Electromagnetism & Relativity

[PHYS10093] Semester 1, 2019/20

Brian Pendleton

- Email: bjp@ph.ed.ac.uk
- Office: JCMB 4413
- Phone: 0131-650-5241
- Web: http://www.ph.ed.ac.uk/~bjp/emr/

January 14, 2020
Books

The course should be self-contained, but it’s always good to read textbooks to broaden your education.

- David J Griffiths,
  Introduction to Electrodynamics, (Prentice Hall)

- John R Reitz, Frederick J Milford, Robert W Christy,
  Foundations of Electromagnetic Theory, (Addison-Wesley)

- JD Jackson,
  Classical Electrodynamics (Wiley) – advanced, good for next year.

- KF Riley, MP Hobson and SJ Bence,

- PC Matthews,

- ML Boas,

- GB Arfken and HJ Weber,

- DE Bourne and PC Kendall,

Griffiths is the main text for electromagnetism; Reitz, Milford & Christy is a standard electromagnetism text; Jackson is pretty advanced, but it will also be good for Classical Electrodynamics next year.

The other books are useful for the first part of the course, which will introduce tensors and their applications in physics.

Acknowledgements: Many thanks to Richard Ball, Martin Evans, Roger Horsley and Donal O’Connell for providing copies of their lecture notes and tutorial sheets, from which I have sampled with (apparent) impunity.
# Contents

1 Revision of vector algebra ......................................................................................... 1
   1.1 Revision of vectors & index/suffix notation ......................................................... 1
      1.1.1 Position vector ............................................................................................. 1
      1.1.2 Cartesian coordinates .................................................................................... 1
      1.1.3 The Kronecker delta symbol $\delta_{ij}$ ............................................................ 2
      1.1.4 The Levi-Civita or epsilon symbol $\epsilon_{ijk}$ ................................................ 2
      1.1.5 Scalar product ............................................................................................... 2
      1.1.6 Vector product ............................................................................................. 3
      1.1.7 Scalar triple product ..................................................................................... 3
      1.1.8 Vector triple product .................................................................................... 3
   1.2 Linear transformation of basis .............................................................................. 4
   1.3 Transformation properties of vectors and scalars ................................................. 4
      1.3.1 Transformation of vector components ............................................................ 4
      1.3.2 Scalar and vector products ......................................................................... 4

2 Cartesian tensors ....................................................................................................... 5
   2.1 Definition and transformation properties ............................................................ 5
      2.1.1 Definition of a tensor ..................................................................................... 5
      2.1.2 Fields ............................................................................................................ 6
      2.1.3 Dyadic notation ........................................................................................... 6
      2.1.4 Internal consistency in the definition of a tensor .......................................... 7
      2.1.5 Properties of Cartesian tensors – tensor algebra ......................................... 7
      2.1.6 The quotient theorem .................................................................................. 9
   2.2 Revision of matrices and determinants ................................................................. 9
      2.2.1 Matrices ....................................................................................................... 9
      2.2.2 Determinants ............................................................................................... 11
      2.2.3 Linear equations ....................................................................................... 13
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.4</td>
<td>Orthogonal matrices</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>Pseudotensors, pseudovectors &amp; pseudoscalars</td>
<td>16</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Vectors (revision)</td>
<td>17</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Pseudovectors</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Some physical examples</td>
<td>19</td>
</tr>
<tr>
<td>2.5</td>
<td>Invariant/Isotropic Tensors</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Rotation, reflection &amp; inversion tensors</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>The rotation tensor</td>
<td>22</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Rotation about an arbitrary axis</td>
<td>22</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Some important properties of the rotation tensor</td>
<td>23</td>
</tr>
<tr>
<td>3.2</td>
<td>Reflections and inversions</td>
<td>24</td>
</tr>
<tr>
<td>3.3</td>
<td>Projection operators</td>
<td>24</td>
</tr>
<tr>
<td>3.4</td>
<td>Active and passive transformations</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>The inertia tensor</td>
<td>26</td>
</tr>
<tr>
<td>4.1</td>
<td>Angular momentum and kinetic energy</td>
<td>26</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Angular momentum</td>
<td>27</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Kinetic energy</td>
<td>27</td>
</tr>
<tr>
<td>4.1.3</td>
<td>The parallel axes theorem</td>
<td>28</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Diagonalisation of rank-two tensors</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>Taylor expansions</td>
<td>34</td>
</tr>
<tr>
<td>5.1</td>
<td>The one-dimensional case</td>
<td>34</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Examples</td>
<td>36</td>
</tr>
<tr>
<td>5.1.2</td>
<td>A precursor to the three-dimensional case</td>
<td>37</td>
</tr>
<tr>
<td>5.2</td>
<td>The three-dimensional case</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Curvilinear coordinates</td>
<td>39</td>
</tr>
<tr>
<td>6.1</td>
<td>Orthogonal curvilinear coordinates</td>
<td>40</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Scale factors and basis vectors</td>
<td>40</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Examples of orthogonal curvilinear coordinates (OCCs)</td>
<td>41</td>
</tr>
<tr>
<td>6.2</td>
<td>Elements of length, area and volume</td>
<td>42</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Vector length</td>
<td>42</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Arc length and metric tensor</td>
<td>43</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Vector area</td>
<td>44</td>
</tr>
<tr>
<td>6.2.4 Volume</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>6.3 Components of a vector field in curvilinear coordinates</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>6.4 Div, grad, curl and the Laplacian in OCCs</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>6.4.1 Gradient</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>6.4.2 Divergence</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>6.4.3 Curl</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>6.4.4 Laplacian of a scalar field</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>6.4.5 Laplacian of a vector field</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>7 Electrostatics</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>7.1 The Dirac delta function in three dimensions</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>7.2 Coulomb’s law</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>7.3 The electric field</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>7.3.1 Field lines</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>7.3.2 The principle of superposition</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>7.4 The electrostatic potential for a point charge</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>7.5 The static Maxwell equations</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>7.5.1 The curl equation</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>7.5.2 Conservative fields and potential theory</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>7.5.3 The divergence equation</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>7.6 Electric dipole</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>7.6.1 Potential and electric field due to a dipole</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>7.6.2 Force, torque and energy</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>7.7 The multipole expansion</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>7.7.1 Worked example</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>7.7.2 Interaction energy of a charge distribution</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>7.7.3 A brute-force calculation - the circular disc</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>7.8 Gauss’ law</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>7.9 Boundaries</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>7.9.1 Normal component</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>7.9.2 Tangential component</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>7.9.3 Conductors</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>7.10 Poisson’s equation</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>7.10.1 Uniqueness of solution</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>
7.10 Methods of solution ........................................... 73
7.10.2 Methods of solution ........................................... 73
7.10.3 The method of images ........................................... 74
7.11 Electrostatic energy .............................................. 77
7.11.1 Electrostatic energy of a general charge distribution ....... 77
7.11.2 Field energy .............................................. 78
7.12 Capacitors (condensers) and capacitance ......................... 79
7.12.1 Parallel-plate capacitor ..................................... 79
7.12.2 Concentric conducting spheres ................................ 80
7.12.3 Energy stored in a capacitor ................................ 80

8 Magnetostatics 81

8.1 Currents ......................................................... 81
8.1.1 Charge and current conservation .............................. 82
8.1.2 Conduction current ......................................... 83
8.2 Forces between currents (Ampère, 1821) ......................... 84
8.2.1 The Lorentz force ......................................... 85
8.3 Biot-Savart law .................................................. 86
8.3.1 Long straight wire .......................................... 87
8.3.2 Two long parallel wires ..................................... 87
8.3.3 Current loop ............................................. 88
8.4 Divergence and curl of the magnetic field: Gauss and Ampère laws .............................................. 88
8.4.1 Ampère’s law (1826) ....................................... 90
8.4.2 Conducting surfaces ....................................... 91
8.5 The vector potential ........................................... 92
8.6 Magnetic dipoles ................................................ 94
8.6.1 Magnetic moment and angular momentum .................. 96
8.6.2 Force and torque on a dipole in an external field ........... 96
8.7 Summary: electrostatics and magnetostatics ...................... 99

9 Electrostatics and magnetostatics in materials 100

9.1 Dielectrics ....................................................... 100
9.1.1 Boundary conditions between two insulators ................. 101
9.2 Linear Dielectrics ............................................. 102
9.3 Magnetisation ................................................ 103
9.3.1 Boundary conditions between two magnetic materials ...... 104
9.4 Diamagnets, Paramagnets and Ferromagnets . . . . . . . . . . . . . . . 105
9.5 Summary: the static Maxwell equations for (linear) media . . . . . . . . 106
Chapter 1

Revision of vector algebra

There should be nothing new in this (brief) chapter. It’s intended to revise concepts and notation, and to refresh your memory. The style and content should be similar to Martin Evans’ *Fields* notes, to which you should refer for detailed expositions.

1.1 Revision of vectors & index/suffix notation

In this chapter we will work entirely in three-dimensional real space, \( \mathbb{R}^3 \).

1.1.1 Position vector

The *position vector* is bound to some origin \( O \) and gives the position of a point relative to that origin; it will generally be denoted by \( \underline{x} \) or \( \underline{r} \).

1.1.2 Cartesian coordinates

Let \( \{\underline{e}_i\}, i = 1, 2, 3 \), be a set of orthonormal Cartesian basis vectors in \( \mathbb{R}^3 \). By definition, these satisfy

\[
\underline{e}_1 \cdot \underline{e}_1 = \underline{e}_2 \cdot \underline{e}_2 = \underline{e}_3 \cdot \underline{e}_3 = 1 \\
\underline{e}_1 \cdot \underline{e}_2 = \underline{e}_2 \cdot \underline{e}_3 = \underline{e}_3 \cdot \underline{e}_1 = 0
\]

The position vector \( \underline{x} \) may be expressed in terms of Cartesian coordinates \( (x_1, x_2, x_3) \):

\[
\underline{x} = x_1 \underline{e}_1 + x_2 \underline{e}_2 + x_3 \underline{e}_3
\]

Similarly, for an arbitrary vector \( \underline{a} \) in \( \mathbb{R}^3 \), we write

\[
\underline{a} = a_1 \underline{e}_1 + a_2 \underline{e}_2 + a_3 \underline{e}_3 = \sum_{i=1}^{3} a_i \underline{e}_i \equiv a_i \underline{e}_i
\]

In the last expression, we’ve used the *Einstein summation convention* to denote an *implicit sum* over the repeated or *dummy* index \( i = 1, 2, 3 \).
1.1.3 The Kronecker delta symbol $\delta_{ij}$

Define $\delta_{ij}$, where $i$ and $j$ can take on the values 1, 2, 3, such that

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

so that $\delta_{11} = \delta_{22} = \delta_{33} = 1$ and $\delta_{12} = \delta_{13} = \delta_{23} = \cdots = 0$.

The equations satisfied by the orthonormal basis vectors $\{\mathbf{e}_i\}$ may then be written succinctly as

$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$

The components $a_i$ of $\mathbf{a}$ in an orthonormal basis may be obtained using orthonormality

$a_i \cdot \mathbf{e}_i = a_j \mathbf{e}_j \cdot \mathbf{e}_i = a_j \delta_{ij} = a_i$ \quad [or occasionally $(a)_i$]

In this equation $i$ is a ‘free’ index and it takes on all three values $i = 1, 2, 3$.

1.1.4 The Levi-Civita or epsilon symbol $\epsilon_{ijk}$

Define $\epsilon_{ijk}$, where $i, j$ and $k$ can each take on the values 1, 2, 3, such that

$\epsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is an even permutation of (123)} \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of (123)} \\ 0 & \text{otherwise (i.e. when 2 or more indices are the same)} \end{cases}$

In other words:

$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$; $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$; all others = 0.

Note the cyclic symmetry

$\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}$

The Kronecker delta and epsilon symbols generalise straightforwardly to any number of dimensions.

1.1.5 Scalar product

The scalar product, also known as the inner product, or just the “dot” product, of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$\mathbf{a} \cdot \mathbf{b} = a b \cos \theta_{ab}$

where $a \equiv |a|$ and $b \equiv |b|$ are the lengths of the two vectors, and $\theta_{ab}$ is the angle between them. In Cartesian coordinates

$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \equiv a_i b_i$

where we used the summation convention in the last expression.
1.1.6 Vector product

The vector product (or cross product) is defined as
\[ \mathbf{a} \times \mathbf{b} = \mathbf{a} \mathbf{b} \sin \theta \mathbf{n} \]
where \(\mathbf{n}\) is a unit vector orthogonal to both \(\mathbf{a}\) and \(\mathbf{b}\) and the vectors \(\{\mathbf{a}, \mathbf{b}, \mathbf{n}\}\) form a right-handed set.

In Cartesian coordinates, the \(i^{th}\) component of the vector product of \(\mathbf{a}\) and \(\mathbf{b}\) may be written
\[ (\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k \]
where there is an implicit sum over both of the repeated indices \(j \& k\):
\[ \sum_{j=1}^{3} \sum_{k=1}^{3} \]

Similarly, we can write the vector product as
\[ \mathbf{a} \times \mathbf{b} = \epsilon_i (\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk} a_j b_k \]

The orthonormality relations satisfied by the vector products of the (right-handed) orthonormal basis vectors \(\{\mathbf{e}_i\}\) can be written as:
\[ \mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk} \mathbf{e}_k \quad \forall i, j = 1, 2, 3 \]
where we have again used the summation convention: there is an implicit sum over the repeated index \(k\):
\[ \sum_{k=1}^{3} \]

1.1.7 Scalar triple product

From the results above, we can easily deduce an algebraic definition of the scalar triple product in Cartesian coordinates
\[ (\mathbf{a}, \mathbf{b}, \mathbf{c}) \equiv \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = a_i (\mathbf{b} \times \mathbf{c})_i = \epsilon_{ijk} a_i b_j c_k \]

1.1.8 Vector triple product

The vector triple product \(\mathbf{a} \times (\mathbf{b} \times \mathbf{c})\) satisfies
\[ \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} \]
which may be proved by considering explicit components of each side of the equation.

It may also be obtained from the following expression for the product of two epsilon symbols with one common (summed) index:
\[ \epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \]

You must know this result!

When the epsilon symbols have two common indices, we get another useful result (exercise)
\[ \epsilon_{ijk} \epsilon_{pjk} = 2 \delta_{ip} \]
1.2 Linear transformation of basis

Let \( \{e_i\} \) and \( \{e'_i\} \) be two orthonormal bases with a common origin related by the linear transformation

\[ e'_i = \ell_{ij} e_j \]

where we have again used the summation convention, so the repeated index \( j \) is summed over.

We sometimes refer to the bases \( \{e_i\} \) and \( \{e'_i\} \) as frames \( S \) and \( S' \), respectively.

Using orthonormality of both bases, it is straightforward to show that the nine constants \( \ell_{ij} \) satisfy (exercise)

\[ \ell_{ik} \ell_{jk} = \ell_{ki} \ell_{kj} = \delta_{ij} \iff L L^T = L^T L = I \]

where we have used matrix notation in the 2nd version.

The \( 3 \times 3 \) transformation matrix \( L \) is an orthogonal matrix, whose elements are given by \((L)_{ij} = \ell_{ij} = e'_i \cdot e_j \) (exercise), and \( I \) is the unit matrix.

Every orthogonal matrix \( L \) has determinant \( \det L = \pm1 \).

If \( L \) represents a rotation of the basis \( \{e_i\} \), then \( \det L = 1 \). This is called a proper transformation.

If \( L \) represents an inversion of the basis, i.e. \( e'_i = -e_i \), or a reflection of the basis vectors in some plane, then \( \det L = -1 \). These are called improper transformations.

1.3 Transformation properties of vectors and scalars

1.3.1 Transformation of vector components

Let \( a \) be any vector, with components \( a_i \) in the \( \{e_i\} \) basis and components \( a'_i \) in the \( \{e'_i\} \) basis, so that

\[ a = a_i e_i = a'_i e'_i \]

The vector \( a \) is the same in both bases, but its components in the two bases are related by

\[ a'_i = \ell_{ij} a_j \]

1.3.2 Scalar and vector products

The scalar product of two vectors \( a \) and \( b \) is the same in both bases

\[ a \cdot b = a_i b_i = a'_i b'_i \]

Hence the name scalar product.

The vector product \( a \times b \) is a pseudovector.

The scalar triple product is a pseudoscalar, and the vector triple product is a vector.

We shall revisit all these properties in detail after we introduce tensors in the next chapter...
Chapter 2

Cartesian tensors

2.1 Definition and transformation properties

Consider a rotation of the \( \{e_i\} \) basis (frame \( S \)) to the \( \{e'_i\} \) basis (frame \( S' \)). This is called a passive rotation.

The rotation matrix \( L \), with components \( \ell_{ij} \), satisfies \( LL^T = I = L^T L \), and it has unit determinant \( \det L = +1 \).

The components of two arbitrary vectors \( a \) and \( b \) in the two frames are related by

\[
\begin{align*}
a'_i &= \ell_{ij} a_j \\
b'_i &= \ell_{ij} b_j
\end{align*}
\]

Let us now define a vector \( a \) as an entity whose 3 components \( a_i \) in \( S \) are related to its 3 components \( a'_i \) in \( S' \) by \( a'_i = \ell_{ij} a_j \).

Note that the vector \( a \) is not rotated – it remains fixed in space – only the basis vectors are rotated. Hence the term passive.

Now consider the 9 quantities \( a_i b_j \). Under the change of basis, these transform to

\[
\begin{align*}
a'_i b'_j &= \ell_{ir} a_r \ell_{js} b_s \\
&= (\ell_{ir} \ell_{js})(a_r b_s)
\end{align*}
\]

The 9 quantities \( a_r b_s \) obey a particular transformation law under the change of basis \( \{e_i\} \to \{e'_i\} \) (change of frame \( S \to S' \)). This motivates our definition of a tensor.

2.1.1 Definition of a tensor

Following on from our new definition of a vector, we define a tensor of rank 2, \( T \), as an entity whose \( 3^2 = 9 \) cartesian components \( T_{pq} \) in \( S \) are related to its 9 components \( T'_{ij} \) in \( S' \) by

\[
T'_{ij} = \ell_{ip} \ell_{jq} T_{pq}
\]

where \( L \) is the rotation matrix with components \( \ell_{ij} \) which takes \( S \to S' \). Since there are 2 free indices \( (i & j) \) in the above expression, it represents 9 equations.
Similarly a tensor of rank \( n \), \( T \), is defined to be an entity whose \( 3^n \) components \( T'_{ijk\cdots op} \) \((n\text{-indices})\) in \( S' \) are related to its \( 3^n \) components \( T_{rst\cdots vw} \) \((n\text{-indices})\) in \( S \) by

\[
T'_{ijk\cdots op} = \ell_{ir} \ell_{js} \ell_{kt} \cdots \ell_{ow} \ell_{pw} T_{rst\cdots vw}
\]

where there are \( n \) factors of \( \ell \) on the RHS.

In this new language

- A scalar is a tensor of rank 0 \((i.e. a'_i = a)\).
- A vector is a tensor of rank 1 \((as a'_i = \ell_{ij} a_j)\).

We shall often be sloppy and say \( T_{ijk\cdots rs} \) is a tensor, when what we really mean is that \( T \) is a tensor with components \( T_{ijk\cdots rs} \) in a particular frame \( S \).

The expressions tensor of rank 2 and second-rank tensor are used interchangeably. Similarly for tensor of rank \( n \) and \( n^{th} \)-rank tensor.

**Important:** Note that a rank-\( n \) tensor is more general than the ‘direct product’ or ‘tensor product’ of components of \( n \) vectors, \( i.e., \) not every tensor has components that can be written as a direct product of vectors \( a_i b_j c_k \ldots p_r q_s \). For example, \( a_i b_j + a_j b_i \) is a rank-2 tensor. Another counterexample for \( n = 2 \) is given in section 2.1.5.

### 2.1.2 Fields

A scalar or vector or tensor quantity is called a field when it is a function of position:

- Temperature \( T(r) \) is a scalar field
- The electric field \( E_i(r) \) is a vector field
- The stress-tensor field \( P_{ij}(r) \) is a \((rank \ 2)\) tensor field

In the latter case the transformation law is

\[
P'_{ij}(r) = \ell_{ip} \ell_{jq} P_{pq}(r) \quad \text{or} \quad P'_{ij}(x'_k) = \ell_{ip} \ell_{jq} P_{pq}(x_k) \quad \text{with} \quad x'_k = \ell_{kl} x_l
\]

These two expressions mean the same thing, but the latter form is perhaps better.

### 2.1.3 Dyadic notation

In some (mostly older) books you will see dyadic notation. This is rather clumsy for tensors – although it works well for vectors of course!

<table>
<thead>
<tr>
<th>dyadic notation</th>
<th>index notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{a}{a} )</td>
<td>( a_i )</td>
</tr>
<tr>
<td>( a \cdot b )</td>
<td>( a_i b_i )</td>
</tr>
<tr>
<td>( A )</td>
<td>( A_{ij} ) or ( a_{ij} )</td>
</tr>
<tr>
<td>( \frac{a}{a} A b ) or ( a \cdot A \cdot b )</td>
<td>( a_i A_{ij} b_j )</td>
</tr>
</tbody>
</table>

We will not use dyadic notation for tensors in this course.
2.1.4 Internal consistency in the definition of a tensor

Let $T_{ij}$, $T'_{ij}$, $T''_{ij}$ be the components of a tensor in frames $S$, $S'$, $S''$ respectively. And let

\[
L = \{ \ell_{ij} \}
\]

be the rotation matrix which takes $S \rightarrow S'$, $M = \{ m_{ij} \}$ be the rotation matrix which takes $S' \rightarrow S''$, then

\[
T''_{ij} = m_{ip} m_{jq} T'_{pq} = (ML)_{ik} (ML)_{jl} T_{kl} = n_{ik} n_{jl} T_{kl}
\]

where $N = ML$ is the rotation matrix which takes $S \rightarrow S''$, so the definition of a tensor is self-consistent. A similar result can be derived for vectors (exercise).

2.1.5 Properties of Cartesian tensors – tensor algebra

- **Addition:** if $T_{ij\ldots p}$ and $U_{ij\ldots p}$ are two tensors with the same rank $n$, i.e. both have $n$ indices, then

\[
V_{ij\ldots p} = T_{ij\ldots p} + U_{ij\ldots p}
\]

is also a tensor of rank $n$. The proof is straightforward.

- **Multiplication or tensor product:** if $T_{ijk\ldots s}$ and $U_{lm\ldots r}$ are the components of tensors $T$ and $U$ of rank $n$ and $m$ respectively, then

\[
V_{ij\ldots slm\ldots r} = T_{ij\ldots s} U_{lm\ldots r}
\]

are the components of a tensor $TU$ of rank $n + m$, which has $3^n \times 3^m = 3^{n+m}$ components:

\[
V'_{i\ldots r} = T'_{i\ldots s} U'_{l\ldots r} = \ell_{ia} \ldots \ell_{s\delta} T_{\alpha\ldots \delta} \ell_{t \epsilon} \cdots \ell_{r \rho} U_{\epsilon \ldots \rho} = \ell_{ia} \ldots \ell_{r \rho} V_{\alpha \ldots \rho}
\]

where there are $n + m$ factors of $\ell$ in the last expression.

Note that we sometimes use Greek letters $\alpha$, $\beta$, etc for dummy indices when we have a large number of them.

**Example:** If $U = \lambda$ is a scalar, and $T$ is a tensor of rank $n$, then $\lambda T$ is a tensor of rank $0 + n = n$.

- **Contraction:** if $T_{ijk\ldots s}$ is a tensor of rank $n$, then $T_{iik\ldots s}$ (i.e. $n-2$ free indices) is a tensor of rank $n-2$. The process of setting indices to be equal and summing is called **contraction**.

**Example:** If $T_{ij} = a_i b_j$ is a tensor of rank 2, then $T_{ii} = a_i b_i$ is a tensor of rank 0 (scalar), it’s the usual scalar product.

The process of multiplying two tensors and contracting over a pair (or pairs) of indices on different tensors is (often) called taking the **scalar product**.
• If \( T_{ij} \) is a tensor then so is \( T_{ji} \) (where \( T_{ji} \) are the components of the transpose \( T^T \) of \( T \)).

• If \( T_{ij} = T_{ji} \) in \( S \), then \( T'_i = T'_j \) in \( S' \):
  \[
  T'_{ij} = \ell_{ip} \ell_{jq} T_{pq} = \ell_{jp} \ell_{iq} T_{pq} = T'_{ji}
  \]
  \( T_{ij} \) is a symmetric tensor; the symmetry is preserved under a change of basis.
  [The notation \( p \leftrightarrow q \) refers to relabelling indices.]

Similarly if \( T_{ij} = -T_{ji} \), then \( T'_i = -T'_j \). \( T_{ij} \) is an anti-symmetric tensor.

Given any second rank tensor \( T \), we can always decompose it into symmetric and anti-symmetric parts

\[
T_{ij} = \frac{1}{2} (T_{ij} + T_{ji}) + \frac{1}{2} (T_{ij} - T_{ji})
\]

• We can re-write the tensor transformation law for rank 2 tensors (only) using matrix notation:
  \[
  T'_{ij} = \ell_{ip} \ell_{jq} T_{pq} = (LT^T)_{ij}
  \]

so

\[
T' = LT^T \equiv LT \ell^{-1} \quad \text{(for rank 2 only)}
\]

• Kronecker delta, \( \delta_{ij} \), is a second rank tensor.

To show this, first note that the definition

\[
\delta_{ij} = \begin{cases} 
  +1 & i = j \\
  0 & i \neq j
\end{cases}
\]

holds in all frames, so that, for example,

\[
\mathbf{a} \cdot \mathbf{b} = \delta_{ij} a_i b_j = \delta_{ij} a'_i b'_j
\]

so \( \delta'_{ij} = \delta_{ij} \).

Starting with

\[
\delta'_{ij} = \delta_{ij}
\]

and using the general result \( \ell_{ip} \ell_{jq} = \delta_{ij} \), we obtain

\[
\delta'_{ij} = \delta_{ij} = \ell_{ip} \ell_{jq} \delta_{pq}
\]

which we recognise as the definition of a second rank tensor.

A tensor which has the same components in all frames is called an invariant or isotropic tensor.

\( \delta_{ij} \) is an example of a tensor that can not be written in the form \( T_{ij} = a_i b_j \).

[To show this, assume we can write \( \delta_{ij} = a_i b_j \). The diagonal terms in this expression require non-zero \( a_i \) and \( b_i \) for all \( i \), while the off-diagonal terms require at least 3 of \( a_i \) and \( b_j \) to be zero, which is a contradiction.]

There are no invariant first rank tensors, apart from the zero vector \( \mathbf{0} \).
2.1.6 The quotient theorem

Let $T$ be an entity with 9 components in any frame, say $T_{ij}$ in $S$, and $T'_{ij}$ in $S'$.

Let $a$ be an arbitrary vector and let $b_i = T_{ij} a_j$. If $b$ always transforms as a vector, then $T$ is a second rank tensor.

To prove this, we determine the transformation properties of $T$. In $S'$ we have

$$b'_i = T'_{ij} a'_j = T'_{ij} \ell_{jk} a_k$$

$$\equiv \ell_{ij} b_j = \ell_{ij} T_{jk} a_k$$

Equate the last expression on each line and rearrange

$$(T'_{ij} \ell_{jk} - \ell_{ij} T_{jk}) a_k = 0$$

This expression holds for all vectors $a$, for example $a = (1, 0, 0)$ etc, therefore

$$T'_{ij} \ell_{jk} = \ell_{ij} T_{jk}$$

$$\Rightarrow T'_{ij} \ell_{jk} \ell_{mk} = \ell_{ij} \ell_{mk} T_{jk}$$

[multiplied both sides by $\ell_{mk}$]

$$\Rightarrow T'_{im} = \ell_{ij} \ell_{mk} T_{jk}$$

where we used $\ell_{jk} \ell_{mk} = \delta_{jm}$ in the last step. Thus $T$ transforms as a tensor.

**Important example:** If there is a linear relationship between two vectors $a$ and $b$, so that $a_i = T_{ij} b_j$, it follows from the quotient theorem that $T$ is a tensor.

This is an alternative definition of a second-rank tensor, and it’s often the way that tensors arise in nature.

**Quotient theorem:** Generalising, if $R_{ij \ldots r}$ is an arbitrary tensor of rank-$m$, and $T_{ij \ldots s}$ is a set of $3^n$ numbers (with $n > m$) such that $T_{ij \ldots s} R_{ij \ldots r}$ is a tensor of rank $n - m$, then $T_{ij \ldots s}$ is a tensor of rank-$n$.

Proof is similar to the rank-2 case, but it isn’t very illuminating.

2.2 Revision of matrices and determinants

Before extending the discussion to pseudotensors, we revise and extend some familiar properties of matrices and determinants.

2.2.1 Matrices

An $M \times N$ matrix is a rectangular array of numbers with $M$ rows and $N$ columns,

$$A = \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1,N-1} & a_{1N} \\
  a_{21} & a_{22} & \cdots & a_{2,N-1} & a_{2N} \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  a_{M-1,1} & a_{M-1,2} & \cdots & a_{M-1,N-1} & a_{M-1,N} \\
  a_{M,1} & a_{M,2} & \cdots & a_{M,N-1} & a_{M,N}
\end{pmatrix} \equiv \{a_{ij}\}$$
The quantities \( \{a_{ij}\} \), with \( a_{ij} \equiv (A)_{ij} \) for all \( 1 \leq i \leq M, 1 \leq j \leq N \), are the elements of the matrix.

A square matrix has \( N = M \). We’ll work mostly with \( 3 \times 3 \) matrices, but the majority of what we’ll do generalises to \( N \times N \) matrices rather easily.

- We can add & subtract same-dimensional matrices and multiply a matrix by a scalar. Component forms are obvious, e.g. \( A = B + \lambda C \) becomes \( a_{ij} = b_{ij} + \lambda c_{ij} \) in index notation. Since both \( i \) and \( j \) are free indices, this represents 9 equations.
- The unit matrix, \( I \), defined by
  \[
  I = \begin{pmatrix}
  1 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1 
  \end{pmatrix},
  \]
  has components \( \delta_{ij} \), i.e. \( I_{ij} = \delta_{ij} \)
- The trace of a square matrix is the sum of its diagonal elements
  \[ \text{Tr } A = a_{ii} \]
  Note the implicit sum over \( i \) due to our use of the summation convention.
- The transpose of a square matrix \( A \) with components \( a_{ij} \) is defined by swapping its rows with its columns, so
  \[(A^T)_{ij} = A^T_{ij} = a_{ji} \]
- If \( A = A^T \) then \( a_{ji} = a_{ij} \) \( A \) is symmetric
  
  If \( A = -A^T \) then \( a_{ji} = -a_{ij} \) \( A \) is antisymmetric

**Product of matrices**

We can very easily implement the usual ‘row into column’ matrix multiplication rule in index notation.

If \( A \) (with elements \( a_{ij} \)) is an \( M \times N \) matrix and \( B \) (with elements \( b_{ij} \)) is an \( N \times P \) matrix then \( C = AB \) is an \( M \times P \) matrix with elements \( c_{ij} = a_{ik} b_{kj} \). Since we’re using the summation convention, there is an implicit sum \( k = 1, \ldots N \) in this expression.

Matrix multiplication is associative and distributive

\[
A(BC) = (AB)C \equiv ABC \\
A(B + C) = AB + AC
\]
respectively, but it’s not commutative: $AB \neq BA$ in general.

An important result is

$$(AB)^T = B^T A^T$$

which follows because

$$(AB)^T_{ij} = (AB)_{ji} = a_{jk} b_{ki} = (B^T)_{ik} (A^T)_{kj} = (B^T A^T)_{ij}$$

### 2.2.2 Determinants

The determinant $\det A$ (or $|A|$ or $||A||$) of a $3 \times 3$ matrix $A$ may be defined by

$$\det A = \epsilon_{lmn} a_{1l} a_{2m} a_{3n}$$

This is equivalent to the ‘usual’ recursive definition,

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

where the first index labels rows and the second index columns. Expanding the determinant gives

$$\det A = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= (\epsilon_{123}a_{11}a_{22}a_{33} + \epsilon_{132}a_{11}a_{23}a_{32}) + \ldots$$

$$= \epsilon_{1mn}a_{11}a_{2m}a_{3n} + \ldots$$

$$= \epsilon_{1mn}a_{1l}a_{2m}a_{3n}$$

Thus the two forms are equivalent. The $\epsilon$ form is convenient for derivation of various properties of determinants.

Note that only one term from each row and column appears in the determinant sum, which is why the determinant can be expressed in terms of the $\epsilon$ symbol.

The determinant is only defined for a square matrix, but the definition can be generalised to $N \times N$ matrices,

$$\det A = \epsilon_{i_1 \ldots i_N} a_{i_1} \ldots a_{i_N}$$

where the epsilon symbol with $N$ indices is defined by

$$\epsilon_{i_1 \ldots i_N} = \begin{cases} +1 & \text{if $i_1, \ldots, i_N$ is an even permutation of $1, \ldots, N$} \\ -1 & \text{if $i_1, \ldots, i_N$ is an odd permutation of $1, \ldots, N$} \\ 0 & \text{otherwise} \end{cases}$$

We shall usually consider $N = 3$, but most results generalise to arbitrary $N$. 
We may use these results to derive several alternative (and equivalent) expressions for the determinant.

First define the quantity

\[ X_{ijk} = \epsilon_{lmn} a_{il} a_{jm} a_{kn} \]

It follows that

\[
X_{jik} = \epsilon_{lmn} a_{jl} a_{im} a_{kn} \\
= \epsilon_{mln} a_{jm} a_{il} a_{kn} \quad \text{(where we relabelled } l \leftrightarrow m) \\
= -\epsilon_{lmn} a_{il} a_{jm} a_{kn} = -X_{ijk}
\]

Thus the symmetry of \( X_{ijk} \) is dictated by the symmetry of \( \epsilon_{lmn} \), and we must have

\[ X_{ijk} = c \epsilon_{ijk} \]

where \( c \) is some constant. To determine \( c \), set \( i = 1, j = 2, k = 3 \), which gives \( \epsilon_{123} c = X_{123} \), so \( c = \epsilon_{lmn} a_{1l} a_{2m} a_{3n} = \det A \) and hence

\[ \epsilon_{ijk} \det A = \epsilon_{lmn} a_{il} a_{jm} a_{kn} \]

Multiplying by \( \epsilon_{ijk} \) and using \( \epsilon_{ijk} \epsilon_{ijk} = 6 \) (exercise) gives the symmetrical form for \( \det A \)

\[ \det A = \frac{1}{3!} \epsilon_{ijk} \epsilon_{lmn} a_{il} a_{jm} a_{kn} \]

This elegant expression isn’t of practical use because the number of terms in the sum increases from \( 3^3 \) to \( 3^6 \) (overcounting).

We can obtain a result similar to the boxed expression above by defining

\[ Y_{lmn} = \epsilon_{ijk} a_{il} a_{jm} a_{kn} \]

Using the same argument as before \[\text{tutorial}\] gives

\[ Y_{lmn} = \epsilon_{lmn} \left[ \epsilon_{ijk} a_{i1} a_{j2} a_{k3} \right] \]

Since \( \det A = 1/3! \epsilon_{lmn} Y_{lmn} \) this means that

\[ \det A = \epsilon_{ijk} a_{i1} a_{j2} a_{k3} \]

and

\[ \epsilon_{lmn} \det A = \epsilon_{ijk} a_{il} a_{jm} a_{kn} \]

[\text{tutorial}]
Properties of Determinants

We can easily derive familiar properties of determinants from the definitions above:

- Adding a multiple of one row to another does not change the value of the determinant.
  **Example:** Adding a multiple of the second row to the first row
  \[ \epsilon_{lmn} a_{1l} a_{2m} a_{3n} \rightarrow \epsilon_{lmn} (a_{1l} + \lambda a_{2l}) a_{2m} a_{3n} = \epsilon_{lmn} a_{1l} a_{2m} a_{3n} + 0 \]
  and \( \det A \) is unaltered. The last term is zero because \( \epsilon_{lmn} a_{2l} a_{2m} = 0 \).

- Adding a multiple of one column to another does not change the value of the determinant (use the other form for the determinant, \( \det A = \epsilon_{ijk} a_{i1} a_{j2} a_{k3} \)).

- Interchanging any two rows of a matrix changes the sign of the determinant.
  **Example:** Interchanging the first and second rows gives
  \[ \begin{align*}
  \epsilon_{lmn} a_{2l} a_{1m} a_{3n} &= \epsilon_{mln} a_{2l} a_{1m} a_{3n} = -\epsilon_{lmn} a_{1l} a_{2m} a_{3n} = -\det A
  \end{align*} \]
  In the first step we simply relabelled \( l \leftrightarrow m \).
  When \( A \) has two identical rows, \( \det A = 0 \).

- Interchanging any two columns of a matrix also changes the sign of the determinant (use the other form for the determinant, \( \det A = \epsilon_{ijk} a_{i1} a_{j2} a_{k3} \)).

Finally,
\[ \epsilon_{ijk} \epsilon_{lmn} \det A = \begin{vmatrix}
  a_{il} & a_{im} & a_{in} \\
  a_{jl} & a_{jm} & a_{jn} \\
  a_{kl} & a_{km} & a_{kn}
\end{vmatrix} \]
(2.1)

To derive this, start with the original definition of \( \det A \) as \(|···|\) and permute rows and columns. This produces \( ± \) signs equivalent to \( \epsilon \) permutations.

### 2.2.3 Linear equations

A standard use of matrices & determinants is to solve (for \( x \)) the matrix-vector equation
\[ A x = y \]
where \( A \) is a square \( 3 \times 3 \) matrix. Representing \( x \) and \( y \) by column matrices, and writing out the components, this becomes

\[ \begin{align*}
  a_{11} x_1 + a_{12} x_2 + a_{13} x_3 &= y_1 \\
  a_{21} x_1 + a_{22} x_2 + a_{23} x_3 &= y_2 \\
  a_{31} x_1 + a_{32} x_2 + a_{33} x_3 &= y_3
\end{align*} \]

In index notation
\[ a_{ij} x_j = y_i \]
With a suitable definition of the inverse $A^{-1}$ of the matrix $A$, we can write the solution as
\[ x = A^{-1}y \quad \text{or} \quad x_i = A^{-1}_{ij}y_j \]
where $A^{-1}_{ij} \equiv (A^{-1})_{ij}$ is the $ij$th element of the inverse of $A$.

By explicit multiplication [tutorial] we can show $AA^{-1} = I = A^{-1}A$ as required. Alternatively [tutorial],
\[ A^{-1} = \frac{C^T}{\det A} \]
where $C = \{c_{ij}\}$ is the co-factor matrix of $A$, and $c_{ij} = (-1)^{i+j} \times$ the determinant formed by omitting the row and column containing $a_{ij}$.

Note that a solution exists if and only if $\det A \neq 0$.

These results generalise to $N \times N$ matrices.

**Determinant of the transpose**

\[
\det A^T = \epsilon_{lmn} A^T_{li} A^T_{mj} A^T_{kn} = \epsilon_{lmn} a_{1l} a_{2m} a_{3n}
\]
and hence
\[ \det A^T = \det A \]

**Product of determinants**

If $C = AB$ so that $c_{ij} = a_{ik}b_{kj}$ then
\[
\det C = \epsilon_{ijk} c_{lj} c_{km} = \epsilon_{ijk} a_{1l} b_{li} a_{2m} b_{mj} a_{3n} b_{nk} = [\epsilon_{ijk} b_{li} b_{mj} b_{nk}] a_{1l} a_{2m} a_{3n}\\ \epsilon_{lmn} \det B a_{1l} a_{2m} a_{3n}
\]
and hence
\[ \det AB = \det A \det B \]

**Product of two epsilons**

The product of two epsilon symbols with no identical indices may be written as
\[
\epsilon_{ijk} \epsilon_{lmn} = \left| \begin{array}{ccc} \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{array} \right|
\]
This equation has 6 free indices, so it represents $3^6 = 729$ identities: 18 say ‘1 = 1’, 18 say ‘−1 = 1’, so 693 say ‘0 = 0’.

The proof follows almost trivially by setting $A = I$ in equation (2.1)

$$\epsilon_{ijk} \epsilon_{lmn} \det A = \begin{vmatrix} a_{il} & a_{im} & a_{in} \\ a_{jl} & a_{jm} & a_{jn} \\ a_{kl} & a_{km} & a_{kn} \end{vmatrix},$$

whence $\det I = 1$ and $a_{ij} = \delta_{ij}$. Unfortunately, this result isn’t as useful as one might expect.

If we set $l = k$ and sum over $k$

$$\epsilon_{ijk} \epsilon_{klm} = \begin{vmatrix} \delta_{ik} & \delta_{il} & \delta_{im} \\ \delta_{jk} & \delta_{jl} & \delta_{jm} \\ \delta_{kk} & \delta_{kl} & \delta_{km} \end{vmatrix} = \delta_{ik} (\delta_{jm} \delta_{kl} - \delta_{jl} \delta_{km}) - \delta_{il} (\delta_{jk} \delta_{km} - 3\delta_{jm}) + \delta_{im} (\delta_{jk} \delta_{kl} - 3\delta_{jl})$$

$$= \delta_{im} \delta_{jl} - \delta_{il} \delta_{jm} + 2\delta_{il} \delta_{jm} - 2\delta_{im} \delta_{jl}$$

we obtain an old friend, which we now note can be written as the determinant of a $2 \times 2$ matrix,

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} = \begin{vmatrix} \delta_{il} & \delta_{im} \\ \delta_{jl} & \delta_{jm} \end{vmatrix}$$

2.2.4 Orthogonal matrices

Let $A$ be a $3 \times 3$ matrix with elements $a_{ij}$, and define the row vectors

$$a^{(1)} = (a_{11}, a_{12}, a_{13}), \quad a^{(2)} = (a_{21}, a_{22}, a_{23}), \quad a^{(3)} = (a_{31}, a_{32}, a_{33}),$$

so that $(a^{(i)})_j = a_{ij}$. If we choose the vectors $a^{(i)}$ to be orthonormal:

$$a^{(1)} \cdot a^{(2)} = a^{(2)} \cdot a^{(3)} = a^{(3)} \cdot a^{(1)} = 0,$$

and

$$|a^{(1)}|^2 = |a^{(2)}|^2 = |a^{(3)}|^2 = 1,$$

i.e.

$$a^{(i)} \cdot a^{(j)} = \delta_{ij},$$

then $A$ is an orthogonal matrix. The rows of $A$ form an orthonormal triad.

Properties of orthogonal matrices

- Consider

$$AA^T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

Therefore

$$AA^T = I$$
• Taking the determinant of both sides of $AA^T = I$ gives $\det A \det A^T = 1$.

Since $\det A^T = \det A$, then $(\det A)^2 = 1$, and therefore

$$\det A = \pm 1$$

• Since $\det A \neq 0$, the inverse matrix $A^{-1}$ always exists, and therefore

$$A^{-1} = A^T$$

Multiplying this equation on the right by $A$ gives

$$A^T A = I$$

so the columns of $A$ also form an orthonormal triad.

2.3 Pseudotensors, pseudovectors & pseudoscalars

Armed with properties of determinants, let’s now return to basis transformations.

Suppose that we now allow reflection and inversion (as well as rotation) of the basis vectors, and represent them all by a transformation matrix $L$ with

$$\det L = +1 \quad \text{for rotations}$$

$$\det L = -1 \quad \text{for reflections and inversions}$$

In the second case, the handedness of the basis is changed: if $S$ is right-handed (RH), then $S'$ is left-handed (LH), and vice versa.

Before introducing pseudotensors, we note that the basis vectors always transform as

$$e'_i = \ell_{ij} e_j$$

A second-rank tensor or true tensor $T$ obeys the transformation law

$$T'_{ij} = \ell_{ip} \ell_{jq} T_{pq}$$

for all transformations, i.e. rotations, reflections and inversions.

A second-rank pseudotensor $T$ obeys the transformation law

$$T'_{ij} = (\det L) \ell_{ip} \ell_{jq} T_{pq}$$

There is a change of sign for components of a pseudotensor, relative to a true tensor, when we make a reflection or inversion of the reference axes.

We can similarly define pseudovectors, pseudoscalars and rank-$n$ pseudotensors.

Note: $\det L = -1$ for a basis transformation that consists of any combination of an odd number of reflections or inversions, plus any number of rotations.
2.3.1 Vectors (revision)

Inversion of the basis vectors is defined by

\[(\vec{e}_1, \vec{e}_2, \vec{e}_3) \rightarrow (\vec{e}'_1, \vec{e}'_2, \vec{e}'_3) = (-\vec{e}_1, -\vec{e}_2, -\vec{e}_3)\]

so that

\[L = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\]

which has components \(\ell_{ij} = -\delta_{ij}\).

If \(\{\vec{e}_i\}\) is a standard right-handed basis, then \(\{\vec{e}'_i\}\) is left-handed.

Let \(\vec{a}\) be a vector (also called a tensor of rank 1 or a polar vector or a true vector).

We showed previously that the components \(a_i\) transform as \(a'_i = \ell_{ij} a_j\), so we have

\[a'_1 = -a_1 \quad a'_2 = -a_2 \quad a'_3 = -a_3\]

The vector itself therefore transforms as

\[\vec{a}' \equiv a'_i \vec{e}'_i = (a_i)(-\vec{e}_i) = a \vec{e}_i = \vec{a}\]

as expected. Thus a (true/polar) vector remains unchanged by inversion of the basis.

Inversion is sometimes called reflection in the origin - hence the label in the figure above.

2.3.2 Pseudovectors

If \(\vec{a}\) and \(\vec{b}\) are (true) vectors then \(\vec{c} = \vec{a} \times \vec{b}\) is a pseudovector (also known as a pseudotensor of rank 1 or an axial vector).

Let’s illustrate this by considering an inversion.

\[c'_1 = a'_2 b'_3 - a'_3 b'_2 = (-a_2)(-b_3) - (-a_3)(-b_2) = +c_1 \quad etc\]

Therefore

\[\vec{c}' \equiv c'_i \vec{e}'_i = (a_i)(-\vec{e}_i) = -c_i \vec{e}_i = -\vec{c}\]

Thus the direction of \(\vec{c} = \vec{a} \times \vec{b}\) is reversed in the new LH basis.
The vectors $a$, $b$ and $a \times b$ form a LH triad in the $S'$ basis, i.e. they always have the same 'handedness' as the underlying basis.

Now consider the general case. We define the epsilon symbol in all bases, regardless of their handedness, in the same way:

$$
\epsilon_{ijk} = \begin{cases} 
+1 & \text{if } ijk \text{ is an even permutation of 123} \\
-1 & \text{if } ijk \text{ is an odd permutation of 123} \\
0 & \text{otherwise}
\end{cases}
$$

The components of $c = a \times b$ are $c_i = \epsilon_{ijk} a_j b_k$ in $S$ and $c'_i = \epsilon_{ijk} a'_j b'_k$ in $S'$.

We can now determine the components of $c = a \times b$ in $S'$

$$
c'_i = \epsilon_{ijk} a'_j b'_k = \epsilon_{ijk} \ell_{jp} a_p \ell_{kq} b_q = \epsilon_{mjk} \ell_{ir} \ell_{mr} \ell_{jp} \ell_{kq} a_p b_q = \det L \ell_{ir} \epsilon_{pq} a_p b_q = \det L \ell_{ir} \epsilon_{r} = (a \times b)_r
$$

where we used $\epsilon_{ijk} = \epsilon_{mjk} \delta_{im} = \epsilon_{mjk} \ell_{ir} \ell_{mr}$ to get to the third line, and the determinant result $\epsilon_{pq} \det L = \epsilon_{mjk} \ell_{mr} \ell_{jp} \ell_{kq}$ to get the fourth line. So, finally

$$
c'_i = \det L \ell_{ir} c_r
$$

This is our definition of a pseudovector. Equivalently, since $e'_i = \ell_{ij} e_j$ then

$$
e'_i \equiv c'_i e'_i = (\det L \ell_{ir} c_r) \left( \ell_{ij} e_j \right) = \det L c_j e_j = \det L c
$$

Therefore a pseudovector changes sign under any improper transformation such as inversion of the basis or reflection of the basis in a plane.

**Pseudotensors:** The $\epsilon$-symbol is a pseudotensor of rank 3. The proof is simple and uses the fact that $\epsilon$ is the same in all bases, together with $(\det L)^2 = 1$

$$
\epsilon'_{ijk} \equiv \epsilon_{ijk} = \det L \det L \epsilon_{ijk} = \det L \ell_{ip} \ell_{jq} \ell_{kr} \epsilon_{pqr}
$$

where we used the definition of the determinant in the last step.
Furthermore since $\epsilon$ is the same in all bases, it’s an isotropic or invariant rank-3 tensor.

We can build pseudotensors of higher rank using a combination of vectors, tensors, $\delta$ and $\epsilon$. For example:

- If $a_i$ and $b_i$ are rank 1 tensors (i.e., vectors), then the $3^5$ quantities $\epsilon_{ijk} a_l b_m$ are the components of a pseudotensor of rank 5.
- $\epsilon_{ijk} \delta_{pq}$ is a pseudotensor of rank 5.

In general:

- The product of two tensors is a tensor
- The product of two pseudotensors is a tensor
- The product of a tensor and a pseudotensor is a pseudotensor

### 2.4 Some physical examples

- Velocity  \[ \mathbf{v} = \frac{d\mathbf{r}(t)}{dt} \] vector
- Acceleration  \[ \mathbf{a} = \dot{\mathbf{v}} \] vector
- Force  \[ \mathbf{F} = m\mathbf{a} \] vector
- Electric field  \[ \mathbf{E} = \mathbf{F} / q \] vector
  (F is the force on a test charge $q$)
- Torque  \[ \mathbf{G} = \mathbf{r} \times \mathbf{F} \] pseudovector
- Angular velocity ($\omega$)  \[ \mathbf{\omega} = \mathbf{r} \times \mathbf{v} \] pseudovector
- Angular momentum  \[ \mathbf{L} = \mathbf{r} \times \mathbf{p} \] pseudovector
- Magnetic field ($\mathbf{B}$)  \[ \mathbf{E} = q\mathbf{v} \times \mathbf{B} \] pseudovector

\[
\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \oint_C I \mathbf{dr}' \times (\mathbf{r} - \mathbf{r}') / |\mathbf{r} - \mathbf{r}'|^3
\]

(Biot–Savart Law – more later)

- $\mathbf{E} \cdot \mathbf{B}$ pseudoscalar

A pseudoscalar changes sign under an improper transformation: $a' = (\det L) a$.

In this case, one can easily show that

\[(\mathbf{E} \cdot \mathbf{B})' = \det L (\mathbf{E} \cdot \mathbf{B})\]
• Helicity \[ h = \mathbf{p} \cdot \mathbf{s} \]

The pseudovector \( \mathbf{s} \) is the angular momentum, or spin, of a particle/ball. As the ball spins, a point on it traces out a RH helix. In the figure, \( \mathbf{p} \) is parallel to \( \mathbf{s} \).

• Inertia tensor \( (I) \) \[ L_i = I_{ij} \omega_j \] rank 2 tensor

• Conductivity tensor \( (\sigma) \) \[ J_i = \sigma_{ij} E_j \] rank 2 tensor

• Stress tensor \( (P) \) \[ dF_i = P_{ij} dS_j \] rank 2 tensor

2.5 Invariant/Isotropic Tensors

A tensor \( T \) is invariant or isotropic\(^1\) if it has the same components, \( T_{ijk...} \) in any Cartesian basis (or frame of reference), so that
\[
T_{ijk...} = \ell_{i\alpha} \ell_{j\beta} \ell_{k\gamma} \cdots T_{\alpha\beta\gamma...}
\]
for every (orthogonal) transformation matrix \( L = \{\ell_{ij}\} \).
Similarly, \( T \) is an invariant pseudotensor if
\[
T_{ijk...} = \det L \ell_{i\alpha} \ell_{j\beta} \ell_{k\gamma} \cdots T_{\alpha\beta\gamma...}
\]

Theorem: If \( a_{ij} \) is a second-rank invariant tensor, then
\[
a_{ij} \text{ invariant} = \lambda \delta_{ij}
\]

Proof: For a rotation of \( \pi/2 \) about the \( z \)-axis
\[
L = \begin{pmatrix}
0 & 1 & 0 \\
-1 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}
\]
Since the only non-zero elements are \( \ell_{12} = 1, \ell_{21} = -1, \ell_{33} = 1 \), using \( a_{ij} = \ell_{i\alpha} \ell_{j\beta} a_{\alpha\beta} \), we find
\[
a_{11} = \ell_{1\alpha} \ell_{1\beta} a_{\alpha\beta} = \ell_{12} \ell_{12} a_{22} = a_{22}
\]
\[
a_{13} = \ell_{1\alpha} \ell_{3\beta} a_{\alpha\beta} = \ell_{12} \ell_{33} a_{23} = a_{23}
\]
\[
a_{23} = \ell_{2\alpha} \ell_{3\beta} a_{\alpha\beta} = \ell_{21} \ell_{33} a_{13} = -a_{13}
\]

\(^1\)Isotropic means “the same in all directions”.
Therefore
\[ a_{11} = a_{22}, \quad a_{13} = 0 = a_{23} \]

Similarly, for a rotation of \( \pi/2 \) about the \( y \)-axis, we find (exercise)
\[ a_{11} = a_{33}, \quad a_{12} = 0 = a_{32} \]

Finally, for a rotation of \( \pi/2 \) about the \( x \)-axis
\[ a_{22} = a_{33}, \quad a_{21} = 0 = a_{31} \]

The only solution to these equations is \( a_{ij} = \lambda \delta_{ij} \). We’ve already shown that \( \delta_{ij} \) is an invariant second-rank tensor, therefore \( \lambda \delta_{ij} \) is the most general invariant second-rank tensor.

One can use a similar argument to show that the only invariant vector is the zero vector. It’s obvious that no non-zero vector has the same components in all bases!

**Theorem:** There is no invariant tensor of rank 3. The most general invariant pseudotensor of rank 3 has components

\[ a_{ijk} \overset{\text{invariant}}{=} \lambda \epsilon_{ijk} \]

**Proof:** This is similar to the rank 2 case [tutorial].

**Theorem:** The most general 4th rank invariant tensor has components

\[ a_{ijkl} \overset{\text{invariant}}{=} \lambda \delta_{ij} \delta_{kl} + \mu \delta_{ik} \delta_{jl} + \nu \delta_{il} \delta_{jk} + \ldots \]

The proof is long and not very illuminating. See, for example, Matthews, Vector Calculus, (Springer) or Jeffreys, Cartesian Tensors, (CUP). However, it’s easy to show that the expression above is indeed an invariant tensor.

**Invariant tensors of higher rank**

- The most general rank-5 invariant pseudotensor has components

\[ a_{ijklm} \overset{\text{invariant}}{=} \lambda \delta_{ij} \epsilon_{klm} + \mu \delta_{ik} \epsilon_{jlm} + \nu \delta_{il} \epsilon_{jkm} + \ldots \]

where the dots indicate all terms consisting of a constant \( \times \) further permutations of the 5 indices.

- The most general rank-6 invariant tensor has components

\[ a_{ijklmn} \overset{\text{invariant}}{=} \lambda \delta_{ij} \delta_{kl} \delta_{mn} + \ldots \]

Note that invariant tensors involving products of two epsilons can always be rewritten as sums of products of \( \delta \)s using the expression for the product of two epsilons.

- The most general invariant tensor of rank \( 2n \) is a sum of products of constants times \( n \) Kronecker deltas. There is no invariant pseudotensor of rank \( 2n \).

- Similarly, the most general invariant pseudotensor of rank \( 2n + 1 \) is a sum of products of constants times one \( \epsilon \) and \( n - 1 \) Kronecker deltas. There is no invariant tensor of rank \( 2n + 1 \).
Chapter 3

Rotation, reflection & inversion tensors

The transformations of the previous chapters in which we rotate (or reflect or invert) the basis vectors keeping the vector fixed are called passive transformations.

Alternatively we can keep the basis fixed and rotate (or reflect or invert) the vector. These are called active transformations and are represented by second-rank tensors.

3.1 The rotation tensor

3.1.1 Rotation about an arbitrary axis

Consider a rotation of a rigid body, through angle $\theta$, about an axis which points in the direction of the unit vector $\hat{n}$. The axis passes through a fixed origin $O$.

The point $P$ is rotated to $Q$. The position vector $\vec{x}$ is rotated to $\vec{y}$.

In the first diagram, $\vec{OS}$ is the projection of $\vec{x}$ onto the $\hat{n}$ direction, i.e. $(\vec{x} \cdot \hat{n}) \hat{n}$.

In the second diagram, $n \times x$ is parallel to $\vec{TQ}$.

\[
\vec{y} = \vec{OS} + \vec{ST} + \vec{TQ} \\
= (\vec{x} \cdot \hat{n}) \hat{n} + \underbrace{SQ \cos \theta}_{|\vec{ST}|} \underbrace{(\vec{x} - (\vec{x} \cdot \hat{n}) \hat{n})}_{\vec{SP}} + SQ \sin \theta \underbrace{\frac{n \times x}{|\vec{TQ}|}}_{|\vec{TQ}|} \\
= (\vec{x} \cdot \hat{n}) \hat{n} + \cos \theta (\vec{x} - (\vec{x} \cdot \hat{n}) \hat{n}) + \sin \theta \hat{n} \times \vec{x}
\]

(as $SP = SQ$ and $|\hat{n} \times \vec{x}| = SP = SQ$).
This gives the important result

\[ y = x \cos \theta + (1 - \cos \theta) (n \cdot x) n + (n \times x) \sin \theta \]

In index notation, this is

\[ y_i = x_i \cos \theta + (1 - \cos \theta) n_j x_j n_i + \epsilon_{ijk} n_k x_j \sin \theta \]

which we write as

\[ y_i = R_{ij} (\theta, n) x_j \]

This is a linear relation between the vectors \( \underline{x} \) and \( \underline{y} \), so \( R (\theta, n) \) is a rank 2 tensor, the rotation tensor, with components

\[ R_{ij} (\theta, n) = \delta_{ij} \cos \theta + (1 - \cos \theta) n_i n_j - \epsilon_{ijk} n_k \sin \theta \]

3.1.2 Some important properties of the rotation tensor

(i) In any basis \( R \) is represented by an orthogonal matrix, with components \( R_{ij} \), and \( \det R = 1 \). Proofs [tutorial]

(ii) It is straightforward to show that [tutorial]

\[ \text{Tr} R = R_{ii} = 1 + 2 \cos \theta \quad \text{[independent of} \ n\text{]} \]

\[ -\frac{1}{2} \epsilon_{kij} R_{ij} = n_k \sin \theta \]

If \( R \) is known, then the angle \( \theta \) and the axis of rotation \( n \) can be determined.

Note: \( R \) has \( 1 + (3 - 1) = 3 \) independent parameters.

(iii) The product of two rotations \( \underline{x} \overset{R}{\rightarrow} \underline{y} \overset{S}{\rightarrow} \underline{z} \) is given by \( \underline{z} = SR \underline{x} \).

(iv) Consider a small (infinitesimal) rotation \( \delta \theta \), for which \( \cos \delta \theta = 1 + O(\delta \theta^2) \) and \( \sin \delta \theta = \delta \theta + O(\delta \theta^3) \), then

\[ R_{ij} = \delta_{ij} - \epsilon_{ijk} n_k \delta \theta \]

A quicker (and sufficient) graphical proof follows directly from the diagram on the right, which gives

\[ \underline{y} - \underline{x} = n \times \underline{x} \delta \theta \]

from which the result above follows directly.

(v) For \( \theta \neq 0, \pi \), \( R \) has only one real eigenvalue \( +1 \), with one real eigenvector \( n \). [tutorial]
3.2 Reflections and inversions

Consider reflection of a vector $\vec{x} \rightarrow \vec{y}$ in a plane with unit normal $\vec{n}$.

From the figure $\vec{y} = \vec{x} - 2(\vec{x} \cdot \vec{n}) \vec{n}$

In index notation, this becomes

$$y_i = \sigma_{ij} x_j$$

where $\sigma_{ij} = \delta_{ij} - 2 n_i n_j$

The quantities $\sigma_{ij}$ are the components of the reflection tensor $\sigma$.

Inversion of a vector in the origin is given by $\vec{y} = -\vec{x}$. This defines the parity operator $P$:

$$y_i = P_{ij} x_j$$

where $P_{ij} = -\delta_{ij}$

The parity operator is an invariant tensor with components $-\delta_{ij}$.

For reflections and inversions, respectively, $\det \sigma = \det P = -1$.

Note that for reflections and inversions, performing the operation twice yields the original vector, i.e. $\sigma^2 = I$, $P^2 = I$.

3.3 Projection operators

$P$ is a parallel projection operator onto a vector $\vec{u}$ if

$$P \vec{u} = \vec{u} \quad \text{and} \quad P \vec{v} = 0$$

where $\vec{v}$ is any vector orthogonal to $\vec{u}$, i.e. $\vec{v} \cdot \vec{u} = 0$. Similarly $Q$ is an orthogonal projection to $\vec{u}$ if

$$Q \vec{u} = 0 \quad \text{and} \quad Q \vec{v} = \vec{v}$$

so that $Q = I - P$. Suitable operators are (exercise: check this!)

$$P_{ij} = \frac{u_i u_j}{u^2} \quad \text{and} \quad Q_{ij} = \delta_{ij} - \frac{u_i u_j}{u^2}$$

These have the properties

$$P^2 = P, \quad Q^2 = Q, \quad PQ = QP = 0$$

They’re also unique. For example, if there exists another operator $T$ with the same properties as $P$, i.e. $T \vec{u} = \vec{u}$ and $T \vec{v} = 0$, then for any vector $\vec{w} \equiv \mu \vec{u} + \nu \vec{v} + \lambda \vec{u} \times \vec{v}$ we have

$$(P - T) \vec{w} = (\mu \vec{u} + 0 + 0) - (\mu \vec{u} + 0 + 0) = 0$$

because

$$P (\vec{u} \times \vec{v}) = P_{ij} (\vec{u} \times \vec{v}) = (u_i u_j / u^2) \epsilon_{jkl} u_k v_l = 0$$

This holds for all vectors $\vec{w}$, so $T = P$. 

3.4 Active and passive transformations

- Rotation of a vector in a fixed basis is called an *active transformation*
  \[ \mathbf{x} \to \mathbf{y} \text{ with } y_i = R_{ij} x_j \text{ in the } \{e_i\} \text{ basis} \]

- Rotation of the basis whilst keeping the *vector fixed* is called a *passive transformation*
  \[ \{e_i\} \to \{e'_i\} \text{ and } x_i \to x'_i = \ell_{ij} x_j \]

If we set \( R_{ij} = \ell_{ij} \), then numerically \( y_i = x'_i \).

Consider a simple example of both types of rotation:

**Rotation of a vector about the z-axis**

\[
R_{ij}(\theta, e_3) = \begin{pmatrix}
\delta_{ij} \cos \theta + (1 - \cos \theta) \delta_{i3} \delta_{j3} - \epsilon_{ijk} \delta_{k3} \sin \theta \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

where we used \( n_i = \delta_{i3} \).

This is an *active* rotation through angle \( \theta \).

**Rotation of the basis about the z-axis**

\[
e'_i = \ell_{ij} e_j \equiv R_{ij} e_j
\]

In components

\[
e'_1 = \cos \theta e_1 - \sin \theta e_2 \\
e'_2 = \sin \theta e_1 + \cos \theta e_2 \\
e'_3 = e_3
\]

This is a *passive* rotation through angle \( -\theta \).

We conclude that

An active rotation of the vector \( \mathbf{x} \) through angle \( \theta \) is equivalent to a passive rotation of the basis vectors by an *equal and opposite* amount.

Colloquially, rotating a vector in one direction is equivalent to rotating the basis in the opposite direction.

The general case can be built from three rotations (Euler angles).
Chapter 4
The inertia tensor

4.1 Angular momentum and kinetic energy

Suppose a rigid body of arbitrary shape rotates with angular velocity \( \omega = \omega n \) about a fixed axis, parallel to the unit vector \( n \), which passes through the origin.

Consider a small element of mass \( dm \) in volume \( dV \) at the point \( P \), with position vector \( \vec{r} \) relative to \( O \).

If the rigid body has density (mass per unit volume) \( \rho(\vec{r}) \), then \( dm = \rho dV \).

The velocity of the element is
\[
\vec{v} = \omega \times \vec{r}
\]

We can see this geometrically from the figure.

The distance \( |\delta r| \) moved in time \( \delta t \), is
\[
|\delta r| = r \sin \phi \delta \theta = \delta \theta |n \times r|
\]

So its velocity is
\[
\vec{v} = \frac{\delta \vec{r}}{\delta t} = \omega \times \vec{r} \quad \text{where} \quad \omega = \frac{\delta \theta}{\delta t} n
\]

Alternatively, we can use the rotation tensor \( R(\theta, n) \)
\[
\delta x_i = R_{ij}(\delta \theta, n) x_j - x_i = [-\epsilon_{ijk} n_k \delta \theta + O((\delta \theta)^2)] x_j
\]
\[
\Rightarrow \quad \frac{\delta x_i}{\delta t} = (n \times r)_i \frac{\delta \theta}{\delta t}
\]

which again gives \( \vec{v} = \omega \times \vec{r} \).
4.1.1 Angular momentum

The angular momentum $L$ of a point particle of mass $m$ at position $\mathbf{r}$ moving with velocity $\mathbf{v} = \dot{\mathbf{r}}$ is

$$L = \mathbf{r} \times \mathbf{p},$$

where the momentum $\mathbf{p} = m\mathbf{v}$.

The angular momentum $dL$ of an element of mass $dm = \rho \, dV$ at $\mathbf{r}$ is

$$dL = \rho(r) \, dV \, r \times v.$$ 

The angular momentum of the entire rotating body is then

$$L = \int_{\text{body}} \rho \mathbf{r} \times (\mathbf{\omega} \times \mathbf{r}) \, dV.$$ 

In components

$$L_i = \int_{\text{body}} \rho \epsilon_{ijk} x_j (\epsilon_{klm} \omega_l x_m) \, dV$$

$$= \int_{\text{body}} \rho (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) x_j \omega_l x_m \, dV$$

$$= \int_{\text{body}} \rho (r^2 \omega_i - x_i x_j \omega_j) \, dV $$ 

Thus

$$L_i = I_{ij} \omega_j \quad \text{with} \quad I_{ij} = \int_{\text{body}} \rho (r^2 \delta_{ij} - x_i x_j) \, dV$$

The geometric quantity $I(O)$ (where $O$ refers to the origin) is called the inertia tensor. It is a tensor because $L$ is pseudovector, $\mathbf{\omega}$ is a pseudovector, and hence from the quotient theorem $I$ is a tensor.

$I_{ij}$ is symmetric and independent of the rotation axis $\mathbf{n}$, it’s a property of the body.

4.1.2 Kinetic energy

The kinetic energy, $dT$, of an element of mass $dm$ is

$$dT = \frac{1}{2} (\rho \, dV) \left| \mathbf{\omega} \times \mathbf{r} \right|^2.$$ 

The kinetic energy of the body is then

$$T = \frac{1}{2} \int_{\text{body}} \rho (\epsilon_{ijk} \omega_j x_k) (\epsilon_{ilm} \omega_l x_m) \, dV$$

$$= \frac{1}{2} \int_{\text{body}} \rho (\delta_{ij} \delta_{km} - \delta_{jm} \delta_{il}) \omega_j x_k \omega_l x_m \, dV$$

$$= \frac{1}{2} \int_{\text{body}} \rho (\omega^2 r^2 - (\mathbf{r} \cdot \mathbf{\omega})^2) \, dV$$

$$= \frac{1}{2} \int_{\text{body}} \rho (r^2 \delta_{ij} - x_i x_j) \, dV \omega_i \omega_j$$

which gives

$$T = \frac{1}{2} \, I_{ij} \omega_i \omega_j = \frac{1}{2} \frac{L \cdot \omega}{m}$$

\(^1\)Some of my colleagues insist that $I(O)$ should be called the moment of inertia tensor. You will see both names in books.
Alternative (more familiar) forms

Recalling that the angular velocity may be written as \( \omega = \omega_n \), consider

\[
\mathbf{n} \cdot \mathbf{L} = n_i I_{ij} \omega_j = I_{ij} n_i n_j \omega \equiv I^{(n)} \omega
\]

\( L \cdot \mathbf{n} \) is the component of angular momentum parallel to the axis \( \mathbf{n} \), and

\[
I^{(n)} = I_{ij} n_i n_j = \int_{\text{body}} \rho \left(r^2 - (r \cdot \mathbf{n})^2\right) \, dV \equiv \int_{\text{body}} \rho r_{\perp}^2 \, dV
\]

is the moment of inertia about \( \mathbf{n} \), with \( r_{\perp} \) the perpendicular distance from the \( \mathbf{n} \)-axis.

Similarly for the kinetic energy, so that

\[
\begin{align*}
L^{(n)} &= I^{(n)} \omega \\
\frac{T}{L} &= \frac{1}{2} I^{(n)} \omega^2
\end{align*}
\]

with \( L^{(n)} = L \cdot \mathbf{n} \), \( I^{(n)} = I_{ij} n_i n_j \)

**Example:** Consider a cube of side \( a \) of constant density \( \rho \) and mass \( M = \rho a^3 \)

\[
I_{ij}(O) = \int \rho \left(r^2 \delta_{ij} - x_i x_j\right) \, dV
\]

In this case

\[
\begin{align*}
I_{11} &= \rho \int_0^a \int_0^a \int_0^a \left\{x^2 + y^2 + z^2 - x^2\right\} \, dx \, dy \, dz \\
&= \rho \left[ \frac{1}{3} y^3 xz + \frac{1}{3} z^3 xy \right]_0^a = \frac{2}{3} \rho a^5 = \frac{2}{3} Ma^2 \\
I_{12} &= \rho \int_0^a \int_0^a \int_0^a (-xy) \, dx \, dy \, dz \\
&= -\rho \left[ \frac{1}{2} x^2 y^2 \right]_0^a = -\frac{1}{4} \rho a^5 = -\frac{1}{4} Ma^2
\end{align*}
\]

Similarly for the other components. By symmetry,

\[
I(O) = Ma^2 \begin{pmatrix}
\frac{2}{3} & -\frac{1}{4} & -\frac{1}{4} \\
-\frac{1}{4} & \frac{2}{3} & -\frac{1}{4} \\
-\frac{1}{4} & -\frac{1}{4} & \frac{2}{3}
\end{pmatrix}
\]

**4.1.3 The parallel axes theorem**

It’s often more useful, and also simpler, to find the inertia tensor about the centre of mass \( G \), rather than about an arbitrary point \( O \). There is, however, a simple relationship between them.

Taking \( O \) to be the origin, and \( \overrightarrow{OG} = \mathbf{R} \), we have \( \mathbf{r} = \mathbf{R} + \mathbf{r}' \), giving

\[
I_{ij}(O) = \int \rho(r') \left(r'^2 \delta_{ij} - x_i x_j\right) \, dV
\]

\[
= \int \rho'(r') \left\{|\mathbf{R} + \mathbf{r}'|^2 \delta_{ij} - (X_i + x_i') (X_j + x_j')\right\} \, dV'
\]
In the above, \( \rho(r) = \rho(R + r') \equiv \rho'(r') \), and we changed integration variables to \( r' \).

Expanding the integrand and using the definition of \( G \), namely \( \int \rho'(r') r' \, dV' = 0 \), we get

\[
I_{ij}(O) = \int \rho'(r') \left\{ (R^2 + r'^2) \delta_{ij} - X_i X_j - x_i' x_j' \right\} \, dV'
\]

Hence

\[
I_{ij}(O) = I_{ij}(G) + M (R^2 \delta_{ij} - X_i X_j)
\]

where \( M = \int \rho'(r') \, dV' \) is the total mass of the body. This is a general result; given \( I(G) \) we can easily find the inertia tensor about any other point.

The general result above is often called the parallel axes theorem. However, the parallel axes theorem technically refers to the inertia tensor about an axis \( n \) through \( G \), which is parallel to the original axis \( n \) through \( O \), as in the figure above, so that

\[
I^{(n)}(O) = I^{(n)}(G) + M R_\perp^2
\]

where \( R_\perp \equiv \sqrt{R^2 - (R \cdot n)^2} \) is the \( \perp \) distance of \( G \) from the \( n \) axis through \( O \).

**Example (revisited):** In our previous example, the centre of mass \( G \) is at the centre of the cube with position vector \( R = \left( \frac{1}{2}a, \frac{1}{2}a, \frac{1}{2}a \right) \). Using Cartesian coordinates *with their origin at* \( G \)

\[
I_{11}(G) = \rho \int_{-a/2}^{a/2} dx \, dy \, dz \left\{ (x^2 + y^2 + z^2) - x^2 \right\}
\]

\[
= \rho \left\{ \frac{1}{3} \left[ y^3 \right]_{-a/2}^{a/2} - \frac{1}{2} \left[ x^2 \right]_{-a/2}^{a/2} \left[ z \right]_{-a/2}^{a/2} + \frac{1}{3} \left[ z^3 \right]_{-a/2}^{a/2} - \frac{1}{2} \left[ x \right]_{-a/2}^{a/2} \left[ y \right]_{-a/2}^{a/2} \right\}
\]

\[
= \rho \left\{ \frac{1}{3} \cdot 2(a/2)^2 2(a/2) 2(a/2) \cdot 2 \right\} = \frac{1}{6} \rho a^5 = \frac{1}{6} Ma^2
\]

and

\[
I_{12}(G) = \rho \int_{-a/2}^{a/2} dx \, dy \, dz (-xy)
\]

\[
= 0 \quad \text{because} \quad \left[ \frac{1}{2} x^2 \right]_{-a/2}^{a/2} = 0
\]

Similarly for all the other components.

The inertia tensor about the centre of mass is then

\[
I_{ij}(G) = Ma^2 \left( \begin{array}{ccc} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{6} \end{array} \right)_{ij} = \frac{Ma^2}{6} \delta_{ij}
\]

and

\[
M (R^2 \delta_{ij} - X_i X_j) = Ma^2 \left( \begin{array}{ccc} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{2} \end{array} \right)_{ij}
\]

Since \( \frac{1}{2} + \frac{1}{6} = \frac{2}{3} \), this reproduces our previous result for \( I_{ij}(O) \).
4.1.4 Diagonalisation of rank-two tensors

Question: are there any directions for \( \omega \) such that \( L \) is parallel to \( \omega \)?

If so, then \( L = \lambda \omega \), and hence

\[
(I_{ij} - \lambda \delta_{ij}) \omega_j = 0
\]

For a non-trivial solution of these three simultaneous linear equations (i.e. \( \omega \neq 0 \)), we must have \( \det (I_{ij} - \lambda \delta_{ij}) = 0 \). Expanding the determinant, or writing it as

\[
\det (I_{ij} - \lambda \delta_{ij}) = \frac{1}{6} \epsilon_{ijk} \epsilon_{lmn} (I_{il} - \lambda \delta_{il}) (I_{jm} - \lambda \delta_{jm}) (I_{kn} - \lambda \delta_{kn}) = 0
\]

and then expanding gives

\[
P - Q\lambda + R\lambda^2 - \lambda^3 = 0
\]

where

\[
P = \frac{1}{6} \epsilon_{ijk} \epsilon_{lmn} I_{il} I_{jm} I_{kn}
\]
\[
Q = \frac{1}{6} \epsilon_{ijk} \epsilon_{lmn} (\delta_{il} I_{jm} I_{kn} + I_{il} \delta_{jm} I_{kn} + I_{il} I_{jm} \delta_{kn})
\]
\[
= \frac{1}{6} (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) I_{jm} I_{kn} \times 3
\]
\[
= \frac{1}{2} [\text{Tr} I]^2 - \text{Tr} (I^2)
\]
\[
R = \frac{1}{6} \epsilon_{ijk} \epsilon_{lmn} (\delta_{il} \delta_{jm} I_{kn} + \delta_{il} I_{jm} \delta_{kn} + I_{il} \delta_{jm} \delta_{kn})
\]
\[
= \frac{1}{6} 2 \delta_{kn} I_{kn} \times 3
\]
\[
= \text{Tr} I
\]

For any tensor \( A \), we know that \( \det A \) and \( \text{Tr} A \) are invariant (the same in any basis), hence the quantities \( P, Q, R \) are invariants of the tensor \( I \) (i.e. their values are also the same in any basis).

The three values of \( \lambda \) (i.e. the solutions of the cubic equation) are the eigenvalues of the rank-two tensor, and the vectors \( \omega \) are its eigenvectors.\(^2\)

We will generally call the eigenvectors \( e \).

**Eigenvectors and eigenvalues:** If we take \( I \omega = \lambda \omega \) (or \( I_{ij} \omega_j = \lambda \omega_i \)), and multiply on the left by a transformation matrix \( L \), we obtain (in matrix notation)

\[
(L I L^T)_{ij} \omega_j = \lambda (L \omega)_i \quad \Rightarrow \quad I'_{ij} (L \omega)_j = \lambda (L \omega)_i
\]

In the primed basis, we have by definition \( I'_{ij} \omega'_j = \lambda' \omega'_i \). Comparing with the second equation above, we see that eigenvectors \( \omega \) are indeed vectors, i.e. they transform as vectors: \( \omega'_i = \ell_{ij} \omega_j \).

Similarly, eigenvalues are scalars, i.e. they transform as scalars: \( \lambda' = \lambda \).

Note that only the direction (up to a \( \pm \) sign) of the eigenvectors is determined by the eigenvalue equation, the magnitude is arbitrary.

The answer to our question is that we must find the eigenvalues \( \lambda^{(i)}, i = 1, 2, 3 \) and the corresponding eigenvectors \( \omega^{(i)} \), whence \( L^{(i)} = \lambda^{(i)} \omega^{(i)} \) (no sum on \( i \)).

\(^2\)Yes, the language is indeed the same as for matrices.
Eigenvalues and eigenvectors of a real symmetric tensor

Theorems

(i) All of the eigenvalues of a real symmetric matrix are real.

(ii) The eigenvectors corresponding to distinct eigenvalues are orthogonal.

If a subset of the eigenvalues is degenerate (eigenvalues are equal), the corresponding eigenvectors can be chosen to be orthogonal because:

(a) the eigenvector subspace corresponding to the degenerate eigenvalues is orthogonal to the other eigenvectors;

(b) within this subspace, the eigenvectors can be chosen to be orthogonal by the Gram-Schmidt procedure.

Proofs will not be given here – see books or lecture notes from mathematics courses.

Diagonalisation of a real symmetric tensor

Let $T$ be a real second-rank symmetric tensor with real eigenvalues $\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}$ and orthonormal eigenvectors $\ell^{(1)}, \ell^{(2)}, \ell^{(3)}$, so that $T \ell^{(i)} = \lambda^{(i)} \ell^{(i)}$ (no summation) and $\ell^{(i)} \cdot \ell^{(j)} = \delta_{ij}$. Let the matrix $L$ have elements

$$
\ell_{ij} = \ell^{(i)}_j = \begin{pmatrix}
\ell^{(1)}_1 & \ell^{(1)}_2 & \ell^{(1)}_3 \\
\ell^{(2)}_1 & \ell^{(2)}_2 & \ell^{(2)}_3 \\
\ell^{(3)}_1 & \ell^{(3)}_2 & \ell^{(3)}_3
\end{pmatrix}
$$

I.e the $i$th row of $L$ is the $i$th eigenvector of $T$. $L$ is an orthogonal matrix

$$(LL^T)_{ij} = \ell_{im} \ell_{jm} = \ell^{(i)}_m \ell^{(j)}_m = \delta_{ij}$$

We can always choose the normalised eigenvectors $\ell^{(i)}$ to form a right-handed basis:

$$\det L = \epsilon_{ijk} \ell_{i1} \ell_{j2} \ell_{k3} = \ell^{(3)} \cdot (\ell^{(1)} \times \ell^{(2)}) = +1$$

With this choice, $L$ is a rotation matrix which transforms $S$ to $S'$.

The tensor $T$ transforms as (summing over the indices $p, q$ only)

$$T'_{ij} = \ell_{ip} \ell_{jq} T_{pq} = \ell^{(i)}_p T_{pq} \ell^{(j)}_q = \ell^{(i)}_p \lambda^{(j)} \ell^{(j)}_p$$

or

$$T'_{ij} = \lambda^{(j)} \delta_{ij} = \begin{pmatrix}
\lambda^{(1)} & 0 & 0 \\
0 & \lambda^{(2)} & 0 \\
0 & 0 & \lambda^{(3)}
\end{pmatrix}
$$

Thus we have found a basis or frame, $S'$, in which the tensor $T$ takes a diagonal form; the diagonal elements are the eigenvalues of $T$.

\[3\] Thus tensors may be diagonalised in the same way as matrices.
The inertia tensor

When studying rigid body dynamics, it’s (usually) best to work in a basis in which the inertia tensor is diagonal. The eigenvectors of $I$ define the principal axes of the tensor. In this (primed) basis

$$ I' = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix} $$

where the (positive) quantities $A$, $B$, $C$ are called the principal moments of inertia.

In this basis, the angular momentum and kinetic energy take the form

$$ L = A \omega'_1 e'_1 + B \omega'_2 e'_2 + C \omega'_3 e'_3 $$
$$ T = \frac{1}{2} \left( A \omega'^2_1 + B \omega'^2_2 + C \omega'^2_3 \right) $$

For a free body (i.e. no external forces), $L$ and $T$ are conserved (time-independent), but $\omega$ will in general be time dependent.

Note that the angular momentum $L$ is parallel to the angular velocity $\omega$ when both are parallel to one of the principal axes of the inertia tensor. For example, if $\omega = \omega'_1 e'_1$ then

$$ L = A \omega'_1 e'_1 $$

In this case, the body is rotating about a principal axis which passes through its centre of mass.

Notes:

- The principal axes basis is used in the Lagrangian Dynamics course to study the rotational motion of a free rigid body in the Newtonian approach to dynamics, and the motion of a spinning top with principal moments $(A, A, C)$ in the Lagrangian approach.

- The principal axes basis/frame is ‘fixed to the body’, i.e. it moves with the rotating body, and is therefore a non-inertial frame.

The inertia ellipsoid

The remainder of this chapter is non-examinable, but it may be of interest to those taking the course Lagrangian Dynamics.

A geometrical picture is provided by the inertia ellipsoid, which is defined by

$$ I_{ij} \omega_i \omega_j = 1 $$

A factor of $\sqrt{2T}$ is absorbed into $\omega$ by convention: $\omega_i \rightarrow \omega_i / \sqrt{2T}$. 
In the principal axes basis, where $\omega'_i = \ell_{ij} \omega_j$, we have

$$A \omega'_1^2 + B \omega'_2^2 + C \omega'_3^2 = 1$$

This is called the normal form. It describes an ellipsoid because $A$, $B$, $C$ are all positive. (This follows from the definition, for example $A = \int \rho (y^2 + z^2) \, dV$.)

[The angular momentum $L$ is labelled $\tilde{h}$ in the figure on the right. To be fixed ...]

In any basis, a small displacement $\omega \to \omega + d\omega$ on the ellipsoidal surface at the point $P$, with normal $\vec{n}$, obeys

$$I_{ij} \omega_i \, d\omega_j = L_j \, d\omega_j = 0$$

for all $d\omega$.

Therefore $L$ is orthogonal to $d\omega$ and parallel to $\vec{n}$, i.e. $L$ is always orthogonal to the surface of the ellipsoid at $P$.

In the principal axes basis

$$L = A \omega'_1 e'_1 + B \omega'_2 e'_2 + C \omega'_3 e'_3$$

The directions for which $L$ is parallel to $\omega$ are obviously the directions of the principal axes of the ellipsoid. For example, if $\omega = \omega'_1 e'_1$, then

$$L = A \omega'_1 e'_1$$

In this case, the body is rotating about a principal axis which passes through its center of mass.

This gives a ‘geometrical’ answer to our original question.

Notes:

- If two principal moments are identical ($A, A, C$), the ellipsoid becomes a spheroid.
- If all three principal moments are identical the ellipsoid becomes a sphere, and $L$ is always parallel to $\omega$. 

33
Chapter 5

Taylor expansions

Taylor expansion is one of the most important and fundamental techniques in the physicist’s toolkit. It allows a differentiable function to be expressed as a power series in its argument(s). This is useful when approximating a function, it often allows the problem to be ‘solved’ in some range of interest, and it’s used in deriving fundamental differential equations. We shall use the expression ‘Taylor’s Theorem’ interchangeably with ‘Taylor expansion’.

We assume familiarity with Taylor expansions of functions of one variable, so we won’t go through this material on the blackboard in the lectures. However, for completeness, we include some comprehensive notes that you should read and/or work through in your own time.

You will most likely have seen the multivariate Taylor expansion beyond leading order in previous courses, but probably not in the way it’s done here...

5.1 The one-dimensional case

Let \( f(x) \) have a continuous \( m^{\text{th}} \)-order derivative \( f^{(m)}(x) \) in \( a \leq x \leq b \), so that

\[
\int_a^{x_1} f^{(m)}(x_0) \, dx_0 = f^{(m-1)}(x_1) - f^{(m-1)}(a)
\]

Integrating a total of \( m \) times gives

\[
\int_a^{x_m} \cdots \int_a^{x_1} f^{(m)}(x_0) \, dx_0 \cdots dx_{m-1}
\]

\[
= \int_a^{x_m} \cdots \int_a^{x_2} \left[ f^{(m-1)}(x_1) - f^{(m-1)}(a) \right] \, dx_1 \cdots dx_{m-1}
\]

\[
= \int_a^{x_m} \cdots \int_a^{x_3} \left[ f^{(m-2)}(x_2) - f^{(m-2)}(a) - (x_2 - a)f^{(m-1)}(a) \right] \, dx_2 \cdots dx_{m-1}
\]

\[
= f(x_m) - f(a) - (x_m - a)f'(a) - \frac{1}{2!}(x_m - a)^2 f''(a) - \cdots
\]

\[
- \frac{1}{(m-1)!}(x_m - a)^{m-1} f^{(m-1)}(a)
\]
where we used the basic integral
\[ \int_a^x (y - a)^{n-1} \, dy = \frac{1}{n} (x - a)^n \]
Now let \( x_m \to x \), which gives
\[
f(x) = f(a) + (x - a)f'(a) + \frac{1}{2!}(x - a)^2 f'' + \ldots + \frac{1}{(m-1)!} (x - a)^{m-1} f^{(m-1)}(a) + R_m ,
\]
where is \( n! = n(n-1) \cdots 1 \) is the usual factorial function, with \( 0! = 1 \), and the remainder \( R_m \) is
\[
R_m = \int_a^x \cdots \int_a^{x_1} f^{(m)}(x_0) \, dx_0 \cdots dx_{m-1}
\]
But from the mean value theorem applied to \( f^{(m)} \), we have
\[
\int_a^x f^{(m)}(x_0) \, dx_0 = (x - a) f^{(m)}(\xi) , \quad a \leq \xi \leq x
\]
which gives the “Lagrange form” for the remainder
\[
R_m(x) = \frac{1}{m!} (x - a)^m f^{(m)}(\xi) , \quad a \leq \xi \leq x
\]
Notes
- We can repeat the proof above for \( \int_c^a \cdots \) where \( x \in [c, a] \) with \( c \leq a \leq b \).
  Since nothing changes, we can talk about expansion in a region about \( x = a \).
- If \( \lim_{m \to \infty} R_m \to 0 \) (as usually assumed here), we have an infinite series
\[
f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a) (x - a)^n
\]
This is the Taylor expansion of \( f(x) \) about \( x = a \).
The set of \( x \) values for which the series converges is called the region of convergence of the Taylor expansion.
- If \( a = 0 \), then
\[
f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n
\]
The Taylor expansion about \( x = 0 \) is called the Maclaurin expansion.
**Physicist’s “proof”** We can bypass the formal proof above by assuming that a power series expansion of \( f(x) \) exists (i.e. the polynomials \( x^n \) form a complete basis) so that
\[
f(x) = \sum_{n=0}^{\infty} a_n x^n.
\]
Now differentiate \( p \) times, set \( x = 0 \), and equate coefficients for each \( p \), to obtain
\[
f^{(p)}(0) = 0 + \cdots + p! a_p + \cdots + 0 \quad \text{which gives} \quad a_p = \frac{1}{p!} f^{(p)}(0) \quad \text{(as before)}
\]
5.1.1 Examples

Example 1: Expand the function

\[ f(x) = \sin x \]

about \( x = 0 \). We need

\[
\begin{align*}
  f^{(2n)}(0) &= (-1)^n \sin 0 = 0 \\
  f^{(2n+1)}(0) &= (-1)^n \cos 0 = (-1)^n
\end{align*}
\]

Now, since \(|f^{(m)}(\xi)| \leq 1\), then, for fixed \( x \),

\[
|R_m| = \frac{1}{m!} x^m |f^{(m)}(\xi)| \leq \frac{1}{m!} x^m \to 0
\]

Therefore

\[
\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots
\]

This ‘small \( x \)’ expansion is shown in the figure.

Example 2: Expand the function

\[ f(x) = (1 + x)^\alpha \]

about \( x = 0 \). In this case

\[
f^{(n)}(0) = \alpha(\alpha - 1) \cdots (\alpha - n + 1) \equiv \frac{\alpha!}{(\alpha - n)!}
\]

giving

\[
f(x) = \sum_{n=0}^{\infty} \frac{\alpha!}{n!(\alpha - n)!} x^n \equiv \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n
\]

The Taylor expansion includes the binomial expansion, \( \alpha \) need not be a +ve integer.

Example 3: a ‘problem’ case Consider, for example, the well-behaved function

\[ f(x) = \exp \left( -\frac{1}{x^2} \right) \]

Now \( f(0) = 0 \), and \( f^{(n)}(0) = 0 \ \forall n \), so

\[
f(x) = 0 + 0 + 0 + \ldots = 0 \ \forall x
\]

Beware of essential singularities – not all functions with an infinite number of derivatives can be expressed as a Taylor series. See Laurent Series in Honours Complex Variables next semester.
5.1.2 A precursor to the three-dimensional case

If we regard \( f(x + a) \equiv g(a) \) temporarily as a function of \( a \) only, we can write \( g(a) \) as a Maclaurin series in powers of \( a \)

\[
f(x + a) \equiv g(a) = \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0) a^n, \quad g^{(n)}(0) = f^{(n)}(x)
\]

which we can rewrite as

\[
f(x + a) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) a^n \equiv \exp \left( a \frac{d}{dx} \right) f(x)
\]

The differential operator \( \exp \left( a \frac{d}{dx} \right) \) is defined by its power-series expansion. This is the form that we shall generalise to three dimensions in an elegant way.

It can also be obtained by first defining

\[
F(t) \equiv f(x + at)
\]

which we regard as a function of \( t \). We need the expansion in powers of \( t \) about \( t = 0 \), namely

\[
F(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} F^{(n)}(0)
\]  \hspace{1cm} (5.1)

Noting that \( F^{(n)}(0) = a^n f^{(n)}(x) \), and setting \( t = 1 \), we find

\[
f(x + a) = F(1) = \sum_{n=0}^{\infty} \frac{1}{n!} a^n f^{(n)}(x)
\]

as before.

5.2 The three-dimensional case

With this trick, we can use the one-dimensional result to find the Taylor expansion of \( \phi(r + a) \) in powers of \( a \) about the point \( r \). Let

\[
F(t) \equiv \phi(x + ta) \equiv \phi(u) \quad \text{(where we defined} \quad u = r + ta) \quad \Rightarrow \quad F(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} F^{(n)}(0)
\]

where we used equation (5.1) above. We want \( \phi(r + a) \) which is \( F(1) \).

Using the chain rule, the first derivative of \( F(t) \) with respect to \( t \) is

\[
F^{(1)}(t) = \frac{\partial \phi(u)}{\partial u_i} \frac{\partial u_i}{\partial t} = \frac{\partial \phi}{\partial u_i} a_i = (a \cdot \nabla u) \phi(u)
\]
where we used
\[ \frac{\partial u_i}{\partial t} = \frac{\partial}{\partial t} (x_i + ta_i) = a_i \]
and defined \( a \cdot \nabla u \equiv a_i \frac{\partial \phi}{\partial u_i} \)

The \( n \)th derivative of \( F(t) \) is
\[ F^{(n)}(t) = (a \cdot \nabla u)^n \phi(u) \]
and hence \( F^{(n)}(0) = (a \cdot \nabla_r)^n \phi(r) \) (5.2)

For \( F(1) \) we have
\[
\phi(r + a) = \sum_{n=0}^{\infty} \frac{1}{n!} (a \cdot \nabla_r)^n \phi(r) \equiv \exp (a \cdot \nabla_r) \phi(r)
\]

This is the Taylor expansion of a scalar field \( \phi(r) \) in three dimensions, in a rather elegant form.

Generalisation to an arbitrary tensor field is easy. Simply replace \( \phi(r) \) by \( T_{ij} \ldots (r) \) in the above expression.

**Example:** Find the Taylor expansion of \( \phi(r + a) = \frac{1}{|r + a|} \) for \( r \gg a \).

Since \( \phi(r) = 1/|r| = 1/r \), we have
\[
\frac{1}{|r + a|} = \sum_{n=0}^{\infty} \frac{1}{n!} (a \cdot \nabla_r)^n \frac{1}{r}
\]
\[
= \frac{1}{r} + \frac{1}{1!} (a_i \partial_i) \frac{1}{r} + \frac{1}{2!} (a_i \partial_i) (a_j \partial_j) \frac{1}{r} + \cdots
\]
\[
= \frac{1}{r} - \frac{a \cdot r}{r^3} + \frac{3(a \cdot r)^2 - a^2 r^2}{2r^5} + O \left( \frac{1}{r^4} \right)
\]

Exercise: check this explicitly.

This result is used in the multipole expansion in electrostatics and magnetostatics.
Chapter 6
Curvilinear coordinates

As you have seen in previous courses, it is often convenient to work with coordinate systems other than Cartesian coordinates \( \{x_i\} \), i.e. \((x_1, x_2, x_3)\) or \((x, y, z)\). Since most of the material in this chapter should be familiar, albeit in a slightly different formalism, we won’t go through it on the blackboard – as suggested by the 2018/19 class in their course feedback! Instead, you should read through it on your own.

For example, spherical polar coordinates \((r, \theta, \phi)\) are defined by:

\[
\begin{align*}
  x &= r \sin \theta \cos \phi \\
  y &= r \sin \theta \sin \phi \\
  z &= r \cos \theta
\end{align*}
\]

We shall set up a formalism to deal with rather general coordinate systems, of which spherical polars are a very important example.

Suppose we make a transformation from the Cartesian coordinates \((x_1, x_2, x_3)\) to the variables \((u_1, u_2, u_3)\), which are functions of the \(\{x_i\}\)

\[
\begin{align*}
  u_1 &= u_1(x_1, x_2, x_3) \\
  u_2 &= u_2(x_1, x_2, x_3) \\
  u_3 &= u_3(x_1, x_2, x_3)
\end{align*}
\]

If the variables \(\{u_i\}\) are single-valued functions of the variables \(\{x_i\}\), then we can make the inverse transformations,

\[
x_i = x_i(u_1, u_2, u_3) \quad \text{for } i = 1, 2, 3,
\]

except possibly at certain points.

A point may be specified by its Cartesian coordinates \(\{x_i\}\), or its curvilinear coordinates \(\{u_i\}\). We may define the curvilinear coordinates by equations giving \(\{x_i\}\) as functions of \(\{u_i\}\), or vice-versa.\(^1\)

\(^1\)Sometimes the curvilinear coordinates are called \((u, v, w)\), just as Cartesians are called \((x, y, z)\).
• For Cartesian coordinates, the surfaces ‘\(x_i = \text{constant}\)’ (\(i = 1, 2, 3\)) are \textit{planes}, with (constant) normal vectors \(\mathbf{e}_i\) (the Cartesian basis vectors) \textit{intersecting at right angles}.

• For general curvilinear coordinates, the surfaces ‘\(u_i = \text{constant}\)’ do \textit{not} have constant normal vectors, nor do they intersect at right angles. For example, in 2-D, we might have the situation illustrated in the diagram on the right.

From the definition of spherical polar coordinates \((r, \theta, \phi)\), we have

\[
\begin{align*}
  r &= \sqrt{x^2 + y^2 + z^2} \\
  \theta &= \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \\
  \phi &= \tan^{-1}\left(\frac{y}{x}\right).
\end{align*}
\]

The surfaces of constant \(r\), \(\theta\), and \(\phi\) are

- \(r = \text{constant} \Rightarrow \) spheres centred at the origin
- \(\theta = \text{constant} \Rightarrow \) cones of semi-angle \(\theta\) and axis along the \(z\)-axis
- \(\phi = \text{constant} \Rightarrow \) planes passing through the \(z\)-axis

Not all of these surfaces are planes, but they \textit{do} intersect at right angles. The angle \(\theta\) is undefined at the origin; \(\phi\) is undefined on the \(z\)-axis.

6.1 Orthogonal curvilinear coordinates

If the coordinate surfaces (surfaces of constant \(u_i\)), intersect at right angles, as in the example of spherical polars, the curvilinear coordinates are said to be \textit{orthogonal}.

6.1.1 Scale factors and basis vectors

Let point \(P\) have position vector \(\mathbf{r} = r(u_1, u_2, u_3)\). If we change \(u_1\) by \(du_1\) (with \(u_2\) and \(u_3\) fixed), then \(\mathbf{r} \to \mathbf{r} + d\mathbf{r}\), with

\[
dr = \frac{\partial r}{\partial u_1} du_1 = h_1 e_1 du_1
\]

where we have defined the \textit{scale factor} \(h_1\) and the \textit{unit vector} \(e_1\) by

\[
h_1 = \left| \frac{\partial r}{\partial u_1} \right| \quad \text{and} \quad e_1 = \frac{1}{h_1} \frac{\partial r}{\partial u_1}
\]

- The \textit{scale factor} \(h_1\) gives the \textit{length} \(h_1 du_1\) of \(d\mathbf{r}\) when we change \(u_1 \to u_1 + du_1\).
- \(e_1\) is a \textit{unit vector} in the direction of increasing \(u_1\) (with \textit{fixed} \(u_2\) and \(u_3\)).

Similarly, we can define \(h_i\) and \(e_i\) for \(i = 2\) and \(3\).
In general, if we change a single $u_i$, keeping the other two fixed, we have

$$\frac{\partial r}{\partial u_i} = h_i \xi_i \quad i = 1, 2, 3 \quad \text{[no sum on $i$]}$$

- The unit vectors $\{\xi_i\}$ are in general not constant vectors – their directions depend on the position vector $r$, and hence on the curvilinear coordinates $\{u_i\}$. [They should perhaps be called $\{\varepsilon_{-u_i}\}$ or $\{\varepsilon_u, \varepsilon_v, \varepsilon_w\}$ to avoid confusion with Cartesian basis vectors.]

- If the curvilinear unit vectors satisfy $\xi_i \cdot \xi_j = \delta_{ij}$, the $\{u_i\}$ are said to be orthogonal curvilinear coordinates, and the three unit vectors $\{\xi_i\}$ form an orthonormal basis. We will always choose the $u_i$ to give a right-handed basis.

6.1.2 Examples of orthogonal curvilinear coordinates (OCCs)

**Cartesian coordinates:**

$$r = x \xi_x + y \xi_y + z \xi_z \quad \Rightarrow \quad h_x \xi_x = \frac{\partial r}{\partial x} = \xi_x, \text{ etc.}$$

The scale factors are all unity, and each individual unit vector points in the same direction everywhere.

**Spherical polar coordinates:** $u_1 = r, u_2 = \theta, u_3 = \phi$ (in that order)

$$r = r \sin \theta \cos \phi \xi_x + r \sin \theta \sin \phi \xi_y + r \cos \theta \xi_z$$

$$\frac{\partial r}{\partial r} = \sin \theta \cos \phi \xi_x + \sin \theta \sin \phi \xi_y + \cos \theta \xi_z \quad \Rightarrow \quad h_r = \left| \frac{\partial r}{\partial r} \right| = 1$$

$$\frac{\partial r}{\partial \theta} = r \cos \theta \cos \phi \xi_x + r \cos \theta \sin \phi \xi_y - r \sin \theta \xi_z \quad \Rightarrow \quad h_{\theta} = \left| \frac{\partial r}{\partial \theta} \right| = r$$

$$\frac{\partial r}{\partial \phi} = -r \sin \theta \sin \phi \xi_x + r \sin \theta \cos \phi \xi_y \quad \Rightarrow \quad h_{\phi} = \left| \frac{\partial r}{\partial \phi} \right| = r \sin \theta$$

Hence the unit vectors for spherical polars are

$$\xi_r = \sin \theta \cos \phi \xi_x + \sin \theta \sin \phi \xi_y + \cos \theta \xi_z = \xi / r$$

$$\xi_{\theta} = \cos \theta \cos \phi \xi_x + \cos \theta \sin \phi \xi_y - \sin \theta \xi_z$$

$$\xi_{\phi} = -\sin \phi \xi_x + \cos \phi \xi_y$$
These unit vectors are normal to the surfaces described above (spheres, cones and planes).

They are orthogonal:
\[ \mathbf{e}_r \cdot \mathbf{e}_\theta = \mathbf{e}_r \cdot \mathbf{e}_\phi = \mathbf{e}_\theta \cdot \mathbf{e}_\phi = 0 \]

And they form a right-handed orthonormal basis:
\[ \mathbf{e}_r \times \mathbf{e}_\theta = \mathbf{e}_\phi, \quad \mathbf{e}_\theta \times \mathbf{e}_\phi = \mathbf{e}_r, \quad \mathbf{e}_\phi \times \mathbf{e}_r = \mathbf{e}_\theta. \]

See also tutorial sheet.

\[ 6.2 \quad \text{Elements of length, area and volume} \]

\[ 6.2.1 \quad \text{Vector length} \]

If we change \( u_1 \to u_1 + du_1 \), keeping \( u_2 \) and \( u_3 \) fixed, then \( r \to r + dr \) with \( dr = h_1 du_1 \). The infinitesimal element of length along \( \mathbf{e}_1 \) is \( h_1 du_1 \).

Similarly, the infinitesimal lengths along the curvilinear basis vectors \( \mathbf{e}_2, \mathbf{e}_3 \), are \( h_2 du_2, h_3 du_3 \) respectively.

If we change \( u_i \to u_i + du_i \), for all \( i = 1, 2, 3 \), then
\[ dr = h_1 du_1 \mathbf{e}_1 + h_2 du_2 \mathbf{e}_2 + h_3 du_3 \mathbf{e}_3 = \sum_{i=1}^{3} h_i du_i \mathbf{e}_i \]

Note: We are not using the summation convention here – because the expression for \( dr \) above contains 3 identical indices! Great care is needing when using the summation convention with OCCs.
6.2.2 Arc length and metric tensor

Defining \( ds \) as the length of the infinitesimal vector \( dr \), we have \((ds)^2 = dr \cdot dr\)

In Cartesian coordinates: \((ds)^2 = (dx)^2 + (dy)^2 + (dz)^2\)

(i) For an arbitrary set of curvilinear coordinates (not necessarily orthogonal), letting \( u_k \to u_k + du_k \), \( \forall k = 1, 2, 3 \), gives

\[
(dr)_k = dx_k = \frac{\partial x_k}{\partial u_1} du_1 + \frac{\partial x_k}{\partial u_2} du_2 + \frac{\partial x_k}{\partial u_3} du_3 = \frac{\partial x_k}{\partial u_i} du_i
\]

The square \((ds)^2\) of the length \( ds \) is then

\[
(ds)^2 = dx_k dx_k = \frac{\partial x_k}{\partial u_i} \frac{\partial x_k}{\partial u_j} du_i du_j = g_{ij} du_i du_j \tag{6.1}
\]

where we defined the \(3^2 = 9\) components \(g_{ij}\) of the metric tensor by

\[
g_{ij} = g_{ji} = \frac{\partial x_k}{\partial u_i} \frac{\partial x_k}{\partial u_j} = \frac{\partial r}{\partial u_i} \cdot \frac{\partial r}{\partial u_j}
\]

The 9 quantities \(g_{ij}\) are the components of a symmetric second-rank tensor. For an arbitrary set of curvilinear coordinates, \(g_{ij}\) is not diagonal in general.

(ii) In terms of scale factors\(^2\)

\[
(ds)^2 = \left( \sum_i h_i \xi_i du_i \right) \cdot \left( \sum_j h_j \xi_j du_j \right) = \sum_{ij} h_i h_j (\xi_i \cdot \xi_j) du_i du_j \tag{6.2}
\]

For orthogonal curvilinear coordinates, we have \(\xi_i \cdot \xi_j = \delta_{ij}\), and hence

\[
(ds)^2 = \sum_i h_i^2 (du_i)^2 = h_1^2 du_1^2 + h_2^2 du_2^2 + h_3^2 du_3^2 \tag{6.3}
\]

Comparing equations (6.1) and (6.2), the metric tensor can be written as

\[
g_{ij} = h_i h_j (\xi_i \cdot \xi_j)
\]

For orthogonal curvilinear coordinates, the metric tensor is diagonal

\[
g_{ij} = h_i^2 \delta_{ij} \quad (\text{no sum on } i)
\]

and it’s simplest to use equation (6.3).

**Example:** For spherical polars, we showed that

\[
h_r = 1 \quad h_\theta = r \quad h_\phi = r \sin \theta
\]

In this case

\[
(ds)^2 = (dr)^2 + r^2(d\theta)^2 + r^2 \sin^2 \theta (d\phi)^2
\]

\(^2\)The shorthand \(\sum_i\) means \(\sum_{i=1}^3\) etc.
6.2.3 Vector area

If we let \( u_1 \rightarrow u_1 + d\mathbf{u}_1 \), with \( u_2, u_3 \) fixed, then

\[
\mathbf{r} \rightarrow \mathbf{r} + d\mathbf{r}_1 \quad \text{with} \quad d\mathbf{r}_1 = h_1 \mathbf{e}_1 d\mathbf{u}_1
\]

Similarly, if we let \( u_2 \rightarrow u_2 + d\mathbf{u}_2 \), with \( u_1, u_3 \) fixed,

\[
\mathbf{r} \rightarrow \mathbf{r} + d\mathbf{r}_2 \quad \text{with} \quad d\mathbf{r}_2 = h_2 \mathbf{e}_2 d\mathbf{u}_2
\]

The vector area of the infinitesimal parallelogram whose sides are the vectors \( d\mathbf{r}_1 \) and \( d\mathbf{r}_2 \) is

\[
dS = (d\mathbf{r}_1) \times (d\mathbf{r}_2) = (h_1 du_1 \mathbf{e}_1) \times (h_2 du_2 \mathbf{e}_2)
\]

For OCCs, this simplifies to

\[
dS = h_1 h_2 du_1 du_2 \mathbf{e}_3,
\]

since \( \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3 \) for orthogonal systems. \( dS_3 \) points in the direction of \( \mathbf{e}_3 \), which is normal to the surfaces \( u_3 = \text{constant} \), and the infinitesimal area is a rectangle.

The vector areas \( dS_1 \) and \( dS_2 \) are defined similarly.

**Example:** For spherical polars, if we vary \( \theta \) and \( \phi \), keeping \( r \) fixed, we easily obtain a familiar result

\[
dS_r = (h_\theta d\theta \mathbf{e}_\theta) \times (h_\phi d\phi \mathbf{e}_\phi) = h_\theta h_\phi dr d\phi d\mathbf{e}_r = r^2 \sin \theta d\theta d\phi d\mathbf{e}_r
\]

Similarly, if we vary \( \phi \) and \( r \), keeping \( \theta \) fixed, we obtain the vector element of area on a cone of semi-angle \( \theta \), with its axis along the \( z \) axis

\[
dS_\theta = (h_\phi d\phi \mathbf{e}_\phi) \times (h_r dr \mathbf{e}_r) = h_\phi h_r d\phi dr \mathbf{e}_\theta = r \sin \theta dr d\phi d\mathbf{e}_\theta
\]

Similarly for \( dS_\phi \) [tutorial].

6.2.4 Volume

The volume of the infinitesimal parallelepiped with edges \( d\mathbf{r}_1, d\mathbf{r}_2 \) and \( d\mathbf{r}_3 \) is

\[
dV = |(dr_1 \times dr_2) \cdot dr_3| = |(h_1 du_1 \mathbf{e}_1) \times (h_2 du_2 \mathbf{e}_2)) \cdot (h_3 du_3 \mathbf{e}_3)|
\]

For OCCs, we have \( (\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 = 1 \) (in a RH basis), hence

\[
dV = h_1 h_2 h_3 du_1 du_2 du_3
\]

In this case, the infinitesimal volume is a cuboid.

For spherical polars, we have

\[
dV = h_r h_\theta h_\phi dr d\theta d\phi = r^2 \sin \theta dr d\theta d\phi
\]

For a general set of curvilinear coordinates

\[
dV = \sqrt{g} du_1 du_2 du_3
\]

where \( g \) is the determinant of the metric tensor.
6.3 Components of a vector field in curvilinear coordinates

A vector field \( \mathbf{a}(\mathbf{r}) \) can be expressed in terms of \textit{curvilinear components} \( a_i \), defined by

\[
\mathbf{a}(\mathbf{r}) = \sum_{i=1}^{3} a_i(u_1, u_2, u_3) \mathbf{e}_i
\]

where \( \mathbf{e}_i \) is the \( i \)th \textit{curvilinear} basis vector (which again should really be called \( \mathbf{e}_{u_i} \) to avoid confusion with the Cartesian basis vectors.)

For \textit{orthogonal} curvilinear coordinates, the component \( a_i \) can be obtained by taking the scalar product of \( \mathbf{a} \) with the \( i \)th curvilinear basis vector \( \mathbf{e}_i \)

\[
a_i = \mathbf{a}(\mathbf{r}) \cdot \mathbf{e}_i
\]

NB \( a_i \) must be expressed in terms of \( u_1, u_2, u_3 \) (not \( x, y, z \)) when working in the \{\( u_i \}\) basis.

\textbf{Example:} If \( \mathbf{a} = a \mathbf{e}_x \) in Cartesians, then in spherical polars

\[
a_r = \mathbf{a} \cdot \mathbf{e}_r = (a \mathbf{e}_x) \cdot (\sin \theta \cos \phi \mathbf{e}_x + \sin \theta \sin \phi \mathbf{e}_y + \cos \theta \mathbf{e}_z) = a \sin \theta \cos \phi
\]

Similarly, \( a_\theta = \mathbf{a} \cdot \mathbf{e}_\theta \) and \( a_\phi = \mathbf{a} \cdot \mathbf{e}_\phi \), and we obtain \( \mathbf{a} \) in the spherical-polar basis (exercise)

\[
\mathbf{a}(r, \theta, \phi) = a \left( \sin \theta \cos \phi \mathbf{e}_r + \cos \theta \cos \phi \mathbf{e}_\theta - \sin \phi \mathbf{e}_\phi \right)
\]

- You can often spot the curvilinear components “by inspection”.

- In general, one chooses the set of coordinates which matches most closely the \textit{symmetry} of the problem.
6.4 Div, grad, curl and the Laplacian in OCCs

6.4.1 Gradient

The gradient of a scalar field \( f(r) \) is defined in terms of the change in the field \( df(r) \) when \( r \to r + dr \):

\[
df(r) = \nabla f(r) \cdot dr
\]  
(6.4)

Now write \( f(r) \) in terms of orthogonal curvilinear coordinates: \( f(r) = f(u_1, u_2, u_3) \).

As usual, we denote the \emph{curvilinear} basis vectors by \( \{e_1, e_2, e_3\} \).

Let \( u_1 \to u_1 + du_1, \ u_2 \to u_2 + du_2, \) and \( u_3 \to u_3 + du_3 \).

Using the chain rule, we have

\[
df = \frac{\partial f}{\partial u_1} du_1 + \frac{\partial f}{\partial u_2} du_2 + \frac{\partial f}{\partial u_3} du_3
\]  
(6.5)

We need to rewrite the RHS of this equation in the form of equation (6.4). Start with

\[
dr = h_1 du_1 e_1 + h_2 du_2 e_2 + h_3 du_3 e_3
\]

and use orthogonality of the \emph{curvilinear} basis vectors, \( e_i \cdot e_j = \delta_{ij} \), to rewrite equation (6.5) as

\[
df = \left( \frac{\partial f}{\partial u_1} e_1 + \frac{\partial f}{\partial u_2} e_2 + \frac{\partial f}{\partial u_3} e_3 \right) \cdot \left( \frac{1}{h_1} \frac{\partial f}{\partial u_1} du_1 + \frac{1}{h_2} \frac{\partial f}{\partial u_2} du_2 + \frac{1}{h_3} \frac{\partial f}{\partial u_3} du_3 \right)
\]  

Comparing this result with equation (6.4), which holds for all \( dr \), gives us \( \nabla f \) in orthogonal curvilinear coordinates

\[
\nabla f = \sum_{i=1}^{3} \frac{1}{h_i} \frac{\partial f}{\partial u_i} e_i
\]

For \emph{spherical polars}, \( h_r = 1, h_\theta = r, h_\phi = r \sin \theta, \) and we have

\[
\nabla f(r, \theta, \phi) = \frac{\partial f}{\partial r} e_r + \frac{1}{r} \frac{\partial f}{\partial \theta} e_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} e_\phi
\]

\(^3\)We use \( f(r) \) rather than \( \phi(r) \) in order to avoid confusion with the angle \( \phi \) in spherical polars.
6.4.2 Divergence

Let \( \mathbf{a}(\mathbf{r}) \) be a vector field, which we write in orthogonal curvilinear coordinates as

\[
\mathbf{a}(\mathbf{r}) = \sum_{i=1}^{3} a_i(u_1, u_2, u_3) \mathbf{e}_i
\]

where \( a_i \) are the components of \( \mathbf{a} \) in the curvilinear basis, \( \mathbf{e}_i \) is the \( i \)th curvilinear basis vector, and \( \mathbf{a} \) is continuously differentiable.

We obtain \( \nabla \cdot \mathbf{a} \) in orthogonal curvilinear using the integral definition of divergence

\[
\nabla \cdot \mathbf{a} = \lim_{\delta V \to 0} \frac{1}{\delta V} \int_{\delta S} \mathbf{a} \cdot d\mathbf{S},
\]

where \( \delta S \) is the closed surface bounding \( \delta V \).

Let the point \( P \) have curvilinear coordinates \( (u_1, u_2, u_3) \).

Choose \( \delta V \) to be a small “cuboid” with its three edges \( \{ \delta r_i \} \) along the basis vectors \( \{ \mathbf{e}_i \} \) at \( P \):

\[
\begin{align*}
\delta r_1 &= h_1 \delta u_1 \mathbf{e}_1 \\
\delta r_2 &= h_2 \delta u_2 \mathbf{e}_2 \\
\delta r_3 &= h_3 \delta u_3 \mathbf{e}_3
\end{align*}
\]

The outward element of area on the face \( ABCD \) is \( dS = +h_2 h_3 \, du_2 \, du_3 \, \mathbf{e}_1 \)

The outward element of area on the face \( PQRS \) is \( dS = -h_2 h_3 \, du_2 \, du_3 \, \mathbf{e}_1 \)

The contributions to the surface integral from the faces \( ABCD \) and \( PQRS \) are then

\[
\int_{u_3}^{u_3+\delta u_3} \int_{u_2}^{u_2+\delta u_2} \left\{ [a_1 h_2 h_3]_{ABCD} - [a_1 h_2 h_3]_{PQRS} \right\} du_2' du_3' = \int_{u_3}^{u_3+\delta u_3} \int_{u_2}^{u_2+\delta u_2} \delta u_1 \left[ \frac{\partial}{\partial u_1} (a_1 h_2 h_3) \right]_{(u_1, u_2', u_3')} \, du_2' du_3' \quad \text{(by Taylor’s theorem)}
\]

In the last step, we assumed that \( \delta V \) is small enough that the integrand is approximately constant over the range of integration. We can then approximate the integrals over \( u_2' \) and \( u_3' \) by the integrand evaluated at the point \( P \),

\[
\delta u_1 \left[ \frac{\partial}{\partial u_1} (a_1 h_2 h_3) \right]_{(u_1, u_2, u_3)}
\]

\[47\]
multiplied by the range of integration $\delta u_2 \delta u_3$. [This is a rather crude use of the mean value theorem!]

The contributions of the other four faces to the integral over $\delta S$ can be obtained similarly, or by cyclic permutations of the indices $\{1, 2, 3\}$ in equation (6.6).

Finally, divide by the volume of the cuboid $\delta V = h_1 h_2 h_3 \delta u_1 \delta u_2 \delta u_3$, whereupon all the factors of $\delta u_i$ cancel, and we obtain our final expression for $\nabla \cdot a$ in orthogonal curvilinear coordinates

$$\nabla \cdot a = \frac{1}{h_1 h_2 h_3} \left\{ \frac{\partial}{\partial u_1} (a_1 h_2 h_3) + \frac{\partial}{\partial u_2} (a_2 h_3 h_1) + \frac{\partial}{\partial u_3} (a_3 h_1 h_2) \right\}$$

For Cartesian coordinates, the scale factors are all unity, and we recover the usual expression for $\nabla \cdot a$ in Cartesians.

For spherical polars we have

$$\nabla \cdot a(r, \theta, \phi) = \frac{1}{r^2 \sin \theta} \left\{ \frac{\partial}{\partial r} (r^2 \sin \theta a_r) + \frac{\partial}{\partial \theta} (r \sin \theta a_\theta) + \frac{\partial}{\partial \phi} (r a_\phi) \right\}$$

$$= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 a_r) + \frac{1}{r \sin \theta} \left\{ \frac{\partial}{\partial \theta} (\sin \theta a_\theta) + \frac{\partial}{\partial \phi} (a_\phi) \right\}$$

where $a_r$, $a_\theta$, and $a_\phi$ are the components of the vector field $a$ in the basis $\{e_r, e_\theta, e_\phi\}$.

### 6.4.3 Curl

We obtain $\nabla \times a$ in orthogonal curvilinear coordinates using the line integral definition of curl.

The component of $\nabla \times a$ in the direction of the unit vector $n$ is

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

where $\delta S$ is a small planar surface, with unit normal $n$, bounded by the closed curve $\delta C$.

Let $\delta S$ be a small rectangular surface parallel to the $\mathbf{e}_2 - \mathbf{e}_3$ plane with one corner at $r(u_1, u_2, u_3)$, and with edges

$$\delta r_2 = h_2 \delta u_2 \mathbf{e}_2 \quad \text{and} \quad \delta r_3 = h_3 \delta u_3 \mathbf{e}_3$$

which lie along the basis vectors, so that $n = \mathbf{e}_1$.

The line integral around the curve $\delta C$ is the sum of the line integrals along the lines

---

48
\[ \oint_{\delta C} a \cdot \mathbf{dr} = \int_{u_2}^{u_2+\delta u_2} [a_2 h_2]_{(u_1, u_2, u_3)} \, du'_2 + \int_{u_3}^{u_3+\delta u_3} [a_3 h_3]_{(u_1, u_2+\delta u_2, u'_3)} \, du'_3 \]

\[ - \int_{u_2}^{u_2+\delta u_2} [a_2 h_2]_{(u_1, u'_2, u_3+\delta u_3)} \, du'_2 - \int_{u_3}^{u_3+\delta u_3} [a_3 h_3]_{(u_1, u_2, u'_3)} \, du'_3 \]

Using Taylor’s theorem, we can write this as

\[ \oint_{\delta C} a \cdot \mathbf{dr} = \int_{u_2}^{u_2+\delta u_2} \left\{ \delta u_2 \left[ \frac{\partial}{\partial u_2} (a_3 h_3) \right]_{(u_1, u_2, u'_3)} \right\} \, du'_3 \]

\[ - \int_{u_2}^{u_2+\delta u_2} \left\{ \delta u_3 \left[ \frac{\partial}{\partial u_3} (a_2 h_2) \right]_{(u_1, u_2, u'_3)} \right\} \, du'_2 \]

In each case, we approximate the integrals over \( u'_3 \) and \( u'_2 \) by the product of the integrand and the integration ranges \( \delta u_3 \) and \( \delta u_2 \), respectively. Hence

\[ \oint_{\delta C} a \cdot \mathbf{dr} = \frac{\partial}{\partial u_2} (a_3 h_3) \delta u_2 \delta u_3 - \frac{\partial}{\partial u_3} (a_2 h_2) \delta u_3 \delta u_2 \]

where all the \( \{a_i\} \) and \( \{h_i\} \) are evaluated at \( r(u_1, u_2, u_3) \).

Finally, we divide by the area of the rectangle \( \delta S = h_2 h_3 \delta u_2 \delta u_3 \), whereupon all the factors of \( \delta u_i \) cancel, and we obtain

\[ \varepsilon_1 \cdot (\nabla \times a) = (\nabla \times a)_1 = \frac{1}{h_2 h_3} \left\{ \frac{\partial}{\partial u_2} (a_3 h_3) - \frac{\partial}{\partial u_3} (a_2 h_2) \right\} \]

The components of \( \nabla \times a \) in the directions of the curvilinear basis vectors \( \varepsilon_2 \) and \( \varepsilon_3 \) may be obtained similarly, or by cyclic permutations of the indices.

It is convenient to write the final result in the form

\[
\begin{vmatrix}
\n\n\n\end{vmatrix} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix}
\n\n\end{vmatrix}
\]

For spherical polars we have

\[
\begin{vmatrix}
\n\n\end{vmatrix} = \frac{1}{r^2 \sin \theta} \begin{vmatrix}
\n\n\end{vmatrix}
\]

49
6.4.4 Laplacian of a scalar field

The action of the Laplacian operator on a scalar field \( f(r) \) is defined by \( \nabla^2 f = \nabla \cdot (\nabla f) \).

Using the expression for \( \nabla \cdot a \), with \( a = \nabla f \), derived above, we find

\[
\nabla^2 f = \frac{1}{h_1 h_2 h_3} \left\{ \frac{\partial}{\partial u_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left( \frac{h_3 h_1}{h_2} \frac{\partial f}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial u_3} \right) \right\}
\]

In spherical polars, we have

\[
\nabla^2 f(r, \theta, \phi) = \frac{1}{r^2 \sin \theta} \left\{ \frac{\partial}{\partial r} \left( r^2 \sin \theta \frac{\partial f}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{\partial}{\partial \phi} \left( \frac{1}{\sin \theta} \frac{\partial f}{\partial \phi} \right) \right\}
\]

\[
= \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \left\{ \sin \theta \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{\partial^2 f}{\partial \phi^2} \right\}
\]

The expression for the Laplacian of a scalar field in spherical polars is one of the most useful results in the course, with applications in electromagnetism, quantum mechanics, optics, elasticity, fluid mechanics, meteorology, general relativity, cosmology, . . .

6.4.5 Laplacian of a vector field

The Laplacian of a vector field \( a(r) \) in curvilinear coordinates is defined by means of the identity

\[
\nabla \times (\nabla \times a) = \nabla (\nabla \cdot a) - \nabla^2 a
\]

in the form\(^4\)

\[
\nabla^2 a = \nabla (\nabla \cdot a) - \nabla \times (\nabla \times a)
\]

The quantities on the right hand side are evaluated using the expressions for grad, div and curl derived above.

\(^4\)Remember the mnemonic ‘grad-div-minus-curl-curl’ or GDMCC, pronounced ‘guddumk’. Thanks to Peter Boyle for teaching me this!
Chapter 7

Electrostatics

7.1 The Dirac delta function in three dimensions

Consider the mass of a body with density $\rho(r)$. The mass of the body is

$$M = \int_V \rho(r) \, dV$$

How can we use this general expression for the case of a single particle? What is the ‘density’ of a single ‘point’ particle with mass $M$ at $r_0$?

We need a ‘function’ $\rho(r)$ with the properties

$$\rho(r) = 0 \quad \forall r \neq r_0$$

$$M = \int_V \rho(r) \, dV \quad r_0 \in V$$

$$\rho(r) = M \delta(r - r_0)$$

Generalising slightly, we define the delta function to pick out the value of the function $f(r_0)$ at one point $r_0$ in the range of integration, so that

$$\int_V dV f(r) \delta(r - r_0) = \begin{cases} f(r_0) & r_0 \in V \\ 0 & \text{otherwise} \end{cases}$$

Similarly, the total charge on a body with charge density (charge per unit volume) $\rho(r)$ is

$$Q = \int_V \rho(r) \, dV$$

The one dimensional delta function

The delta function may be defined by a sequence of functions $\delta_{\epsilon}(x - a)$, each of ‘area’ unity, which have the desired limit when integrated over. We give a number of examples of how this may be done.

$^1 f(r_0) = M$ in the example above
• Top hat

\[ \delta_\epsilon(x - a) = \begin{cases} 
\frac{1}{2\epsilon} & a - \epsilon < x < a + \epsilon \\
0 & \text{otherwise} 
\end{cases} \]

• Witch’s hat

\[ \delta_\epsilon(x - a) = \begin{cases} 
\frac{1}{\epsilon^2} \left( \epsilon - |x - a| \right) & a - \epsilon < x < a + \epsilon \\
0 & \text{otherwise} 
\end{cases} \]

• Gaussian

\[ \delta_\epsilon(x - a) = \frac{1}{\epsilon \sqrt{\pi}} \exp \left\{ -\frac{(x - a)^2}{\epsilon^2} \right\} \]

In each case

\[ \int_{-\infty}^{+\infty} dx \, f(x) \, \delta_\epsilon(x - a) = \int_{-\infty}^{+\infty} dx \, f(x + a) \, \delta_\epsilon(x) \]

\[ = \int_{-\infty}^{+\infty} dx \left[ f(a) + x f'(a) + x^2/2 f''(a) + \ldots \right] \delta_\epsilon(x) \]

where we shifted the integration variable in the first line, and Taylor-expanded the integrand in the second. The function \( f(x) \) is a ‘good’ test function, i.e. one for which the integral is convergent for all \( \epsilon \).
For the top hat, we need to evaluate

\[ \int_{-\infty}^{+\infty} dx \ x^n \delta_{\epsilon}(x) = \int_{-\epsilon}^{+\epsilon} dx \ x^n \ \frac{1}{2\epsilon} \]

\[ = \begin{cases} \frac{1}{n+1} \epsilon^n & n = 0, 2, 4, \ldots \\ 0 & n = 1, 3, 5, \ldots \end{cases} \to \begin{cases} 1 & n = 0 \\ 0 & \text{otherwise} \end{cases} \]

Hence

\[ \int_{-\infty}^{+\infty} dx \ f(x) \delta_{\epsilon}(x-a) \to f(a) \quad \text{i.e.} \quad \delta_{\epsilon}(x-a) \to \delta(x-a) \]

Similarly for the other representations. The Gaussian representation is the cleanest, because it’s a smooth function.

**Notes:**

(i) The Dirac delta ‘function’ isn’t a function, it’s a *distribution* or *generalised* function.

(ii) Colloquially, it’s an infinitely-tall infinitely-thin spike of unit area.

(iii) The delta function is the continuous-variable analogue of the Kronecker delta symbol. If we let \( i \to x \)

\[ u_i \delta_{ij} = u_j \to \int dx \ x \delta(x-x_0) = x_0 \]

(iv) An important identity is

\[ \int_{-\infty}^{+\infty} dx \ f(x) \delta(g(x)) = \sum_i \frac{f(x_i)}{|g'(x_i)|} \]

where \( g(x_i) = 0 \), i.e. \( x_i \) are the *simple* zeroes of \( g(x) \)  [tutorial].

**The three dimensional delta function**

In Cartesian coordinates \((x, y, z)\),

\[ \delta^{(3)}(r-r_0) \equiv \delta(r-r_0) = \delta(x-x_0) \delta(y-y_0) \delta(z-z_0) \]

In orthogonal curvilinear co-ordinates \((u_1, u_2, u_3)\),

\[ \delta(r-a) = \frac{1}{h_1 h_2 h_3} \delta(u_1-a_1) \delta(u_2-a_2) \delta(u_3-a_3) \]

where \( h_1, h_2, h_3 \) are the usual scale factors  [tutorial].

(In the last equation, we set \( r_0 = a \) to avoid double subscripts on the RHS.)
7.2 Coulomb’s law

Experimentally, the force between two point charges $q$ and $q_1$ at positions $r$ and $r_1$, respectively, is given by Coulomb’s law

$$F_1 = \frac{1}{4\pi\epsilon_0} \frac{q q_1 (r - r_1)}{|r - r_1|^3}$$

$F_1$ is the force on the charge $q$ at $r$, produced by the charge $q_1$ at $r_1$.

Charges can be positive or negative. For $qq_1 > 0$ we have repulsion, and for $qq_1 < 0$ we have attraction: like charges repel and opposite charges attract.

In SI units, charge is measured in Coulombs ($C$). The proton charge is defined to be exactly $1.602176634 \times 10^{-19} C$, so that $1C$ is the charge of $0.62415 \ldots \times 10^{19}$ protons.

The permittivity of free space is measured as $\epsilon_0 = 8.85418781 \ldots \times 10^{-12} C^2 N^{-1} m^{-2}$.

Aside: Similarly for Newton’s law of gravitation,

$$F_1 = -G \frac{mm_1 (r - r_1)}{|r - r_1|^3}$$

which is always attractive (hence the negative sign, so that $G$, $m$, $m_1$ are all positive).

In SI units: $G = 6.672 \times 10^{-11} N m^2 kg^{-2}$.

7.3 The electric field

The electric field $E$ is ‘produced’ by a charge configuration, and is defined in terms of the force on a small positive test charge,

$$E(r) = \lim_{q \to 0} \frac{1}{q} F$$

Clearly, $E$ is a vector field.

7.3.1 Field lines

Field lines are the ‘lines of force’ on the test charge.

Newton’s equations imply that the motion of a (test) particle is unique, which implies that the field lines do not cross, and thus that they are well-defined and can be measured.

Thus for our two charges $q$ and $q_1$ we have

$$F_1 = q E(r)$$

i.e. particle 1 ‘produces’ an electrostatic field $E(r)$. The diagram shows the field lines produced by a negative charge.
The particle at $P$ ‘feels’ the electrostatic field as a force $qE(r)$ with

$$E(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 (r - r_1)}{|r - r_1|^3}$$  \hfill (7.1)

### 7.3.2 The principle of superposition

Consider a set of charges $q_i$ situated at $r_i$

The principle of superposition is motivated by experiment; it states that the total electric field at $r$ is the vector sum of the fields due to the individual charges at $r_i$

$$E(r) = \frac{1}{4\pi\epsilon_0} \sum_i q_i \frac{(r - r_i)}{|r - r_i|^3}$$

In the limit of (infinitely) many charges, we introduce a continuous charge density (charge/volume) $\rho(r')$, so that the charge in $dV'$ at position $r'$ is $\rho(r') dV'$. The electric field is then

$$E(r) = \frac{1}{4\pi\epsilon_0} \int_V dV' \rho(r') \frac{(r - r')}{|r - r'|^3}$$

To return to our original example of a single charge $q_1$ at position $r_1$, we simply set the charge density $\rho(r') = q_1 \delta(r' - r_1)$, which recovers the result in equation (7.1).

### 7.4 The electrostatic potential for a point charge

Since

$$\nabla \left( \frac{1}{|r - r_1|} \right) = -\frac{(r - r_1)}{|r - r_1|^3}$$  \hfill (7.2)

where $\nabla$ operates on $r$ (not $r_1$), then for a point charge $q_1$ at $r_1$

$$E(r) = \frac{q_1}{4\pi\epsilon_0} \frac{(r - r_1)}{|r - r_1|^3} = -\frac{q_1}{4\pi\epsilon_0} \nabla \left( \frac{1}{|r - r_1|} \right)$$

i.e. we may write

$$E(r) = -\nabla \phi(r) \quad \text{with} \quad \phi(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1}{|r - r_1|}$$  \hfill (7.3)

$\phi(r)$ is the electrostatic potential for the electric field $E(r)$. 

55
7.5 The static Maxwell equations

7.5.1 The curl equation

For a continuous charge distribution, we again use equation (7.2) to write the electric field as a gradient

\[ E(r) = -\frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(r')}{|r - r'|} \nabla \frac{1}{|r - r'|} = -\nabla \left( \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(r')}{|r - r'|} \right) \]

(7.4)

Note that \( \nabla \) operates on \( r \) (not \( r' \)) so we can take it out of the volume integral over \( r' \). Therefore

\[ \nabla \times E = -\nabla \times \nabla \left( \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(r')}{|r - r'|} \right) \]

But the curl of the gradient of a scalar field is always zero, which implies

\[ \nabla \times E = 0 \]

for all static electric fields. This is (the static version of) Maxwell’s third equation.

7.5.2 Conservative fields and potential theory

A vector field that satisfies \( \nabla \times E = 0 \) is said to be conservative (or irrotational).

Consider the integral of \( \nabla \times E \) over an open surface \( S \) bounded by the closed curve \( C_1 - C_2 \). Using Stokes’ theorem

\[ 0 = \int_S (\nabla \times E) \cdot dS = \oint_{C_1 - C_2} E \cdot dr \]

Therefore

\[ \int_{C_1} E \cdot dr = \int_{C_2} E \cdot dr \]

Since the line integral is independent of the path from \( a \) to \( b \), it can only depend on the end points. So, for some scalar field \( \phi \), we must have

\[ -\int_a^b E \cdot dr = \phi(b) - \phi(a) \]

Now let \( a = r \) and \( b = r + \delta r \), where \( \delta r \) is small, so we can approximate the integral

\[ -E(r) \cdot \delta r + \ldots = \phi(r + \delta r) - \phi(r) = \nabla \phi \cdot \delta r + \ldots \]

where we used the definition of the gradient in the last step. Since this holds \( \forall \delta r \), we can always write

\[ E(r) = -\nabla \phi(r) \]

The scalar field \( \phi(r) \) is called the potential for the vector field \( E(r) \).
An explicit expression for $\phi(r)$ can be obtained from (7.4). We have $E = -\nabla \phi$ with

$$
\phi(r) = \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(r')}{|r - r'|} \, dV'
$$

This is linear superposition for potentials.

As in the case of the electric field, if we set $\rho(r') = q_1 \delta(r' - r_1)$, we recover the potential for a single charge (equation (7.3))

$$
\phi(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1}{|r - r_1|}
$$

Notes:

- For a surface charge distribution, with charge/unit-area $\sigma(r)$, the electric field produced is

$$
E(r) = \frac{1}{4\pi\epsilon_0} \int_S dS' \sigma(r') \frac{(r - r')}{|r - r'|^3}
$$

and

$$
\phi(r) = \frac{1}{4\pi\epsilon_0} \int_S dS' \frac{\sigma(r')}{|r - r'|}
$$

where $dS'$ is the infinitesimal (scalar) element of area at $r'$ on the surface $S$.

- For a line distribution of charge, with charge/unit-length $\lambda(r)$

$$
E(r) = \frac{1}{4\pi\epsilon_0} \int_C dl' \lambda(r') \frac{(r - r')}{|r - r'|^3}
$$

and

$$
\phi(r) = \frac{1}{4\pi\epsilon_0} \int_C dl' \frac{\lambda(r')}{|r - r'|}
$$

where $dl'$ is the infinitesimal element of length along the line (or curve) $C$.

- In SI units, the potential is measured in Volts $V$. In terms of other units $V = C/(C^2N^{-1}m^{-1}) = NmC^{-1} = JC^{-1}$.

- Field lines are perpendicular to surfaces of constant potential $\phi$, called equipotentials or equipotential surfaces.

Let $dr$ be a small displacement of the position vector $r$ of a point in the equipotential surface $\phi = \text{constant}$.

Therefore

$$
0 = d\phi = \nabla \phi \cdot dr
$$

so $E = -\nabla \phi$ is perpendicular to $dr$.

Thus electric field lines $E$ are everywhere perpendicular to the surfaces $\phi = \text{constant}$.
• The potential \( \phi \) is only defined up to an overall constant. If we let \( \phi \to \phi + c \), the electric field \( \nabla \phi \) (and hence the force) is unchanged. So only potential differences have physical significance. In most physical situations, \( \phi \to \) constant as \( r \to \infty \), and we usually choose the constant to be zero.

• So far we’ve defined the potential in purely mathematical terms. Physically, the potential difference, \( V_{AB} \), between two points \( A \) and \( B \) is defined as the energy per unit charge required to move a small test charge \( q \) from \( A \) to \( B \):

\[
V_{AB} \equiv \lim_{q \to 0} \frac{1}{q} W_{AB}
\]

\[
= -\frac{1}{q} \int_{C} F \cdot dr = - \int_{C} E \cdot dr
\]

\[
= \int_{C} \nabla \phi \cdot dr = \int_{A}^{B} d\phi
\]

\[
= \phi_B - \phi_A
\]

The \(-\)ve sign is because this is the work done against the force \( F \). Since the field is conservative, the integral is independent of the path – it depends only on the end points.

• The potential energy of a point charge \( q \) at position \( r \) in an external electrostatic field \( E_{\text{ext}}(r) = -\nabla \phi_{\text{ext}}(r) \) is therefore given by \( q \phi_{\text{ext}}(r) \).

We may generalise this to a charge distribution in an external electric field \( E_{\text{ext}}(\vec{r}) = -\nabla \phi_{\text{ext}} \). In this case, the (interaction) energy is

\[
W = \int_{V} dV \rho(\vec{r}) \phi_{\text{ext}}(\vec{r})
\]

Note that this does not include the the \textit{self-energy} of the charge distribution. To emphasize this we write \( \phi_{\text{ext}} \). [More on this later]

7.5.3 The divergence equation

Let’s return to the potential for an arbitrary charge distribution

\[
\phi(r) = \frac{1}{4\pi\epsilon_0} \int_{V} dV' \frac{\rho(r')}{|r - r'|}
\]

Since \( E = -\nabla \phi \), we have \( \nabla \cdot E = -\nabla^2 \phi \), and hence

\[
\nabla \cdot E(r) = -\frac{1}{4\pi\epsilon_0} \int_{V} dV' \rho(r') \nabla^2 \frac{1}{|r - r'|}
\]

(7.5)

Note that \( \nabla^2 \) acts only on \( r \) (not on \( r' \)), so we can take it inside the integral over \( r' \).
Theorem
\[ \nabla^2 \left( \frac{1}{r} \right) = -4\pi \delta(r) \quad \forall r \]

Proof: We first prove it for \( r \neq 0 \)
\[ \nabla^2 \left( \frac{1}{r} \right) \bigg|_{r \neq 0} = -\partial_i \left( \frac{x_i}{r^3} \right) = -\left( \frac{3}{r^3} - x_i \frac{3}{2} r^{-5} 2x_i \right) = 0 \]

To prove the result for \( r = 0 \), we integrate \( \nabla^2 \left( \frac{1}{r} \right) \) over an arbitrary volume \( V \) containing the origin \( r = 0 \).
\[
\int_V \nabla^2 \left( \frac{1}{r} \right) \ dV = \int_{V_\varepsilon} \nabla^2 \left( \frac{1}{r} \right) \ dV = -\int_{V_\varepsilon} \nabla \cdot \left( \frac{r}{r^3} \right) \ dV \\
= -\int_{S_\varepsilon} \frac{r}{r^3} \cdot dS = -\frac{\varepsilon}{\varepsilon^3} 4\pi \varepsilon^2 = -4\pi 
\]

In the first line, we used our previous result that \( \nabla^2 \left( \frac{1}{r} \right) = 0 \) everywhere away from the origin to write the original integral as an integral over a sphere of radius \( \varepsilon \) centred on the origin, with volume \( V_\varepsilon \) and area \( S_\varepsilon \) respectively.

We then used the divergence theorem to obtain the first result on the second line. On the surface \( S_\varepsilon \), we have \( r = \varepsilon \hat{e}_r \) and \( dS = \hat{e}_r dS \), where \( \hat{e}_r \) is a unit vector in the direction of \( r \), so the integral over the surface of the sphere is straightforward – check it!

[Alternatively, we may write the surface integral as an integral over solid angle \( \int_S \frac{r}{r^3} \cdot dS = \int_S d\Omega = 4\pi \).]

We can now take the limit \( \varepsilon \to 0 \), which simply shrinks the sphere down to the origin, leaving the integral unchanged.

Since our result for the integral holds for an arbitrary volume \( V \) centred on the origin, and \( \int_V \delta(\varepsilon) \ dV = 1 \), we deduce that \( \nabla^2 \left( \frac{1}{r} \right) = -4\pi \delta(r) \). Similarly
\[ \nabla^2 \left( \frac{1}{|r - r'|} \right) = -4\pi \delta(r - r') \]
Substituting this result into equation (7.5) gives
\[
\nabla \cdot E(r) = -\frac{1}{4\pi \epsilon_0} \int_V dV' \rho(r') \left( -4\pi \delta \left( r - r' \right) \right) 
\]

Using the delta function to perform the integral on the right hand side, we get Maxwell’s first equation
\[ \nabla \cdot E(r) = \frac{\rho(r)}{\epsilon_0} \]

We now have the two electrostatic Maxwell equations
\[ \nabla \cdot E = \frac{\rho}{\epsilon_0} \quad \nabla \times E = 0 \]

In terms of the potential
\[ E = -\nabla \phi \quad \nabla^2 \phi = -\frac{\rho}{\epsilon_0} \]

The second equation is called Poisson’s equation.
7.6 Electric dipole

Physically, an electric dipole consists of two nearby equal and opposite (point) charges, with charge \(-q\) situated at \(r_0\) and charge \(+q\) at \(r_0 + d\).

Define the dipole moment \(p = qd\).

It will turn out to be useful to consider the dipole limit, in which

\[
p = \lim_{d \to 0} qd
\]

with \(p\) finite (and constant). This is sometimes called a point dipole or an ideal dipole.

7.6.1 Potential and electric field due to a dipole

**Dipole potential** The electrostatic potential \(\phi(r)\) produced by the dipole is

\[
\phi(r) = \frac{q}{4\pi \epsilon_0} \left[ \frac{1}{|r - r_0 - d|} - \frac{1}{|r - r_0|} \right]
\]

\[
= \frac{q}{4\pi \epsilon_0} \left[ \frac{1}{|r - r_0|} + \frac{d \cdot (r - r_0)}{|r - r_0|^3} + O(d^2) - \frac{1}{|r - r_0|} \right]
\]

where we Taylor (or binomial) expanded the first term about \(r - r_0\) [tutorial].

In the dipole limit, the terms of \(O(qd^2)\) vanish, and the potential is simply

\[
\phi(r) = \frac{1}{4\pi \epsilon_0} \frac{p \cdot (r - r_0)}{|r - r_0|^3}
\]

For a dipole at the origin we have

\[
\phi(r) = \frac{1}{4\pi \epsilon_0} \frac{p \cdot r}{r^3}
\]

Note that \(\phi(r)\) falls off as \(1/r^2\).

**Electric field** The \(i\)th component of the electric field produced by (or due to) a dipole of moment \(p\) situated at the origin is

\[
E_i(r) = -\partial_i \phi = - \frac{1}{4\pi \epsilon_0} \partial_i \left( \frac{p_j x_j}{r^3} \right)
\]

\[
= - \frac{1}{4\pi \epsilon_0} p_j \left[ \delta_{ij} \frac{x_j}{r^3} + x_j \left( -\frac{3}{2} \frac{r^{-5}}{r^5} \right) 2x_i \right]
\]

Therefore

\[
E(r) = \frac{1}{4\pi \epsilon_0} \left[ \frac{3p \cdot r}{r^5} r - \frac{p}{r^3} \right]
\]

which falls off as \(1/r^3\).
Spherical polar coordinates

Consider spherical polar coordinates \((r, \theta, \chi)\), with the z-axis chosen parallel to the dipole moment, \(\mathbf{p} = p \mathbf{e}_z\), so that \(\mathbf{p} \cdot \mathbf{r} = pr \cos \theta\).

[We use \(\chi\) instead of \(\phi\) for the azimuthal angle in order to avoid confusion with the potential \(\phi\).]

Then \(\phi(r) = \frac{p}{4\pi \epsilon_0} \frac{1}{r^2} \cos \theta\)

\[E(r) = \frac{p}{4\pi \epsilon_0} \frac{1}{r^3} \left[ 3 \cos \theta \mathbf{e}_r - \mathbf{e}_z \right]\]

The last expression uses a “mixed coordinate basis” \((r, \theta, z)\), which is useful at times.

We can also obtain this result using the expression for \(\nabla \phi\) in polar co-ordinates

\[
\nabla \phi = -\frac{p}{4\pi \epsilon_0} \left( \frac{2}{r^3} \cos \theta \mathbf{e}_r - \frac{\sin \theta}{r^3} \mathbf{e}_\theta \right)
\]

The second form can be obtained from the first by substituting \(\mathbf{e}_z = \mathbf{e}_r \cos \theta - \mathbf{e}_\theta \sin \theta\) into the latter. [Exercise: verify this.]

The sketch shows the electric field (full lines) and the potential (dashed lines) for the dipole.

This picture holds in the dipole limit, but it’s also valid when \(r \gg d\), the ‘far zone’.

7.6.2 Force, torque and energy

Force on a dipole

The force on a dipole at position \(\mathbf{r}\) due to an external electric field \(E_{\text{ext}}\) is

\[
F(r) = -q E_{\text{ext}}(r) + q E_{\text{ext}}(r + d)
= -q E_{\text{ext}}(r) + q \left[ E_{\text{ext}}(r) + (d \cdot \nabla) E_{\text{ext}}(r) + \cdots \right]
\]

In the point dipole limit

\[
F(r) = (p \cdot \nabla) E_{\text{ext}}(r)
\]
Torque on a dipole

The torque (or moment of the force, or couple) on a dipole, about the point \( r \) where the dipole is located, due to the external electric field is

\[
G(r) = -q \vec{0} \times \vec{E}_{\text{ext}}(r) + q (\vec{0} + d) \times \vec{E}_{\text{ext}}(r + d)
\]

\[
= q d \times [\vec{E}_{\text{ext}}(r) + (d \cdot \nabla) \vec{E}_{\text{ext}}(r) + \cdots]
\]

Taking the dipole limit (i.e. ignoring terms of order \( O(qd^2) \)), we find

\[
G(r) = p \times \vec{E}_{\text{ext}}(r)
\]

Energy of a dipole

The energy of a dipole in an external electric field \( \vec{E}_{\text{ext}} \) is

\[
W = -q \phi_{\text{ext}}(r) + q \phi_{\text{ext}}(r + d)
\]

\[
= -q \phi_{\text{ext}}(r) + q \left[ \phi_{\text{ext}}(r) + (d \cdot \nabla) \phi_{\text{ext}}(r) + \cdots \right]
\]

In the dipole limit, using \( \vec{E}_{\text{ext}} = -\nabla \phi_{\text{ext}} \), we find

\[
W = -p \cdot \vec{E}_{\text{ext}}
\]

How is this expression for the energy of the dipole related to the force on the dipole, namely \( F = (p \cdot \nabla) \vec{E}_{\text{ext}} \)?

Recall the following identity for vector fields \( a \) and \( b \)

\[
\nabla (a \cdot b) = (a \cdot \nabla) b + (b \cdot \nabla) a + a \times (\nabla \times b) + b \times (\nabla \times a)
\]

If we set \( a = p = \) constant, and \( b = \vec{E}_{\text{ext}} \), then since \( (\vec{E} \cdot \nabla) p = 0 \) and \( \nabla \times p = 0 \), we find

\[
\nabla (p \cdot \vec{E}_{\text{ext}}) = (p \cdot \nabla) \vec{E}_{\text{ext}} + p \times (\nabla \times \vec{E}_{\text{ext}})
\]

But \( \nabla \times \vec{E}_{\text{ext}} = 0 \), so \( p \times (\nabla \times \vec{E}_{\text{ext}}) = 0 \), and hence

\[
F(r) = (p \cdot \nabla) \vec{E}_{\text{ext}} = \nabla (p \cdot \vec{E}_{\text{ext}}) = -\nabla W
\]

The force on the dipole is the gradient of the potential energy, as one would expect.

Examples

- For the case of a homogeneous (i.e. constant, independent of \( r \)) external field, \( \vec{E}_{\text{ext}}(r) = \vec{E}_0 \), we have

\[
F = 0 \quad \text{and} \quad G(r) = p \times \vec{E}_0
\]

Since \( W = -p \cdot \vec{E}_0 \), a stable or equilibrium position (i.e. position of minimum energy) occurs when \( p \) is parallel to \( \vec{E}_0 \). Colloquially, dipoles “like to align with the field”.

62
• If the dipole at \( r \) has dipole moment \( p_1 \), and the electric field \( E_{\text{ext}}(r) \) is due to a second dipole of moment \( p_2 \) at the origin, then

\[
W = -p_1 \cdot E_{\text{ext}}(r) \quad \text{with} \quad E_{\text{ext}}(r) = \frac{1}{4\pi\varepsilon_0} \left[ \frac{3 (p_2 \cdot r)}{r^5} - \frac{r - p_2}{r^3} \right]
\]

Therefore

\[
W = \frac{1}{4\pi\varepsilon_0} \left[ \frac{p_1 \cdot p_2}{r^3} - \frac{3 (r \cdot p_1)(r \cdot p_2)}{r^5} \right]
\]

The interaction energy is not only dependent on the distance between the dipoles, but also on their relative orientations.

7.7 The multipole expansion

Consider the case of a charge distribution, \( \rho(r) \), localised in a volume \( V \). For convenience we will take the origin inside \( V \).

The potential at the point \( P \) is

\[
\phi(r) = \frac{1}{4\pi\varepsilon_0} \int_V \frac{dV'}{|r - r'|} \rho(r')
\]

For \( |r| \) much larger than the extent of \( V \), \( i.e. \) \( |r| \gg |r'| \) for all \( |r'| \) such that \( \rho(r') \neq 0 \), we can expand the denominator using the binomial theorem

\[
(1 + x)^n = 1 + nx + \frac{n(n-1)}{2} x^2 + O(x^3)
\]

\[
|r - r'|^{-1} = \left\{ |r|^2 - 2 r \cdot r' + |r'|^2 \right\}^{-1/2}
\]

\[
= |r|^{-1} \left\{ 1 - 2 \frac{r \cdot r'}{|r|^2} + \frac{|r'|^2}{|r|^2} \right\}^{-1/2}
\]

\[
= \frac{1}{r} \left\{ 1 + \frac{r \cdot r'}{r^2} - \frac{1}{2} \frac{r'^2}{r^2} + 3 \left( -\frac{2 r \cdot r'}{r^2} \right)^2 + O \left( \frac{r'}{r} \right)^3 \right\}
\]

This can also be obtained by Taylor expansion [exercise]. Then

\[
\phi(r) = \frac{1}{4\pi\varepsilon_0} \int_V dV' \rho(r') \left[ \frac{1}{r} + \frac{r' \cdot r}{r^3} + \frac{3 (r' \cdot r)^2}{2r^5} - \frac{2 r \cdot r'^2}{2r^5} + \ldots \right]
\]

This gives the multipole expansion for the potential

\[
\phi(r) = \frac{1}{4\pi\varepsilon_0} \frac{Q}{r} + \frac{1}{4\pi\varepsilon_0} \frac{p \cdot r}{r^3} + \frac{1}{4\pi\varepsilon_0} \frac{Q_{ij} x_i x_j}{2r^5} + \ldots
\]
where

\[ Q = \int_V \mathrm{d}V' \rho(r') \]

is the total charge within \( V \)

\[ p = \int_V \mathrm{d}V' r' \rho(r') \]

is the dipole moment about the origin

\[ Q_{ij} = \int_V \mathrm{d}V' \left( 3x'_i x'_j - r'^2 \delta_{ij} \right) \rho(r') \]

is the quadrupole tensor

The multipole expansion is valid in the far zone, i.e. when \( r \gg r_0 \), with \( r_0 \) the size of the charge distribution.

- If \( Q \neq 0 \), the monopole term dominates
  \[ \phi(r) = \frac{1}{4\pi\epsilon_0} \frac{Q}{r} \]
  and in the far zone, \( r \gg r_0 \), the \( E \) field is that of a point charge at the origin.

- When the total charge \( Q = 0 \) the dipole term dominates
  \[ \phi(r) = \frac{1}{4\pi\epsilon_0} \frac{p \cdot r}{r^3} \]

If the charge density is given by two equal but oppositely-charged particles close together, i.e. \( \rho(r') = q \left[ \delta(r' - d) - \delta(r') \right] \), then

\[ p = \int \mathrm{d}V' r' q \left[ \delta(r' - d) - \delta(r') \right] = q d \]

which is the dipole moment as defined previously, and hence justifies the name.

- If \( Q = 0 \) and \( p = 0 \), the quadrupole term dominates
  \[ \phi(r) = \frac{1}{4\pi\epsilon_0} \frac{Q_{ij} x_i x_j}{2r^5} \]

The quadrupole tensor \( Q_{ij} \) is symmetric, \( Q_{ij} = Q_{ji} \), and traceless, \( Q_{ii} = 0 \).

**Why quadrupole?**

A simple linear quadrupole is defined by placing two dipoles (so four charges) ‘back to back’ with equal and opposite dipole moments, as shown.

From the figure

\[ \phi(r) = \frac{1}{4\pi\epsilon_0} \left( \frac{q}{|r - r' - d|} + \frac{q}{|r - r' + d|} + \frac{-2q}{|r - r'|} \right) \]
Expanding the denominators in the usual way, and defining $\rho = r - r'$, the leading $(1/\rho)$ and dipole $(1/\rho^2)$ terms cancel, so for large $\rho$

$$\phi(r) = \frac{1}{4\pi\epsilon_0} q \left( \frac{3(r \cdot d)^2 - d^2 r^2}{\rho^5} \right) = \frac{1}{4\pi\epsilon_0} \frac{Q_{ij} \rho_i \rho_j}{\rho^5}$$

where $Q_{ij} = q (3d_i d_j - \delta_{ij} d^2)$ is the (traceless, symmetric) quadrupole tensor – as above.

The quadrupole moment is sometimes defined to be $Q = 2qd^2$, where the ‘2’ is conventional.

### 7.7.1 Worked example

The region inside the sphere $r < a$, contains a charge density

$$\rho(x, y, z) = f z (a^2 - r^2)$$

where $f$ is a constant. Show that at large distances from the origin the potential due to the charge distribution is given approximately by

$$\phi(r) = \frac{2fa^7}{105\epsilon_0} \frac{z}{r^3}$$

The multipole expansion gives

$$\phi(r) = \frac{1}{4\pi\epsilon_0} \left( \frac{Q}{r} + \frac{P \cdot r}{r^3} + O \left( \frac{1}{r^3} \right) \right)$$

In spherical polars $(r, \theta, \chi)$,

$$x = r \sin \theta \cos \chi, \quad y = r \sin \theta \sin \chi, \quad z = r \cos \theta$$

The total charge $Q$ is (we drop the primes in this calculation for brevity):

$$Q = \int_V \rho(r) dV = \int_0^{2\pi} \int_0^{\pi} \int_0^a \left( f r \cos \theta (a^2 - r^2) \right) r^2 \sin \theta dr d\theta d\chi = 0.$$

This integral vanishes because $\int_0^{\pi} \cos \theta \sin \theta d\theta = \int_0^{\pi} \frac{1}{2} \sin(2\theta) d\theta = 0$.

The total dipole moment $P$ about the origin is

$$P = \int_V r \rho(r) dV = \int_V r \varepsilon_{\dot{r}} \rho(r) dV$$

$$= \int_0^{2\pi} \int_0^{\pi} \int_0^a r (\sin \theta \cos \chi \varepsilon_1 + \sin \theta \sin \chi \varepsilon_2 + \cos \theta \varepsilon_3)$$

$$\left( f r \cos \theta (a^2 - r^2) \right) r^2 \sin \theta dr d\theta d\chi$$

The $x$ and $y$ components of the $\chi$ integral vanish. The $z$ component factorises:

$$P_z = f \int_0^{2\pi} d\chi \int_0^{\pi} \sin \theta \cos^2 \theta d\theta \int_0^a r^4 (a^2 - r^2) dr = f \frac{2a^7}{3}$$

Putting it all together, we obtain

$$\phi(r) = \frac{1}{4\pi\epsilon_0} \frac{8\pi a^7 f \varepsilon_3 \cdot r}{105 \epsilon_0} = \frac{2fa^7}{105\epsilon_0} \frac{z}{r^3}$$
7.7.2 Interaction energy of a charge distribution

Let’s consider the interaction energy $W$ of an arbitrary (but bounded) charge distribution in an external electric field $E_{\text{ext}} = -\nabla \phi_{\text{ext}}$.

$$W = \int dV \rho(r) \phi_{\text{ext}}(r)$$

For a charge distribution localised around the origin, we Taylor-expand $\phi_{\text{ext}}(r)$ about $r = 0$

$$\phi_{\text{ext}}(r) = \phi_{\text{ext}}(0) + (r \cdot \nabla) \phi_{\text{ext}}(0) + \frac{1}{2!} (r \cdot \nabla)^2 \phi_{\text{ext}}(0) + \ldots$$

The last term may be re-written as $-\frac{1}{6} (3x_i x_j - r^2 \delta_{ij}) \partial_j E_{\text{ext},i}(0)$. The additional term with the $\delta_{ij}$ vanishes because an external field satisfies $\nabla \cdot E_{\text{ext}} = 0$ in the region around the origin. Therefore

$$W = Q \phi_{\text{ext}}(0) - p \cdot E_{\text{ext}}(0) - \frac{1}{6} Q_{ij} \partial_j E_{\text{ext},i}(0) + \ldots$$

The physical picture is that the total charge couples to the external potential $\phi_{\text{ext}}$, the dipole moment with the external field $E_{\text{ext}}$, and the quadrupole moment with the spatial derivative of the external field.

7.7.3 A brute-force calculation - the circular disc

The electric field $E$ and potential $\phi$ can be evaluated exactly for a number of interesting symmetric charge distributions. We give one example using cylindrical coordinates before moving on to more powerful techniques. A circular disc of radius $a$ carries uniform surface charge density $\sigma$. Find the electric field and the potential due to the disc on the axis of symmetry.

**Electric field:** Start with the general expression

$$E(r) = \frac{1}{4\pi \epsilon_0} \int_S dS' \sigma(r') \frac{(r - r')}{|r - r'|^3}$$

Choose the $z$ axis parallel to the axis of symmetry with the origin at the centre of the disc, so that $r$ lies on the $z$ axis and $r'$ lies in the $x - y$ plane.

In cylindrical coordinates $(\rho, \chi, z)$, we have

$$r = z \xi_z \quad \text{and} \quad r' = \rho \xi_\rho$$

Therefore

$$r - r' = z \xi_z - \rho \xi_\rho \quad \text{and} \quad |r - r'| = (\rho^2 + z^2)^{1/2}$$
Using the expression for $e_\rho$ in terms of $e_x$ and $e_y$, we have

$$E(z) = \frac{\sigma}{4\pi\varepsilon_0} \int_0^a \int_0^{2\pi} \rho d\rho d\chi \frac{z e_x - \rho (\cos \chi e_x + \sin \chi e_y)}{(\rho^2 + z^2)^{3/2}}$$

$$= \frac{\sigma}{4\pi\varepsilon_0} 2\pi \int_0^a \frac{z \rho d\rho}{(\rho^2 + z^2)^{3/2}} e_z + 0 = \frac{\sigma z}{2\varepsilon_0} \left[ -1 \right]_{\rho=0}^{\rho=a}$$

$$= \frac{\sigma z}{2\varepsilon_0} \left[ \frac{1}{|z|} - \frac{1}{(a^2 + z^2)^{1/2}} \right] e_z = \frac{\sigma z}{2\varepsilon_0} \left[ \frac{z}{|z|} - \frac{z}{(a^2 + z^2)^{1/2}} \right] e_z$$

The electric field on the $z$ axis is parallel to the $z$ axis because of symmetry. The sum of all the contributions to $E_x$ and $E_y$ cancel, i.e. they integrate to zero. So we really only needed to calculate $E_z$!

Consider two limits:

(i) $z \gg a$ Expand the second term in square brackets

$$\frac{1}{(a^2 + z^2)^{1/2}} = \frac{1}{|z|} \left( 1 + a^2/z^2 \right)^{-1/2} = \frac{1}{|z|} \left[ 1 - \frac{1}{2} \frac{a^2}{z^2} + O\left( \frac{a^4}{z^4} \right) \right]$$

Keeping only the leading term, we have

$$E(z) = \text{sgn}(z) \frac{\sigma a^2}{4\pi\varepsilon_0 z^2} e_z = \text{sgn}(z) \frac{Q}{4\pi\varepsilon_0} \frac{e_z}{z^2}$$

where the signum (or sign) function $\text{sgn}(z) \equiv z/|z|$ is +1 for $z > 0$, and −1 for $z < 0$; and $Q = \sigma \pi a^2$ is the total charge on the disc. In the far zone, we recover the field for a point charge, as expected.

(ii) $a \gg z$ In this case, the leading behaviour is obtained by dropping the second term in square brackets:

$$E(z) = \text{sgn}(z) \frac{\sigma}{2\varepsilon_0} e_z$$

This is the electric field due to an infinite charged surface – see later.

**Potential:** Start with the general expression

$$\phi(r) = \frac{1}{4\pi\varepsilon_0} \oint_S dS' \frac{\sigma(r')}{|r - r'|}$$

For the disc, we have

$$\phi(z) = \frac{\sigma}{4\pi\varepsilon_0} \int_0^a \int_0^{2\pi} \rho d\rho d\chi \frac{(\rho^2 + z^2)^{1/2}}{(\rho^2 + z^2)^{1/2}} = \frac{\sigma}{2\varepsilon_0} \left[ (\rho^2 + z^2)^{1/2} \right]_{\rho=0}^{\rho=a}$$

$$= \frac{\sigma}{2\varepsilon_0} \left[ (a^2 + z^2)^{1/2} - |z| \right]$$

Note: it’s often much easier to find $\phi$ than $E$!
(i) $z \gg a$ Expanding as before, we find (exercise – important!)

$$\phi(z) = \frac{\sigma}{4\epsilon_0} \frac{a^2}{|z|} = \frac{Q}{4\pi\epsilon_0 |z|}$$

as expected.

(ii) $a \gg z$ In this case we find a linear potential

$$\phi(z) = \frac{\sigma}{2\epsilon_0} [a - |z|] = -\frac{\sigma}{2\epsilon_0} |z| + \text{constant}$$

Exercise: Check that $E(z) = -\frac{\partial \phi}{\partial z} \epsilon_z$ in each case.

### 7.8 Gauss’ law

We showed previously that the electric field, $E(r)$, due to a charge distribution, $\rho(r)$, satisfies $\nabla \cdot E = \rho/\epsilon_0$. This is the differential form of Maxwell’s first equation.

Integrating $\nabla \cdot E = \rho/\epsilon_0$ over a volume $V$, bounded by a closed surface $S$, and using the divergence theorem

$$\int_V \nabla \cdot E \, dV = \int_S E \cdot dS$$

gives Gauss’ law

$$\int_S E \cdot dS = \frac{1}{\epsilon_0} \int_V \rho \, dV = \frac{Q_{\text{enc}}}{\epsilon_0}$$

where $Q_{\text{enc}}$ is the total charge enclosed by the volume $V$. [For brevity, we will often drop the subscript ‘enc’.]

Gauss’ Law (also known as Gauss’ Theorem) is the integral form of Maxwell’s first equation. It’s extremely useful, particularly for problems with a symmetry\(^2\), for solving problems in potential theory, and for determining the behaviour of fields at boundaries.

**Examples**

- Consider a sphere of radius $a$, carrying charge $Q$, centred on the origin, with uniform charge density $\rho_0 = Q/(4\pi a^3)$.

  By symmetry, the electric field will point radially outwards, so that $E(r) = E_r(r) \epsilon_r$.

  Integrating over a sphere of radius $r$, we find

  $$\int_S E \cdot dS = E_r(r) 4\pi r^2 = \begin{cases} 
  \frac{4\pi r^3 \rho_0}{\epsilon_0} & r \leq a \\
  \frac{4\pi a^3 \rho_0}{\epsilon_0} & r \geq a
\end{cases}$$

\(^2\)This goes against the general rule that it is easier to compute the potential (a scalar) first rather than the electric field (a vector)!
Therefore

\[ E(r) = \begin{cases} 
\frac{Q}{4\pi \epsilon_0} \frac{r}{a^3} & r \leq a \\
\frac{Q}{4\pi \epsilon_0} \frac{r}{r^3} & r \geq a 
\end{cases} \]

Outside the sphere, the electric field appears to come from a point source, and inside it increases linearly with \( r \).

We can obtain the electrostatic potential from

\[ E(r) = E_r(r) e_r = -\nabla \phi(r) = -\frac{\partial \phi(r)}{\partial r} e_r \]

Integrating with respect to \( r \) gives

\[ \phi(r) = \begin{cases} 
-\frac{Q}{4\pi \epsilon_0} \left( \frac{r^2}{2a^3} + C_1 \right) & r \leq a \\
\frac{Q}{4\pi \epsilon_0} \left( \frac{1}{r} + C_2 \right) & r \geq a 
\end{cases} \]

For \( r > a \) we chose the constant of integration \( C_2 \) to be zero so that \( \phi \to 0 \) as \( r \to \infty \). The potential outside the sphere is again that of a point charge.

For \( r < a \), we chose the constant of integration \( C_1 = -3/(2a) \) so that \( \phi \) is continuous across the boundary at \( r = a \). [Exercise: check this calculation.]

Note: \( E \) has a cusp, so the derivative \( \partial E_r / \partial r \) is discontinuous at the boundary.

- Consider a long (infinite) straight wire with constant charge/unit length \( \lambda \).

Using cylindrical coordinates with the \( z \) axis parallel to the wire, we integrate over a cylinder of length \( L \) and radius \( \rho \) with its axis along the wire. By symmetry we must have \( E = E_\rho(\rho) e_\rho \).

Using Gauss’ Law, we get

\[ \int_S E \cdot dS = E_\rho(\rho) 2\pi \rho L + \left. \left. 0 \right|_{\text{ends}} \right| = \frac{1}{\epsilon_0} \lambda L \]

This gives

\[ E(\rho) = \frac{\lambda}{2\pi \epsilon_0} \frac{1}{\rho} e_\rho \]

The potential can be found by integrating \( E_\rho = -\partial \phi / \partial \rho \), which gives

\[ \phi(\rho) = -\frac{\lambda}{2\pi \epsilon_0} \ln \rho + \text{constant} = -\frac{\lambda}{2\pi \epsilon_0} \ln \left( \frac{\rho}{\rho_0} \right) \]

where we chose the constant of integration to give \( \phi(r) = 0 \) when \( \rho = \rho_0 \), where \( \rho_0 \) is a constant.
• Infinite flat sheet of charge with constant charge density $\sigma$ per unit area. Integrate over a cylindrical ‘Gaussian pill box’ with axis perpendicular to the sheet. See tutorial, and below.

7.9 Boundaries

Useful results for the changes in the normal and tangential components of the electric field across a boundary may be obtained using Gauss’ Law and Stokes’ theorem.

Consider a surface carrying surface charge density $\sigma$. The electric field on one side of the boundary is $E_1$, and on the other $E_2$. The unit normal to the surface is $\mathbf{n}$.

7.9.1 Normal component

Consider a small cylindrical ‘Gaussian pillbox’ of area $A$ and negligible height $\delta l$ (so that $\delta l \ll \sqrt{A}$), which straddles the surface. If $A$ is sufficiently small, $E(r)$ is approximately constant over $A$, but due to the charge density $\sigma$ on the surface, $E(r)$ will be different on the top and bottom of the pillbox. Apply Gauss’ Law

$$\int_S E \cdot dS = \frac{1}{\varepsilon_0} \int \rho dV$$

and recall $dS = n \, dS$, then for small $A$

$$(E_2 - E_1) \cdot n \, A = (E_{2\perp} - E_{1\perp}) \, A = \frac{\sigma A}{\varepsilon_0}$$

where $E_{1\perp}$ and $E_{2\perp}$ are the components of the electric field perpendicular to the surface. The factors of $A$ cancel, hence

$$n \cdot (E_2 - E_1) = \frac{\sigma}{\varepsilon_0}$$

Thus the normal component of the electric field $E$, is discontinuous across the boundary when $\sigma \neq 0$.

In this case the discontinuity is proportional to the surface charge density.
7.9.2 Tangential component

Consider a small rectangle of length \( l \), and negligible width \( \delta l \) (so that \( \delta l \ll l \)), which straddles the surface. Applying Stokes’ theorem

\[
\int_S \nabla \times \vec{E} \cdot d\vec{S} = \oint_C \vec{E} \cdot d\vec{r},
\]

we can ignore contributions from the ends with length \( \delta \ell \). Since \( \nabla \times \vec{E} = 0 \), we get

\[
0 = \oint_C \vec{E} \cdot d\vec{r} = (E_{2||} - E_{1||}) l
\]

where \( E_{1||} \) and \( E_{2||} \) are the components of the electric field parallel to the boundary.

This can be written as

\[
\vec{n} \times \vec{E}_1 = \vec{n} \times \vec{E}_2
\]

where we used the fact that the cross product of the electric field \( \vec{E} \) with \( \vec{n} \) picks out the tangential component \( E_\parallel \) of the electric field, since

\[
\vec{E} = E_\perp \vec{n} + E_\parallel
\]

Thus the tangential component of \( \vec{E} \) is continuous across a charged boundary.

7.9.3 Conductors

Physically, a conductor is a material in which ‘free’, or ‘nearly free’ or ‘surplus’ electrons can move (or flow) freely when an electric field is applied.

In Electrostatics

- For a conductor in equilibrium, where all charges are at rest, all the charge resides on the surface of the conductor, i.e. \( \rho = 0 \) inside a conductor.

This holds because if \( \rho \neq 0 \), then due to Maxwell’s first equation \( \nabla \cdot \vec{E} = \rho/\epsilon_0 \) (or Gauss’ law), so we must have \( \vec{E} \neq 0 \), and hence the charge would move and we wouldn’t have equilibrium – a contradiction. So \( \vec{E} = 0 \) and hence \( \phi = \text{constant} \) everywhere inside a conductor.

Physically, the charges repel and move to the surface.

- The electric field on the surface of a conductor is normal to the surface, i.e. \( \vec{E} \parallel \vec{n} \), otherwise charge would move along the surface.

Thus if \( d\vec{r} \) is a displacement on the surface of a conductor, \( \vec{E} \cdot d\vec{r} = -d\phi = 0 \), so \( \phi = \text{constant} \) on the surface of a conductor, i.e. an equipotential.
Therefore, on the *surface* of a conductor,

\[
E_t = 0, \quad E_n = \frac{\sigma}{\epsilon_0}
\]

The external electric field *induces* a charge on the surface of the conductor, which in turn deforms the external field so that it is perpendicular to the conductor surface. In the case of a conductor, the surface charge is calculated from the electric field (and not vice-versa as is the usual case).

For *insulators* we have the opposite situation – the charges are *fixed* and we must calculate the electric field and the potential from the charge density.

### 7.10 Poisson’s equation

In section (7.5.3), we showed that \( \nabla \times E = 0 \), and hence there exists a potential \( \phi \) such that \( E = -\nabla \phi \). Substituting this into Maxwell’s first equation \( \nabla \cdot E = \rho / \epsilon_0 \) gives Poisson’s equation

\[
\nabla^2 \phi = -\frac{\rho}{\epsilon_0}
\]

If we know \( \rho \), this may be solved for \( \phi \) (or vice versa), given appropriate boundary conditions (\( bcs \)). In a charge-free region (i.e. \( \rho = 0 \)) Poisson’s equation becomes *Laplace’s equation*

\[
\nabla^2 \phi = 0
\]

### 7.10.1 Uniqueness of solution

Partial differential equations have many linearly-independent solutions. How do we know which is the correct one for a given physical system?

**Theorem:** For a volume \( V \) bounded by a surface (or set of surfaces) \( S \), if we are given a set of *boundary conditions*:

*Either* on the potential \( \phi \),

\[
\phi(r) \quad \text{for} \quad r \in S \quad \text{Dirichlet bcs}
\]

Or the normal component of \( E \), i.e. \( E_n = -n \cdot \nabla \phi \),

\[
\frac{\partial \phi}{\partial n} = n \cdot \nabla \phi(r) \quad \text{for} \quad r \in S \quad \text{Neumann bcs}
\]

where \( n \) is the unit normal to \( S \), then the solution of Poisson’s equation in \( V \) is *unique.*
Proof: Let \( \phi_1(r) \) and \( \phi_2(r) \) be two solutions of Poisson’s equation, both of which satisfy the boundary conditions, and let \( \psi = \phi_1 - \phi_2 \), so that \( \nabla^2 \psi = 0 \) for both sets of boundary conditions.

Applying the divergence theorem to the LHS of the vector calculus identity
\[
\nabla \cdot (\psi \nabla \psi) = \psi \nabla^2 \psi + \nabla \psi \cdot \nabla \psi
\]
gives
\[
\int_S \psi (\nabla \psi \cdot n) \, dS = \int_V \left( \psi \nabla^2 \psi + |\nabla \psi|^2 \right) \, dV
\]
Since either \( \psi = 0 \) (Dirichlet) or \( \partial \psi / \partial n = 0 \) (Neumann) on \( S \), then
\[
\int_V |\nabla \psi|^2 \, dV = 0
\]
This implies that \( \nabla \psi = 0 \) everywhere in \( V \), which integrates to \( \psi = \text{constant} \).

Consider the Dirichlet and Neumann boundary condition cases separately:

**Dirichlet:** In this case \( \psi = 0 \) on \( S \), so the constant is zero and the two solutions are equal
\[
\phi_1 = \phi_2
\]

**Neumann:** In this case
\[
\phi_1 = \phi_2 + \text{constant}
\]
Since the potential is only defined up to a constant, the constant may be disregarded.

Thus both types of boundary conditions give a unique solution of Poisson’s equation.

Notes:

- We can specify either Dirichlet or Neumann boundary conditions at each point on the boundary, but not both. To specify both is generally inconsistent, since the solution is then overdetermined.
- However, we can specify either Dirichlet or Neumann boundary conditions on different parts of the surface.
- The uniqueness theorem means we can use any method we wish to obtain the solution - if it satisfies the correct boundary conditions, and is a solution of the equation, then it is the correct solution.

7.10.2 Methods of solution

The theorem is useful: if you find a solution somehow, it is the solution. For example:

(i) Guesswork

(ii) Numerical methods
(iii) Direct integration of Poisson’s equation
(iv) Gauss’ law plus symmetry
(v) Method of images
(vi) Separation of variables
(vii) Green-function method

Method (iii) uses ‘direct’ integration to find the (unique) solution of Poisson’s equation, \( \phi(r) = \frac{1}{4\pi\varepsilon_0} \int_V \frac{\rho(r')}{|r-r'|} \, dV \)
which is okay for simple situations with a bounded potential.

We have studied method (iv) already.

### 7.10.3 The method of images

In the method of images, we add fictitious charges outside the volume under consideration in such a way that the system including the fictitious charges satisfies Poisson’s equation in the region of interest, plus the correct boundary conditions.

\[ \rho + \text{bcs} \iff \rho + \text{mirror charges in unphysical region to mimic bcs} \]

**Example: Point charge and a planar conducting surface**

In the figure, the point charge \( q \) is at position \((0, 0, z_0)\) above an infinite conducting surface in the \( x-y \) plane (on the left in the figure). The region of interest \( V \) is the half-space \( z > 0 \). The image charge \(-q\) is at position \( (0, 0, -z_0) \) below the \( x-y \) plane, so it’s not in the physical region.

The potential due to the pair of charges is

\[ \phi(r) = \frac{1}{4\pi\varepsilon_0} \left[ \frac{q}{(x^2 + y^2 + (z - z_0)^2)^{\frac{1}{2}}} - \frac{q}{(x^2 + y^2 + (z + z_0)^2)^{\frac{1}{2}}} \right] \]
Since $\phi$ satisfies Poisson’s equation everywhere in the half-space ($z > 0$), and

$$\phi(r)|_{z=(x,y,0)} = 0$$

then the surface of the conductor (at $z = 0$) is an equipotential, i.e. $\phi$ satisfies the boundary conditions, so $\phi$ is the unique solution.

**Electric field:** We may calculate the electric field from the potential

$$E(r) = -\nabla \phi = \frac{q}{4\pi \varepsilon_0} \left[ \frac{(x, y, z - z_0)}{(x^2 + y^2 + (z - z_0)^2)^{3/2}} - \frac{(x, y, z + z_0)}{(x^2 + y^2 + (z + z_0)^2)^{3/2}} \right]$$

On the surface of the conductor ($z = 0$) this becomes

$$E(r)|_{z=(x,y,0)} = -\frac{q}{2\pi \varepsilon_0} \frac{z_0}{(x^2 + y^2 + z_0^2)^{3/2}} \hat{e}_z$$

i.e. the field at the surface of the conductor is perpendicular to the surface, as it must be.

The **surface charge density** on the conductor is

$$\sigma(r) = \varepsilon_0 E(r)|_{z=(x,y,0)} = -\frac{q}{2\pi} \frac{z_0}{(x^2 + y^2 + z_0^2)^{3/2}}$$

The **total charge** induced on the conducting surface is, using polar coordinates $(\rho, \chi)$,

$$Q = \int_{z=0} \sigma \, dS$$

$$= -\frac{q}{2\pi} z_0 \int_0^{2\pi} d\chi \int_0^\infty \rho \, d\rho \frac{1}{(\rho^2 + z_0^2)^{3/2}} = -q$$

which is just the mirror charge, as one would expect.

**Force** The force on the positive charge due to the conductor is that between two point charges separated by distance $2z_0$

$$F = -\frac{q^2}{4\pi \varepsilon_0 (2z_0)^2} \hat{e}_z$$
Example: point charge and a metal sphere

Consider an earthed sphere of radius $a$ centred on the origin, together with a point charge $q$ at point $P$ which lies a distance $b$ from the centre of the sphere. The volume $V$ is the space $r \geq a$ and we choose the mirror charge $q' = -qa/b$, at $P'$ where $OP' = a^2/b$, in order to satisfy the boundary conditions at $r = a$.

[The potential on an earthed conductor is defined to be $\phi = 0$ everywhere on and inside the conductor – the potential of the ‘earth’.

The potential of the two-charge system is

$$\phi(r) = \frac{1}{4\pi\epsilon_0} \left[ \frac{q}{r_1} + \frac{q'}{r_2} \right]$$

$$\phi(r) = \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{(r^2 + b^2 - 2br\cos\theta)^{1/2}} - \frac{a/b}{(r^2 + (a^4/b^2) - 2(a^3/b) r\cos\theta)^{1/2}} \right]$$

On the boundary ($r = a$)

$$\phi(r)|_{r=a} = \frac{q}{4\pi\epsilon_0} \left[ \frac{1}{(a^2 + b^2 - 2ab\cos\theta)^{1/2}} - \frac{a/b}{(a^2 + (a^4/b^2) - 2(a^3/b) \cos\theta)^{1/2}} \right] = 0$$

and hence by the uniqueness theorem, this is the correct solution.

Example: Earthed conducting sphere in a uniform electric field

Consider an earthed conducting sphere of radius $a$ centred on the origin, and subject to a uniform external electric field $E$.

$$\phi(r) = -Ez + \frac{(Be_z) \cdot r}{r^3} = -Er\cos\theta + \frac{B}{r^3}\cos\theta$$

For a constant electric field $E = (0, 0, E)$ we have $\phi = -Ez = -Er\cos\theta$. We anticipate that the field will induce a dipole moment parallel to $e_z$ on the sphere, so we use a linear superposition as an ansatz (guess) for the potential:
This is a solution of Laplace’s equation outside the sphere, where $\rho = 0$. As $r \to \infty$, we have $\phi \to -Er \cos \theta$. The boundary condition on the surface of the sphere is

$$\phi|_{r=a} = -Ea \cos \theta + \frac{B}{a^2} \cos \theta = 0$$

which gives $B = a^3 E$

Therefore the (unique) solution for this problem is

$$\phi(r, \theta) = -Er \cos \theta + \frac{a^3 E}{r^2} \cos \theta$$

Alternatively, we can use the method of images, as illustrated below. The solution to the constant-external-field problem is obtained in the limit that the charges outside the sphere move to infinity, and the ones inside move inwards to form a dipole at the centre.

7.11 Electrostatic energy

7.11.1 Electrostatic energy of a general charge distribution

Recall that the work done in moving a point charge $q$ in a field $E$ from $r_a$ to $r_b$ is

$$W_e = -q \int_{r_a}^{r_b} E \cdot dr = q \int_{r_a}^{r_b} \nabla \phi \cdot dr = \int_{r_a}^{r_b} d\phi = q \left( \phi(r_b) - \phi(r_a) \right)$$

Thus the work done to bring in a point charge $q$ from infinity (where $\phi = 0$) to position $r$ is $W = q \phi(r)$.

Let’s assemble a system of $n$ charges $q_i$ from $\infty$ to positions $r_i$:

$$W_1 = 0 \quad \text{[nothing else is there]}$$

$$W_2 = q_2 \frac{1}{4\pi\varepsilon_0} \frac{q_1}{r_1 - r_2} \quad \text{[the work done in the field of } q_1 \text{]}$$

$$W_3 = q_3 \frac{1}{4\pi\varepsilon_0} \left[ \frac{q_1}{r_1 - r_3} + \frac{q_2}{r_2 - r_3} \right] \quad \text{[the work done in the field of } q_1 \text{ and } q_2 \text{]}$$

and so on. Therefore the work done to bring in the $i^{th}$ charge $q_i$ to position $r_i$ is

$$W_i = \frac{q_i}{4\pi\varepsilon_0} \sum_{j=1}^{i-1} \frac{q_j}{|r_j - r_i|}$$
The total work done is

\[ W_e = \sum_{i} W_i = \sum_{i} \sum_{j=1}^{i-1} \frac{1}{4\pi \epsilon_0} \frac{q_i q_j}{r_{ij}} = \frac{1}{2} \sum_{i,j,(i \neq j)}^{n} \frac{1}{4\pi \epsilon_0} \frac{q_i q_j}{|r_j - r_i|} \equiv \frac{1}{2} \sum_i q_i \phi_i \]

which we may write as

\[ W_e = \frac{1}{2} \sum_i q_i \phi_i \quad \text{with} \quad \phi_i = \phi(r_i) = \sum_{j,(j \neq i)} \frac{q_j}{4\pi \epsilon_0 |r_j - r_i|} \]

where \( \phi_i \) is the potential felt by \( q_i \) due to all the other charges.\(^3\)

In the limit of a continuous charge distribution \( \rho(r) \), we have

\[ W_e = \frac{1}{2} \int_V dV \rho(r) \phi(r) \]

where the integral is over all space.

### 7.11.2 Field energy

We can write \( W_e \) in terms of the (total) electric field \( E \) using Maxwell’s first equation in the form \( \rho = \epsilon_0 \nabla \cdot E \). Then \( W_e \) becomes

\[ W_e = \frac{\epsilon_0}{2} \int dV \phi (\nabla \cdot E) \]

Now use the product rule to rewrite

\[ \phi (\nabla \cdot E) = \nabla \cdot (\phi E) - (\nabla \phi) \cdot E = \nabla \cdot (\phi E) + |E|^2 \]

so that

\[ W_e = \frac{\epsilon_0}{2} \int dV \left( \nabla \cdot (\phi E) + |E|^2 \right) \]

Applying the divergence theorem to the integral of the first term, taking \( V \) to be all space, and \( S \) to be a large sphere of radius \( R \to \infty \), this integral becomes

\[ \int_V \nabla \cdot (\phi E) \, dV = \int_S (\phi E) \cdot dS \approx O \left( \frac{1}{R} \frac{1}{R^2} 4\pi R^2 \right) \to 0 \text{ as } R \to \infty \]

The total energy stored in the electric field is then

\[ W_e = \frac{\epsilon_0}{2} \int dV |E(r)|^2 \]

\(^3\)The factor of 1/2 ensures that we don’t count the energy required to bring in charge \( j \) from \( \infty \) in the presence of charge \( i \) at \( r_i \) and the energy required to bring in charge \( i \) from \( \infty \) in the presence of charge \( j \) at \( r_j \). Note there is no factor of 1/2 in our previous expression for the energy of a charge in an external electrostatic field.
and the energy density (energy/unit volume) at \( r \) is

\[
 w_e(r) = \frac{\epsilon_0}{2} |E(r)|^2
\]

**Notes:** The general result for \( W_e \) was derived for a continuous charge distribution. When there are point charges we have to be careful with self-energy contributions which should be excluded from the integral because they lead to divergences.

The two boxed expressions for \( W_e \) are complementary. We can think of the electrostatic energy as lying in the charge distribution or as being stored in the \( E \) field.

Finally, note that since the energy density \( w_e \) is quadratic in the electric field, we don’t have superposition of energy density.

### 7.12 Capacitors (condensers) and capacitance

A capacitor is formed from a pair of conductors 1 and 2 carrying equal and opposite charges, \( Q \) and \( -Q \). The potentials on the conductors are \( \phi_1 \) and \( \phi_2 \), so the potential difference is \( V = \phi_1 - \phi_2 \). Clearly, \( \phi_1 \) and \( \phi_2 \) are proportional to \( Q \) (up to a common constant), so \( V \propto Q \), and we define capacitance

\[
 C = \frac{Q}{V}
\]

which depends only on the geometry of the capacitor. The SI unit of capacitance is the farad or Coulomb per Volt, \( 1 \text{F} = 1 \text{CV}^{-1} \).

#### 7.12.1 Parallel-plate capacitor

The simplest example is the parallel plate capacitor

Two parallel plates of area \( A \) have a separation \( a \) (with \( a \ll \sqrt{A} \)), and carry surface charge densities \( +\sigma \) and \( -\sigma \) on their inner surfaces (because of the attractive force between the charges on the two plates).

We can obtain the electric field using Gauss’ Law. First take a pillbox that straddles the inner surface of the upper plate (say). \( E = 0 \) inside the plate, and between the two plates the field is normal to the inner surface of the upper plate, which gives

\[
 E_{\text{inside}} = \left( \frac{+\sigma}{\epsilon_0} \right) (-\varepsilon_z) = \frac{-\sigma}{\epsilon_0} \varepsilon_z
\]
Now take a pillbox that straddles the outer surface of the upper plate. Since there is no charge on the upper surface and $E = 0$ inside the plate, we have $E = 0$ at the outer surface. Similarly for the lower plate. Therefore

$$E_{\text{outside}} = 0 \quad E_{\text{inside}} = -\frac{\sigma}{\epsilon_0} e_z$$

We can obtain the potential between the plates in the usual way

$$E_z = -\frac{\sigma}{\epsilon_0} = -\frac{\partial \phi}{\partial z} \quad \Rightarrow \quad \phi(z) = \frac{\sigma z}{\epsilon_0} + \text{constant}$$

so the potential difference between the plates is

$$V = \frac{\sigma a}{\epsilon_0} = \frac{Q a}{A \epsilon_0}$$

and the capacitance is

$$C = \frac{Q}{V} = \frac{A \epsilon_0}{a}$$

which is a purely geometrical property of the plates, as expected.

### 7.12.2 Concentric conducting spheres

Consider a capacitor consisting of two concentric spheres, radii $a$ and $b$, with $b > a$, centred on the origin. The outer sphere carries charge $+Q$ and the inner one $-Q$.

By Gauss' law, the electric field is zero inside the inner sphere, and outside the outer one [exercise].

By Gauss' law, between the spheres, we have

$$\phi(r) = -\frac{Q}{4\pi \epsilon_0} \frac{1}{r} + \text{constant} \quad \text{for } a < r < b$$

The potential difference is

$$V = \phi_b - \phi_a = -\frac{Q}{4\pi \epsilon_0} \left( \frac{1}{b} - \frac{1}{a} \right) = \frac{Q}{4\pi \epsilon_0} \frac{b - a}{ab}$$

so the capacitance is

$$C = 4\pi \epsilon_0 \frac{ab}{b - a}$$

### 7.12.3 Energy stored in a capacitor

Using our first expression for the energy of a charge distribution, we have

$$W = \frac{Q}{2} (\phi_1 - \phi_2) = \frac{QV}{2} = \frac{CV^2}{2} = \frac{Q^2}{2C}$$

which holds for a capacitor of any shape. For parallel plates

$$W_e = \frac{\epsilon_0}{2} \int |E|^2 \, dV = \frac{\epsilon_0}{2} (Aa) \left( \frac{\sigma}{\epsilon_0} \right)^2 = \frac{a}{2\epsilon_0} \frac{Q^2}{A} = \frac{Q^2}{2C}$$

as before.
Chapter 8

Magnetostatics

In the previous chapter, we studied static charge distributions, which lead to an electric field. In this chapter we study the case of steady currents (also known as time-independent currents), which lead to a magnetic field.

8.1 Currents

A current is created by moving charges: a charge \( q \) moving at velocity \( v \) gives an ‘elementary’ current \( \vec{j} = q \vec{v} \), which is a vector quantity.

Consider a volume charge density \( \rho(r) \) moving with velocity \( \vec{v}(r) \). This defines a (bulk) current density,

\[
\vec{J}(r) = \rho(r) \vec{v}(r)
\]

We can also have a surface current density,

\[
\vec{K}(r) = \sigma(r) \vec{v}(r)
\]

where \( \sigma(r) \) is the surface charge density (charge/area), and a line current

\[
\vec{I}(r) = \lambda(r) \vec{v}(r)
\]

where \( \lambda(r) \) is the line charge density (charge/length).

We define a current element \( d\vec{I} \) for each of these three cases by

\[
d\vec{I}(r) = \begin{cases} 
J(r) \, dV & \text{current element in the bulk} \\
K(r) \, dS & \text{current element in a surface} \\
I(r) \, dr & \text{current element along a line (or wire)}
\end{cases}
\]

**Note:** Care is needed with current elements. For example, \( \vec{K} \, dS \neq K \, d\vec{S} \) since the left hand side points in the direction of the current vector on the surface, but the right hand side points normal to the surface. On the other hand \( \vec{I} \, dr = I \, dr \) since a line current element always points along the wire in the direction \( dr \).

**Units:** From the definition of line current, the dimension of current \( I \) is

\[
[I] = (C \, m^{-1}) \, (m \, s^{-1}) = (C \, s^{-1}) \equiv A \, \text{(Ampères)}
\]
The unit of current \( I \) is called the Ampère (A), \( 1 \text{A} = 1 \text{C} \, \text{s}^{-1} \) (Coulombs/second). Similarly

\[
[K] = \text{A} \, \text{m}^{-1} \quad [J] = \text{A} \, \text{m}^{-2} \quad [dI] = \text{A} \, \text{m}
\]

Note that current \( I \) and current element \( dI \) have different units,\(^1\) and that none of these current ‘densities’ has units of current/volume! This is a little confusing, but it’s standard . . .

The total (scalar) current \( I \) passing through a surface \( S \) is the flux of \( J \) through \( S \)

\[
I = \int_S J \cdot d\mathbf{S}
\]

where \( d\mathbf{S} \) is normal to the surface. Similarly, the flux of \( K \) across the curve \( C \) is

\[
I = \int_C K \cdot n' \, dr
\]

where the unit vector \( n' \) is normal to \( C \), in the plane of \( K \).

8.1.1 Charge and current conservation

Experiment: total (net) charge is conserved; charge can’t be created or destroyed.

Consider a volume \( V \) bounded by a closed surface \( S \). The charge \( Q \) in the volume \( V \) changes due to current flowing across the surface \( S \), with outward vector element of area \( d\mathbf{S} \):

\[
\frac{\partial Q(t)}{\partial t} = \frac{\partial}{\partial t} \int_V \rho(r, t) \, dV = \int_V \frac{\partial \rho(r, t)}{\partial t} \, dV = -\int_S J(r, t) \cdot d\mathbf{S} = -\int_V \nabla \cdot J(r, t) \, dV
\]

where \( \int_S J(r, t) \cdot d\mathbf{S} \) is the total current flow across \( S \). The minus sign is due to current flowing out of the surface reducing the charge in \( V \), and we used the divergence theorem in the last step. This equation holds for all volumes \( V \), so

\[
\frac{\partial \rho(r, t)}{\partial t} + \nabla \cdot J(r, t) = 0
\]

which describes charge conservation locally at the point \( r \).

In static situations, \( \partial \rho/\partial t = 0 \) (by definition), so \( \rho(r, t) = \rho(r) \) is independent of \( t \), and \( \nabla \cdot J = 0 \), which is called current conservation.

**Example:** Consider a disc, carrying uniform charge density \( \sigma \), rotating about its axis at angular velocity \( \omega \).

The current density on the disc is

\[
K = \sigma v = \sigma \omega \times r
\]

Hence

\[
\nabla \cdot K = \sigma \nabla \cdot (\omega \times r) = 0
\]

\(^1\)Some authors use different fonts \( I \) and \( I \) for these quantities, but this is hard to do on the blackboard!
8.1.2 Conduction current

A common situation is where we have no net electric charge, but a current exists because positive and negative charges move with different velocities

\[ \mathbf{J} = \rho^+ \mathbf{v}^+ + \rho^- \mathbf{v}^- \quad \text{with} \quad \rho = \rho^+ + \rho^- = 0, \quad \text{but} \quad |\mathbf{v}^+| \neq |\mathbf{v}^-| \]

**Examples**

**Metal:** nuclei are fixed, but electrons move: \( \mathbf{v}^+ = 0 \), so \( \mathbf{J} = \rho^- \mathbf{v}^- \)

**Electrolyte:** positive and negative ions move with different velocities: \( \mathbf{J} = \rho^+ (\mathbf{v}^+ - \mathbf{v}^-) \)

Current is often created by an electric field, which forces charge carriers to move. A simple linear model, which agrees with experiment in common situations is *Ohm’s law*

\[ \mathbf{J} = \sigma \mathbf{E} \]

where \( \sigma \) is called the conductivity\(^2\).

In some materials \( \mathbf{J} \) is not in general parallel to \( \mathbf{E} \), and the linear model must be generalised to

\[ J_i = \sigma_{ij} E_j \]

where \( \sigma_{ij} \) are the components of a second-rank tensor, the *conductivity tensor*.

**Notes**

(i) When \( \rho = 0 \), we have \( \nabla \cdot \mathbf{E} = 0 \) (using Maxwell’s first equation) and \( \nabla \cdot \mathbf{J} = 0 \) (current conservation).

(ii) We cannot have a *static* closed current loop. Using Stokes’ theorem

\[ \oint_{C} \mathbf{E} \cdot d\mathbf{r} = \int_{S} (\nabla \times \mathbf{E}) \cdot d\mathbf{S} = 0 \]

because \( \nabla \times \mathbf{E} = 0 \) for static electric fields. Ohm’s law, \( \mathbf{J} = \sigma \mathbf{E} \), then implies

\[ \oint_{C} \mathbf{J} \cdot d\mathbf{r} = 0 \]

so \( I = 0 \) for a closed loop.

So we must have a battery to keep current flowing. [But see next semester for non-static situations.]

In static situations, we can write \( \mathbf{E} = -\nabla \phi \), so

\[ \int_{V_1}^{V_2} \mathbf{E} \cdot d\mathbf{r} = \phi(r_1) - \phi(r_2) = V_{12} \]

*i.e.* the potential supplied by the battery. This is called the *electromotive force* or *emf* \( \mathcal{E} \).

This is not a sensible name because emf is *not* a force, but we’re stuck with it.

\(^2\)Not to be confused with the surface charge density!
(iii) In a perfect conductor (for example a superconductor), $\sigma \to \infty$, so to keep $J$ finite, we must have $E = 0$ (as in electrostatics).

(iv) For an insulator, $\sigma = 0$.

### 8.2 Forces between currents (Ampère, 1821)

Consider two parallel wires, of length $L$, a distance $d$ apart, carrying currents $I_1$ and $I_2$.

From experiment, we find:

- For perpendicular currents, there is no force.
- Otherwise, the force between current elements $dI_1$ and $dI_2$ is a central force with an inverse square law. The force on infinitesimal current element 1 due to infinitesimal current element 2 in SI units is

$$dF_{12} = -\frac{\mu_0}{4\pi} \frac{dI_1 \cdot dI_2}{|r_1 - r_2|^3} = -\frac{\mu_0}{4\pi} dI_1 \cdot dI_2 \frac{\hat{r}_{12}}{r_{12}^2}$$

where $\hat{r}_{12} = \frac{r_1 - r_2}{|r_1 - r_2|}$.

The constant $\mu_0$ is measured to be $\mu_0 = 1.2566370614 \ldots \times 10^{-6}$ N A$^{-2} = 1.0000000082(20) \times 4\pi \times 10^{-7}$ N A$^{-2}$ (or N C$^{-2}$ S$^2$), and is called the permeability of the vacuum (or of free space), or just the magnetic constant.

For two current loops $C_1$ and $C_2$ carrying currents $I_1$ and $I_2$
Linear superposition (which comes from experiment) gives

\[ F_{12} = -\frac{\mu_0}{4\pi} \oint_{C_1} \oint_{C_2} (I_1 \, dr_1) \cdot (I_2 \, dr_2) \frac{r_{12}}{r_{12}^3} \]

where we used the expression for current elements \( dI_1 \cdot dI_2 = I_1 \, dr_1 \cdot I_2 \, dr_2 \).

This is the analog for currents of Coulomb’s law. Note that \( F_{21} = -F_{12} \) as expected.

We now separate this into two parts. The loop carrying current \( I_2 \) produces a magnetic field \( B(r) \), which in turn produces a force on the loop carrying \( I_1 \). Start by writing

\[ dr_1 \times (dr_2 \times r_{12}) = dr_2 (dr_1 \cdot r_{12}) - (dr_1 \cdot dr_2) r_{12} \]

Using Stokes’ theorem,

\[ \oint_{C_1} dr_1 \cdot \nabla \left( \frac{1}{r_{12}} \right) = -\oint_{C_1} dS_1 \cdot \left( \nabla \times \nabla \left( \frac{1}{r_{12}} \right) \right) = 0 \]

where \( \nabla \) is the gradient with respect to \( r_1 \), and curl-grad is always zero. Therefore

\[ F_{12} = \frac{\mu_0}{4\pi} I_1 I_2 \left\{ \oint_{C_1} \oint_{C_2} \frac{dr_1 \times (dr_2 \times r_{12})}{r_{12}^3} - \oint_{C_2} dr_2 \oint_{C_1} \frac{dr_1 \cdot r_{12}}{r_{12}^3} \right\} \]

\[ = \oint_{C_1} I_1 \, dr_1 \times B(r_1) \]

which is the force on \( C_1 \) due to \( B(r_1) \), where

\[ B(r_1) = \frac{\mu_0}{4\pi} \oint_{C_2} \frac{I_2 \, dr_2 \times r_{12}}{r_{12}^3} = \frac{\mu_0}{4\pi} \oint_{C_2} \frac{I_2 \, dr_2 \times (r_1 - r_2)}{|r_1 - r_2|^3} \]

is the magnetic field\(^3 \) at \( r_1 \) due to the current in \( C_2 \). This is the Bio-Savart law, (approx 1820).

### 8.2.1 The Lorentz force

Generalising, the force on a current element \( dI \) due to a magnetic field \( B \) is (locally)

\[ dF = dI \times B \]

so

\[ F = \begin{cases} \int_C I \, dr \times B & \text{for a line current} \\ \int_S K \times B \, dS & \text{for a surface current} \\ \int_V J \times B \, dV & \text{for a bulk current} \end{cases} \]

\(^3\)In some textbooks, \( B \) is called the magnetic induction.
The force on a charge distribution \( \rho \) due to an electric field \( E \) is \( \int_V \rho E \, dV \), so the force density (force/unit volume) at \( r \) is

\[
\mathbf{f}(\mathbf{r}) = \rho(\mathbf{r}) \mathbf{E}(\mathbf{r})
\]

Similarly, for a bulk current density in a magnetic field, we often write \( \mathbf{f} = \mathbf{J} \times \mathbf{B} \)

Therefore, the combined force density is

\[
\mathbf{f} = \mathbf{\rho E} + \mathbf{J} \times \mathbf{B}
\]

which is called the Lorentz force (strictly, the force density).

For a moving point charge \( q \) at \( r' \), we have \( \rho(\mathbf{r}) = q \delta(\mathbf{r} - \mathbf{r}') \) and \( \mathbf{J}(\mathbf{r}) = q \mathbf{v} \delta(\mathbf{r} - \mathbf{r}') \).

Integrating the Lorentz force density over a volume containing the charge gives

\[
\mathbf{F} = q \left( \mathbf{E} + \mathbf{v} \times \mathbf{B} \right)
\]

We may regard these as the definitions of the fields \( \mathbf{E} \) and \( \mathbf{B} \).

Note that magnetic fields do no work. If a charge \( q \) in a magnetic field \( \mathbf{B} \) moves a distance \( \mathbf{dr} = \mathbf{v} \, dt \), the work done by the field is

\[
dW = \mathbf{F} \cdot d\mathbf{r} = q \left( \mathbf{v} \times \mathbf{B} \right) \cdot \mathbf{dr} = q \left( \mathbf{v} \times \mathbf{B} \right) \cdot \mathbf{v} \, dt = q \left( \mathbf{v} \times \mathbf{v} \right) \cdot \mathbf{B} \, dt = 0
\]

So magnetic fields change the direction of charged particles, but do not accelerate them to higher (or lower) speeds.

Units: \( [\mathbf{B}] = \text{NC}^{-1}\text{m}^{-1}\text{s} = \text{NA}^{-1}\text{m}^{-1} = \text{T} \) (Tesla). \( [\mathbf{E}] \) and \( [\mathbf{B}] \) have different units.

### 8.3 Biot-Savart law

The magnetic field \( d\mathbf{B}(\mathbf{r}) \) at \( \mathbf{r} \) due to a current element \( I \, d\mathbf{r}' \) at \( \mathbf{r}' \) is

\[
d\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{I \, d\mathbf{r}' \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}
\]

\( d\mathbf{B}(\mathbf{r}) \) is orthogonal to \( d\mathbf{r}' \) and \( (\mathbf{r} - \mathbf{r}') \), and tangential to a circle centred on an axis through \( d\mathbf{r}' \). The field lines of \( d\mathbf{B} \) ‘wrap’ around this axis. This is the right-hand grip rule.

The magnetic field at \( \mathbf{r} \) due to a current loop carrying current \( I \) is

\[
\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \oint_C \frac{I \, d\mathbf{r}' \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}
\]

For a bulk current density \( \mathbf{J} \), with \( d\mathbf{I} = \mathbf{J} \, dV \), we have

\[
\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_V d\mathbf{V}' \frac{\mathbf{J}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}
\]

Similarly for a surface current density

We can use the Bio-Savart law to compute the magnetic field directly for some simple current distributions.
8.3.1 Long straight wire

Consider a long (infinite) straight wire along the z axis. Take the position vector \( r \) of the point \( P \) to be perpendicular to the z axis at the origin. From the diagram, using cylindrical coordinates, we have \( r = \rho \, e_\rho \) and \( r' = z' \, e_z \). Hence

\[
\begin{align*}
    r - r' &= \rho \, e_\rho - z' \, e_z \\
    dr' &= dz' \, e_z
\end{align*}
\]

Therefore

\[
\begin{align*}
    dr' \times (r - r') &= \rho \, dz' \left( e_z \times e_\rho \right) = \rho \, dz' \, e_\phi
\end{align*}
\]

The magnetic field due to the wire is then

\[
\begin{align*}
    B(r) &= \frac{\mu_0}{4\pi} \int_{-\infty}^{\infty} \frac{I \, dz'}{\sqrt{(r - r')^2}} = \frac{\mu_0}{4\pi} I \rho \int_{-\infty}^{\infty} \frac{dz'}{(\rho^2 + z'^2)^{3/2}} \, e_\phi
\end{align*}
\]

To evaluate this integral, we use the substitution \( z' = \rho \tan \theta \), so \( dz' = \rho \sec^2 \theta \, d\theta \) and \( \rho^2 + z'^2 = \rho^2 (1 + \tan^2 \theta) = \rho^2 \sec^2 \theta \), which gives

\[
\begin{align*}
    B(r) &= \frac{\mu_0}{4\pi} \frac{I}{\rho} \int_{-\pi/2}^{\pi/2} \cos \theta \, d\theta \, e_\phi = \frac{\mu_0 I}{2\pi \rho} \, e_\phi
\end{align*}
\]

So \( B \) is inversely proportional to the distance \( \rho \) of the point \( P \) from the wire, and it points in the direction of increasing \( \phi \), i.e. its field lines ‘wrap around’ the wire in a circle. This is the right-hand grip rule again.

8.3.2 Two long parallel wires

Now consider the force on a current element of a second parallel wire (at distance \( d \) from the first) coming from the magnetic field \( B_1(r) \) due to the first wire. Again we choose coordinates so that this current element lies at \( r = \rho \, e_\rho \) in cylindrical coordinates. The force \( dF \) on the current element \( I_2 \, dr_2 \) at \( r_2 \) due to the magnetic field \( B_1 \) produced by wire 1 is:

\[
\begin{align*}
    dF &= I_2 \, dr_2 \times B_1(r_2) = I_2 \, dz \, e_z \times \frac{\mu_0}{2\pi d} I_1 e_\phi = -\frac{\mu_0 I_1 I_2}{2\pi d} \, dz \, e_\phi
\end{align*}
\]

There is an attractive (for \( I_1 I_2 > 0 \)) force per unit length between two parallel infinitely long straight wires:

\[
    f = \frac{\mu_0 I_1 I_2}{2\pi d}
\]

---

\( ^4 \) Until 20 May 2019, this was the basis of the definition of the Ampère, which was defined so that if the wires were 1m apart and the current in each wire was 1A, the force per unit length between the two wires was exactly \( 2 \times 10^{-7} \text{Nm}^{-1} \). Since 1A = 1C s\(^{-1} \), it also defined the Coulomb.
8.3.3 Current loop

Consider a current loop, radius \(a\), carrying current \(I\). Find the magnetic field on the axis through the centre of the loop. From the figure:

\[
r = z \mathbf{e}_z, \quad r' = a \mathbf{e}_\rho, \quad |r - r'| = (a^2 + z^2)^{1/2}
\]

Hence

\[
dr' \times (r - r') = a \mathbf{e}_\phi \times (z \mathbf{e}_z - a \mathbf{e}_\rho) = (az \mathbf{e}_\rho - a^2 (-\mathbf{e}_z)) \, d\phi
\]

So

\[
B(r) = \frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{(az \mathbf{e}_\rho + a^2 \mathbf{e}_z)}{(a^2 + z^2)^{3/2}} \, d\phi = \frac{\mu_0 I}{2} \frac{a^2}{(a^2 + z^2)^{3/2}} \mathbf{e}_z
\]

where we used \(\int_0^{2\pi} \mathbf{e}_\rho \, d\phi = \int_0^{2\pi} (\cos \phi \mathbf{e}_x + \sin \phi \mathbf{e}_y) \, d\phi = 0\). The components of the magnetic field perpendicular to \(\mathbf{e}_z\) cancel due to symmetry as we integrate around the loop.

8.4 Divergence and curl of the magnetic field: Gauss and Ampère laws

The magnetic field produced by a bulk current density \(\mathbf{J}\) is

\[
B(r) = \frac{\mu_0}{4\pi} \int_V dV' \frac{\mathbf{J}(r') \times (r - r')}{|r - r'|^3} = -\frac{\mu_0}{4\pi} \int_V dV' \mathbf{J}(r') \times \nabla \left( \frac{1}{|r - r'|} \right)
\]

where \(\nabla\) acts only on \(r\) (not \(r'\)). Now

\[
\nabla \cdot \left\{ \mathbf{J} \times \nabla \left( \frac{1}{|r - r'|} \right) \right\} = -\mathbf{J} \cdot \left\{ \nabla \times \nabla \left( \frac{1}{|r - r'|} \right) \right\} = 0
\]

because ‘curl grad’ is always zero. Note that \(\mathbf{J} = \mathbf{J}(r')\), so it’s treated as a constant vector when calculating the gradient with respect to \(r\). Therefore

\[
\nabla \cdot B(r) = \frac{\mu_0}{4\pi} \int_V dV' \mathbf{J}(r') \cdot \left\{ \nabla \times \nabla \left( \frac{1}{|r - r'|} \right) \right\} = 0
\]

and we obtain Maxwell’s second equation

\[
\nabla \cdot B = 0
\]
Since the second equation always holds, it tells us that there are no magnetic charges or ‘monopoles’. Compare this with Maxwell’s first equation, $\nabla \cdot E = \rho/\epsilon_0$, which states that electric charge density $\rho$ is the source of electric field [strictly, electric flux – see next semester].

Similarly

$$\nabla \times \left\{ J(r') \times \nabla \left( \frac{1}{|r - r'|} \right) \right\} = \nabla^2 \left( \frac{1}{|r - r'|} \right) J - (J \cdot \nabla) \nabla \left( \frac{1}{|r - r'|} \right)$$

$$= -4\pi \delta (r - r') J - (J \cdot \nabla) \nabla \left( \frac{1}{|r - r'|} \right)$$

so

$$\nabla \times B(r) = -\frac{\mu_0}{4\pi} \int_V dV' \left\{ (J(r') \cdot \nabla) \nabla \left( \frac{1}{|r - r'|} \right) \right\}$$

The integral over the volume in the first term can be performed (trivially) with the delta function, whilst the second term vanishes (see below), and we obtain.

$$\nabla \times B = \mu_0 J$$

This is the differential form of Ampère’s law for steady currents, also known as the static form of Maxwell’s fourth equation.\(^5\)

To show that the second term vanishes, we take the second $\nabla$ out of the integral to give (up to constants)

$$\nabla \int_V dV' \frac{J(r') \cdot \nabla}{|r - r'|}$$

The integral may be written in terms of the gradient $\nabla'$, wrt $r'$

$$\int_V dV' \frac{J(r') \cdot \nabla}{|r - r'|} = -\int_V dV' \frac{J(r') \cdot \nabla'}{|r - r'|}$$

$$= \int_V dV' \left\{ \nabla' \cdot \left( \frac{J(r')}{|r - r'|} \right) - (\nabla' \cdot J(r')) \frac{1}{|r - r'|} \right\}$$

$$= -\int_S \frac{J(r')}{|r - r'|} \cdot dS' + 0$$

where we used the divergence theorem on the first term, which is zero provided that the current density $J(r') \to 0$ at infinity, and the second term is zero because $\nabla \cdot J = 0$ in magnetostatics, where we have steady currents.

In fact, $\nabla \times B = \mu_0 J$ implies $\nabla \cdot J = 0$, since $\nabla \cdot (\nabla \times B) \equiv 0$.

\(^5\)In electrodynamics, there is an additional term on the RHS, due to time-dependent charge distributions and therefore non-steady currents. See next semester.
8.4.1 Ampère’s law (1826)

The fundamental laws of magnetostatics are
\[ \nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \]

Applying the divergence theorem to a closed surface \( S \), which bounds the volume \( V \), in the first equation gives
\[ \int_S \mathbf{B} \cdot dS = \int_V \nabla \cdot \mathbf{B} \, dV = 0 \]
i.e. the magnetic flux (flux of \( \mathbf{B} \)) through any closed surface is zero – Gauss’ law of magnetostatics.

Applying Stokes’ theorem for a closed curve \( C \) bounding an open surface \( S \) to the second equation gives
\[ \oint_C \mathbf{B} \cdot d\mathbf{r} = \int_S \nabla \times \mathbf{B} \cdot dS = \mu_0 \int_S \mathbf{J} \cdot dS = \mu_0 I \]

In words: the circulation of the magnetic field \( \mathbf{B} \) around a closed loop \( C \) is equal to \( \mu_0 \) times the total current \( I \) flowing through the loop.

This is the integral form of Ampère’s law, which is extremely useful in finding \( \mathbf{B} \) in symmetric situations, just like Gauss’ law in electrostatics.

Example: long straight wire Consider a long straight wire lying along the \( z \) axis and use cylindrical coordinates. By symmetry, \( \mathbf{B}(\mathbf{r}) \) will be independent of \( z \), its magnitude will be independent of \( \phi \), and it must point in the \( \phi \) direction. Hence
\[ \mathbf{B}(\mathbf{r}) = B_\phi(\rho) \hat{\phi} \]

Consider a circle of radius \( \rho \) with its centre on the \( z \) axis, and apply Ampère’s law
\[ \oint_C \mathbf{B} \cdot d\mathbf{r} = \mu_0 I \quad \Rightarrow \quad B_\phi(\rho) 2\pi \rho = \mu_0 I \]

which reproduces our previous result with very little effort,
\[ \mathbf{B}(\mathbf{r}) = \frac{\mu_0 I}{2\pi \rho} \hat{\phi} \]
**Example: coil/solenoid**  Consider a long coil\(^6\) of radius \(a\), with \(N\) tightly-wound turns/unit length, centred on the \(z\) axis. Again, we use cylindrical coordinates.

By symmetry \(B(r)\) must be in the \(z\) direction, and independent of \(\phi\) and \(z\), so

\[
\overrightarrow{B(r)} = B_z(\rho) \, \mathbf{e}_z
\]

Away from the wire, we have \(\nabla \times \overrightarrow{B} = 0\), so

\[
\frac{\partial B_z}{\partial \rho} = 0 \quad \Rightarrow \quad B_z = \text{constant}
\]

Since the constant is zero as \(\rho \rightarrow \infty\), then \(B\) must be zero everywhere outside the coil, which is a remarkable result!

So we have

\[
B_z = \begin{cases} 
B : & \rho < a \\
0 : & \rho > a
\end{cases}
\]

Now apply Ampère’s law to the rectangular ‘loop’ of length \(L\) shown in the figure, which has one long side inside the coil, and one outside. Therefore

\[
\overrightarrow{B} \cdot L - 0 \cdot L = \mu_0 N I
\]

The magnetic field is **constant** inside the coil, with

\[
\overrightarrow{B} = \mu_0 N I \mathbf{e}_z
\]

### 8.4.2 Conducting surfaces

Consider a conducting surface carrying current density \(\overrightarrow{K}\). What are the boundary conditions on the magnetic field at the surface?

**Normal component:** Use a small Gaussian pill-box of negligible height with a circular surface of area \(A\), which straddles the surface as shown.

\[
0 = \int_V \nabla \cdot \overrightarrow{B} \, dV = \int_S \overrightarrow{B} \cdot d\overrightarrow{S} = (B_2 - B_1) \cdot n \, A
\]

So

\[
B_2 \cdot n = B_1 \cdot n
\]

The normal component of the magnetic field is **continuous** across the boundary.

---

\(^6\)A ‘long’ coil means its length is much greater than its radius, so we can treat it as being effectively infinite.
**Tangential component:** Use a small rectangular Ampère loop of length $L$ and negligible width straddling the surface, which has normal $n$, as shown in the figure.

Let the surface spanning the loop have normal $n'$, so the long side of the loop has direction $n' \times n$.

\[
\oint_C B \cdot dr = (B_2 - B_1) \cdot (n' \times n) L = \int_S \mu_0 J \cdot dS = \mu_0 \int_C K \cdot n' \cdot dr = \mu_0 K \cdot n' L
\]

where the line $C$ is along that part of the surface contained within the loop.

So

\[
\left[ n \times (B_2 - B_1) \right] \cdot n' = \mu_0 K \cdot n'
\]

This holds for all $n'$ tangential to the surface, therefore

\[
- (B_2 - B_1) = \mu_0 K
\]

There is a *discontinuity* in the tangential components of the magnetic field due to the surface current density.

### 8.5 The vector potential

A vector field $\mathbf{B}(r)$ that satisfies $\nabla \cdot \mathbf{B} = 0$ can be written as the curl of a *vector potential*\(^7\)

\[
\nabla \cdot \mathbf{B} = 0 \iff \exists \mathbf{A} \text{ such that } \mathbf{B} = \nabla \times \mathbf{A}
\]

The vector potential is not unique, it’s defined up to the gradient of an arbitrary scalar field $\chi(r)$. If we make the *gauge transformation* $\mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla \chi$

then

\[
\nabla \times \mathbf{A} \rightarrow \nabla \times \mathbf{A}' = \nabla \times \mathbf{A} + \nabla \times \nabla \chi = \nabla \times \mathbf{A} + 0
\]

The magnetic field $\mathbf{B}$ (and hence the physics) is unchanged by the transformation.

To *fix the gauge* uniquely we may choose to add an additional constraint on $\mathbf{A}$, so that

\[
\nabla \cdot \mathbf{A} = 0
\]

which is called *Coulomb gauge*. If $\nabla \cdot \mathbf{A} = \psi \neq 0$, we can always find a gauge transformation $\chi$ such that $\nabla \cdot \mathbf{A}' = 0$, *i.e.*

\[
\nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} + \nabla^2 \chi = 0 \Rightarrow \nabla^2 \chi = -\psi
\]

This is just Poisson’s equation, which can be solved for $\chi$ in the standard way.

---

\(^7\)Compare this with the case of electrostatics, where $\nabla \times \mathbf{E} = 0 \Rightarrow \text{the electric field can be expressed as the gradient of a scalar potential } \phi$, and *vice versa*, *i.e.* $\nabla \times \mathbf{E} = 0 \iff \exists \phi$ such that $\mathbf{E} = -\nabla \phi$
Now if \( B = \nabla \times A \), with \( \nabla \cdot A = 0 \), then from the differential form of Ampère’s law,
\[
\mu_0 J = \nabla \times B = \nabla \times (\nabla \times A) = \nabla (\nabla \cdot A) - \nabla^2 A = -\nabla^2 A
\]
So
\[
\nabla^2 A = -\mu_0 J
\]
This has the form of Poisson’s equation for \( A \) (rather 3 equations for \( A_i \)), with solution
\[
A(r) = \frac{\mu_0}{4\pi} \int_V \frac{J(r')}{|r - r'|} \, dV' = \Phi
\]
which satisfies the boundary condition \( A(r) \to 0 \) as \( r \to \infty \).
If we take the curl of this equation, we recover the Biot-Savart law [exercise].
We can write similar expressions for line and surface currents [exercise].

**Example:** For a magnetic field \( B \), which is constant everywhere,
\[
A = \frac{1}{2} (B \times r)
\]
because
\[
\nabla \times \frac{1}{2} (B \times r) = \frac{1}{2} B (\nabla \cdot r) - \frac{1}{2} (B \cdot \nabla) r = \frac{3}{2} B - \frac{1}{2} B = B
\]
If we know \( B \), we can find \( A \) using Stokes’ theorem in problems with a lot of symmetry.
\[
\oint_C A \cdot dr = \int_S (\nabla \times A) \cdot dS = \int_S B \cdot dS = \Phi
\]
\( \Phi \) is the magnetic flux crossing the surface \( S \) bounded by the closed curve \( C \).

**Example:** For a solenoid centred on the \( z \) axis, we showed previously that
\[
B = \begin{cases} 
B \varepsilon_z & : \quad \rho < a \\
0 & : \quad \rho > a 
\end{cases}
\]
By symmetry, \( A(r) = A_\phi(\rho) \varepsilon_\phi \).
Choosing \( S \) to be a horizontal disc of radius \( \rho \), as shown
\[
\rho > a \quad : \quad 2\pi \rho A_\phi = B \pi a^2 \quad \Rightarrow \quad A_\phi(\rho) = \frac{1}{2} B \frac{a^2}{\rho}
\]
\[
\rho < a \quad : \quad 2\pi \rho A_\phi = B \pi \rho^2 \quad \Rightarrow \quad A_\phi(\rho) = \frac{1}{2} B \rho
\]
[The case \( \rho < a \) is as in the first example above.]
Note that \( A \neq 0 \) everywhere outside the solenoid, even though \( B = 0 \) everywhere outside it! This is important in quantum mechanics – see the Quantum Theory course.

---

\(^8\)Compare this with \( \nabla^2 \phi = -\rho/\varepsilon_0 \) in electrostatics, with solution \( \phi(r) = \frac{1}{4\pi\varepsilon_0} \int_V \frac{\rho(r')}{|r - r'|} \, dV' \).
8.6 Magnetic dipoles

Consider a circular wire, radius $a$, in the $x - y$ plane, carrying current $I$. What is the form of the vector potential at points with $r \gg a$?

$$A(r) = \frac{\mu_0}{4\pi} \int_{C} \frac{I \, dr'}{|r - r'|}$$

Wlog take $r$ in the $x - z$ plane, so $r = x e_x + z e_z$ and

$$r' = a (\cos \phi e_x + \sin \phi e_y)$$
$$dr' = a (-\sin \phi e_x + \cos \phi e_y) \, d\phi$$

For $r \gg a$

$$\frac{1}{|r - r'|} = \frac{1}{r} \left( 1 + \frac{r \cdot r'}{r^2} + \ldots \right) \quad \text{with} \quad r - r' = ax \cos \phi$$

Then

$$A(r) = \frac{\mu_0 I}{4\pi} \int_{0}^{2\pi} d\phi \, a \left( -\sin \phi e_x + \cos \phi e_y \right) \frac{1}{r} \left( 1 + \frac{ax \cos \phi}{r^2} + \ldots \right)$$

$$\approx \frac{\mu_0 I}{4\pi} \frac{\pi a^2 x}{r^3} e_y = \frac{\mu_0}{4\pi} \frac{m \times r}{r^3}$$

where we used $\int_{0}^{2\pi} \cos^2 \theta = \pi$, with all the other integrals being zero, and defined

$$m = \pi a^2 I e_z = I S$$

where $S$ is the vector area of the loop. Now

$$\nabla \times \left( \frac{m \times r}{r^3} \right) = \frac{1}{r^3} \nabla \times (m \times r) + \nabla \left( \frac{1}{r^3} \right) \times (m \times r) = \frac{3 (m \cdot r) r - r^2 m}{r^5}$$

So that

$$B(r) = \nabla \times A(r) = \frac{\mu_0}{4\pi} \frac{3 (m \cdot r) r - r^2 m}{r^5}$$

which is the field of a magnetic dipole with dipole moment $m$.

Compare this with the expression for the electric field for an electric dipole of moment $p$ at the origin:

$$E(r) = \frac{1}{4\pi \varepsilon_0} \left[ \frac{3 p \cdot r}{r^5} r - \frac{p}{r^3} \right]$$
As we shall now show, we get the same result far away from any current loop, and the field lines for electric dipoles, magnetic dipoles, and bar magnets (see next chapter) are therefore the same in the far zone (far-field limit).

Multipole expansion  For any current loop near the origin, we expand the magnetic vector potential $A$ as above,

$$A(r) = \frac{\mu_0 I}{4\pi} \oint_C \frac{1}{r} \left( 1 + \frac{r \cdot r'}{r^2} + \ldots \right)$$

Applying Stokes’ theorem to an arbitrary constant vector $c$ gives

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

Since this holds for all vectors $c$, then $\oint_C \mathbf{d}r' = 0$ for any closed curve $C$, so the first term in the expansion of $A$ always vanishes. Hence there are no magnetic monopoles, as expected.

To evaluate the second term, first apply Stokes’ theorem to $c(r \cdot r')$.

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

$$= \int_S (r \cdot c) \cdot \mathbf{d}S' = c \cdot \int_S \mathbf{d}S' \times r = c \cdot (S \times r)$$

where $S = \int_S \mathbf{d}S'$ is the vector area of the current loop. Since this holds for all constant vectors $c$, we find

$$\oint_C \mathbf{d}r' \cdot (r \cdot r') = S \times r$$

Therefore the first non-zero term in the multipole expansion of $A(r)$ is

$$A(r) = \frac{\mu_0 I}{4\pi} \frac{S \times r}{r^3} = \frac{\mu_0}{4\pi} \frac{m \times r}{r^3}$$

as in our explicit example of the circular wire above. This holds for any current loop whose size is much less than $r$.

If we go further in the expansion, we get a magnetic quadrupole term, etc.
8.6.1 Magnetic moment and angular momentum

As system of charged particles flowing in a closed current loop has angular momentum.

For a constant vector \( \mathbf{c} \), consider

\[
\mathbf{c} \cdot \oint_{C} \mathbf{r} \times d\mathbf{r} = \oint_{C} (\mathbf{c} \times \mathbf{r}) \cdot d\mathbf{r} = \int_{S} \nabla \times (\mathbf{c} \times \mathbf{r}) \cdot d\mathbf{S} \\
= \int_{S} \{ \mathbf{c} (\nabla \cdot \mathbf{r}) - (\mathbf{c} \cdot \nabla) \mathbf{r} \} \cdot d\mathbf{S} = 2 \mathbf{c} \cdot \int_{S} d\mathbf{S}
\]

Since this holds for all constant vectors \( \mathbf{c} \), we have

\[
\int_{S} d\mathbf{S} = S = \frac{1}{2} \oint_{C} \mathbf{r} \times d\mathbf{r}
\]

We can then rewrite the dipole moment as

\[
m = I \int_{S} d\mathbf{S} = \frac{1}{2} I \oint_{C} \mathbf{r} \times d\mathbf{r} = \frac{1}{2} \oint_{C} \mathbf{r} \times d\mathbf{I}
\]

where \( d\mathbf{I} \) is the current element.

For a bulk current \( \mathbf{dI} = \mathbf{J} dV \), this becomes

\[
m = \frac{1}{2} \int_{V} (\mathbf{r} \times \mathbf{J}) dV,
\]

and since \( \mathbf{J} = \rho \mathbf{v} \), we have

\[
m = \frac{1}{2} \int_{V} (\mathbf{r} \times \mathbf{v}) \rho dV
\]

Now suppose the charge density \( \rho \) is made up of particles each with charge \( q \) and mass \( m \), then

\[
\rho = (q/m) \rho_{m}
\]

where \( \rho \) is the charge density (charge/unit volume), and \( \rho_{m} \) is the mass density (mass/unit volume). Then

\[
m = \frac{q}{2m} \int_{V} \mathbf{r} \times \mathbf{v} \rho_{m} dV = \frac{q}{2m} \int_{V} \mathbf{r} \times p dV = \frac{q}{2m} L = g L
\]

where \( p \) is the momentum density (momentum/unit volume), so \( L \) is the angular momentum, and \( g \) is called the gyromagnetic ratio of the particles.

It turns out that this derivation is wrong by a factor of two for an isolated spin-\( \frac{1}{2} \) particle such as an electron (with charge \(-e\)), for which \( g = -e/m \). The extra factor of two requires relativity.

8.6.2 Force and torque on a dipole in an external field

Force on the dipole

Consider a current loop in an external magnetic field \( \mathbf{B}(\mathbf{r}) \). The Lorentz force on the loop is

\[
\mathbf{F} = I \oint_{C} d\mathbf{r}' \times \mathbf{B}(\mathbf{r}')
\]

\[
= I \oint_{C} d\mathbf{r}' \times \{ \mathbf{B}(\mathbf{r}) + ((\mathbf{r}' - \mathbf{r}) \cdot \nabla) \mathbf{B}(\mathbf{r}) + \ldots \}
\]
where \( r' \) is a point on the loop, and we Taylor-expanded \( B(r') \) about a point \( r \) inside (or close to) the loop. We’ve already shown that \( \oint_C \overrightarrow{dr'} = 0 \), so the terms in the integrand that don’t depend on \( r' \) integrate to zero, and hence

\[
\overrightarrow{F} = I \oint_C \overrightarrow{dr'} \times (\overrightarrow{r'} \cdot \nabla) B(r) + \ldots
\]

We showed above that

\[
\oint_C \overrightarrow{dr'} (\overrightarrow{r'} \cdot \overrightarrow{r}) = \overrightarrow{S} \times \overrightarrow{r}
\]

The derivation above works for any vector field \( \overrightarrow{a}(r) \) [check it!], so

\[
\oint_C \overrightarrow{dr'} (\overrightarrow{r'} \cdot \overrightarrow{a}(r)) = \overrightarrow{S} \times \overrightarrow{a(r)} \quad (8.1)
\]

Then

\[
F_i = \sum_{ij} \epsilon_{ijk} \oint_C \overrightarrow{dx_j} (\overrightarrow{r'} \cdot \nabla_j) B_k = \sum_{ij} \epsilon_{ijk} (\overrightarrow{S} \times \nabla_j) B_k = I ((\overrightarrow{S} \times \nabla) \times \overrightarrow{B}),
\]

so that

\[
\overrightarrow{F} = (\overrightarrow{m} \times \nabla) \times \overrightarrow{B} = -m (\nabla \cdot \overrightarrow{B}) + \nabla (m \cdot \overrightarrow{B})
\]

But \( \nabla \cdot \overrightarrow{B} = 0 \), therefore

\[
\overrightarrow{F} = \nabla (m \cdot \overrightarrow{B}) \equiv -\nabla W_{\text{dip}}
\]

So the energy of the dipole in the external field is

\[
W_{\text{dip}} = -(m \cdot \overrightarrow{B})
\]

Note: this is not the total energy of the dipole. It also takes energy to make the current flow in the loop. [This is covered next semester.]

Note that these results mirror those for the electric dipole where \( \overrightarrow{F} = \nabla (p \cdot \overrightarrow{E}) \) and \( W = -p \cdot \overrightarrow{E} \).

Torque on the dipole

\[
G = \oint_C \overrightarrow{r'} \times \left[ I \overrightarrow{dr'} \times \overrightarrow{B(r')} \right] = I \oint_C \left[ \overrightarrow{dr'} (\overrightarrow{r'} \cdot \overrightarrow{B(r)}) - \overrightarrow{B(r)} (\overrightarrow{r'} \cdot \overrightarrow{dr'}) \right]
\]

where we approximated \( \overrightarrow{B(r')} \approx \overrightarrow{B(r)} \) on the RHS. The second term vanishes on applying Stokes’ theorem (because \( \overrightarrow{S}' \times \overrightarrow{r'} = 0 \)), and we again apply the result of equation (8.1) to the first term, which gives

\[
G = I \oint_C \overrightarrow{dr'} (\overrightarrow{r'} \cdot \overrightarrow{B}) = IS \times \overrightarrow{B} = m \times \overrightarrow{B}
\]

which again mirrors the result \( G = p \times \overrightarrow{E} \) for an electric dipole.

The results are the same as for the electric dipole because one can think of a magnetic dipole as made up of two magnetic monopoles.
This is a useful way to remember the formulae, but magnetic monopoles (probably!) do not exist.

**Example: compass needle**  We can measure the magnetic field $B$ via oscillations of a compass needle (which can be regarded as a magnetic dipole).

Since torque is the rate of change of angular momentum, $G = \frac{dL}{dt}$, we have

$$\frac{dL}{dt} = m \times B$$

For a needle with moment of inertia $I$, this becomes

$$\frac{d}{dt} \left( I \dot{\phi} \right) = -mB \sin \phi$$

where $\phi$ is the angle between $m$ and $B$ (see diagram).

For small oscillations, where $\phi \ll 1$, we have

$$\ddot{\phi} + \frac{mB}{I} \phi = 0$$

which gives an oscillation frequency of

$$\omega = \sqrt{\frac{mB}{I}}$$
8.7 Summary: electrostatics and magnetostatics

**Electrostatics:** Stationary charges \( \frac{\partial \rho}{\partial t} = 0 \) are the source of electric fields

\[
\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \quad \text{[ME1]}
\]

Coulomb’s law (field due to point charge) leads to

\[
\nabla \times \mathbf{E} = 0 \quad \text{[ME3 (static)]} \quad \Rightarrow \quad \mathbf{E} = -\nabla \phi
\]

In turn, the above lead to Poisson’s equation for the scalar potential \( \phi \)

\[
\nabla^2 \phi = -\frac{\rho}{\varepsilon_0}
\]

**Magnetostatics:** Steady current loops \( \nabla \cdot \mathbf{J} = 0 \) produce magnetic fields [no magnetic monopoles].

\[
\nabla \cdot \mathbf{B} = 0 \quad \text{[ME2]} \quad \Rightarrow \quad \mathbf{B} = \nabla \times \mathbf{A}
\]

The Biot-Savart law (field due to current element) leads to

\[
\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad \text{[ME4 (static)]}
\]

In turn, in Coulomb gauge, \( \nabla \cdot \mathbf{A} = 0 \), the above lead to a vector Poisson equation for \( \mathbf{A} \)

\[
\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}
\]

**Lorentz force:** On a point particle

\[
\mathbf{F} = q \left( \mathbf{E} + \mathbf{v} \times \mathbf{B} \right)
\]

or the Lorentz force density

\[
\mathbf{f} = \rho \mathbf{E} + \mathbf{J} \times \mathbf{B}
\]

Next semester, you’ll see how the Maxwell equations involving curls, (ME3 and ME4), need to be modified when *time-varying currents and fields* are present.

In the final (short!) chapter, we will study magnetostatics in dielectric and magnetic materials.
Chapter 9

Electrostatics and magnetostatics in materials

Our goal in this chapter is to understand a little more about the behaviour of electric and magnetic fields in materials.

Maxwell’s equations (and the Lorentz force) are fundamental microscopic equations that correctly describe the electric and magnetic fields in materials. But materials are made up of atoms and/or molecules, which are in turn made up of electrons and nuclei. The fields therefore vary very rapidly at short distances (for example, near an atomic nucleus.) To understand the large scale properties of materials, we will average over the microscopic fields, and deduce a set of equations which describe electrostatics and magnetostatics at the macroscopic level.

9.1 Dielectrics

So far, we’ve considered electric and magnetic fields in a vacuum, and we have touched on (Ohmic) conductors where \( J = \sigma E \). Now we’ll consider a third class of materials: insulators.

In an insulator, the conductivity, \( \sigma = 0 \), so the current density \( J = 0 \) and \( \frac{\partial \rho}{\partial t} = 0 \), i.e. charges are fixed. Consequently, insulators can carry static charge density \( \rho(\mathbf{r}) \).

The most important effect of the microscopic charges at large scales is that they can be polarised by an applied external electric field—that is, the applied field will induce a molecular electric dipole moment \( \mathbf{p} \).

Suppose there are \( N \) molecules per unit volume; then an applied electric field will lead to an average electric polarisation density \( \mathbf{P}(\mathbf{r}) = N \langle \mathbf{p} \rangle \) where \( \langle \mathbf{p} \rangle \) is the spatial average of the molecules’ electric dipole moments.

We know that \( \nabla \times \mathbf{E} = 0 \) for static electric fields, so \( \mathbf{E} = -\nabla \phi \). In a polarised

\[ \phi(\mathbf{r}) = \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi \varepsilon_0 r^3} = -\mathbf{p} \cdot \nabla \left( \frac{1}{4\pi \varepsilon_0 \mathbf{r}} \right). \]

\[ ^1 \text{It’s helpful to recall from electrostatics that the potential of an elementary dipole with moment } \mathbf{p} \text{ is } \phi(\mathbf{r}) = \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi \varepsilon_0 r^3} = -\mathbf{p} \cdot \nabla \left( \frac{1}{4\pi \varepsilon_0 \mathbf{r}} \right). \]
material, the (averaged) potential has two contributions

\[ \phi(r) = \frac{1}{4\pi\epsilon_0} \int_V dV' \left[ \frac{\rho(r')}{|r - r'|} + \frac{P(r') \cdot (r - r')}{|r - r'|^3} \right]. \]

But

\[ \frac{P(r') \cdot (r - r')}{|r - r'|^3} = P(r') \cdot \nabla' \left( \frac{1}{|r - r'|} \right) = \nabla' \cdot \left( \frac{P(r')}{|r - r'|} \right) - \frac{\nabla' \cdot P(r')}{|r - r'|}, \]

so (using the divergence theorem)

\[ \phi(r) = \frac{1}{4\pi\epsilon_0} \int_V dV' \left[ \frac{\rho(r') - \nabla' \cdot P(r')}{|r - r'|} \right] + \frac{1}{4\pi\epsilon_0} \int_S dS' \cdot \frac{P(r')}{|r - r'|}. \quad (9.1) \]

We can interpret these two terms as follows: \( [\rho(r') - \nabla' \cdot P(r')] \) is an effective (volume) charge density, while, in the second term on the right hand side, \( P(r') \) plays the role of an effective surface charge density (more on this later). So, inside this material, known as a dielectric,

\[ \rho \rightarrow \rho - \nabla \cdot P \]

i.e., \( \nabla \cdot E = \frac{\rho - \nabla \cdot P}{\epsilon_0} \)

We define the electric displacement \( D(r) \) by

\[ D = \epsilon_0 E + P \]

\[ \therefore \nabla \cdot D = \rho. \]

This equation replaces Maxwell’s first (divergence) equation, and hence the Gauss law, inside dielectrics.

### 9.1.1 Boundary conditions between two insulators

We still have Maxwell’s third (curl) equation

\[ \nabla \times E = 0. \]

As before, we consider a narrow rectangular loop of very small area, and use

\[ \oint E \cdot dr = 0 \]
If we shrink the loop so that it just crosses the boundary between the dielectrics, we find that the tangential component of the $E$ field is continuous, precisely as we showed in electrostatics previously (see Section 7.9.2 on page 71).

To find out what happens to the normal component of the electric field, we repeat our previous derivation using Gauss’s law, but with $D$ rather than $E$, which gives (see Section 7.9.1 on page 70)

$$ \mathbf{n} \cdot (D_2 - D_1) = \sigma $$

where $\sigma$ is the surface charge density. Thus, if there is no surface charge, the normal component of $D$ is continuous.

Example: Polarised insulator in vacuum

![Diagram of polarised insulator in vacuum]

We take medium 1 to be the insulator, while medium 2 is vacuum. Therefore

$$ D_1 = \epsilon_0 E_1 + P_1, \quad D_2 = \epsilon_0 E_2 $$

Continuity of the normal component of $D$ tells us that

$$ \mathbf{n} \cdot (\epsilon_0 E_2 - \epsilon_0 E_1 - P_1) = 0 $$

$$ \Rightarrow \quad \mathbf{n} \cdot (E_2 - E_1) = \frac{n \cdot P_1}{\epsilon_0}. $$

Thus, we have an effective surface bound charge density $\sigma_b = n \cdot P_1$ due to the polarisation of the insulator, as we might have anticipated from the surface integral term in equation (9.1).

### 9.2 Linear Dielectrics

For many (isotropic, homogeneous) materials, we can assume (for weak external fields)

$$ P = \epsilon_0 \chi_e E, $$

that is, we can assume that the electric polarisation is proportional to the applied field. The coefficient of proportionality is chosen so that $\chi_e$ is a dimensionless number, called the electric susceptibility of the medium.

Then,

$$ D = \epsilon_0 E + P = \epsilon_0 (1 + \chi_e) E \equiv \epsilon E, $$

where $\epsilon = \epsilon_0 (1 + \chi_e)$ is the permittivity of the medium. The quantity $K = (1 + \chi_e) = \epsilon/\epsilon_e$ is called the dielectric constant of the medium.
Typical electric susceptibilities are:

Vacuum \( \chi_e = 0 \)
Gases \( \chi_e \sim 10^{-3} - 10^{-4} \)
Solids/liquids \( \chi_e \sim 1 - 100 \)

So it's easy to write equations in a linear, homogeneous dielectric: simply replace \( \epsilon_0 \) with \( \epsilon \)!

For example, the energy density (energy/volume) \( w(r) \) for an electrostatic field becomes

\[
\frac{1}{2} \epsilon_0 E^2 \rightarrow \frac{1}{2} \epsilon E^2 = \frac{1}{2} E \cdot D.
\]

**Example: Capacitors**

It’s useful to separate the plates of a capacitor with a dielectric material.

\[
\begin{array}{c|c|c}
\epsilon_0 & E = 0 & +Q \\
\hline \epsilon & D & a \\
\hline \epsilon_0 & E = 0 & -Q
\end{array}
\]

The field between the plates is obtained from our previous result by simply replacing \( \epsilon_0 \) by \( \epsilon \)

\[
D = \epsilon E = -\sigma e_z
\]

and therefore the potential difference is

\[
V = \frac{\sigma a}{\epsilon}
\]

\[
\therefore C = \frac{Q}{V} = \frac{A\epsilon}{a} > \frac{A\epsilon_0}{a}
\]

So the capacitance increases.

**9.3 Magnetisation**

Just as molecules in materials can pick up induced electric dipole moments \( p \), they can also gain induced magnetic dipole moments \( m \). Again, we average over many molecules: there can be a net magnetic moment density \( \vec{M} = N\langle \vec{m} \rangle \), which is usually called the magnetisation.

Since \( \nabla \cdot B = 0 \), we have \( B = \nabla \times A \). The vector potential has two contributions

\[
A(r) = \frac{\mu_0}{4\pi} \int dV' \left[ \frac{J(r')}{|r-r'|} + \frac{M(r') \times (r - r')}{|r - r'|^3} \right].
\]
Working in basically the same way as for dielectrics, we write
\[
\frac{M(r') \times (r-r')}{|r-r'|^3} = M(r') \times \nabla' \frac{1}{|r-r'|}
\]
\[
= -\nabla' \times \left( \frac{M(r')}{|r-r'|} \right) + \frac{\nabla' \times M(r')}{|r-r'|},
\]
so (using the divergence theorem – exercise)
\[
A(r) = \mu_0 \frac{1}{4\pi} \int_V \frac{\left[ J(r') + \nabla' \times M(r') \right]}{|r-r'|} - \mu_0 \frac{1}{4\pi} \int_S \frac{dS' \times M(r')}{|r-r'|}.
\]
We can interpret these two terms as follows: \(J(r') + \nabla' \times M(r')\) is an effective current density, while, in the second term on the right hand side, \(M(r')\) plays the role of an effective surface current density. So, inside this material \(J \rightarrow J + \nabla \times M\) and Maxwell’s fourth equation (the differential form of Ampère’s law) becomes
\[
\nabla \times B = \mu_0 J + \mu_0 \nabla \times M.
\]
We define
\[
H \equiv \frac{1}{\mu_0} B - M
\]
so that
\[
\nabla \times H = J.
\]
This equation replaces Ampère’s law in magnetic materials.

Some people call \(H\) the “magnetic field” or the “magnetic intensity”, and \(B\) the “magnetic induction”. So there’s potential for confusion! Fortunately, everyone agrees on the symbols. For us, \(E\) and \(B\) are the fundamental, microscopic fields while \(D\) and \(H\) are phenomenological, macroscopic fields, which are obtained by averaging the microscopic fields \(E\) and \(B\).

### 9.3.1 Boundary conditions between two magnetic materials

The calculation of the boundary conditions satisfied by \(B\) and \(H\) is similar to the calculation we did for \(E\) and \(D\) in dielectrics.

Since \(\nabla \cdot B = 0\), it follows (by considering a Gaussian surface as we did previously) that
\[
B_1 \cdot n_1 = B_2 \cdot n_2.
\]
Finally, the boundary condition for the tangential components of \(H\) follow from Ampère’s law \(\nabla \times H = J\) or \(\oint_C H \cdot d\mathbf{r} = I\). Consider a small Ampérian loop across the interface between media 1 and 2. As the loop area shrinks to zero, we learn that the tangential components of \(H\) can be discontinuous only in the presence of a surface current density \(K\):
\[
\mathbf{n} \times (H_2 - H_1) = K.
\]

104
Exercise: fill in the details by comparing with the case of conducting surfaces in Section 8.4.2.

Notice that each Maxwell equation leads to a boundary condition; equations which are of the form $\nabla \cdot \text{(something)}$ provide information on normal components, while the two equations containing $\nabla \times \text{(something)}$ tell us about the tangential components.

### 9.4 Diamagnets, Paramagnets and Ferromagnets

Many magnetic materials are isotropic and homogeneous, and have the property that for weak external fields the magnetisation is linear in $H$:

$$M = \chi_m H,$$

where the constant $\chi_m$ is called the magnetic susceptibility. Hence,

$$B = \mu_0 (H + M) = \mu_0 (1 + \chi_m) H \equiv \mu H,$$

where $\mu = \mu_0 (1 + \chi_m)$ is the permeability. Magnetic susceptibilities are typically small $\sim 10^{-4} - 10^{-5}$. However, unlike electric susceptibility, it is common to encounter materials which have magnetic susceptibilities of either sign. Materials with $\chi_m > 0$ (that is, $\mu > \mu_0$) are called paramagnets while materials with $\chi_m < 0 \Rightarrow \mu < \mu_0$ are called diamagnets.

Some materials, such as iron, can be magnetised even if there is no external field: these are called ferromagnets. They can have $\mu \sim 10^4$ so there are very large effects! Iron inside a coil greatly increases the magnetic field.

Just like dielectrics, it’s easy to write the Maxwell equations inside a linear magnetic material: simply replace $\mu_0 \rightarrow \mu$. For example, you will see next semester that the magnetic energy density is

$$w_m = \frac{1}{2\mu_0} B^2 \mapsto \frac{1}{2\mu} B^2 = \frac{1}{2} B \cdot H,$$

and the Poynting vector is

$$S = \frac{1}{\mu_0} E \times B \mapsto \frac{1}{\mu} E \times B = E \times H.$$
9.5 Summary: the static Maxwell equations for (linear) media

We have seen that the static Maxwell’s equations for dielectric and diamagnetic or paramagnetic materials are\(^2\)

\[
\begin{align*}
\nabla \cdot D &= \rho \\
\nabla \cdot B &= 0 \\
\nabla \times E &= 0 \\
\nabla \times H &= J,
\end{align*}
\]

To solve these equations, we also need relations between \( D, H \) and \( E, B \). These are generally known as constitutive relations; in this course, we will restrict our attention to linear, homogeneous, isotropic, static, nondispersive media for which we can take

\[ D = \epsilon E, \quad B = \mu H. \]

To specify the physics completely we need the boundary conditions between media. We’ve already worked these out; they can be written as

\[
\begin{align*}
\mathbf{n} \times (E_2 - E_1) &= 0 \\
\mathbf{n} \cdot (D_2 - D_1) &= \sigma \\
\mathbf{n} \cdot (B_2 - B_1) &= 0 \\
\mathbf{n} \times (H_2 - H_1) &= K.
\end{align*}
\]

[End of Semester I]

\(^2\)These are the static forms of the “equations in the floor” on Level 3 of JCMB. In Semester 2, you will see that in electrodynamics the curl equations acquire time-derivatives of the \( B \) and \( D \) fields on their right-hand sides.