Mathematical Methods in Physics I

Based on lectures by Dr. Ritam Mallick
Notes taken by Rwik Dutta

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.[1]

Contents

[1] This is how Dexter Chua describes his lecture notes from Cambridge. I could not have described mine in any better way.

1 Vector Analysis

Let $\{\hat{\mathbf{e}}_\mathbf{i}\}$ be an orthonormal basis, i.e., $\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}_\mathbf{j} = \delta_{ij}$(where δ_{ij} represents the Kroneker delta function), of our vector space. For any vector $\mathbf{A}$ we have

$$\mathbf{A} = \underbrace{A_i \hat{\mathbf{e}}_\mathbf{i}}_{\text{Einstein summation}} \tag{1.0.1}$$

where A_i belong to the scalar field over which the vector space is defined, like $\mathbb{R}$ or $\mathbb{C}$.

1.1 Products

Definition 1 (Scalar product)**.**

$$\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}_\mathbf{j} = \delta_{ij} \tag{1.1.1}$$

The result is a scalar, hence, the name. Using this and the fact that scalar product is distributive over addition we arrive at

$$\mathbf{A} \cdot \mathbf{B} = A_i B_i \tag{1.1.2}$$

Definition 2 (Vector Product)**.**

$$(\hat{\mathbf{e}}_\mathbf{i} \times \hat{\mathbf{e}}_\mathbf{j})_k = \varepsilon_{ijk} \tag{1.1.3}$$

The result is a vector, hence, the name. Using this and the fact that vector product is distributive over addition we arrive at

$$C_i = (\mathbf{A} \times \mathbf{B})_i = \varepsilon_{ijk} A_j B_k \tag{1.1.4}$$

Using (1.1.4) we can find

Theorem 1 (Scalar triple product)**.**

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = A_i \ \varepsilon_{ijk} B_j C_k \tag{1.1.5}$$

This is also represented as $[\mathbf{ABC}]$.

Theorem 2 (Vector triple product)**.**

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \tag{1.1.6}$$

The vector product is not associative, hence the position of the parenthesis is important in the triple product.

(1.1.6) is derived using (1.1.4) and the identity (1.1.8).

Theorem 3 (Product of levi-cevitas)**.**

$$\varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix} = \delta_{il}\left(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\right) - \delta_{im}\left(\delta_{jl}\delta_{kn} - \delta_{jn}\delta_{kl}\right) + \delta_{in}\left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) \tag{1.1.7}$$

$$\varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} \tag{1.1.8}$$

$$\varepsilon_{ijk}\varepsilon_{ijn} = 2\delta_{kn} \tag{1.1.9}$$

Exercise 1. Show that

$$(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) = [\mathbf{ABD}]\mathbf{C} - [\mathbf{ABC}]\mathbf{D} \tag{1.1.10}$$

1.2 Rotation Transformation

The rotation of the 2D coordinate axes by an angle ϕ, keeping the origin fixed, leads to the

Definition 3 (Rotation transformation).

$$\mathbf{A}' = S\mathbf{A} \tag{1.2.1}$$

where S is the rotation transformation represented by the matrix

$$S = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{1.2.2}$$

in our orthonormal basis. It is clear that $SS^T = I$, i.e., the rotation transformation is orthogonal.

It takes the same form in higher dimensions, i.e., rotation of vectors is an orthogonal transformation. This is a special property of vectors in physics. There is another kind of quantity called

Definition 4 (Pseudovectors).

$$\mathbf{A}' = |S|S\mathbf{A} \tag{1.2.3}$$

This is how these quantities transform under rotation, where, S is again an orthogonal transformation and $|S|$ represents the determinant of S.

1.3 Differential Calculus

We require the del operator

$$\nabla_i \equiv \frac{\partial}{\partial e_i} \tag{1.3.1}$$

in Cartesian coordinates. Using this we can define quantities like

Definition 5 (Gradient).

$$\text{Grad } f = \boldsymbol{\nabla} f$$

A small change in a scalar field along $\mathbf{r}$ is given by

$$df = \mathbf{r} \cdot \boldsymbol{\nabla} f \tag{1.3.2}$$

Say, the direction of $\boldsymbol{\nabla} f$ is given by some $\hat{\mathbf{r}}$. We have the most rapid increase in f along $\hat{\mathbf{r}}$ and $|\boldsymbol{\nabla} f| = \frac{Df}{D\hat{\mathbf{r}}}$ is the directional derivative of f along $\hat{\mathbf{r}}$.

Definition 6 (Divergence).

$$\text{Div } \mathbf{A} = \boldsymbol{\nabla} \cdot \mathbf{A}$$

This gives a measure of the accumulation or depletion of $\mathbf{A}$ at a point. In other words, it finds how strongly a point acts as a source or sink. It is also known as the *source density* of the vector field.

Definition 7 (Curl).

$$\text{Curl } \mathbf{A} = \boldsymbol{\nabla} \times \mathbf{A}$$

This gives a measure of the circulation of $\mathbf{A}$ at a point. It is also known as the *circulation density* of the vector field.

Definition 8 (Laplacian).

$$\nabla^2 f = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) \tag{1.3.3}$$

The Laplacian is also defined for a vector field as

$$(\nabla^2 \mathbf{A})_i = \frac{\partial^2 A_i}{\partial e_i^2} \tag{1.3.4}$$

These can also be calculated in curvilinear coordinate systems as given in this Wikipedia article.

Some useful properties of vector derivatives are

Theorem 4.

$$\nabla \times (\nabla f) = 0 \tag{1.3.5}$$
$$\nabla \cdot (\nabla \times \mathbf{A}) = 0 \tag{1.3.6}$$
$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \tag{1.3.7}$$

(1.3.7) is only valid in Cartesian coordinates.

Theorem 5 (Exact differential)**.** $\sum_i A_i \; di$ is an exact differential iff $\nabla \times \mathbf{A} = 0$.

Two special kinds of vector fields that often appear in physics are

Definition 9 (Solenoidal)**.**

$$\nabla \cdot \mathbf{A} = 0 \tag{1.3.8}$$

Definition 10 (Irrotational)**.**

$$\nabla \times \mathbf{A} = 0 \tag{1.3.9}$$

Vector fields can be decomposed into solenoidal and irrotational components using

Theorem 6 (Helmholtz)**.** Suppose, $\nabla \cdot \mathbf{F}$ and $\nabla \times \mathbf{F}$ vanish at infinity and $\mathbf{F}$ is smooth. $\exists f, \mathbf{A}$ such that

$$\mathbf{F} = \underbrace{-\nabla f}_{\text{irrotational}} + \underbrace{\nabla \times \mathbf{A}}_{\text{solenoidal}} \tag{1.3.10}$$

1.4 Integral Calculus

A very useful quantity for integration is

Theorem 7 (Infintesimal length)**.**

$$\begin{aligned} d\mathbf{l} &= dx \; \hat{\mathbf{e_x}} + dy \; \hat{\mathbf{e_y}} + dz \; \hat{\mathbf{e_z}} & (1.4.1) \\ &= dr \; \hat{\mathbf{e_r}} + r \; d\theta \; \hat{\mathbf{e}}_\theta + r \sin\theta \; d\phi \; \hat{\mathbf{e}}_\phi & (1.4.2) \\ &= ds \; \hat{\mathbf{e_s}} + s \; d\phi \; \hat{\mathbf{e}}_\phi + dz \; \hat{\mathbf{e_z}} & (1.4.3) \end{aligned}$$

Now, we may define some of the most important vector integrals in physics.

Definition 11 (Line integral)**.**

$$\int_C \mathbf{A} \cdot d\mathbf{l}$$

Definition 12 (Surface integral)**.**

$$\iint_S \mathbf{A} \cdot d\boldsymbol{\sigma}$$

The contour enclosing a surface S is represented by ∂S.

Definition 13 (Volume integral).

$$\iiint_V \mathbf{A}\ d\tau$$

The surface enclosing a volume V is represented by ∂V.

Some important theorems that will help simplify integrals are

Theorem 8 (Conservation of gradient).

$$\int_C (\boldsymbol{\nabla} f)\cdot d\mathbf{l} = f(\mathbf{b}) - f(\mathbf{a}) \tag{1.4.4}$$

where $\mathbf{a}, \mathbf{b}$ are the endpoints of C(which is directed from $\mathbf{a}$ to $\mathbf{b}$).

Theorem 9 (Green).

$$\oint_{\partial S} P(x,y)\ dx + Q(x,y)\ dy = \iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ dA \tag{1.4.5}$$

Theorem 10 (Gauss).

$$\oiint_{\partial V} \mathbf{A}\cdot d\boldsymbol{\sigma} = \iiint_V \boldsymbol{\nabla}\cdot\mathbf{A}\ d\tau \tag{1.4.6}$$

Theorem 11 (Stokes).

$$\oint_{\partial S} \mathbf{A}\cdot d\mathbf{l} = \iint_S (\boldsymbol{\nabla}\times\mathbf{A})\cdot d\boldsymbol{\sigma} \tag{1.4.7}$$

Note that C, S, V need to be "sufficiently nice" for these theorems to be valid, which is often the case in physics.

1.5 Orthogonal Curvilinear Coordinate Systems

Let $\mathbf{\hat{q}_i}$ be the unit vectors of our generalized system. We would like to transform among coordinate systems.

Note. A vector $\mathbf{A}$ can be written as $A_i\mathbf{\hat{q}_i}$ in the generalized coordinate system. However, the position $\mathbf{r} \neq q_i\mathbf{\hat{q}_i}$, in general, though the equality holds in Cartesian system.

The square of the infinitesimal arc length is given by $d\mathbf{r}\cdot d\mathbf{r} = ds^2 = dx^2 + dy^2 + dz^2$ in Cartesian coordinates. This arc length is invariant of our coordinate system. In the generalized system, we have

$$ds^2 = g_{ij}dq_i dq_j \tag{1.5.1}$$

The g_{ij} depend on the geometry of the coordinate system and are given by

$$g_{ij} = \frac{\partial \mathbf{r}}{\partial q_i}\cdot\frac{\partial \mathbf{r}}{\partial q_j} \tag{1.5.2}$$

For orthogonality,

$$g_{ij} = 0,\ i \neq j$$
$$\mathbf{\hat{q}_i}\cdot\mathbf{\hat{q}_j} = \delta_{ij}$$

We can take $g_{ii} = h_i^2 > 0$(not Einstein summation). Now, to represent $d\mathbf{r}$ in our generalized orthogonal coordinate system, we use

Theorem 12 (Scale factors).

$$ds^2 = \sum_i (h_i\ dq_i)^2 \tag{1.5.3}$$

$$d\mathbf{r} = \sum_i h_i\ dq_i\mathbf{\hat{q}_i} \tag{1.5.4}$$

$$\frac{\partial \mathbf{r}}{\partial q_i} = h_i \mathbf{\hat{q}_i} \tag{1.5.5}$$

(1.5.5) does not involve Einstein summation.

Thus, we can define our differential operators as

Definition 14 (Gradient)**.**

$$\boldsymbol{\nabla}\phi = \sum_i \frac{1}{h_i} \frac{\partial \phi}{\partial q_i} \mathbf{\hat{q}_i} \tag{1.5.6}$$

Definition 15 (Divergence)**.**

$$\boldsymbol{\nabla} \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3} \sum_{\text{cyclic } ijk} \frac{\partial A_i}{\partial q_i} h_j h_j \tag{1.5.7}$$

Definition 16 (Curl)**.**

$$\boldsymbol{\nabla} \times \mathbf{A} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \mathbf{\hat{q_1}} & h_2 \mathbf{\hat{q_2}} & h_3 \mathbf{\hat{q_3}} \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix} \tag{1.5.8}$$

2 Tensor Analysis

Scalars and vectors are special cases of tensors. We will consider Cartesian coordinate systems.

2.1 Rank 1 tensors

Rank 1 tensors are called vectors. These are of two types.
Position vector $\mathbf{A}$ transforms in the following way

$$A'_i = (\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}_\mathbf{j})A_j = a_{ij}A_j$$

This a_{ij} can be found by writing down the transformation of the differential element using (1.3.2)

$$dx'_i = \frac{\partial x'_i}{\partial x_j} dx_j$$

Now, we can set $a_{ij} = \frac{\partial x'_i}{\partial x_j}$. Vectors that transform like this are called *contravarient vectors* and their indices are writen as superscripts. The cartesian coordinates are an example of contravarient vectors. Thus, we have

Definition 17 (Contravarient vector).

$$A'^i = a^{ij}A^j \tag{2.1.1}$$

where

$$a^{ij} = \frac{\partial x'^i}{\partial x^j} \tag{2.1.2}$$

Note that we have defined the contravarient vectors using a transformation law. This is the standard way of defining tensors in physics.

The gradient of a scalar function ϕ is given by

$$\begin{aligned}
(\boldsymbol{\nabla}\phi)'_i &= \frac{\partial \phi}{\partial x'_i} \\
&= \frac{\partial x_j}{\partial x'_i}\frac{\partial \phi}{\partial x_j} && \text{(chain rule)} \\
&= \frac{\partial x_j}{\partial x'_i}(\boldsymbol{\nabla}\phi)_j \\
&= b_{ji}(\boldsymbol{\nabla}\phi)_j
\end{aligned}$$

Vectors that transform like this are called *covarient vectors* and their indices are writen as subscripts. Thus we have

Definition 18 (Covarient vector).

$$B'_i = b_{ji}B_j \tag{2.1.3}$$

where

$$b_{ji} = \frac{\partial x^j}{\partial x'^i} \tag{2.1.4}$$

2.2 Rank 2 tensors

Rank 2 tensors are of 3 types.

Definition 19 (Contravarient).

$$A'^{ij} = a^{ik}a^{jl}A^{kl} \tag{2.2.1}$$

Definition 20 (Mixed).

$$B'^i_j = a^{ik}b_{lj}B^k_l \tag{2.2.2}$$

Definition 21 (Covarient).

$$C'_{ij} = b_{ki}b_{lj}C_{kl} \tag{2.2.3}$$

Example 1 (Kronecker delta)**.** The familiar Kronecker delta function is a mixed rank 2 tensor δ^i_j. We have

$$\begin{aligned}
a^{ik}b_{lj}\delta^k_l &= a^{ik}b_{kj} && \text{(property of Kronecker delta function)} \\
&= \frac{\partial x'^i}{\partial x^k}\frac{\partial x^k}{\partial x' j} \\
&= \frac{\partial x'^i}{\partial x'^j} \\
&= \delta'^i_j && (x_i, x_j \text{ are independent if } i \neq j)
\end{aligned}$$

Thus, it transforms like (2.2.2).

The order in which the indices appear in our description of a tensor is important. In general, A^{mn} is independent of A^{nm}, but there are some cases of special interest.

Definition 22 (Symmetric)**.**

$$A^{mn} = A^{nm}, \forall m, n \tag{2.2.4}$$

Definition 23 (Antisymmetric)**.**

$$A^{mn} = -A^{nm}, \forall m, n \tag{2.2.5}$$

Theorem 13. Every second rank tensor can be decomposed into a symmetric and antisymmetric tensor.

$$A^{mn} = \frac{1}{2}(A^{mn} + A^{nm}) + \frac{1}{2}(A^{mn} - A^{nm}) \tag{2.2.6}$$

Definition 24 (Contraction)**.**

$$B'^i_i = B^k_k \tag{2.2.7}$$

The contracted tensor is invariant and therefore a scalar(rank 0). In general, we set one contravarient index equal to a covarient index and sum over the repeated indices. Contraction reduces the rank of a tensor by 2.

Definition 25 (Direct product)**.** It can be shown that A_iB^j is a second rank mixed tensor

$$A'_iB'^j = b_{ki}a^{jl}A_iB^j \tag{2.2.8}$$

This is called the direct product which is a technique for creating new, higher-rank tensors. If we contract A_iB^j, we get

$$A'_iB'^i = A_kB^k$$

which is the familiar scalar product.

2.3 Quotient Rule

The quotient rule gives us a way to determine the rank of tensors in a given equation of the form

$$\mathbf{KA} = \mathbf{B}$$

where $\mathbf{K}$ is the unknown tensor. The rank of $\mathbf{K}$ is given by the following equations

Theorem 14 (Quotient rule)**.**

$$K_iA_i = B \tag{2.3.1}$$
$$K_{ij}A_j = B_i \tag{2.3.2}$$
$$K_{ij}A_{jk} = B_{kl} \tag{2.3.3}$$
$$K_{ijkl}A_{ij} = B_{kl} \tag{2.3.4}$$
$$K_{ij}A_k = B_{ijk} \tag{2.3.5}$$

We can go to higher ranks but this will be enough for our purposes. As mentioned before, the number of indices is the rank of the tensor.