 \mathbf{e}_3 \\ \nabla \times \mathbf{V} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_{x_1} & \partial_{x_2} & \partial_{x_3} \\ x_1^2 x_2 & x_1 x_2^2 & x_1 x_2 x_3 \end{vmatrix} \\ &= \mathbf{e}_1 (x_1 x_3 - 0) - \mathbf{e}_2 (x_2 x_3 - 0) + \mathbf{e}_3 (x_2^2 - x_1^2) \\ &= x_1 x_3 \mathbf{e}_1 - x_2 x_3 \mathbf{e}_2 + (x_2^2 - x_1^2) \mathbf{e}_3 \end{aligned}$$

## 6 Geometric Interpretation and the Laplacian

### 6.1 Specific Examples

Consider the vector field  $\mathbf{a}(\mathbf{r}) = \mathbf{r}$ . The divergence and curl are

$$\nabla \cdot \mathbf{r} = 3, \quad \nabla \times \mathbf{r} = \mathbf{0}$$

Consider the vector field  $\mathbf{b}(\mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}$  for constant  $\boldsymbol{\omega}$ . The divergence is

$$\nabla \cdot \mathbf{b} = \nabla \cdot (\boldsymbol{\omega} \times \mathbf{r}) = \nabla \cdot (\varepsilon_{ijk} \mathbf{e}_i \omega_j x_k) = \partial_{x_i} \varepsilon_{ijk} \omega_j x_k = \varepsilon_{ijk} \omega_j \partial_{x_i} x_k = \varepsilon_{ijk} \omega_j \delta_{ik} = \varepsilon_{ijk} \delta_{ik} \omega_j = 0$$

since the contraction of a symmetric and antisymmetric tensor is zero. The curl of this field is

$$\begin{aligned} \nabla \times \mathbf{b} &= \nabla \times (\boldsymbol{\omega} \times \mathbf{r}) \\ &= \varepsilon_{ijk} \mathbf{e}_i \partial_{x_j} (\boldsymbol{\omega} \times \mathbf{r})_k \\ &= \varepsilon_{ijk} \mathbf{e}_i \partial_{x_j} (\varepsilon_{klm} \omega_l x_m) \\ &= \varepsilon_{ijk} \varepsilon_{klm} \mathbf{e}_i \omega_l \partial_{x_j} x_m \\ &= \varepsilon_{ijk} \varepsilon_{klm} \mathbf{e}_i \omega_l \delta_{jm} \\ &= \varepsilon_{ijk} \varepsilon_{klm} \delta_{jm} \mathbf{e}_i \omega_l \\ &= \varepsilon_{ijk} \varepsilon_{klj} \mathbf{e}_i \omega_l \\ &= \varepsilon_{ijk} \varepsilon_{ljk} \mathbf{e}_i \omega_l \end{aligned}$$

We can use the Lagrange identity to write the Levi-Civita tensors as Kronecker deltas:

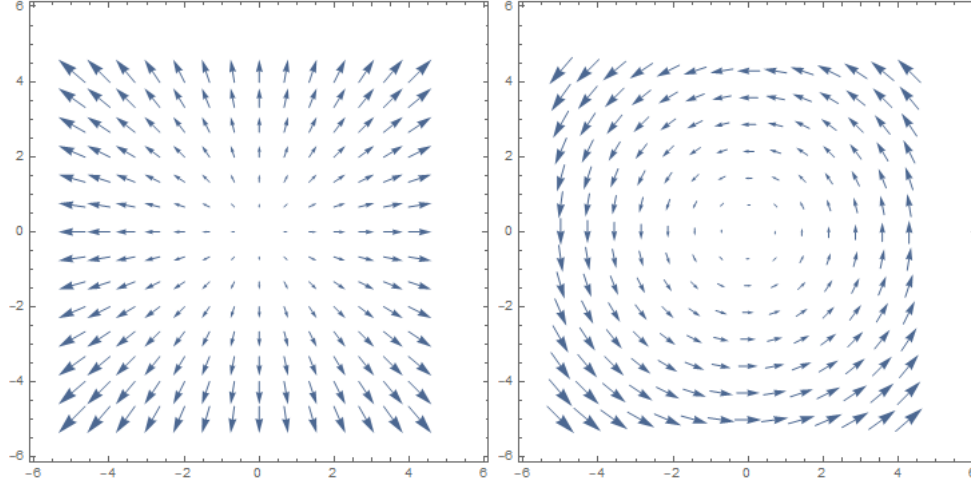
$$\begin{aligned} \varepsilon_{ijk} \varepsilon_{ljk} &= \delta_{il} \delta_{jj} - \delta_{ij} \delta_{lj} \\ &= 3\delta_{il} - \delta_{il} \\ &= 2\delta_{il} \end{aligned}$$

substituting back into the curl

$$\begin{aligned} \nabla \times \mathbf{b} &= 2\delta_{il} \mathbf{e}_i \omega_l \\ &= 2\mathbf{e}_i \omega_i \\ &= 2\boldsymbol{\omega} \end{aligned}$$

We can draw these two vector fields in the  $\mathbf{e}_1 - \mathbf{e}_2$  plane if we take  $\boldsymbol{\omega} = \omega \mathbf{e}_3$ . This is done in figure 6.1.

<sup>9</sup>Strictly speaking curl is actually defined by an integral around a closed loop as the area enclosed tends to zero. See the section on the Stokes' theorem for more.

Figure 6.1: Vector fields  $\mathbf{a} = \mathbf{r}$  and  $\mathbf{b} = \boldsymbol{\omega} \times \mathbf{r}$  in the  $\mathbf{e}_1 - \mathbf{e}_2$  plane

## 6.2 Geometric Interpretation

The divergence at a point can be thought of as a measure of how much the flux at that point acts like a source. If there are more field lines coming out of that point than going in the divergence is positive. If more are going in than coming out then the divergence is negative. Finally if the same number of field lines is going in as coming out then the field has zero divergence at that point. A field where the divergence is zero everywhere is known as solenoidal.

The curl at a point can be thought of as a measure of how much the field would cause a small ball at that point to rotate if the field represented the velocity of a fluid. The direction of the curl follows the right hand grip rule. In the special case where the curl is zero everywhere the field is said to be irrotational.

## 6.3 Laplacian

The Laplacian operator  $\nabla^2$  can be thought of as  $\nabla \cdot \nabla$ . In Cartesian coordinates it is

$$\begin{aligned}
 \nabla^2 &= \nabla \cdot \nabla \\
 &= \mathbf{e}_i \partial_{x_i} \cdot \mathbf{e}_j \partial_{x_j} \\
 &= \mathbf{e}_i \cdot \mathbf{e}_j \partial_{x_i} \partial_{x_j} \\
 &= \delta_{ij} \partial_{x_i} \partial_{x_j} \\
 &= \partial_{x_i} \partial_{x_i} \\
 &= \sum_i \partial_{x_i}^2
 \end{aligned}$$

Here we have used the fact that  $\partial_{x_i} \mathbf{e}_j = 0$  and the notation:

$$\partial_{x_i}^n = \frac{\partial^n}{\partial x_i^n}$$

The Laplacian is a scalar operator and as such can act on either a scalar or vector field. If it acts on a scalar field it returns a scalar field and if it acts on a vector field it returns a vector field. In general if the Laplacian acts on an  $n$ -tensor field it returns an  $n$ -tensor field. The Laplacian can be calculated as the divergence of the gradient but it is often easier to calculate directly.

$$\nabla^2 \varphi = (\nabla \cdot \nabla) \varphi = \nabla \cdot (\nabla \varphi)$$

**Example 6.1**

$$\begin{aligned}\nabla^2 r^2 &= \nabla \cdot (\nabla r^2) = \nabla \cdot (2r) = 2\nabla \cdot r = 2 \cdot 3 = 6 \\ \nabla^2(\mathbf{a} \cdot \mathbf{r}) &= \nabla \cdot (\nabla(\mathbf{a} \cdot \mathbf{r})) = \nabla \cdot \mathbf{a} = 0\end{aligned}$$

for constant  $\mathbf{a}$ .

**6.4 Vector Operator Identities****6.4.1 Distributive Laws**

$$\nabla \cdot (\mathbf{a} + \mathbf{b}) = \nabla \cdot \mathbf{a} + \nabla \cdot \mathbf{b} \quad (\text{A})$$

This follows from the distributive property of partial derivatives:

$$\begin{aligned}[\nabla \cdot (\mathbf{a} + \mathbf{b})]_i &= \partial_{x_i}(a_i + b_i) = \partial_{x_i}a_i + \partial_{x_i}b_i = [\nabla \cdot \mathbf{a}]_i + [\nabla \cdot \mathbf{b}]_i \\ \nabla \times (\mathbf{a} + \mathbf{b}) &= \nabla \times \mathbf{a} + \nabla \times \mathbf{b}\end{aligned} \quad (\text{B})$$

This follows from the distributive property of partial derivatives:

$$[\nabla \times (\mathbf{a} + \mathbf{b})]_i = \varepsilon_{ijk}\partial_{x_j}(a_k + b_k) = \varepsilon_{ijk}\partial_{x_j}a_k + \varepsilon_{ijk}\partial_{x_j}b_k = [\nabla \times \mathbf{a}]_i + [\nabla \times \mathbf{b}]_i$$

**6.4.2 Product Laws (One Vector Field)**

$$\nabla \cdot (\varphi \mathbf{a}) = (\nabla \varphi) \cdot \mathbf{a} + \varphi(\nabla \cdot \mathbf{a}) = \nabla \varphi \cdot \mathbf{a} + \varphi \nabla \cdot \mathbf{a} \quad (\text{C})$$

This follows from the product law for partial derivatives:

$$\begin{aligned}[\nabla \cdot \varphi \mathbf{a}]_i &= \partial_{x_i}\varphi a_i = (\partial_{x_i}\varphi)a_i + \varphi(\partial_{x_i}a_i) = [\nabla \varphi]_i a_i + \varphi[\nabla \cdot \mathbf{a}]_i = [\nabla \varphi \cdot \mathbf{a}]_i + [\varphi \nabla \cdot \mathbf{a}]_i \\ \nabla \times (\varphi \mathbf{a}) &= (\nabla \varphi) \times \mathbf{a} + \varphi(\nabla \times \mathbf{a}) = \nabla \varphi \times \mathbf{a} + \varphi \nabla \times \mathbf{a}\end{aligned} \quad (\text{D})$$

This follows from the product law for partial derivatives:

$$\begin{aligned}[\nabla \times \varphi \mathbf{a}]_i &= \varepsilon_{ijk}\partial_{x_j}\varphi a_k \\ &= \varepsilon_{ijk}(\partial_{x_j}\varphi)a_k + \varepsilon_{ijk}\varphi(\partial_{x_j}a_k) \\ &= \varepsilon_{ijk}[\nabla \varphi]_j a_k + \varphi[\nabla \times \mathbf{a}]_i \\ &= [\nabla \varphi \times \mathbf{a}]_i + [\varphi \nabla \times \mathbf{a}]_i\end{aligned}$$

**7 More Identities****7.1 Product Laws (Two Vector Fields)**

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b}) \quad (\text{E})$$

This is easiest to prove using the index notation:

$$\begin{aligned}\nabla \cdot (\mathbf{a} \times \mathbf{b}) &= \varepsilon_{ijk}\partial_{x_i}a_j b_k \\ &= \varepsilon_{ijk}(\partial_{x_i}a_j)b_k + \varepsilon_{ijk}a_j(\partial_{x_i}b_k) \\ &= b_k \varepsilon_{kij}\partial_{x_i}a_j - a_j \varepsilon_{jik}\partial_{x_i}b_k \\ &= b_k(\nabla \times \mathbf{a})_k - a_j(\nabla \times \mathbf{b})_j \\ &= \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b}) \\ \nabla \times (\mathbf{a} \times \mathbf{b}) &= (\nabla \cdot \mathbf{b})\mathbf{a} + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\nabla \cdot \mathbf{a})\mathbf{b} - (\mathbf{a} \cdot \nabla)\mathbf{b}\end{aligned} \quad (\text{F})$$

This is easiest to prove using the vector triple product

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

Note that since  $\nabla$  doesn't commute with vector fields it is important that the vectors come after the dot products and not before.

$$\begin{aligned} \nabla \times (\mathbf{a} \times \mathbf{b}) &= \partial_{x_i} b_i \mathbf{a} - \partial_{x_i} a_i \mathbf{b} \\ &= (\partial_{x_i} b_i) \mathbf{a} + b_i (\partial_{x_i} \mathbf{a}) - (\partial_{x_i} a_i) \mathbf{b} - a_i (\partial_{x_i} \mathbf{b}) \\ &= (\nabla \cdot \mathbf{b}) \mathbf{a} + (\mathbf{b} \cdot \nabla) \mathbf{a} - (\nabla \cdot \mathbf{a}) \mathbf{b} - (\mathbf{a} \cdot \nabla) \mathbf{b} \\ \nabla(\mathbf{a} \cdot \mathbf{b}) &= (\mathbf{a} \cdot \nabla) \mathbf{b} + (\mathbf{b} \cdot \nabla) \mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) \end{aligned} \tag{G}$$

## 7.2 Identities with two $\nabla$

$$\nabla \times \nabla \varphi = 0 \tag{H}$$

Consider one component of the curl of the gradient:

$$[\nabla \times \nabla \varphi]_i = \varepsilon_{ijk} \partial_{x_j} \partial_{x_k} \varphi = 0$$

This is because the partial derivatives commute (for sufficiently smooth  $\varphi$ ) so we can consider them to form a symmetric 2-tensor  $t_{jk} = \partial_{x_j} \partial_{x_k}$ . This means that we are then doing a contraction between a symmetric and antisymmetric tensor which is always zero.

$$\nabla \cdot (\nabla \times \mathbf{a}) = 0 \tag{I}$$

Consider one component of the divergence of the curl:

$$[\nabla \cdot (\nabla \times \mathbf{a})]_i = \varepsilon_{ijk} \partial_{x_i} \partial_{x_j} a_k = 0$$

This is zero for the same reason as above.

$$\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a} \tag{J}$$

Identities A-D and H-I should be known.

We can use J to find the Laplacian in curvilinear coordinates if we already know the curl, divergence and gradient.

### Example 7.1

Let  $\mathbf{c}$  be a constant vector. What the curl of the following vector potential?

$$\mathbf{A} = (\mathbf{c} \times \mathbf{r}) r^{-3}$$

$$\begin{aligned} \nabla \times \mathbf{A} &= \nabla \times [(\mathbf{c} \times \mathbf{r}) r^{-3}] \\ &= \nabla r^{-3} \times (\mathbf{c} \times \mathbf{r}) + r^{-3} \nabla \times (\mathbf{c} \times \mathbf{r}) && \text{Identity D} \\ &= -3r^{-5} \mathbf{r} \times (\mathbf{c} \times \mathbf{r}) + 2r^{-3} \mathbf{c} && \nabla r^n = nr^{n-2} \mathbf{r}, \nabla \times (\mathbf{c} \times \mathbf{r}) = 2\mathbf{c} \\ &= -3r^{-5} [\mathbf{c}(\mathbf{r} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{r} \cdot \mathbf{c})] + 2r^{-3} \mathbf{c} \\ &= -3r^{-5} [r^2 \mathbf{c} - (\mathbf{r} \cdot \mathbf{c}) \mathbf{r}] + 2r^{-3} \mathbf{c} \\ &= -3r^{-3} \mathbf{c} + 3r^{-5} (\mathbf{r} \cdot \mathbf{c}) \mathbf{r} + 2r^{-3} \mathbf{c} \\ &= 3r^{-5} (\mathbf{r} \cdot \mathbf{c}) \mathbf{r} - r^{-3} \mathbf{c} \end{aligned}$$

### 7.3 Summary of Elementary Results

$\mathbf{a}$  is a constant vector.

$$\begin{aligned}\nabla r &= \hat{\mathbf{r}}, & \nabla r^n &= nr^{n-2}\mathbf{r} = nr^{n-1}\hat{\mathbf{r}}, & \nabla \cdot \mathbf{r} &= 3, & \nabla \times \mathbf{r} &= 0 \\ \nabla(\mathbf{a} \cdot \mathbf{r}) &= \mathbf{a}, & (\mathbf{a} \cdot \nabla)\mathbf{r} &= \mathbf{a}, & \nabla \times (\mathbf{a} \times \mathbf{r}) &= 2\mathbf{a}, & \nabla \cdot (\mathbf{a} \times \mathbf{r}) &= 0\end{aligned}$$

## Part II

# Vector Integral Calculus

## 8 One Dimensional Integrals

### 8.1 Ordinary Integrals

The area under a function  $f$ , defined over a domain  $[a, b]$ , can be split into  $n$  strips. The leftmost  $x$  coordinate of the  $j^{\text{th}}$  strip is  $x_j$ . For each strip choose a value  $x_j^*$  such that  $x_j \leq x_j^* \leq x_{j+1}$ . We define the Riemann sum as

$$S_n = \sum_{j=0}^n f(x_j^*)(x_{j+1} - x_j)$$

If the limit exists and is the same no matter what value of  $x_j^*$  is picked in the range then

$$\lim_{n \rightarrow \infty} S_n = \int_a^b f(x) dx$$

Almost all functions that we come across in physics are nicely behaved enough that it isn't a problem. One function that isn't is

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \\ 1 & x \notin \mathbb{Q} \end{cases}$$

### 8.2 Motivation

The work done under the influence of a force  $\mathbf{F}$  depends on the path  $C$  chosen. The work done  $dW$  in a small step  $d\mathbf{r}$  along that path is  $dW = \mathbf{F} \cdot d\mathbf{r}$ . The total work done is then given by

$$W = \int_C \mathbf{F} \cdot d\mathbf{r}$$

To find what this is we need to ask several questions:

- How can we parametrise  $C$ ?
- What is  $d\mathbf{r}$ ?
- When is  $W$  path-independent?

### 8.3 Parametrising A Curve

We will work in 2 dimensions for ease in this section although all concepts can easily be extended to 3 or more dimensions. In Cartesian coordinates a curve can be described as a set of points  $(x_1, x_2)$ . Since a curve is one dimensional it is possible to parametrise this curve with only one variable  $\lambda$  such that our coordinates can be written as functions of  $\lambda$  so

$$x_1 = x_1(\lambda) \quad \& \quad x_2 = x_2(\lambda)$$

This allows us to write the position vector in terms of  $\lambda$ :

$$\mathbf{r}(\lambda) = x_1(\lambda)\mathbf{e}_1 + x_2(\lambda)\mathbf{e}_2$$

We can use the chain rule to find  $d\mathbf{r}$ :

$$d\mathbf{r} = \partial_\lambda \mathbf{r}(\lambda) d\lambda$$

In components this is

$$\begin{aligned} d\mathbf{r} &= \mathbf{e}_1 dx_1 + \mathbf{e}_2 dx_2 \\ &= \mathbf{e}_1 dx_1(\lambda) + \mathbf{e}_2 dx_2(\lambda) \\ &= \mathbf{e}_1 \partial_\lambda x_1(\lambda) d\lambda + \mathbf{e}_2 \partial_\lambda x_2(\lambda) d\lambda \\ &= (\mathbf{e}_1 \partial_\lambda x_1 + \mathbf{e}_2 \partial_\lambda x_2) d\lambda \\ &= \partial_\lambda \mathbf{r} d\lambda \end{aligned}$$

### Example 8.1

$$\mathbf{a} = -kx_1^3\mathbf{e}_1 + kx_2^3\mathbf{e}_2$$

The path  $C$  is from  $(1, 0)$  to  $(0, 1)$  by the shortest path on the unit circle. Evaluate

$$\int_C \mathbf{a} \cdot d\mathbf{r}$$

The path  $C$  can be parametrised by

$$x_1 = \cos \lambda, \quad x_2 = \sin \lambda$$

where  $\lambda \in [0, \pi/2]$ . This allows us to write  $\mathbf{a}$  as

$$\mathbf{a} = -k \cos^3 \lambda \mathbf{e}_1 + k \sin^3 \lambda \mathbf{e}_2$$

We can also write  $d\mathbf{r}$  as

$$d\mathbf{r} = -\mathbf{e}_1 \sin \lambda d\lambda + \mathbf{e}_2 \cos \lambda d\lambda$$

This gives

$$\begin{aligned} \mathbf{a} \cdot d\mathbf{r} &= (\sin^4 \lambda + \cos^4 \lambda) d\lambda \\ \int_C \mathbf{a} \cdot d\mathbf{r} &= k \int_0^{\pi/2} (\sin^4 \lambda + \cos^4 \lambda) d\lambda = 2kI \\ I &= \int_0^{\pi/2} \sin^4 \lambda d\lambda \end{aligned}$$

Here we have used the fact that the integral of  $\sin$  or  $\cos$ , to an even power, over a range  $[a, a + \pi/2]$  is the same which is easily verified by looking at the graphs. This means that the integral of  $\sin^4 \lambda + \cos^4 \lambda$  is equal to two times the integral of  $\sin^4 \lambda$  which is  $2I$ .

$$\begin{aligned} I &= \int_0^{\pi/2} \sin^4 \lambda d\lambda \\ &= \int_0^{\pi/2} \sin^2 \lambda (1 - \cos^2 \lambda) d\lambda \\ &= \underbrace{\int_0^{\pi/2} \sin^2 \lambda d\lambda}_{\tilde{I}} - \underbrace{\int_0^{\pi/2} \sin^2 \lambda \cos^2 \lambda d\lambda}_{I'} \end{aligned}$$

We can use the same trick again in reverse to write that

$$\tilde{I} = \int_0^{\pi/2} \sin^2 \lambda d\lambda = \frac{1}{2} \int_0^{\pi/2} (\sin^2 + \cos^2) d\lambda = \int_0^{\pi/2} d\lambda = \frac{\pi}{4}$$

$$I = \frac{\pi}{4} - I'$$

The second integral,  $I'$ , must be done by integration by parts.

$$I' = \int_0^{\pi/2} \sin^2 \lambda \cos^2 \lambda d\lambda$$

$$u = \cos \lambda \quad v = \frac{1}{3} \sin^3 \lambda$$

$$u' = -\sin \lambda \quad v' = \cos \lambda \sin^2 \lambda$$

$$I' = \left[ \frac{1}{3} \cos \lambda \sin^3 \lambda \right]_0^{\pi/2} + \frac{1}{3} \int_0^{\pi/2} \sin^4 \lambda d\lambda = 0 + \frac{I}{3}$$

$$I = \frac{\pi}{4} - \frac{I}{3}$$

$$\frac{4}{3}I = \frac{\pi}{4}$$

$$I = \frac{3\pi}{16}$$

Finally we can find the integral that we actually want:

$$\int_C \mathbf{a} \cdot d\mathbf{r} = 2kI = \frac{3k\pi}{8}$$

## 8.4 Different Types of Line Integral

There are many different ways to integrate over a line.  $\varphi$  is a scalar field and  $\mathbf{a}$  is a vector field over a region  $R$  and  $C$  is a curve in  $R$ . We define  $d\mathbf{r}$  as an infinitesimal step along  $C$ . We define  $ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}}$ . We can find various integrals along  $C$  such that we get a scalar result:

$$\int \varphi ds, \quad \int \mathbf{a} \cdot d\mathbf{r}$$

or a vector result:

$$\int \mathbf{a} ds, \quad \int \varphi d\mathbf{r}, \quad \int \mathbf{a} \times d\mathbf{r}$$

If  $C$  is a close path (ie it starts and ends in the same place) then it is common to write the integral sign with a circle to show this, for example:

$$\oint \mathbf{F} \cdot d\mathbf{r}$$

is the work done around a closed path (zero for a conservative force).

### Example 8.2

$$\mathbf{a} = (3x^2 + 6y)\mathbf{e}_1 - 14yz\mathbf{e}_2 + 20xz^2\mathbf{e}_3$$

Find the value of

$$I = \int_C \mathbf{a} \cdot d\mathbf{r}$$

where  $C$  is the path of straight lines connecting:

$$(0, 0, 0) \rightarrow (1, 0, 0) \rightarrow (1, 1, 0) \rightarrow (1, 1, 1)$$

We can split  $I$  in to three integrals:

$$I_1 = \int_{(0,0,0)}^{(1,0,0)} \mathbf{a} \cdot d\mathbf{r} \Big|_{y=z=0}$$

$$I_2 = \int_{(1,0,0)}^{(1,1,0)} \mathbf{a} \cdot d\mathbf{r} \Big|_{x=1, z=0}$$

$$I_3 = \int_{(1,1,0)}^{(1,1,1)} \mathbf{a} \cdot d\mathbf{r} \Big|_{x=y=1}$$

Such that  $I = I_1 + I_2 + I_3$ . It is easy to find  $d\mathbf{r}$  for each integral just by considering the geometry. For the three integrals  $d\mathbf{r}$  is given by  $dx\mathbf{e}_1$ ,  $dy\mathbf{e}_2$  and  $dz\mathbf{e}_3$  in turn. We can then find  $\mathbf{a} \cdot d\mathbf{r}$  for each integral and use the conditions on which variables are constant to get

$$I_1 = \int_0^1 3x^2 dx = 1$$

$$I_2 = \int_0^1 0 dy = 0$$

$$I_3 = \int_0^1 20z^2 dz = \frac{20}{3}$$

Finally

$$I = 1 + 0 + \frac{20}{3} = \frac{23}{3}$$

Now find the value of

$$I' = \int_{C'} \mathbf{a} \cdot d\mathbf{r}$$

where  $C'$  is the curve parametrised by

$$x = \lambda, \quad y = \lambda^2, \quad z = \lambda^3$$

where  $\lambda \in [0, 1]$ .

$$\begin{aligned} d\mathbf{r} &= \partial_\lambda \mathbf{r} d\lambda = \mathbf{e}_1 + 2\lambda\mathbf{e}_2 + 3\lambda^2\mathbf{e}_3 \\ \mathbf{a}(\mathbf{r}(\lambda)) &= 9\lambda^2\mathbf{e}_1 - 14\lambda^5\mathbf{e}_2 + 20\lambda^7\mathbf{e}_3 \\ \mathbf{a} \cdot d\mathbf{r} &= 9\lambda^2 - 28\lambda^6 + 30\lambda^9 \\ I' &= \int_{C'} \mathbf{a} \cdot d\mathbf{r} = \int_0^1 9\lambda^2 - 28\lambda^6 + 30\lambda^9 d\lambda \\ &= [3\lambda^3 - 4\lambda^7 + 6\lambda^{10}]_0^1 = 3 - 4 + 6 = 5 \end{aligned}$$

So  $I \neq I'$ . In general the integral along two different paths will not be equal. However if the integral is path-independent then the vector field is called conservative as if it were a force field the force would be conservative. We will see later that this occurs when the field has no curl.

## 9 Surface Integrals and Polar Coordinates

### 9.1 Surface Elements

A surface element is defined as  $d\mathbf{S} = dS\hat{\mathbf{n}}$  where  $\hat{\mathbf{n}}$  is normal to the surface and  $dS$  is an infinitesimal with the same magnitude as the area. There is a choice for every surface in the direction of  $\hat{\mathbf{n}}$  since the normal can point “up” or “down” from the surface. The convention is that for a closed surface always take the normal to be out of volume enclosed by the surface and for an open surface just pick a direction and be consistent.



## 9.2 Parametric Form

A surface can, since it is two dimensional, be described by only two parameters, say  $u$  and  $v$ . This allows the position vector to be written as

$$\mathbf{r} = \mathbf{r}(u, v) = x_1(u, v)\mathbf{e}_1 + x_2(u, v)\mathbf{e}_2 + x_3(u, v)\mathbf{e}_3$$

This gives line and surface elements:

$$d\mathbf{r} = \partial_u \mathbf{r} du + \partial_v \mathbf{r} dv$$

$$d\mathbf{S} = \partial_u \mathbf{r} du \times \partial_v \mathbf{r} dv$$

since the cross product gives the normal to the surface with the magnitude of the area. The choice in the order in the terms in the cross product is due to the choice in orientation of the surface normal. The coordinates  $u$  and  $v$  are orthogonal if  $\partial_u \mathbf{r}$  and  $\partial_v \mathbf{r}$  are orthogonal.

## 9.3 Plane Polar Coordinates

Plane polar coordinates are another coordinate system to describe a point in the plane. Instead of giving an  $x$  and  $y$  component to each point its distance,  $\rho$ , from the origin and the anticlockwise angle,  $\varphi$ , from some predefined 0 is given.  $\rho \in [0, \infty)$  and  $\varphi \in [0, 2\pi)$ . The Cartesian coordinates are given by

$$x_1(\rho, \varphi) = \rho \cos \varphi, \quad x_2(\rho, \varphi) = \rho \sin \varphi$$

Which gives the relations

$$\rho = \sqrt{x_1^2 + x_2^2}, \quad \tan \varphi = \frac{x_2}{x_1}$$

A position vector with Cartesian unit vectors is

$$\mathbf{r} = \rho \cos \varphi \mathbf{e}_1 + \rho \sin \varphi \mathbf{e}_2$$

We can also find unit vectors in this new coordinate system:

$$\mathbf{e}_\rho = \frac{\partial_\rho \mathbf{r}}{|\partial_\rho \mathbf{r}|} = \cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2 = \frac{1}{\rho} \mathbf{r}$$

since  $|\partial_\rho \mathbf{r}| = 1$ .

$$\mathbf{e}_\varphi = \frac{\partial_\varphi \mathbf{r}}{|\partial_\varphi \mathbf{r}|} = -\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2$$

since  $|\partial_\varphi \mathbf{r}| = \rho$ . We can check that these are orthonormal by computing the dot products:

$$\mathbf{e}_\rho \cdot \mathbf{e}_\varphi = -\cos \varphi \sin \varphi + \sin \varphi \cos \varphi = 0$$

$$\mathbf{e}_\rho \cdot \mathbf{e}_\rho = \cos^2 \varphi + \sin^2 \varphi = 1$$

$$\mathbf{e}_\varphi \cdot \mathbf{e}_\varphi = \sin^2 \varphi + \cos^2 \varphi = 1$$

As expected  $\mathbf{e}_\rho \cdot \mathbf{e}_\rho = \mathbf{e}_\varphi \cdot \mathbf{e}_\varphi = 1$  and  $\mathbf{e}_\rho \cdot \mathbf{e}_\varphi = 0$ . A vector,  $\mathbf{a}$ , in this orthonormal basis set,  $\{\mathbf{e}_\rho, \mathbf{e}_\varphi\}$ , can be written as

$$\mathbf{a} = a_\rho \mathbf{e}_\rho + a_\varphi \mathbf{e}_\varphi$$

where  $a_\rho = \mathbf{a} \cdot \mathbf{e}_\rho$  and  $a_\varphi = \mathbf{a} \cdot \mathbf{e}_\varphi$ . Note that  $\mathbf{e}_\rho$  and  $\mathbf{e}_\varphi$  are position dependent. A position vector in this basis is

$$\mathbf{r} = x_\rho \mathbf{e}_\rho + x_\varphi \mathbf{e}_\varphi$$

$$x_\varphi = \mathbf{r} \cdot \mathbf{e}_\varphi = \rho(\cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2) \cdot (-\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2) = 0$$

This makes sense as  $\mathbf{r}$  has no  $\varphi$  dependence.

$$x_\rho = \mathbf{r} \cdot \mathbf{e}_\rho = \rho(\cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2) \cdot (\cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2) = \rho$$

So a position vector can be written as

$$\mathbf{r} = \rho \mathbf{e}_\rho$$

## 9.4 Cylindrical Coordinates

Cylindrical coordinates extend plane polar coordinates into the third dimension. They are the same as plane polar coordinates but introduce a new  $z$  component in the  $\mathbf{e}_3$  direction. The relation of this  $z$  component to Cartesian coordinates is simply  $x_3 = z$ . This means that  $z \in (-\infty, \infty)$  and  $\mathbf{e}_z = \mathbf{e}_3$ . Extending position vector into cylindrical coordinates we get

$$\mathbf{r} = \rho \mathbf{e}_\rho + z \mathbf{e}_z$$

$\{\mathbf{e}_\rho, \mathbf{e}_\varphi, \mathbf{e}_z\}$  is an orthonormal, right handed, basis set:

$$\mathbf{e}_z = \mathbf{e}_\rho \times \mathbf{e}_\varphi$$

A vector,  $\mathbf{a}$ , in this basis is given by

$$\mathbf{a} = a_\rho \mathbf{e}_\rho + a_\varphi \mathbf{e}_\varphi + a_z \mathbf{e}_z$$

where  $a_\rho = \mathbf{a} \cdot \mathbf{e}_\rho$ ,  $a_\varphi = \mathbf{a} \cdot \mathbf{e}_\varphi$  and  $a_z = \mathbf{a} \cdot \mathbf{e}_z$ .

## 9.5 Spherical Polar Coordinates

Spherical coordinates are composed of a distance,  $r$ , from the origin, an angle,  $\vartheta$ , from the  $x_3$  axis and an angle,  $\varphi$ , as in plane polar coordinates. The Cartesian components are given by

$$x_1 = r \sin \vartheta \cos \varphi, \quad x_2 = r \sin \vartheta \sin \varphi, \quad x_3 = r \cos \vartheta$$

$$\begin{aligned} r^2 &= x_1^2 + x_2^2 + x_3^2 = r^2 \sin^2 \vartheta \cos^2 \varphi + r^2 \sin^2 \vartheta \sin^2 \varphi + r^2 \cos^2 \vartheta \\ &= r^2 [\sin^2 \vartheta (\cos^2 \varphi + \sin^2 \varphi) + \cos^2 \vartheta] = r^2 (\sin^2 \vartheta + \cos^2 \vartheta) = r^2 \end{aligned}$$

$$\cos \vartheta = \frac{x_3}{r}, \quad \tan \varphi = \frac{x_2}{x_1}$$

$r \in [0, \infty)$ ,  $\vartheta \in [0, \pi)$  and  $\varphi \in [0, 2\pi)$ . The unit vectors are

$$\mathbf{e}_r = \frac{\partial_r \mathbf{r}}{|\partial_r \mathbf{r}|} = \sin \vartheta \cos \varphi \mathbf{e}_1 + \sin \vartheta \sin \varphi \mathbf{e}_2 + \cos \vartheta \mathbf{e}_3$$

since  $|\partial_r \mathbf{r}| = 1$ .

$$\mathbf{e}_\vartheta = \frac{\partial_\vartheta \mathbf{r}}{|\partial_\vartheta \mathbf{r}|} = \cos \vartheta \cos \varphi \mathbf{e}_1 + \cos \vartheta \sin \varphi \mathbf{e}_2 + \sin \vartheta \mathbf{e}_3$$

since  $|\partial_\vartheta \mathbf{r}| = r$

$$\mathbf{e}_\varphi = \frac{\partial_\varphi \mathbf{r}}{|\partial_\varphi \mathbf{r}|} = -\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2$$

since  $|\partial_\varphi \mathbf{r}| = r \sin \vartheta$

# 10 More Surface Integrals and Flux

## 10.1 Flux

The flux of a vector field through a surface is a measure of how much that field flows through the surface in a direction normal to the surface. Let  $\mathbf{V}(\mathbf{r})$  be a vector field. The flux of  $\mathbf{V}$  through a surface  $S$  with surface elements  $d\mathbf{S}$  is defined as

$$I = \int_S \mathbf{V} \cdot d\mathbf{S} = \int_S V \cos \vartheta dS$$

If the field is entirely parallel to the surface, or zero for all points on the surface, then the flux is zero. The magnitude of the flux is maximal when the field is normal to the surface. Whether the flux is positive or negative depends on the direction chosen for the surface normal. For a closed surface with an outwards

pointing surface normal the flux being positive means that there is a net flow out of the volume and likewise a negative flux means a net flow into the volume. This sounds similar to the divergence's geometric definition and we shall see why when we look at the divergence theorem and the integral definition of the divergence.

### Example 10.1

One common physics example of flux is for a vector field  $\mathbf{v}$  representing the velocity of a fluid the flux across a surface gives the volume of flow through that surface in one second.

Another common physics example is that the flux of an electric field through a closed surface is equal to  $Q/\epsilon_0$  where  $Q$  is the charge enclosed in the surface. Likewise for a magnetic field the flux through a closed surface is always zero. Combined these are the first two of Maxwell's equations in integral form.

### Example 10.2

Let the surface  $S$  be the sphere of radius  $R$  centred on the origin. A point  $(x_1, x_2, x_3)$  is on this sphere if and only if

$$x_1^2 + x_2^2 + x_3^2 = R^2$$

Let  $\mathbf{a}(\mathbf{r}) = r^{-3}\mathbf{r}$  be a vector field. Note that  $a = r^{-2}$  so  $\mathbf{a}$  follows an inverse square law like gravity or the Coulomb force for a point charge. This sphere can be parametrised using spherical coordinates. The position vector,  $\mathbf{r}$ , is given by

$$\mathbf{r} = r(\sin \vartheta \cos \varphi \mathbf{e}_1 + \sin \vartheta \sin \varphi \mathbf{e}_2 + \cos \vartheta \mathbf{e}_3)$$

$$\partial_{\vartheta} \mathbf{r} = r(\cos \vartheta \cos \varphi \mathbf{e}_1 + \cos \vartheta \sin \varphi \mathbf{e}_2 - \sin \vartheta \mathbf{e}_3) = r \mathbf{e}_{\vartheta}$$

$$\partial_{\varphi} \mathbf{r} = r(-\sin \vartheta \sin \varphi \mathbf{e}_1 + \cos \vartheta \cos \varphi \mathbf{e}_2) = r \sin \vartheta \mathbf{e}_{\varphi}$$

The surface normal is then

$$\partial_{\vartheta} \mathbf{r} \times \partial_{\varphi} \mathbf{r} = r^2 \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos \vartheta \cos \varphi & \cos \vartheta \sin \varphi & -\sin \vartheta \\ -\sin \vartheta \sin \varphi & \cos \vartheta \cos \varphi & 0 \end{vmatrix}$$

With a lot of work this can be shown to be  $r^2 \sin \vartheta \mathbf{e}_r$ . It is easier to find the surface normal using spherical coordinates

$$\partial_{\vartheta} \mathbf{r} \times \partial_{\varphi} \mathbf{r} = r \mathbf{e}_{\vartheta} \times r \sin \vartheta \mathbf{e}_{\varphi} = r^2 \sin \vartheta \mathbf{e}_r$$

So a surface element is

$$d\mathbf{S} = (\partial_{\vartheta} \mathbf{r} \times \partial_{\varphi} \mathbf{r}) d\vartheta d\varphi = r^2 \sin \vartheta d\vartheta d\varphi \mathbf{e}_r$$

## 10.2 Other Surface Integrals

Like with line integrals there are different ways to integrate over a surface. We define  $dS = \sqrt{d\mathbf{S} \cdot d\mathbf{S}}$ .  $\varphi$  and  $\mathbf{a}$  are a scalar and vector field respectively over a region  $R$  and  $S$  is a surface in  $R$  then we can find various integrals over  $S$  such that we get a scalar result:

$$\int \varphi dS, \quad \int \mathbf{a} \cdot d\mathbf{S}$$

or a vector result:

$$\int \mathbf{a} dS, \quad \int \varphi d\mathbf{S}, \quad \int \mathbf{a} \times d\mathbf{S}$$

One common integral is

$$\int_S dS$$

which gives the surface area of  $S$ . Another useful result is that for a closed surface  $S$

$$\oint_S d\mathbf{S} = \mathbf{0}$$

since all of the components cancel to zero by the symmetry of the situation.

## 11 Volume Integrals

The volume integral of a scalar field  $f$  over a volume  $V$  is denoted as below and in Cartesian coordinates calculated as

$$\int_V f(\mathbf{r}) dV = \iiint_V f(\mathbf{r}) dx dy dz$$

Where the limits of the three integrals on the right hand side are chosen such that the entire volume is integrated over. Note that if  $f(\mathbf{r}) = 1$  then the integral over  $V$  gives the volume of  $V$ .

### Example 11.1

Integrate  $f(x, y, z) = xyz^2$  over the cube with diagonally opposite corners  $(0, 0, 0)$  and  $(a, b, c)$ .

$$I = \int_V f(x, y, z) dV = \int_0^c \int_0^b \int_0^a xyz^2 dx dy dz = \int_0^c z^2 dz \int_0^b y dy \int_0^a z dx = \left[ \frac{z^3}{3} \right]_0^c \left[ \frac{y^2}{2} \right]_0^b \left[ \frac{x}{2} \right]_0^a = \frac{a^2 b^2 c^3}{12}$$

Unfortunately not all integrals are this easily evaluated. Consider the octant of the unit sphere:

$$V = \{(x, y, z) | x^2 + y^2 + z^2 \leq 1, x, y, z \geq 0\}$$

The integral over this volume of a function,  $f$ , is

$$\int_V f dV = \int_0^1 \int_0^{\sqrt{1-z^2}} \int_0^{\sqrt{1-z^2-y^2}} f dx dy dz$$

Even for simple functions  $f$  the limits on the first integral makes the later integrals difficult to calculate.

Many volume integrals can be simplified by parametrisation. This is equivalent to a change of coordinates. In three dimensions three parameters are needed to describe a volume. If a position vector can be parametrised by  $u, v$  and  $w$  then  $\mathbf{r} = \mathbf{r}(u, v, w)$ . In this case

$$d\mathbf{r} = \partial_u \mathbf{r} du + \partial_v \mathbf{r} dv + \partial_w \mathbf{r} dw$$

This expands out into nine terms. The  $\mathbf{e}_i$  term is

$$d\mathbf{r} \cdot \mathbf{e}_i = \partial_u x_i du + \partial_v x_i dv + \partial_w x_i dw$$

The volume element  $dV$  is given by a scalar triple product:

$$dV = |(\partial_u \mathbf{r}, \partial_v \mathbf{r}, \partial_w \mathbf{r})| du dv dw$$

This is the same as the result using the Jacobian

$$J = |(\partial_u \mathbf{r}, \partial_v \mathbf{r}, \partial_w \mathbf{r})| = |\det(\partial_u x_i)|$$

where  $\mathbf{r} = x_i(u_1, u_2, u_3)\mathbf{e}_i$ .

### Example 11.2

A cylinder of height  $c$  and radius  $a$  is placed around the  $z$  axis with one face in the plane  $z = 0$ . Find the volume of the cylinder.

A volume element is given by

$$dV = |\partial_\rho \mathbf{r} \cdot (\partial_\varphi \mathbf{r} \times \partial_z \mathbf{r})| d\rho d\varphi dz = \rho d\rho d\varphi dz$$

$$V = \int_V dV = \int_0^c \int_0^{2\pi} \int_0^a \rho d\rho d\varphi dz = \pi a^2 c$$

which is what we would expect

### Example 11.3

Find the volume of a sphere radius  $r$ . A volume element is given by

$$dV = |\partial_r \mathbf{r} \cdot (\partial_\varphi \mathbf{r} \times \partial_\vartheta \mathbf{r})| d\rho d\varphi d\vartheta = r^2 \sin \vartheta dr d\vartheta d\varphi$$

$$V = \int_V dV = \int_0^\pi \int_0^{2\pi} \int_0^r r^2 \sin \vartheta dr d\vartheta d\varphi = 2\pi [-\cos \vartheta]_0^\pi \frac{r^3}{3} = \frac{4}{3}\pi r^3$$

### 11.1 Integrals Over Vector Fields

A volume integral can also be done over a vector field  $\mathbf{a}$ . In this case the result is an integral

$$\mathbf{I} = \int_V \mathbf{a} dV = \mathbf{e}_i \iiint_V a_i dx_1 dx_2 dx_3$$

So there are three integrals over the three basis vectors resulting in nine total integrals.

## 12 Divergence Theorem

### 12.1 Integral Definition of Divergence

Let  $\mathbf{a}$  be a vector field and  $\delta V$  be a volume with boundary  $\delta S$ . Let  $\mathbf{P}$  be a point in  $\delta V$ . We define the divergence of  $\mathbf{a}$  at point  $\mathbf{P}$  as

$$\operatorname{div} \mathbf{a}|_{\mathbf{P}} = \lim_{\delta V \rightarrow 0} \frac{1}{\delta V} \int_{\delta S} \mathbf{a} \cdot d\mathbf{S}$$

It can be shown that this is independent of the shape chosen for  $\delta V$ . We will show that this is the same as the definition of  $\nabla \cdot \mathbf{a}$  in Cartesian coordinates which will show that they are the same operator.

### 12.2 Showing Equivalence

Since we can choose any shape for  $\delta V$  we pick a cube with diagonally opposite vertices at  $(0,0,0)$  and  $(\delta_x, \delta_y, \delta_z)$ . The point  $\mathbf{P} = (x_0, y_0, z_0)$  is at the centre of the cube so  $\mathbf{P} = (\delta_x/2, \delta_y/2, \delta_z/2)$ . We will start by considering the flux across the planes normal to  $\mathbf{e}_1$ . A surface element on these planes is given by  $d\mathbf{S}_x = \pm dy dz \mathbf{e}_1$  where the  $\pm$  comes from the two faces with surface normals antiparallel. This gives  $\mathbf{a} \cdot d\mathbf{S}_x = \pm a_x dy dz$ . We now compute the right hand side of the divergence definition:

$$\frac{1}{\delta V} \int_{\delta S_x} \mathbf{a} \cdot d\mathbf{S} = \frac{1}{\delta V} \iint dz dy \left[ a_x \left( x_0 + \frac{\delta_x}{2}, y, z \right) - a_x \left( x_0 - \frac{\delta_x}{2}, y, z \right) \right]$$

Where the two  $a_x$  terms are due to the two faces. We can make  $\delta V \rightarrow 0$  by taking  $\delta_x, \delta_y, \delta_z \rightarrow 0$ . Since  $\delta_x$  is small we can Taylor expand this to get

$$\frac{1}{\delta V} \iint dy dz \left[ a_x(x_0, y_0, z_0) + \frac{\delta_x}{2} \partial_x a_x|_{(x_0, y, z)} + \mathcal{O}(\delta_x^2) - \left( a_x(x_0, y_0, z_0) - \frac{\delta_x}{2} \partial_x a_x|_{(x_0, y, z)} + \mathcal{O}(\delta_x^2) \right) \right]$$

Cancelling terms we get

$$\frac{\delta_x}{\delta V} \iint dy dz \partial_x a_x$$

We can use the fact that as  $a \rightarrow b$

$$\int_a^b f(x) dx \rightarrow f(x_0)(a - b)$$

for  $x_0 \in [a, b]$ , to find the value of this double integral since both integrals are over vanishing ranges  $\delta_y, \delta_z \rightarrow 0$

$$\frac{\delta_x}{\delta V} \iint dy dz \partial_x a_x = \frac{\delta_x \delta_y \delta_z}{\delta V} \partial_x a_x = \partial_x a_x|_{(x_0, y_0, z_0)}$$

since  $\delta V = \delta_x \delta_y \delta_z$ . Doing the same for the other four faces of the cube we see that

$$\operatorname{div} \mathbf{a} = \partial_x a_x + \partial_y a_y + \partial_z a_z$$

which is the same as the definition of  $\nabla \cdot \mathbf{a}$  in Cartesian coordinates.

### 12.3 Intuitive Understanding of Divergence

This allows us get a better intuitive understanding of divergence as a measure of the flux out of a point. For a vector field,  $\mathbf{a}$ , if  $\nabla \cdot \mathbf{a} > 0$  then there must be a net flux out of the point, if  $\nabla \cdot \mathbf{a} < 0$  then there must be a net flux into the point. This makes these two points a source and a sink respectively. If  $\nabla \cdot \mathbf{a} = 0$  then  $\mathbf{a}$  must be constant at the point so the flux in is equal to the flux out. This agrees with the Cartesian definition where  $\nabla \cdot \mathbf{a}$  is a derivative of  $\mathbf{a}$  so if  $\mathbf{a}$  is constant then this will be zero.

### 12.4 The Divergence Theorem

Let  $\mathbf{a}$  be a vector field and  $V$  be a volume with surface  $S$  (sometimes, confusingly, called  $\delta V$ ). The divergence theorem is

$$\int_V \nabla \cdot \mathbf{a} \, dV = \int_S \mathbf{a} \cdot d\mathbf{S}$$

This is a generalisation to three dimensions of the fundamental theorem of calculus:

$$\int_a^b \partial_x f(x) \, dx = f(b) - f(a)$$

We consider an arbitrary volume  $V$ . This volume can be approximated using small cubes. The  $i^{\text{th}}$  cube has a volume  $\delta V^{(i)}$  and surface  $\delta S^{(i)}$ . It is at position  $\mathbf{r}^{(i)}$ . The divergence at this cube can be calculated as

$$\nabla \cdot \mathbf{a}(\mathbf{r}^{(i)}) = \frac{1}{\delta V^{(i)}} \int_{\delta S^{(i)}} \mathbf{a} \cdot d\mathbf{S} + \varepsilon^{(i)}$$

where  $\varepsilon^{(i)}$  is a correction term to account for the fact that  $\delta V^{(i)}$  is finite. We can then approximate the integral of the divergence as the sum of this over all cubes. Let  $N$  be the number of cubes, then

$$\sum_{i=1}^N \nabla \cdot \mathbf{a}(\mathbf{r}^{(i)}) \delta V^{(i)} = \sum_{i=1}^N \int_{\delta S^{(i)}} \mathbf{a} \cdot d\mathbf{S} + \varepsilon^{(i)} \delta V^{(i)}$$

Consider the adjacent cubes 1 and 2 with volumes  $\delta V^{(1)}$  and  $\delta V^{(2)}$  respectively. Define the total volume of both as  $\delta V^{(1 \cup 2)}$ . The surface area of this volume is then given by the surface areas of the two cubes minus two times the area that they share. We will denote this  $\delta S^{(1 \oplus 2)}$ . Fortunately since both surfaces have equal and opposite surface normals at the intersection of the two surfaces the contribution of this cancels in the sum. This means

$$\int_{\delta S^{(1)}} \mathbf{a} \cdot d\mathbf{S} + \int_{\delta S^{(2)}} \mathbf{a} \cdot d\mathbf{S} = \int_{\delta S^{(1 \oplus 2)}} \mathbf{a} \cdot d\mathbf{S}$$

We now let  $N \rightarrow \infty$  and hence  $\delta V^{(i)} \rightarrow 0$ .

$$\begin{aligned} \sum_{i=1}^N \nabla \cdot \mathbf{a}(\mathbf{r}^{(i)}) \delta V^{(i)} &\rightarrow \int_V \nabla \cdot \mathbf{a} \, dV \\ \sum_{i=1}^N \varepsilon^{(i)} \delta V^{(i)} &\rightarrow 0 \end{aligned}$$

assuming that  $\varepsilon < 1/N$  so shrinks faster than  $N$  grows.

$$\sum_{i=1}^N \int_{\delta S^{(i)}} \mathbf{a} \cdot d\mathbf{S} \rightarrow \int_S \mathbf{a} \cdot d\mathbf{S}$$

From this we get

$$\int_V \nabla \cdot \mathbf{a} \, dV = \int_S \mathbf{a} \cdot d\mathbf{S}$$

which is the divergence theorem.

## 12.5 Volume of a Body Using the Divergence Theorem

Recall that  $\nabla \cdot \mathbf{r} = 3$  and that the volume integral of one is the volume. This means that

$$\frac{1}{3} \int_V \nabla \cdot \mathbf{r} \, dV = \frac{1}{3} \int_V 3 \, dV = \int_V dV = V$$

### Example 12.1

Use this to find the volume of the upper hemisphere with radius  $R$ . We define a coordinate system with the origin at the centre of the flat circular base of the hemisphere which is in the  $x_1$ – $x_2$  plane where  $x_3 = 0$ . We consider two surfaces  $S_c$  and  $S_b$  where  $S_c$  is the curved surface and  $S_b$  is the base. The whole surface is  $S = S_c + S_b$ . A surface element of the base is given by  $d\mathbf{S}_b = -\mathbf{e}_3 dS_b$  hence  $\mathbf{r} \cdot d\mathbf{S}_b = -x_3$  but  $x_3 = 0$  so  $\mathbf{r} \cdot d\mathbf{S}_b = 0$ . A surface element of the curved surface is given in spherical coordinates by

$$d\mathbf{S}_c = R^2 \sin \vartheta d\vartheta d\varphi \mathbf{e}_r$$

so  $\mathbf{r} \cdot d\mathbf{S}_c = R\mathbf{e}_r \cdot R^2 \sin \vartheta d\vartheta d\varphi \mathbf{e}_r = R^3 \sin \vartheta d\vartheta d\varphi$ . Hence the volume of the hemisphere is

$$\begin{aligned} V &= \frac{1}{3} \int_V \nabla \cdot \mathbf{r} \, dV = \int_S \mathbf{r} \cdot d\mathbf{S} = \int_{S_b} \mathbf{r} \cdot d\mathbf{S}_b + \int_{S_c} \mathbf{r} \cdot d\mathbf{S}_c = \frac{R^3}{3} \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin \vartheta \, d\vartheta \\ &= \frac{2}{3} \pi R^3 [-\cos \vartheta]_0^{\pi/2} = \frac{2}{3} \pi R^3 \end{aligned}$$

which is half the volume of a sphere as we would expect.

## 13 Continuity Equation and Gauss's Law

### 13.1 Continuity Equation

A volume  $V$  contains a fluid of density  $\rho(\mathbf{r}, t)$ . The fluid has velocity  $\mathbf{v}(\mathbf{r})$ . The flux per unit time of mass is the loss of mass in volume  $V$  per unit time and is given by

$$\int_S \rho \mathbf{v} \cdot d\mathbf{S}$$

where  $S$  is the surface of  $V$ .  $\rho \mathbf{v} = \mathbf{J}$  is a current density. The loss of mass per unit time can also be calculated as

$$-\partial_t \int_V \rho \, dV$$

since the volume integral of a density is the mass and the  $-\partial_t$  gives the rate of decrease. Hence

$$\int \rho \mathbf{v} \cdot d\mathbf{S} = -\partial_t \int_V \rho \, dV$$

By the divergence theorem

$$\int_V \nabla \cdot (\rho \mathbf{v}) \, dV = -\partial_t \int_V \rho \, dV$$

for a time independent volume this is

$$\int_V \nabla \cdot (\rho \mathbf{v}) \, dV = \int_V -\partial_t \rho \, dV$$

For this to be true for arbitrary volume  $V$  it is necessary and sufficient that the integrands be equal

$$\nabla \cdot (\rho \mathbf{v}) = -\partial_t \rho$$

Hence

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = \partial_t \rho + \nabla \cdot \mathbf{J} = 0$$

This same analysis works for other context for example:

- $\rho$  is thermal energy and  $\mathbf{J}$  is heat current density
- $\rho$  is electric charge density and  $\mathbf{J}$  is electric current density
- $\rho$  is a quantum mechanical probability density and  $\mathbf{J}$  is a probability current density

There are some special cases:

- Mass flux conservation:  $\partial_t \rho = \nabla \cdot \mathbf{J} = 0$ . No net flux means the mass in the volume is constant
- Incompressible fluid:  $\partial_t \rho = 0, \partial_{x_i} \rho = 0 \implies \nabla \cdot \mathbf{J} = \rho \nabla \cdot \mathbf{v} = 0$ . No local change in fluid density means the fluid is incompressible and hence  $\rho$  is a constant.

### 13.2 Sources and Sinks

The amount to which a certain volume,  $V$ , with surface  $S$  is a source of a vector field  $\mathbf{a}$  is characterised by

$$x_s = \int_S \mathbf{a} \cdot d\mathbf{S}$$

if  $x_s > 0$  then the volume is a net source of  $\mathbf{a}$ , if  $x_s < 0$  then the volume is a net sink of  $\mathbf{a}$  and if  $x_s = 0$  then the volume is neither a source nor sink of  $\mathbf{a}$ . This idea can be localised using the integral definition of the divergence.

$$\nabla \cdot \mathbf{a}|_P = \lim_{V \rightarrow 0} \frac{1}{V} \int_S \mathbf{a} \cdot d\mathbf{S}$$

this gives a measure of how the point  $P$  acts as a source or sink with the same sign requirements as the global version.

### 13.3 Electrostatics

The electric field from a point charge at the origin is

$$\mathbf{E}(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{1}{r^3} \mathbf{r}$$

This has a  $r^{-2}$  characteristic and we have already shown that for  $r \neq 0$  this gives a divergence of

$$\nabla \cdot \mathbf{E} = \frac{q}{4\pi\epsilon_0} \nabla \cdot (r^{-3} \mathbf{r}) = \frac{q}{4\pi\epsilon_0} [\nabla r^{-3} \cdot \mathbf{r} + r^{-3} \nabla \cdot \mathbf{r}] = 0$$

Formally if we want to include the point  $r = 0$  then we get

$$\nabla \cdot (r^{-3} \mathbf{r}) = 4\pi\delta^{(3)}(\mathbf{r})$$

where  $\delta^{(3)}$  is the three dimensional Dirac delta function. Let  $S$  be a sphere of radius  $R$  centred at the origin. Then  $d\mathbf{S} = \mathbf{e}_r r^2 d\vartheta d\varphi$

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} \int_S \frac{\mathbf{r} \cdot d\mathbf{S}}{r^3} = \frac{q}{4\pi\epsilon_0} \int_0^{2\pi} \int_0^\pi \frac{\mathbf{r} \cdot \mathbf{e}_r}{r^3} r^2 \sin \vartheta d\vartheta d\varphi = \frac{q}{4\pi\epsilon_0} 2\pi [-\cos]_0^\pi = \frac{q}{\epsilon_0}$$

This also holds for a charge enclosed by  $S$  but not at the origin. We want to show that it holds for an arbitrary surface.

Suppose  $S$  is a surface containing a point charge. We remove a small sphere,  $S_1$  around the point charge so that it is not in the surface  $S$ . The normal vectors for this sphere point into the sphere as they point out of  $S$ . Everywhere in  $S$  but not  $S_1$  the divergence is zero. Thus the flux across the surfaces  $S$  and  $S_1$  (ie the new surface with the sphere removed) is

$$\int_{S+S_1} \mathbf{E} \cdot d\mathbf{S} = \int_{V-V_1} \nabla \cdot \mathbf{E} dV = \int_{V-V_1} 0 dV = 0$$



where the first equality is by the divergence theorem and  $V - V_1$  is the volume contained by  $S$  minus the volume contained by  $S_1$ . The flux can also be calculated as

$$\begin{aligned}\int_{S+S_1} \mathbf{E} \cdot d\mathbf{S} &= \int_S \mathbf{E} \cdot d\mathbf{S} + \int_{S_1} \mathbf{E} \cdot d\mathbf{S} = \int_S \mathbf{e} \cdot d\mathbf{S} - \frac{q}{\varepsilon_0} = 0 \\ \implies \int_S \mathbf{E} \cdot d\mathbf{S} &= \frac{q}{\varepsilon_0}\end{aligned}$$

note the negative sign because of the inwards pointing surface vectors. From this we can see that this holds for any surface.

For  $N$  charges,  $q_i$ , we do the same but cut out  $N$  spheres  $S_i$ . We apply the same logic and the empirical result that electric fields add linearly and we get

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \int_S \sum_{i=1}^N \mathbf{E}_i \cdot d\mathbf{S} = \sum_{i=1}^N \int_{S_i} \mathbf{E}_i \cdot d\mathbf{S} = \sum_{i=1}^N \frac{q_i}{\varepsilon_0} = \frac{Q_{\text{total}}}{\varepsilon_0}$$

where  $Q_{\text{total}}$  is the total charge contained by the surface. This also generalises to charge density distributions  $\rho(\mathbf{r})$  defined such that

$$\int \rho(\mathbf{r}) dV = Q_{\text{total}}$$

in this case we get

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\varepsilon_0} \int \rho(\mathbf{r}) dV = \frac{Q_{\text{total}}}{\varepsilon_0}$$

This is Gauss's law and is the first of Maxwell's laws. This is it in integral form, we can get it in differential form using the divergence theorem:

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{E} dV = \frac{1}{\varepsilon_0} \int \rho(\mathbf{r}) dV$$

For this to be true for arbitrary volumes,  $V$ , it is necessary and sufficient for the integrands to be equal

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$$

This is Gauss's law in differential form.

### 13.4 Corollaries of the Divergence Theorem

If the vector field  $\mathbf{a}$  is a constant then  $\nabla \cdot \mathbf{a} = 0$  so by the divergence theorem

$$0 = \int_V \nabla \cdot \mathbf{a} = \oint_S \mathbf{a} \cdot d\mathbf{S} = \mathbf{a} \cdot \oint_S d\mathbf{S} = 0$$

For this to be true for arbitrary constant  $\mathbf{a}$  then

$$\oint_S d\mathbf{S} = \mathbf{0}$$

that is that the vector integral over any closed surface is zero. It has to be closed as the divergence theorem takes  $S$  as the surface containing  $V$  which must mean that  $S$  is closed.

If instead  $\mathbf{a} = p(\mathbf{r})\mathbf{c}$  for some constant vector  $\mathbf{c}$  then

$$\nabla \cdot \mathbf{a} = \nabla \cdot (p\mathbf{c}) = \nabla p \cdot \mathbf{c} + p\nabla \cdot \mathbf{c} = \nabla p \cdot \mathbf{c}$$

since the divergence of a constant vector is zero. From this and the divergence theorem we get

$$\mathbf{c} \cdot \int_V \nabla p dV = \int_V \mathbf{c} \cdot \nabla p dV = \int_V \nabla \cdot \mathbf{a} dV = \int_S \mathbf{a} \cdot d\mathbf{S} = \int_S p\mathbf{c} \cdot d\mathbf{S} = \mathbf{c} \cdot \int_S p d\mathbf{S}$$

for this to be true for arbitrary constants  $\mathbf{c}$  we require

$$\int \nabla p \, dV = \int p \, d\mathbf{S}$$

If  $\mathbf{a}(\mathbf{r}) = \mathbf{b}(\mathbf{r}) \times \mathbf{c}$  for some vector field  $\mathbf{b}(\mathbf{r})$  and constant vector  $\mathbf{c}$  then

$$\nabla \cdot \mathbf{a} = \mathbf{c} \cdot \nabla \times \mathbf{b} - \mathbf{b} \cdot \nabla \times \mathbf{c} = \mathbf{c} \cdot \nabla \times \mathbf{b}$$

since the curl of a constant vector is zero.

$$\int_V \nabla \cdot \mathbf{a} \, dV = \mathbf{c} \cdot \int_V \nabla \times \mathbf{b} \, dV = \int_S (\mathbf{b} \times \mathbf{c}) \cdot d\mathbf{S} = -\mathbf{c} \cdot \int_S \mathbf{b} \times d\mathbf{S}$$

For this to hold for all  $\mathbf{c}$  we require

$$\int_V \nabla \times \mathbf{b} \, dV = - \int_S \mathbf{b} \times d\mathbf{S}$$

## 14 Stokes' Theorem

### 14.1 Integral Definition of Curl

Let  $\delta C$  be a closed curve enclosing surface  $\delta S$ . The curl of a vector field  $\mathbf{a}$  through the surface satisfies

$$\text{curl}(\mathbf{a}) \cdot \hat{\mathbf{n}} = \lim_{\delta S \rightarrow 0} \frac{1}{\delta S} \oint_{\delta C} \mathbf{a} \cdot d\mathbf{r}$$

where  $\hat{\mathbf{n}}$  is the surface normal to  $\delta S$ . This can be used to define curl. Note that the directions of  $\delta C$  and  $\hat{\mathbf{n}}$  are important. By convention we choose them to be right handed so that if your right thumb points in the  $\hat{\mathbf{n}}$  direction then your fist curls in the direction of  $\delta C$ . It can be shown that  $\text{curl}(\mathbf{a}) \cdot \hat{\mathbf{n}}$  is independent of the shape of  $\delta C$ . We want to show that  $(\nabla \times \mathbf{a}) \cdot \hat{\mathbf{n}} = \text{curl}(\mathbf{a}) \cdot \hat{\mathbf{n}}$ .

### 14.2 Showing Equivalence

Since the curl is independent of the shape of  $\delta C$  we can choose any shape so we choose a square. We start with a rectangle in the  $x$ - $y$  plane. The rectangle has sides  $\delta x$  and  $\delta y$  in the  $x$  and  $y$  directions respectively. We choose to find the curl at a point  $P = (x_0, y_0, z_0) = (\delta x/2, \delta y/2, 0)$  which is in the middle of the rectangle. The path along which we integrate is then from  $A$  to  $B$  to  $C$  to  $D$  and back to  $A$  where

$$A = (\delta x, 0, 0), \quad B = (\delta x, \delta y, 0), \quad C = (0, \delta y, 0), \quad D = (0, 0, 0)$$

The surface bounded by this path is  $\delta S = \delta x \delta y$ . The surface normal is  $\hat{\mathbf{n}} = \mathbf{e}_3$ . First we compute the right hand side of the integral definition of curl. We get

$$\begin{aligned} \int_{\delta C} \mathbf{a} \cdot d\mathbf{r} &= \int_A^B \mathbf{a} \cdot d\mathbf{r} + \int_C^D \mathbf{a} \cdot d\mathbf{r} + \int_B^C \mathbf{a} \cdot d\mathbf{r} + \int_D^A \mathbf{a} \cdot d\mathbf{r} \\ &= \int_A^B \mathbf{a} \cdot d\mathbf{r} - \int_D^C \mathbf{a} \cdot d\mathbf{r} - \int_C^B \mathbf{a} \cdot d\mathbf{r} + \int_D^A \mathbf{a} \cdot d\mathbf{r} \end{aligned}$$

For the first two integrals  $d\mathbf{r} = \mathbf{e}_2 dy$  and for the second two  $d\mathbf{r} = \mathbf{e}_1 dx$

$$= \int_A^B a_y \, dx - \int_D^C a_y \, dy + \int_D^A a_x \, dx - \int_C^B a_x \, dy$$

We call these integrals  $I_3$ ,  $I_4$ ,  $I_1$  and  $I_2$  respectively. First we consider  $I_1$ .

$$I_1 = \int_A^B a_x dx = \int_{x_-}^{x_+} a_x \left( x, y_0 - \frac{\delta y}{2}, z_0 \right) dx$$

where  $x_{\pm} = x_0 \pm \delta x/2$ . We Taylor expand this to get

$$I_1 = \int_{x_-}^{x_+} a_x(x, y_0, z_0) - \partial_y a_x(x, y_0, z_0) \frac{\delta y}{2} + \mathcal{O}(\delta y^2) dx$$

we can also Taylor expand  $I_2$  and we get

$$I_2 = - \int_{x_-}^{x_+} a_x(x, y_0, z_0) + \partial_y a_x(x, y_0, z_0) \frac{\delta y}{2} + \mathcal{O}(\delta y^2) dx$$

In anticipation of our goal of showing equivalence we will calculate  $(I_1 + I_2)/\delta S$ :

$$\begin{aligned} \frac{I_1 + I_2}{\delta S} &= \frac{1}{\delta x \delta y} \int_{x_-}^{x_+} -\partial_y a_x(x, y_0, z_0) \delta y + \mathcal{O}(\delta y^2) dx \\ &= \frac{1}{\delta x} \int_{x_-}^{x_+} -\partial_y a_x(x, y_0, z_0) + \mathcal{O}(\delta y) dx \end{aligned}$$

Recall that for  $x^* \in [a, b]$  as  $a \rightarrow b$

$$\lim_{a \rightarrow b} \int_a^b f(x) dx = f(x^*)(b - a)$$

so as  $\delta x \rightarrow 0$  and  $x_- \rightarrow x_+$  we get

$$\frac{I_1 + I_2}{\delta S} = -\partial_y a_x(x_0, y_0, z_0)$$

Similarly we can show

$$\frac{I_3 + I_4}{\delta S} = -\partial_x a_y(x_0, y_0, z_0)$$

Finally we get

$$\text{curl}(\mathbf{a}) \cdot \mathbf{e}_3 = (\partial_x a_y - \partial_y a_x)|_{(x_0, y_0, z_0)}$$

which agrees with  $\nabla \times \mathbf{a}|_{(x_0, y_0, z_0)} \cdot \mathbf{e}_3$ . By symmetry this argument must hold for all other orientations of the initial surface so we can say that

$$\text{curl}(\mathbf{a}) \cdot \hat{\mathbf{n}} = (\nabla \times \mathbf{a}) \cdot \hat{\mathbf{n}}$$

hence the two definitions of curl are equivalent.

### 14.3 Intuitive Understanding of Curl

This gives us an intuitive understanding of curl as how much the field is in the same direction as  $\delta C$ . The more the directions agree the more positive the curl and if they are mostly in opposite directions then the curl will be negative. If either there happens to be cancellation around  $\delta C$  or if the vector field is perpendicular to  $\delta C$  all around then the curl will be 0.

### 14.4 Stokes' Theorem

Stokes' theorem is

$$\int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_C \mathbf{a} \cdot d\mathbf{r}$$

This follows from the integral definition of curl. Consider a surface  $\delta S$ , it can be approximated as being made of small squares. When two of these squares share an edge  $\delta C$  for each square is in opposite directions along the shared edge so the component of curl from the shared edge cancels to zero. This means that the

curl of a vector field over the surface is the same as the sum of the curl of the vector field over each square. For small, but not infinitesimal squares at positions  $\mathbf{r}^{(i)}$

$$\nabla \times \mathbf{a}(\mathbf{r}^{(i)}) \cdot \hat{\mathbf{n}}^{(i)} = \frac{1}{\delta S^{(i)}} \oint_{\delta C^{(i)}} \mathbf{a} \cdot d\mathbf{r} + \varepsilon^{(i)}$$

where  $\varepsilon^{(i)}$  is an error term that goes to zero as the number of squares goes to infinity. Multiplying through by  $\delta S$  and summing over all the squares we get

$$\sum_{i=1}^N \nabla \times \mathbf{a}(\mathbf{r}^{(i)}) \cdot \hat{\mathbf{n}}^{(i)} \delta S^{(i)} = \sum_{i=1}^N \oint_{\delta S^{(i)}} \mathbf{a} \cdot d\mathbf{r} \delta S^{(i)} + \sum_{i=1}^N \delta S^{(i)} \varepsilon^{(i)}$$

The second term goes to zero as  $N$  goes to infinity and the term on the left becomes an integral with  $\hat{\mathbf{n}} \delta S = \delta \mathbf{S} \rightarrow d\mathbf{S}$ , for this to always be true we require

$$\int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_C \mathbf{a} \cdot d\mathbf{r}$$

which is Stokes' theorem.

## 15 Stokes' Theorem In Use

Stokes' Theorem is a generalisation to three dimensions of

$$\int_y \partial_x f(x) dx = f(y) + C$$

### Example 15.1

Verify Stokes' theorem for a hemisphere of radius  $R$  and the vector field  $\mathbf{a} = 4y\mathbf{e}_1 + x\mathbf{e}_2 + 2z\mathbf{e}_3$ .

$$\nabla \times \mathbf{a} = -3\mathbf{e}_3$$

$$d\mathbf{S} = R^2 \sin \vartheta d\vartheta d\varphi \mathbf{e}_r$$

The left hand side of Stokes' theorem evaluates to

$$\int (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = -3R^2 \int_0^{\pi/2} d\vartheta \int_0^{2\pi} d\varphi \sin \vartheta \mathbf{e}_r \cdot \mathbf{e}_3$$

$$\mathbf{e}_r \cdot \mathbf{e}_3 = \cos \vartheta$$

$$\begin{aligned} &= -3R^2(2\pi) \int_0^1 d(\cos \vartheta) \cos \vartheta \\ &= -6\pi R^2 \int_0^1 \cos \vartheta d(\cos \vartheta) \\ &= -6\pi R^2 \frac{1}{2} \\ &= -3\pi R^2 \end{aligned}$$

Parametrising in terms of  $\varphi$  we get  $x(\varphi) = R \cos \varphi$ ,  $y(\varphi) = R \sin \varphi$  and  $z(\varphi) = 0$  which gives  $\mathbf{r}(\varphi) = R(\cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2)$ .

$$d\mathbf{r} = \partial_\varphi d\varphi = R(-\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2) d\varphi$$

$$\mathbf{a} \cdot d\mathbf{r} = R(-\sin \varphi 4y + \cos \varphi x) d\varphi = R^2(-4 \sin^2 \varphi + \cos^2 \varphi) d\varphi$$

The right hand side of Stokes' theorem evaluates to

$$\begin{aligned} \int_{\delta S} \mathbf{a} \cdot d\mathbf{r} &= R^2 \int_0^{2\pi} (-4 \sin^2 \varphi + \cos^2 \varphi) d\varphi \\ &= \pi R^2 (-4 + 1) \\ &= -3\pi R^2 \end{aligned}$$

which is the same as the left hand side so Stokes' theorem held.

### 15.1 Planar Areas as a Special Case

Consider the vector field

$$\mathbf{a} = \frac{1}{2}(-ye_1 + x \vee 2)$$

The curl of this vector field is  $\nabla \times \mathbf{a} = \mathbf{e}_3$  and  $d\mathbf{S} = \mathbf{e}_3 dS$ . The left hand side of Stokes' theorem is

$$\begin{aligned} \int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} &= \int_S \mathbf{e}_3 \cdot d\mathbf{S} = \int dS = S \\ \mathbf{a} \cdot d\mathbf{r} &= \frac{1}{2}(-ye_1 + xe_2) \cdot (dx e_1 + dy e_2) = \frac{1}{2}(x dy - y dx) \end{aligned}$$

The right hand side of Stokes' theorem is

$$S = \int_C \mathbf{a} \cdot d\mathbf{r} = \frac{1}{2} \oint_C x dy - y dx \quad (15.1)$$

#### Example 15.2

Consider the curve  $x^{2/3} + y^{2/3} = 1$ . This can be parametrised by  $x = \cos^3 \varphi$  and  $y = \sin^3 \varphi$  for  $\varphi \in [0, 2\pi]$ .

$$dx = \partial_\varphi x d\varphi = -3\cos^2 \varphi \sin \varphi d\varphi, \quad dy = \partial_\varphi y d\varphi = 3\sin^2 \varphi \cos \varphi d\varphi$$

By equation 15.1 this gives

$$\begin{aligned} S &= \frac{1}{2} \int_0^\pi 3(\cos^4 \varphi \sin^2 \varphi + \sin^4 \varphi \cos^2 \varphi) d\varphi \\ &= \frac{3}{2} \int_0^\pi \cos^2 \varphi \sin^2 \varphi d\varphi \\ &= \frac{3}{8} \int_0^\pi \sin^2 2\varphi d\varphi \\ &= \frac{3\pi}{8} \end{aligned}$$

### 15.2 Corollaries of Stokes' Theorem

For a constant vector field  $\mathbf{c}$

1. For the case where the vector field  $\mathbf{a} = \mathbf{c}$  we have  $\nabla \times \mathbf{a} = 0$  so

$$\oint_C \mathbf{a} \cdot d\mathbf{r} = \mathbf{c} \cdot \oint_C d\mathbf{r} = \int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = 0 \iff \oint_C d\mathbf{r} = 0$$

2. For a vector field  $\mathbf{a} = -Y(x, y)\mathbf{e}_1 + X(x, y)\mathbf{e}_2$  we have  $\nabla \times \mathbf{a} = (\partial_x X + \partial_y Y)\mathbf{e}_3$ . Also  $d\mathbf{S} = dx dy \mathbf{e}_3$  so

$$\begin{aligned} (\nabla \times \mathbf{a}) \cdot d\mathbf{S} &= (\partial_x X + \partial_y Y) dx dy \\ d\mathbf{r} &= dx \mathbf{e}_1 + dy \mathbf{e}_2 \implies \mathbf{a} \cdot d\mathbf{r} = -Y dx + X dy \\ \int_S (\partial_x X + \partial_y Y) dx dy &= \int_C X dy - Y dx \end{aligned}$$

This is Green's Theorem

3. For the case where the vector field is  $\mathbf{a} = \varphi(\mathbf{r})\mathbf{c}$  then

$$\nabla \times \mathbf{a} = \nabla \varphi \times \mathbf{c} + \varphi \nabla \times \mathbf{c} = \nabla \varphi \times \mathbf{c}$$

since the curl of a constant is zero. Hence

$$\int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_S (\nabla \varphi \times \mathbf{c}) \cdot d\mathbf{S} = \mathbf{c} \cdot \int_S d\mathbf{S} \times \nabla \varphi = \mathbf{c} \cdot \oint_C \varphi d\mathbf{r}$$

To be true for all  $\mathbf{c}$  we require

$$\oint_{\partial S} \varphi d\mathbf{r} = - \int_S \nabla \varphi \times d\mathbf{S}$$

Note that the case where  $\varphi = 1$  is the first item in this list.

### 15.3 Divergence and Stokes' Theorems Synopsis

Let  $S$  be a closed surface containing volume  $V$ . We can consider this to be an open surface with small hole with boundary  $C$  and then let  $C \rightarrow 0$  so the surface becomes closed and also  $\int (\nabla \times \mathbf{a}) \cdot d\mathbf{S} \rightarrow 0$  then

$$\int_C \mathbf{a} \cdot d\mathbf{r} = \int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_V \nabla \cdot (\nabla \times \mathbf{a}) dV = 0$$

Where the first equality is by Stokes' theorem and the second by the divergence theorem and the final equality is true as the integrand is zero since the divergence of the curl is always zero.

## 16 Potentials

### 16.1 Maxwell's Equations

One very common use of vector calculus is in Maxwell's equations. For electric field,  $\mathbf{E}$ , magnetic field,  $\mathbf{B}$ , charge density,  $\rho$ , and current density,  $\mathbf{J}$ , Maxwell's equations in a vacuum are

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0}, & \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} + \partial_t \mathbf{B} &= 0, & \nabla \times \mathbf{B} - \frac{1}{c^2} \partial_t \mathbf{E} &= \mu_0 \mathbf{J} \end{aligned}$$

It can be shown that it is always possible to write the magnetic and electric fields in terms of a vector potential,  $\mathbf{A}$ , and scalar potential,  $\varphi$ , such that

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla \varphi + \partial_t \mathbf{A}$$

Maxwell's equations can also be written relativistically. For this we work in natural units where  $c = 1$  and  $\mu_0 = 1$ . In dot products we replace the Kronecker delta,  $\delta_{ij}$ , with the Minkowski metric,  $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$  where we use the sign convention  $(+ - - -)$ . To write Maxwell's equations we need some 4-vectors:

- The four momentum which is  $p_\mu = (E, \mathbf{p})$  where  $E$  is the energy and  $\mathbf{p}$  is the 3-momentum.
- The four position which is  $x_\mu = (t, -\mathbf{x})$  where  $t$  is the time and  $\mathbf{x}$  is the 3-position.
- The four potential which is  $A_\mu = (\varphi, \mathbf{A})$  where  $\varphi$  is the scalar potential and  $\mathbf{A}$  is the 3-potential.
- The four current density which is  $j_\mu = (\rho, \mathbf{j})$  where  $\rho$  is the charge density and  $\mathbf{j}$  is the 3-current density.

The electric and magnetic fields are combined in the electromagnetic field tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

where  $\partial_\mu$  is the covariant 4-gradient,  $(\partial_t, \nabla)$ . Finally we need to define the Hodge dual which is

$$\tilde{F}_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}$$

where  $F^{\mu\nu} = \eta^{\mu\alpha} F_{\alpha\beta} \eta^{\beta\nu}$  and  $\eta^{\mu\nu} = \eta_{\mu\nu}$ . This allows us to compactly write Maxwell's equations as

$$\partial^\mu F_{\mu\nu} = j_\nu$$

$$\partial^\mu \tilde{F}_{\mu\nu} = 0$$

where  $\partial^\mu$  is the contravariant four gradient,  $\partial^\mu = \eta^{\mu\nu} \partial_\nu = (\partial_t, -\nabla)$ .

## 16.2 Potentials Introduction

Let  $f$  be a sufficiently nice function, then there exists a function  $F$  such that  $f(x) = F'(x)$  for all  $x$ . This makes  $F$  a potential for  $f$ . We want to extend this concept to a vector field. For a vector field it is not always possible to find a potential however Helmholtz's theorem states that for a smooth, twice continuously differentiable vector field  $\mathbf{V}$  then there exists  $\mathbf{D}$  and  $\mathbf{C}$  such that  $\mathbf{V} = \mathbf{D} + \mathbf{C}$  and  $\nabla \cdot \mathbf{D} = 0$  and  $\nabla \times \mathbf{C} = 0$ . That is every (sufficiently smooth) vector field can be written as the sum of a solenoidal vector field and an irrotational vector field. We can show that

$$\nabla \times \mathbf{C} = 0 \iff \exists c \text{ such that } \mathbf{C} = \nabla c$$

$$\nabla \cdot \mathbf{D} = 0 \iff \exists \mathbf{d} \text{ such that } \mathbf{D} = \nabla \times \mathbf{d}$$

hence

$$\mathbf{V} = \nabla c + \nabla \times \mathbf{d}$$

for some  $c$  and  $\mathbf{d}$ .  $c$  is known as the scalar potential for  $\mathbf{V}$  and  $\mathbf{d}$  is known as the vector potential. Note that this is not a unique decomposition since  $\nabla c = \nabla(c + c')$  and  $\nabla \times \mathbf{d} = \nabla \times (\mathbf{d} + \mathbf{d}')$  for constants  $c'$  and  $\mathbf{d}'$ . This allows us to choose a point to define as being equal to 0 for the potential. The question now is how do we find  $c$  and  $\mathbf{d}$ ?

## 16.3 Scalar Potential

We want to show that

$$\int_C \mathbf{a} \cdot d\mathbf{r} \text{ is path independent} \iff \nabla \times \mathbf{a} = 0 \iff \exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi$$

We start by showing

$$\int_C \mathbf{a} \cdot d\mathbf{r} \text{ is path independent} \iff \nabla \times \mathbf{a} = 0$$

Starting with the backwards implication. Consider a domain  $D$  containing two points  $A$  and  $B$ . Consider two paths from  $A$  to  $B$ . These two paths must form a closed curve. This curve is the boundary of a surface  $S$ . Call the two distinct paths  $C_1$  and  $C_2$ . Let  $\mathbf{a}$  be a vector field such that  $\nabla \times \mathbf{a} = 0$  inside  $D$ . Hence

$$0 = \int_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_{C_1} \mathbf{a} \cdot d\mathbf{r} - \int_{C_2} \mathbf{a} \cdot d\mathbf{r}$$

by Stokes' theorem since going from  $A$  to  $B$  along  $C_1$  and then back via  $C_2$  is a closed path around  $S$ . From this we can see that

$$\int_{C_1} \mathbf{a} \cdot d\mathbf{r} = \int_{C_2} \mathbf{a} \cdot d\mathbf{r}$$

hence the value of the integral is independent of the path taken and

$$\nabla \times \mathbf{a} = 0 \implies \int_C \mathbf{a} \cdot d\mathbf{r} \text{ is path independent}$$

Next we show the forward implication. The same argument works in reverse to require  $\nabla \times \mathbf{a} = 0$  assuming path independence of  $\int_C \mathbf{a} \cdot d\mathbf{r}$ . Hence

$$\int_C \mathbf{a} \cdot d\mathbf{r} \text{ is path independent} \implies \nabla \times \mathbf{a} = 0$$

Next we want to show

$$\nabla \times \mathbf{a} = 0 \iff \exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi$$

We start with the backwards implication.

$$\mathbf{a} = \nabla \varphi \implies \nabla \times \mathbf{a} = \nabla \times \nabla \varphi = 0$$

since the curl of the gradient is always zero. Hence

$$\exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi \implies \nabla \times \mathbf{a} = 0$$

For the forwards implication assume that there are two points  $\mathbf{r}_1$  and  $\mathbf{r}_2$  with some path between them. Assuming that  $\nabla \times \mathbf{a} = 0$  we know that the path between  $\mathbf{r}_1$  and  $\mathbf{r}_2$  is not important. Assuming sufficiently smooth  $\mathbf{a}$  we know that there exists some antiderivative  $\varphi$  such that

$$\varphi(\mathbf{r}_1) - \varphi(\mathbf{r}_2) = \int_{\mathbf{r}_2}^{\mathbf{r}_1} \mathbf{a} \cdot d\mathbf{r}'$$

Next we calculate  $\nabla \varphi \cdot d\mathbf{r}$  using a Taylor expansion:

$$\varphi(\mathbf{r} + d\mathbf{r}) \approx \varphi(\mathbf{r}) + \nabla \varphi \cdot d\mathbf{r} + \mathcal{O}(dr^2)$$

$$\nabla \varphi \cdot d\mathbf{r} = \varphi(\mathbf{r} + d\mathbf{r}) - \varphi(\mathbf{r}) + \mathcal{O}(dr^2)$$

Using the definition of  $\varphi$  from above this gives

$$\nabla \varphi \cdot d\mathbf{r} = \int_{\mathbf{r}}^{\mathbf{r}+d\mathbf{r}} \mathbf{a} \cdot d\mathbf{r}' + \mathcal{O}(dr^2)$$

where we have replaced  $\mathbf{r}_2$  with  $\mathbf{r}$  and  $\mathbf{r}_1$  with  $\mathbf{r} + d\mathbf{r}$ . Allowing  $dr \rightarrow 0$  we get

$$\nabla \varphi \cdot d\mathbf{r} \rightarrow \mathbf{a}(\mathbf{r}) \cdot \int_{\mathbf{r}}^{\mathbf{r}+d\mathbf{r}} d\mathbf{r}' = \mathbf{a} \cdot d\mathbf{r}$$

where we have used  $\int_{\mathbf{r}}^{\mathbf{r}+d\mathbf{r}} = \mathbf{r} + d\mathbf{r} - \mathbf{r} = d\mathbf{r}$  and  $\mathcal{O}(dr^2) \rightarrow 0$ . For this to be true we require

$$\nabla \varphi = \mathbf{a}$$

which was what we wanted so we have shown that

$$\nabla \times \mathbf{a} = 0 \implies \exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi$$

Hence

$$\nabla \times \mathbf{a} = 0 \iff \exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi$$

Hence we have shown what we set out to show:

$$\int_C \mathbf{a} \cdot d\mathbf{r} \text{ is path independent} \iff \nabla \times \mathbf{a} = 0 \iff \exists \varphi \text{ such that } \mathbf{a} = \nabla \varphi$$

## 17 Calculating Scalar Potentials

### 17.1 The First Method

To find a scalar potential we make use of the path independence and we take a simple path, a straight line, from the origin to some arbitrary point  $\mathbf{r}$ . This path can be parametrised by  $\mathbf{r}(\lambda) = \lambda \mathbf{r}$  where  $\lambda \in [0, 1]$ . This gives us  $x_i(\lambda) = \lambda x_i$ . Then we define  $\varphi$  such that

$$\varphi(\mathbf{r}) - \varphi(0) = \int_0^1 \mathbf{a} \cdot d\mathbf{r}' = \int_0^1 \mathbf{a}(\lambda \mathbf{r}) \cdot d(\lambda \mathbf{r}) = \int_0^1 \mathbf{a}(\lambda \mathbf{r}) \cdot \mathbf{r} d\lambda$$

It is common to choose to set  $\varphi(0) = 0$  although sometimes this can lead to problems if we have factors of  $1/r$  in which case it may be preferable to set  $\varphi(\infty) = 0$  so that factors of  $1/r$  vanish.

#### Example 17.1

Find the scalar potential for the vector field  $\mathbf{a} = (2xy + z^3)\mathbf{e}_1 + x^2\mathbf{e}_2 + 3xz^2\mathbf{e}_3$ . First we check that  $\nabla \times \mathbf{a} = 0$ :

$$\nabla \times \mathbf{a} = (\partial_y 3xz^2 - \partial_z x^2)\mathbf{e}_1 + (\partial_z (2xy + z^3) - \partial_x 3xz^2)\mathbf{e}_2 + (\partial_x x^2 - \partial_y (2xy + z^3))\mathbf{e}_3$$



$$\begin{aligned}
&= (3z^2 - 3z^2)\mathbf{e}_2 + (2x - 2x)\mathbf{e}_3 \\
&= 0
\end{aligned}$$

Hence we can find  $\varphi$  by

$$\varphi(\mathbf{r}) = \int_0^1 \mathbf{a}(\lambda\mathbf{r}) \cdot \mathbf{r} \, d\lambda$$

$$\begin{aligned}
\mathbf{a}(\lambda\mathbf{r}) \cdot \mathbf{r} &= [(2xy\lambda^2 + z^3\lambda^3)\mathbf{e}_1 + x^2\lambda^2\mathbf{e}_2 + 3xz^2\lambda^3\mathbf{e}_3] \cdot [x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3] \\
&= 2x^2y\lambda^2 + xz^3\lambda^3 + x^2y\lambda^2 + 3xz^3\lambda^3 \\
&= 3x^2y\lambda^2 + 4xz^3\lambda^3
\end{aligned}$$

Note that each term  $x$ ,  $y$  or  $z$  contributes a factor of  $\lambda$  to the  $\mathbf{a}(\lambda\mathbf{r})$  term. Finally we can find  $\varphi$ :

$$\varphi(\mathbf{r}) = \int_0^1 3x^2y\lambda^2 + 4xz^3\lambda^3 \, d\lambda = x^2y + z^3x$$

We can check that this is the correct potential by calculating  $\nabla\varphi$ :

$$\begin{aligned}
\nabla\varphi &= \partial_x(x^2y + z^3x)\mathbf{e}_1 + \partial_y(x^2y + z^3x)\mathbf{e}_2 + \partial_z(x^2y + z^3x)\mathbf{e}_3 \\
&= (2y + z^3)\mathbf{e}_1 + x^2\mathbf{e}_2 + 3xz^2\mathbf{e}_3 \\
&= \mathbf{a}
\end{aligned}$$

so we have found the potential.

### Example 17.2

Find the potential for the vector field  $\mathbf{a} = 2(\mathbf{c} \cdot \mathbf{r})\mathbf{r} + r^2\mathbf{c}$  for a constant vector  $\mathbf{c}$ . First we check that the field is irrotational:

$$\begin{aligned}
\nabla \times \mathbf{a} &= 2\nabla \times [(\mathbf{c} \cdot \mathbf{r})\mathbf{r}] + \nabla \times [r^2\mathbf{c}] \\
&= 2[\underbrace{\nabla(\mathbf{c} \cdot \mathbf{r})}_{=\mathbf{c}} \times \mathbf{r} + (\mathbf{c} \cdot \mathbf{r}) \underbrace{\nabla \times \mathbf{r}}_{=0}] + \underbrace{\nabla r^2}_{=2\mathbf{r}} \times \mathbf{c} + r^2 \underbrace{\nabla \times \mathbf{c}}_{=0} \\
&= 2\mathbf{c} \times \mathbf{r} + 2\mathbf{r} \times \mathbf{c} \\
&= 2\mathbf{c} \times \mathbf{r} - 2\mathbf{c} \times \mathbf{r} \\
&= 0
\end{aligned}$$

so we can find a scalar potential  $\varphi$ .

$$\begin{aligned}
\varphi(\mathbf{r}) &= \int_0^1 \mathbf{a}(\lambda\mathbf{r}) \cdot \mathbf{r} \, d\lambda \\
&= \mathbf{a}(\mathbf{r}) \cdot \mathbf{r} \int_0^1 \lambda^2 \, d\lambda
\end{aligned}$$

Both terms in  $\mathbf{a}$  have a factor of  $r^2$  which contribute  $\lambda^2$  to value of  $\mathbf{a}(\lambda\mathbf{r})$

$$\begin{aligned}
&= [2(\mathbf{c} \cdot \mathbf{r})(\mathbf{r} \cdot \mathbf{r}) + r^2(\mathbf{c} \cdot \mathbf{r})] \left[ \frac{\lambda}{3} \right]_0^1 \\
&= 3r^2(\mathbf{c} \cdot \mathbf{r}) \frac{1}{3} \\
&= r^2(\mathbf{c} \cdot \mathbf{r})
\end{aligned}$$

We can check this is correct by calculating  $\nabla\varphi$ :

$$\nabla\varphi = \nabla[r^2(\mathbf{c} \cdot \mathbf{r})]$$

$$\begin{aligned}
&= (\nabla r^2)(\mathbf{c} \cdot \mathbf{r}) + r^2 \nabla(\mathbf{c} \cdot \mathbf{r}) \\
&= 2\mathbf{r}(\mathbf{c} \cdot \mathbf{r}) + r^2 \mathbf{c} \\
&= \mathbf{a}
\end{aligned}$$

so we found the correct potential.

## 17.2 The Second Method

We can consider each component of  $\nabla\varphi = \mathbf{a}$  as a PDE which gives  $\partial_i\varphi = a_i$ . This gives us

$$\begin{aligned}
\varphi(x, y, z) &= \int_x a_x(x', y, z) dx' + f(y, z) \\
&= \int_y a_y(x, y', z) dy' + g(x, z) \\
&= \int_z a_z(x, y, z') dz' + h(x, y)
\end{aligned}$$

since two of the variables are unbound in each case the “constant” of integration is actually a function of the unbound variables. We can evaluate these integrals to find a series of simultaneous equations for  $f$ ,  $g$  and  $h$ .

### Example 17.3

Consider again the case of  $\mathbf{a} = (2xy + z^3)\mathbf{e}_1 + x^2\mathbf{e}_2 + 3xz^2\mathbf{e}_3$ . Finding all three integrals in turn gives us

$$\begin{aligned}
\varphi &= x^2y + xz^3 + f(y, z) \\
&= x^2y + g(x, z) \\
&= xz^3 + h(x, y)
\end{aligned}$$

By inspection we can see that  $f = 0$ ,  $g = xz^3$  and  $h = x^2y$  is a solution. Hence  $\varphi = x^2y + xz^3$  is a potential which is the same as the answer we got before.

## 17.3 The Third Method

It is also sometimes possible to find the potential by inspection. **Example 17.4**

Consider the case of  $\mathbf{a} = (\mathbf{c} \cdot \mathbf{r})\mathbf{c}$  for some constant vector  $\mathbf{c}$ . We can write  $\mathbf{c} = \nabla(\mathbf{c} \cdot \mathbf{r})$  so we can write

$$\mathbf{a} = (\mathbf{c} \cdot \mathbf{r})\nabla(\mathbf{c} \cdot \mathbf{r})$$

We can recognise this as part of the product rule for the gradient

$$\nabla[(\mathbf{c} \cdot \mathbf{r})^2] = 2(\mathbf{c} \cdot \mathbf{r})\nabla(\mathbf{c} \cdot \mathbf{r})$$

Hence we can write  $\mathbf{a}$  as

$$\mathbf{a} = \frac{1}{2}\nabla[(\mathbf{c} \cdot \mathbf{r})^2] = \nabla\left[\frac{1}{2}(\mathbf{c} \cdot \mathbf{r})^2\right]$$

Noting that we now have something of the form  $\mathbf{a} = \nabla\varphi$  we can see that

$$\varphi = \frac{1}{2}(\mathbf{c} \cdot \mathbf{r})^2$$

### Example 17.5

Consider again the case of  $\mathbf{a} = 2(\mathbf{c} \cdot \mathbf{r})\mathbf{r} + r^2\mathbf{c}$ . We can identify the factor of  $2\mathbf{r}$  in the first term as  $\nabla r^2$  and the factor of  $\mathbf{c}$  in the second term as  $\nabla(\mathbf{c} \cdot \mathbf{r})$  and we can write  $\mathbf{a}$  as

$$\mathbf{a} = (\mathbf{c} \cdot \mathbf{r})\nabla r^2 + r^2\nabla(\mathbf{c} \cdot \mathbf{r})$$

Again we recognise the product rule allows us to write this as

$$\mathbf{a} = \nabla[(\mathbf{c} \cdot \mathbf{r})r^2]$$

and since  $\mathbf{a} = \nabla\varphi$  we identify

$$\varphi = (\mathbf{c} \cdot \mathbf{r})r^2$$

which is the same result as we got before.

## 17.4 Conservative Forces and Conservation of Energy

We have shown that for a vector field  $\mathbf{F}$  if and only if  $\nabla \times \mathbf{F} = 0$  then there exists  $\varphi$  such that  $\mathbf{F} = \nabla\varphi$ . If  $\mathbf{F}$  is a force then we choose to define the potential for that force as  $V$  such that  $\mathbf{F} = -\nabla V$ . The negative sign simply means that the force acts to minimise the potential which makes more sense with our physical intuition of a potential. For a force the potential has units of energy and is referred to as the potential energy. Like with other potentials it is only unique up to a constant term and as such we can choose where to define  $V(\mathbf{r}) = 0$ . Newton's second law

$$m\ddot{\mathbf{r}} = \mathbf{F}$$

can be written in terms of a potential:

$$m\ddot{\mathbf{r}} = -\nabla V$$

$$m\ddot{\mathbf{r}} \cdot d\mathbf{r} = \mathbf{F} \cdot d\mathbf{r} = -\nabla V \cdot d\mathbf{r}$$

This can be integrated between two points  $\mathbf{r}_A$  and  $\mathbf{r}_B$  and we get

$$m \int_{\mathbf{r}_B}^{\mathbf{r}_A} \ddot{\mathbf{r}} \cdot d\mathbf{r} = - \int_{\mathbf{r}_A}^{\mathbf{r}_B} \nabla V \cdot d\mathbf{r} \quad (17.1)$$

We will call the integral on the left hand side  $I_1$  and the integral on the right hand side  $I_2$ .

$$I_1 = \int_{\mathbf{r}_A}^{\mathbf{r}_B} \ddot{\mathbf{r}} \cdot d\mathbf{r} = \int_{t_A}^{t_B} \ddot{\mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} dt = \frac{1}{2} \int_{t_A}^{t_B} \frac{d}{dt}(\dot{\mathbf{r}}^2) dt$$

where we have used the fact that  $\dot{\mathbf{r}}^2 = \dot{\mathbf{r}} \cdot \dot{\mathbf{r}}$  and

$$\frac{d}{dt}(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}) = \dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} + \ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} = 2\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}}$$

Hence we find  $I_1$  is

$$I_1 = \frac{1}{2} \frac{d}{dt}(\dot{\mathbf{r}}^2) dt = \frac{1}{2} [\dot{\mathbf{r}}^2]_{t_A}^{t_B} = \frac{1}{2} (v_B^2 - v_A^2)$$

where  $v_A$  and  $v_B$  are the velocities at times  $t_A$  and  $t_B$  respectively. This means that the left hand side of equation 17.1 is

$$\frac{1}{2} m (v_B^2 - v_A^2) = \Delta T$$

Now we can evaluate  $I_2$ :

$$I_2 = - \int_{\mathbf{r}_A}^{\mathbf{r}_B} \nabla V \cdot d\mathbf{r} = - \int_{\mathbf{r}_A}^{\mathbf{r}_B} dV = V_A - V_B = -\Delta V$$

where we have used  $dV(\mathbf{r}) = \nabla \cdot d\mathbf{r}$ , and  $V_A$  and  $V_B$  are the potentials at points  $\mathbf{r}_A$  and  $\mathbf{r}_B$  respectively. Putting this back into equation 17.1 we get

$$\frac{1}{2} m v_B^2 - \frac{1}{2} m v_A^2 = V_A - V_B$$

$$\frac{1}{2} m v_A^2 + V_A = \frac{1}{2} m v_B^2 + V_B$$

so the total energy  $E = T + V$  is the same at both points  $\mathbf{r}_A$  and  $\mathbf{r}_B$  so energy is conserved for a conservative force.

## 17.5 Gravitation and Electrostatics

Both the force due to gravity and the Coulomb force can be described by

$$\mathbf{F}(\mathbf{r}) = k \frac{\mathbf{r}}{r^3}$$

For gravity  $k = -GMm$  for two masses  $M$  and  $m$ , where we assume  $M \gg m$ . For Coulomb's force  $k = Qq/(4\pi\epsilon_0)$  for two charges  $Q$  and  $q$ . Both can also be described by a vector field, for gravity the gravitational field  $\mathbf{G}$  is such that a test mass  $m$  experiences a force  $\mathbf{F} = m\mathbf{G}$ . For the Coulomb force the electric field  $\mathbf{E}$  is such that a positive test charge experiences a force  $\mathbf{F} = q\mathbf{E}$ . From the initial equation we see that

$$\mathbf{G} = -GM \frac{\mathbf{r}}{r^3}, \quad \mathbf{E} = \frac{Q}{4\pi\epsilon_0} \frac{\mathbf{r}}{r^3}$$

We can show that both of these are conservative fields the same way by showing that their curl is zero. We will do it here for the gravitational field:

$$\begin{aligned} \nabla \times \mathbf{G} &= -GM \nabla \times \frac{\mathbf{r}}{r^3} \\ &= -GM [r^{-3} \nabla \times \mathbf{r} + \nabla r^{-3} \times \mathbf{r}] \\ &= -GM [-3r^{-3} \mathbf{r} \times \mathbf{r}] \\ &= 0 \end{aligned}$$

We can do the same with the electric field or just note that if we replace  $-GM$  with  $Q/(4\pi\epsilon_0)$  we get the electric field so it must also have zero curl since it is proportional to the gravitational field.

We want to find a potential for both fields. Again we will do it here for the gravitational field. We want to find  $\varphi$  such that  $\mathbf{G} = -\nabla\varphi$ .

$$\varphi(\mathbf{r}) = - \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{G}(\mathbf{r}') \cdot d\mathbf{r}'$$

$\mathbf{r}(\lambda) = \lambda\mathbf{r}$  hence  $\lambda_0\mathbf{r} = \mathbf{r}_0$ .

$$\varphi(\mathbf{r}) = - \int_{\lambda_0}^1 \mathbf{G}(\lambda\mathbf{r}) \cdot \mathbf{r} d\lambda = GM \frac{\mathbf{r} \cdot \mathbf{r}}{r^3} \int_{\lambda_0}^1 \frac{1}{\lambda^2} d\lambda = \frac{GM}{r} \left( -1 + \frac{1}{\lambda_0} \right)$$

Where we have used the fact that  $\mathbf{G}(\mathbf{r}) \propto r^{-2}\hat{\mathbf{r}}$  which leaves us with a factor of  $1/\lambda^2$  when we replace  $\mathbf{r}$  with  $\lambda\mathbf{r}$ . If we chose  $\lambda_0 = 0$  then we would have  $\varphi$  be infinite at the origin. Instead we choose to let  $\lambda_0 \rightarrow \infty$  so that

$$\varphi = -\frac{GM}{r}$$

This keeps  $\varphi$  finite and also means that all potentials are negative which represents the fact that gravity is always attractive. We can replace  $-GM$  with  $Q/(4\pi\epsilon_0)$  again to get the potential for the electric field.

## 17.6 Poisson and Laplace Equations

The Poisson and Laplace equations are common second order PDEs which occur a lot in electrostatics. Two of Maxwell's laws are Gauss's law and Faraday's law, in differential form these are

$$\nabla \cdot \mathbf{E} = \frac{\rho(\mathbf{r})}{\epsilon_0}, \quad \nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0$$

In electrostatics  $\partial_t \mathbf{B} = 0$  so  $\nabla \times \mathbf{E} = 0$ . This means that we can find a potential  $\varphi$  such that  $\mathbf{E} = -\nabla\varphi$ . If we do this then Gauss's law becomes

$$\nabla \cdot \mathbf{E} = -\nabla \cdot \nabla\varphi = -\nabla^2\varphi = \frac{\rho(\mathbf{r})}{\epsilon_0}$$

More generally

$$\nabla^2\varphi = f(\mathbf{r})$$

This is Poisson's equation. The special case where  $f(\mathbf{r}) = 0$  is Laplace's equation:

$$\nabla^2\varphi = 0$$

## 18 Vector Potentials

### 18.1 Upcoming Result

Recall that for a vector field  $\mathbf{a}$  if and only if  $\nabla \times \mathbf{a} = 0$  then  $\exists \varphi$  such that  $\mathbf{a} = -\nabla \varphi$  and that we can find  $\varphi$  by evaluating

$$\varphi = \int_0^1 \mathbf{a}(\lambda \mathbf{r}) \cdot \mathbf{r} \, d\lambda$$

It can be shown that similarly for a vector field  $\mathbf{B}$  if and only if  $\nabla \cdot \mathbf{B} = 0$  then  $\exists \mathbf{A}$  such that  $\mathbf{B} = \nabla \times \mathbf{A}$  and  $\mathbf{A}$  can be found by evaluating

$$\mathbf{A} = \int_0^1 \mathbf{B}(\lambda \mathbf{r}) \times \mathbf{r} \lambda \, d\lambda$$

Note the extra factor of  $\lambda$  as opposed to the equation for  $\varphi$ . We have assumed path independence here which can be shown to hold in a similar way to how we showed it for  $\int \mathbf{a} \cdot d\mathbf{r}$  with  $\nabla \times \mathbf{a} = 0$ .

### 18.2 Helmholtz's Theorem

Helmholtz's theorem is that for any twice continuously differentiable vector field  $\mathbf{V}$  there is (non-unique) decomposition into a curl free and a divergence free component:

$$\mathbf{V} = -\nabla \varphi + \nabla \times \mathbf{A}$$

where  $\varphi$  is a scalar potential and  $\mathbf{A}$  is a vector potential. It is easy to show that if this is true  $\nabla \varphi$  is curl free since  $\nabla \times \nabla \varphi = 0$  and  $\nabla \times \mathbf{A}$  is divergence free since  $\nabla \cdot \nabla \times \mathbf{A} = 0$ . Helmholtz's theorem is a three dimensional case of the more general Hodge decomposition which is valid in any dimension on any surface.

#### Example 18.1

Consider the vector field  $\mathbf{B} = \mathbf{c} \times \mathbf{r}$  for constant vector  $\mathbf{c}$ . First we need to show that  $\nabla \cdot \mathbf{B} = 0$ :

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\mathbf{c} \times \mathbf{r}) = \partial_{x_i} \varepsilon_{ijk} c_j x_k = \delta_{ik} \varepsilon_{ijk} c_j = 0$$

since the product of a symmetric and antisymmetric tensor is zero. We can now attempt to find  $\mathbf{A}$  such that  $\mathbf{B} = \nabla \times \mathbf{A}$ :

$$\begin{aligned} \mathbf{a} &= -\mathbf{r} \times \int_0^1 \mathbf{B}(\lambda \mathbf{r}) \lambda \, d\lambda \\ &= -\mathbf{r} \times \int_0^1 (\mathbf{c} \times \lambda \mathbf{r}) \lambda \, d\lambda \\ &= -\mathbf{r} \times (\mathbf{c} \times \mathbf{r}) \int_0^1 \lambda^2 \, d\lambda \\ &= \frac{1}{3} [(\mathbf{c} \cdot \mathbf{r}) \mathbf{r} - r^2 \mathbf{c}] \end{aligned}$$

where we have used the vector triple product in the last step. We can now show that this is the correct potential:

$$\begin{aligned} \nabla \times \mathbf{A} &= \frac{1}{3} \nabla \times [(\mathbf{c} \cdot \mathbf{r}) \mathbf{r} - r^2 \mathbf{c}] \\ &= \frac{1}{3} [\underbrace{\nabla(\mathbf{c} \cdot \mathbf{r})}_{=\mathbf{c}} \times \mathbf{r} + (\mathbf{c} \cdot \mathbf{r}) \underbrace{\nabla \times \mathbf{r}}_{=0} - \underbrace{\nabla r^2}_{=2\mathbf{r}} \times \mathbf{c} - r^2 \underbrace{\nabla \times \mathbf{c}}_{=0}] \\ &= \frac{1}{3} [\mathbf{c} \times \mathbf{r} - 2\mathbf{r} \times \mathbf{c}] \\ &= \frac{1}{3} [\mathbf{c} \times \mathbf{r} + 2\mathbf{c} \times \mathbf{r}] \end{aligned}$$

$$\begin{aligned}
&= \mathbf{c} \times \mathbf{r} \\
&= \mathbf{B}
\end{aligned}$$

We can show that our formula for  $\mathbf{A}$  works by showing that

$$\nabla \times \mathbf{A} = \nabla \times \int_0^1 \mathbf{B}(\lambda \mathbf{r}) \times \mathbf{r} \lambda \, d\lambda = \mathbf{B}(\mathbf{r})$$

The first step in showing this is to notice that

$$\nabla \times \int_0^1 \mathbf{B}(\lambda \mathbf{r}) \times \mathbf{r} \lambda \, d\lambda = \int_0^1 f(\lambda, \mathbf{r}) \lambda \, d\lambda$$

for some function  $f$  given by

$$\begin{aligned}
f(\lambda, \mathbf{r}) &= \nabla \times [\mathbf{B}(\lambda \mathbf{r}) \times \mathbf{r}] \\
&= \underbrace{(\nabla \cdot \mathbf{r}) \mathbf{B}(\lambda \mathbf{r})}_{=3} + (\mathbf{r} \cdot \nabla) \mathbf{B}(\lambda \mathbf{r}) - \underbrace{(\nabla \cdot \mathbf{B}(\lambda \mathbf{r})) \mathbf{r}}_{=0} - \underbrace{(\mathbf{B}(\lambda \mathbf{r}) \cdot \nabla) \mathbf{r}}_{=\mathbf{B}(\lambda \mathbf{r})} \\
&= 3\mathbf{B}(\lambda \mathbf{r}) + (\mathbf{r} \cdot \nabla) \mathbf{B}(\lambda \mathbf{r}) - \mathbf{B}(\lambda \mathbf{r}) \\
&= 2\mathbf{B}(\lambda \mathbf{r}) + (\mathbf{r} \cdot \nabla) \mathbf{B}(\lambda \mathbf{r}) \\
&= 2\mathbf{B}(\lambda \mathbf{r}) + x_i \partial_{x_i} \mathbf{B}(\lambda x_1, \lambda x_2, \lambda x_3) \\
&= 2\mathbf{B}(\lambda \mathbf{r}) + \lambda \partial_\lambda \mathbf{B}(\lambda x_1, \lambda x_2, \lambda x_3)
\end{aligned}$$

where we have used the fact that every  $x_i$  appears with a  $\lambda$  to replace  $x_i \partial_{x_i}$  with  $\lambda \partial_\lambda$ .

$$\nabla \times \mathbf{A} = \int_0^1 \underbrace{2\mathbf{B}(\lambda \mathbf{r}) \lambda}_{=I_1} + \underbrace{\lambda^2 \mathbf{B}(\lambda x_1, \lambda x_2, \lambda x_3)}_{=I_2} \, d\lambda$$

$$I_2 = \int_0^1 \lambda^2 \partial_\lambda \mathbf{B}(\lambda \mathbf{r}) \, d\lambda$$

$$\begin{aligned}
u &= \lambda^2 & v &= \mathbf{B}(\lambda \mathbf{r}) \\
u' &= 2\lambda & v' &= \partial_\lambda \mathbf{B}(\lambda \mathbf{r})
\end{aligned}$$

$$\begin{aligned}
I_2 &= [\lambda^2 \mathbf{B}(\lambda \mathbf{r})]_0^1 - \int_0^1 2\lambda \mathbf{B}(\lambda \mathbf{r}) \, d\lambda \\
&= \mathbf{B}(\mathbf{r}) - I_1
\end{aligned}$$

$$\nabla \times \mathbf{A} = I_1 + I_2 = I_1 + \mathbf{B}(\mathbf{r}) - I_1 = \mathbf{B}(\mathbf{r})$$

which is what we wanted to show.

### 18.3 Gauge Invariance

We can always add the gradient of an arbitrary scalar field  $f$  to vector potential transforming from  $\mathbf{A}$  to  $\mathbf{A}' = \mathbf{A} + \nabla f$  since  $\nabla \times \mathbf{A}' = \nabla \times \mathbf{A} + \nabla \times \nabla f = \nabla \times \mathbf{A}$ . This is called gauge invariance and is one of the most important symmetries in physics.

### 18.4 Physics Examples

Magnetic field lines do not have sources or sinks. This is shown in Maxwell's equations by  $\nabla \cdot \mathbf{B} = 0$ . This means that we can write  $\mathbf{B} = \nabla \times \mathbf{A}$  where  $\mathbf{A}$  is the magnetic vector potential.

For incompressible fluids the mass density,  $\rho$ , is constant so if there are no sources or sinks then it can be shown that the velocity field  $\mathbf{v}(\mathbf{r})$  satisfies  $\nabla \cdot \mathbf{v} = 0$ . In this case there exists a velocity potential  $\Psi(\mathbf{r})$  such that  $\mathbf{v} = \nabla \times \Psi$ .

## Part III

## Orthogonal Curvilinear Coordinates

## 19 OCC line, area and volume elements, grad and divergence

## 19.1 What Are Orthogonal Curvilinear Coordinates?

Orthogonal curvilinear coordinates (OCC) are any set of coordinates  $\{u_i\}$  where all unit vectors are orthogonal to each other. That is  $\mathbf{e}_{u_i} \cdot \mathbf{e}_{u_j} = 0$  for  $i \neq j$ . The curvilinear part means that the coordinate axis may be curved. In general this means that there is a transformation to Cartesian coordinates  $\{x_i\}$  where  $\{u_i\}$  are functions of  $\{x_i\}$ , that is  $u_i = u_i(x_1, x_2, x_3)$ . If  $\{u_i\}$  are single valued functions of variables  $\{x_i\}$  then there is an inverse transformation  $x_i = x_i(u_1, u_2, u_3)$ , except possibly at certain points. A point can be equally well specified in Cartesian coordinates or curvilinear coordinates.

For orthogonal coordinates all surfaces where one coordinate is held constant meet at right angles. The tangent vectors to these surfaces are  $\{\partial_{u_i} \mathbf{r}\}$ . Since these are orthogonal  $\partial_{u_i} \mathbf{r} \cdot \partial_{u_j} \mathbf{r} = 0$  when  $i \neq j$ . We define the scale factor  $h_i = |\partial_{u_i} \mathbf{r}|$ , this gives the length  $h_i du_i$  of  $d\mathbf{r}$  when we change  $u_i$  to  $u_i + du_i$ . We can define the unit vectors for the coordinate system as

$$\mathbf{e}_{u_i} = \frac{1}{h_i} \partial_{u_i} \mathbf{r}$$

and since these are orthogonal and have norm 1 we get  $\mathbf{e}_{u_i} \cdot \mathbf{e}_{u_j} = \delta_{ij}$ . In non-orthogonal curvilinear coordinates (aka skew coordinates)  $\mathbf{e}_{u_i} \cdot \mathbf{e}_{u_j} \neq \delta_{ij}$  but it is possible to define a dual basis  $\{\tilde{\mathbf{e}}_{u_i}\}$  such that  $\tilde{\mathbf{e}}_{u_i} \cdot \mathbf{e}_{u_j} = \delta_{ij}$ .

**Example 19.1**

In Cartesian coordinates:

$$\begin{aligned} \mathbf{r} &= x_i \mathbf{e}_i \\ \partial_{x_i} \mathbf{r} &= \mathbf{e}_i \implies h_i = 1 \end{aligned}$$

**Example 19.2**

In spherical polar coordinates:

$$\begin{aligned} \mathbf{r} &= r(\sin \vartheta \cos \varphi \mathbf{e}_1 + \sin \vartheta \sin \varphi \mathbf{e}_2 + \cos \vartheta \mathbf{e}_3) \\ \partial_r \mathbf{r} &= \sin \vartheta \cos \varphi \mathbf{e}_1 + \sin \vartheta \sin \varphi \mathbf{e}_2 + \cos \vartheta \mathbf{e}_3 \implies h_r = 1 \\ \partial_\vartheta \mathbf{r} &= r(\cos \vartheta \cos \varphi \mathbf{e}_1 + \cos \vartheta \sin \varphi \mathbf{e}_2 - \sin \vartheta \mathbf{e}_3) = r \mathbf{e}_\vartheta \implies h_\vartheta = r \\ \partial_\varphi \mathbf{r} &= r \sin \vartheta (-\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2) = r \sin \vartheta \mathbf{e}_\varphi \implies h_\varphi = r \sin \vartheta \end{aligned}$$

**Example 19.3**

In cylindrical coordinates:

$$\begin{aligned} \mathbf{r} &= \rho \cos \varphi \mathbf{e}_1 + \rho \sin \varphi \mathbf{e}_2 + z \mathbf{e}_3 \\ \partial_\rho \mathbf{r} &= \cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2 \implies h_\rho = 1 \\ \partial_\varphi \mathbf{r} &= -\rho \sin \varphi \mathbf{e}_1 + \rho \cos \varphi \mathbf{e}_2 \implies h_\varphi = \rho \\ \partial_z \mathbf{r} &= \mathbf{e}_3 \implies h_z = 1 \end{aligned}$$

## 19.2 Length, Area and Volume Elements in OCC

## 19.2.1 Length

A vector length element is given by  $d\mathbf{r} = (\partial_{u_i} \mathbf{r}) du_i = h_i \mathbf{e}_{u_i} du_i$ . A scalar length element  $ds$  is then given by

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = h_i \mathbf{e}_{u_i} du_i \cdot h_j \mathbf{e}_{u_j} du_j = h_i du_i h_j du_j \delta_{ij} = h_i^2 du_i^2 = h_1^2 du_1^2 + h_2^2 du_2^2 + h_3^2 du_3^2$$

### 19.2.2 Area

The area element of an area normal to  $\mathbf{e}_{\mathbf{u}_i}$  is

$$d\mathbf{S}_{\mathbf{u}_i} = h_j \mathbf{e}_{\mathbf{u}_j} du_j \times h_k \mathbf{e}_{\mathbf{u}_k} du_k = h_j h_k du_j du_k \mathbf{e}_{\mathbf{u}_i}$$

For example in spherical polar coordinates  $d\mathbf{S}_{\mathbf{r}}$  is given by

$$d\mathbf{S}_{\mathbf{r}} = h_\vartheta d\vartheta \mathbf{e}_\vartheta \times h_\varphi d\varphi \mathbf{e}_\varphi = h_\vartheta h_\varphi d\vartheta d\varphi \mathbf{e}_{\mathbf{r}} = r^2 \sin \vartheta d\vartheta d\varphi \mathbf{e}_{\mathbf{r}}$$

### 19.2.3 Volume

The volume element  $dV$  is defined as

$$dV = |(h_1 \mathbf{e}_{\mathbf{u}_1} du_1, h_2 \mathbf{e}_{\mathbf{u}_2} du_2, h_3 \mathbf{e}_{\mathbf{u}_3} du_3)| = h_1 h_2 h_3 du_1 du_2 du_3$$

since in orthogonal coordinates the parallelepiped with volume  $dV$  is actually a cuboid so its volume is just the product of the side lengths. For example in spherical polar coordinates  $dV$  is given by

$$dV = h_r h_\vartheta h_\varphi dr d\vartheta d\varphi = r^2 \sin \vartheta dr d\vartheta d\varphi$$

## 19.3 Components of a Vector Field

A vector field  $\mathbf{a}(\mathbf{r})$  can be expressed in curvilinear coordinates as  $\mathbf{a}(\mathbf{r}) = a_i(u_1, u_2, u_3) \mathbf{e}_{\mathbf{u}_i}$ . In orthogonal curvilinear coordinates  $a_i$  is given by  $a_i = \mathbf{e}_{\mathbf{u}_i} \cdot \mathbf{a}$ . In non-orthogonal curvilinear coordinates  $a_i$  is instead given by  $a_i = \tilde{\mathbf{e}}_{\mathbf{u}_i} \cdot \mathbf{a}$ .

## 19.4 Gradient in OCC

Recall that we defined the gradient in terms of the change,  $df$ , of a scalar field  $f$  when we let  $\mathbf{r} \rightarrow \mathbf{r} + d\mathbf{r}$ :

$$df(\mathbf{r}) = \nabla f(\mathbf{r}) \cdot d\mathbf{r}$$

We can also find  $df$  as

$$df = (\partial_{u_i} f) du_i$$

We can use the orthogonality of our coordinate system to separate this into

$$df = \nabla f \cdot (h_i \mathbf{e}_i du_i)$$

Equating the two we get that

$$\nabla f = \mathbf{e}_{\mathbf{u}_i} \frac{1}{h_i} \partial_{u_i} f \implies \nabla = \mathbf{e}_{\mathbf{u}_i} \frac{1}{h_i} \partial_{u_i}$$

For example in spherical polar coordinates:

$$\nabla f = (\partial_r f) \mathbf{e}_{\mathbf{r}} + \frac{1}{r} (\partial_\vartheta f) \mathbf{e}_\vartheta + \frac{1}{r \sin \vartheta} (\partial_\varphi f) \mathbf{e}_\varphi$$

## 19.5 Divergence in OCC

For a vector field  $\mathbf{a}(\mathbf{r}) = a_i(u_1, u_2, u_3) \mathbf{e}_{\mathbf{u}_i}$  the divergence is

$$\nabla \cdot \mathbf{a} = \operatorname{div} \mathbf{a} = \lim_{\delta V \rightarrow 0} \frac{1}{\delta V} \underbrace{\int_{\delta S} \mathbf{a} \cdot d\mathbf{S}}_{=I}$$

Since this can be shown to be independent of the shape we choose the shape given by starting at a point  $(u_1, u_2, u_3)$  and then increasing each coordinate individually by  $\delta u_i$  to give something that is approximately



a cuboid (but with potentially curved edges). We will consider only two faces opposite each other and the other faces follow by symmetry. We consider the faces normal to  $\mathbf{e}_{u_1}$ . A surface element is then given by

$$d\mathbf{S} = \pm h_2 h_3 du_3 du_2 \mathbf{e}_{u_1}$$

The integral  $I$  is actually the sum of three integrals  $I_i$  over each opposite pair of faces. So for the case we are considering we need to find  $I_1$  which is

$$I_1 = \int_{u_3}^{u_3+\delta u_3} du_3 \int_{u_2}^{u_2+\delta u_2} du_2 [a_1 h_2 h_3|_{(u_1+\delta u_1, u_2, u_3)} - a_1 h_2 h_3|_{(u_1, u_2, u_3)}]$$

We Taylor expand this about  $(u_1, u_2, u_3)$  and we get

$$I_1 = \int_{u_3}^{u_3+\delta u_3} du_3 \int_{u_2}^{u_2+\delta u_2} du_2 [\delta u_1 \partial_{u_1} (a_1 h_2 h_3)|_{(u_1, u_2, u_3)} + \mathcal{O}(\delta u_1^3)]$$

When  $\delta u_1, \delta u_2 \approx 0$  we get

$$I_1 = \delta u_1 \delta u_2 \delta u_3 \partial_{u_1} |_{(u_1, u_2, u_3)}$$

Then using  $\delta V = h_1 h_2 h_3 \delta u_1 \delta u_2 \delta u_3$  and adding in the integrals over the other surfaces we get

$$\nabla \cdot \mathbf{a} = \frac{1}{h_1 h_2 h_3} [\partial_{u_1} (a_1 h_2 h_3) + \partial_{u_2} (a_2 h_1 h_3) + \partial_{u_3} (a_3 h_1 h_2)]$$

For example in spherical polar coordinates we have  $h_1 h_2 h_3 = r^2 \sin \vartheta$  and we get

$$\begin{aligned} \nabla \cdot \mathbf{a} &= \frac{1}{r^2 \sin \vartheta} [\partial_r (a_r h_\vartheta h_\varphi) + \partial_\vartheta (a_\vartheta h_r h_\varphi) + \partial_\varphi (a_\varphi h_r h_\vartheta)] \\ &= \frac{1}{r^2 \sin \vartheta} [\partial_r (a_r r^2 \sin \vartheta) + \partial_\vartheta (a_\vartheta r \sin \vartheta) + \partial_\varphi (a_\varphi r)] \\ &= \frac{1}{r^2} \partial_r (a_r r^2) + \frac{1}{r \sin \vartheta} [\partial_\vartheta (a_\vartheta \sin \vartheta) + \partial_\varphi a_\varphi] \end{aligned}$$

## 20 Curl and the Laplacian in OCC

### 20.1 Curl in OCC

We use the integral definition of curl:

$$\hat{\mathbf{n}} \cdot (\nabla \times \mathbf{a}) = \hat{\mathbf{n}} \cdot \text{curl}(\mathbf{a}) = \lim_{\delta S \rightarrow 0} \frac{1}{\delta S} \oint_{\delta C} \mathbf{a} \cdot d\mathbf{r}$$

In OCC we have  $\frac{d}{dr} = h_i du_i \mathbf{e}_{u_i}$ . We choose a rectangular surface in the  $u_2$ - $u_3$ . The surface has diagonally opposite corners  $(u_1, u_2, u_3)$  and  $(u_1, u_2 + \delta u_2, u_3 + \delta u_3)$ . We split the integral into four parts for each side of the rectangle.

$$\begin{aligned} \oint_{\delta C} \mathbf{a} \cdot d\mathbf{r} &= \int_{u_2}^{u_2+\delta u_2} [a_2 h_2|_{(u_1, u_2, u_3)} - a_2 h_2|_{(u_1, u_2, u_3+\delta u_3)}] du_2 \\ &\quad - \int_{u_3}^{u_3+\delta u_3} [a_3 h_3|_{(u_1, u_2, u_3)} - a_3 h_3|_{(u_1, u_2+\delta u_2, u_3)}] du_3 \end{aligned}$$

We Taylor expand around  $(u_1, u_2, u_3)$  and we get

$$\begin{aligned} \oint_{\delta C} \mathbf{a} \cdot d\mathbf{r} &= \int_{u_2}^{u_2+\delta u_2} [\delta u_3 \partial_{u_3} a_2 h_2|_{(u_1, u_2, u_3)} + \mathcal{O}(\delta u_3^3)] du_2 \\ &\quad - \int_{u_3}^{u_3+\delta u_3} [\delta u_2 \partial_{u_2} a_3 h_3|_{(u_1, u_2, u_3)} + \mathcal{O}(\delta u_2^3)] du_3 \end{aligned}$$

We then use the fact that for  $a \rightarrow b$  and  $x^* \in [a, b]$

$$\int_a^b f(x) dx = (b - a)f(x^*)$$

Now allowing  $\delta u_i \rightarrow 0$  we note that  $b - a$  corresponds to  $u_i + \delta u_i - u_i = \delta u_i$  so our limits approach each other so we can use this approximation

$$\oint_{\delta C} \mathbf{a} \cdot d\mathbf{r} = \delta u_2 \delta u_3 [-\partial_{u_3}(a_2 h_2) + \partial_{u_2}(a_3 h_3)]$$

Using  $\delta S = h_2 h_3 \delta u_2 \delta u_3$  we get

$$\mathbf{e}_{u_1} \cdot (\nabla \times \mathbf{a}) = (\nabla \times \mathbf{a})_1 = \frac{1}{h_2 h_3} [\partial_{u_2}(a_3 h_3) - \partial_{u_3}(a_2 h_2)]$$

We can do the same for the other components of the curl and we get

$$\nabla \times \mathbf{a} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \mathbf{e}_{u_1} & h_2 \mathbf{e}_{u_2} & h_3 \mathbf{e}_{u_3} \\ \partial_{u_1} & \partial_{u_2} & \partial_{u_3} \\ h_1 a_1 & h_2 a_2 & h_3 a_3 \end{vmatrix}$$

### Example 20.1

In spherical polar coordinates this is

$$\begin{aligned} \nabla \times \mathbf{a} &= \frac{1}{r^2 \sin \vartheta} \begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\vartheta & r \sin \vartheta \mathbf{e}_\varphi \\ \partial_r & \partial_\vartheta & \partial_\varphi \\ a_r & r a_\vartheta & r \sin \vartheta a_\varphi \end{vmatrix} \\ &= \frac{1}{r^2 \sin \vartheta} \{ \mathbf{e}_r [\partial_\vartheta (r \sin \vartheta a_\varphi) - \partial_\varphi (r a_\vartheta)] - r\mathbf{e}_\vartheta [\partial_r (r \sin \vartheta a_\varphi) - \partial_\varphi (a_r)] + r \sin \vartheta \mathbf{e}_\varphi [\partial_r (r a_\vartheta) - \partial_\vartheta (a_r)] \} \\ &= \frac{1}{r \sin \vartheta} \mathbf{e}_r [\partial_\vartheta (\sin \vartheta a_\varphi) - \partial_\varphi (a_\vartheta)] + \frac{1}{r} \mathbf{e}_\vartheta \left[ \frac{\partial_\varphi (a_r)}{\sin \vartheta} - \partial_r (r a_\varphi) \right] + \frac{1}{r} \mathbf{e}_\varphi [\partial_r (r a_\vartheta) - \partial_\vartheta (a_r)] \end{aligned}$$

## 20.2 Laplacian in OCC

### 20.2.1 Laplacian of a Scalar Field

The Laplacian can be thought of as the composition of the divergence and the gradient

$$\nabla^2 f = \nabla \cdot \nabla f$$

Using the previously derived formulae for  $\nabla \cdot$  and  $\nabla$  in OCC we get

$$\nabla^2 f = \frac{1}{h_1 h_2 h_3} \left[ \partial_{u_1} \left( \frac{h_2 h_3}{h_1} \partial_{u_1} f \right) + \partial_{u_2} \left( \frac{h_1 h_3}{h_2} \partial_{u_2} f \right) + \partial_{u_3} \left( \frac{h_1 h_2}{h_3} \partial_{u_3} f \right) \right]$$

### Example 20.2

In spherical polar coordinates

$$\begin{aligned} \nabla^2 f &= \frac{1}{r^2 \sin \vartheta} \left[ \partial_r (r^2 \sin \vartheta \partial_r f) + \partial_\vartheta (\sin \vartheta \partial_\vartheta f) + \partial_\varphi \left( \frac{1}{r \sin \vartheta} \partial_\varphi f \right) \right] \\ \nabla^2 f &= \frac{1}{r^2} \partial_r (r^2 \partial_r f) + \frac{1}{r^2 \sin^2 \vartheta} [\sin \vartheta \partial_\vartheta (\sin \vartheta \partial_\vartheta f) + \partial_\varphi^2 f] \end{aligned}$$

## 20.3 Laplacian of a Vector Field

We can't do the same for a vector field as we did for a scalar field as the gradient of a vector field is undefined (we could however define  $\nabla$  in OCC and then dot it with itself). Fortunately we have an equation relating the Laplacian to other known quantities:

$$\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2 \mathbf{a} \implies \nabla^2 \mathbf{a} = \nabla(\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a})$$

This gets very messy so we won't do anything more with it.

## 20.4 Alternative Derivation of Divergence

We can also define  $\nabla$  and dot it with a vector field  $\mathbf{a}$  to find the divergence of  $\mathbf{a}$ . We will do so here for cylindrical coordinates. Using the gradient in OCC we can see that in cylindrical coordinates

$$\nabla = \mathbf{e}_{u_i} \frac{1}{h_i} \partial_{u_i} = \mathbf{e}_\rho \partial_\rho + \mathbf{e}_\varphi \frac{1}{\rho} \partial_\varphi + \mathbf{e}_z \partial_z$$

We need to be careful when doing the dot product as in general  $\partial_{u_i} \mathbf{e}_{u_j} \neq \mathbf{0}$  as is the case with Cartesian coordinates. In cylindrical coordinates both  $\mathbf{e}_\rho$  and  $\mathbf{e}_\varphi$  have  $\varphi$  dependence.

$$\partial_\varphi \mathbf{e}_\rho = \partial_\varphi (\cos \varphi \mathbf{e}_1 + \sin \varphi \mathbf{e}_2) = -\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2 = \mathbf{e}_\varphi$$

$$\partial_\varphi \mathbf{e}_\varphi = \partial_\varphi (-\sin \varphi \mathbf{e}_1 + \cos \varphi \mathbf{e}_2) = -\cos \varphi \mathbf{e}_1 - \sin \varphi \mathbf{e}_2 = -\mathbf{e}_\rho$$

Using this we can compute the dot product

$$\begin{aligned} \nabla \cdot \mathbf{a} &= \left( \mathbf{e}_\rho \partial_\rho + \mathbf{e}_\varphi \frac{1}{\rho} \partial_\varphi + \mathbf{e}_z \partial_z \right) \cdot (a_\rho \mathbf{e}_\rho + a_\varphi \mathbf{e}_\varphi + a_z \mathbf{e}_z) \\ &= (\mathbf{e}_\rho \partial_\rho) \cdot (a_\rho \mathbf{e}_\rho) + \left( \mathbf{e}_\varphi \frac{1}{\rho} \partial_\varphi \right) \cdot (a_\rho \mathbf{e}_\rho) + \left( \mathbf{e}_\varphi \frac{1}{\rho} \partial_\varphi \right) \cdot (a_\varphi \mathbf{e}_\varphi) + (\mathbf{e}_z \partial_z) \cdot (a_z \mathbf{e}_z) \end{aligned}$$

All other pairs of unit vectors are zero when dotted together and this can be done immediately since they commute with the partial derivatives that aren't  $\partial_\varphi$

$$\begin{aligned} &= \partial_\rho a_\rho + \frac{1}{\rho} \mathbf{e}_\varphi \cdot \partial_\varphi (a_\rho \mathbf{e}_\rho) + \frac{1}{\rho} \mathbf{e}_\varphi \cdot \partial_\varphi (a_\varphi \mathbf{e}_\varphi) + \partial_z a_z \\ &= \partial_\rho a_\rho + \frac{1}{\rho} \mathbf{e}_\varphi \cdot [a_\rho \partial_\varphi \mathbf{e}_\rho + \mathbf{e}_\rho \partial_\varphi a_\rho] + \frac{1}{\rho} \mathbf{e}_\varphi \cdot [a_\varphi \partial_\varphi \mathbf{e}_\varphi + \mathbf{e}_\varphi \partial_\varphi a_\varphi] + \partial_z a_z \\ &= \partial_\rho a_\rho + \frac{1}{\rho} \mathbf{e}_\varphi \cdot [a_\rho \mathbf{e}_\varphi + \mathbf{e}_\rho \partial_\varphi a_\rho] + \frac{1}{\rho} \mathbf{e}_\varphi \cdot [-a_\varphi \mathbf{e}_\rho + \mathbf{e}_\varphi \partial_\varphi a_\varphi] + \partial_z a_z \\ &= \partial_\rho a_\rho + \frac{1}{\rho} a_\rho + \frac{1}{\rho} \partial_\varphi a_\varphi + \partial_z a_z \\ &= \frac{1}{\rho} \partial_\rho (\rho a_\rho) + \frac{1}{\rho} \partial_\varphi a_\varphi + \partial_z a_z \end{aligned}$$