Henri Poincare : Mathematics is the art of giving the same name to different things.

1 What is a Vector ?

1.1 Examples of vectors in geometry.

1. Displacements on a plane :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.1}$$

This space of vectors is determined by two real numbers. This **vector space** is called $\mathbb{R}^2$. We say this space has **dimension** two. The space of all such vectors is **closed under the addition** :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} x_1 + x_1' \\ x_2 + x_2' \end{pmatrix} \tag{1.2}$$

A displacement can be **scaled** by a scale factor λ to give another displacement of different magnitude but with the same direction

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rightarrow \lambda \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} \tag{1.3}$$

2. Displacements in space :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{1.4}$$

This vector space is $\mathbb{R}^3$ and has dimension 3. Again this space is closed under addition.

3. We could consider generalizations $\mathbb{R}^N$ describing

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix} \tag{1.5}$$

This is a vector space of dimension N. It may be viewed as displacements in an N-dimensional space, and underlies the generalization of geometry from two and three dimensions to any dimension.

1.2 Vectors and quantum states

Generalizations of the above with complex numbers replacing real numbers describe the mathematics of quantum states.

1. The space of column vectors with complex numbers ψ_1, ψ_2

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{1.6}$$

describes the quantum states of a **spin-half particle** such as the electron. Any vector in this vector space of quantum states is specified by two numbers. The vector space is called $\mathbb{C}^2$. This space is **closed under addition** . It is a **complex vector space**. But it shares with the real vector spaces of geometry the property of admitting addition. If $\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ and $\Psi' = \begin{pmatrix} \psi'_1 \\ \psi'_2 \end{pmatrix}$ are two state vectors, then the sum is also a state vector

$$\Psi + \Psi' = \begin{pmatrix} \psi_1 + \psi'_1 \\ \psi_2 + \psi'_2 \end{pmatrix} \tag{1.7}$$

That there are **two numbers** here is related to the fact that the measurement of the internal angular momentum of an electron along any axis, e.g the z-axis, can give only two possible values. This is very different from classsical phsyics where a spinning ball can have any value of angular momentum along a chosen axis (More on this in QMA and QMB).

2. More generally we can consider the vector space $\mathbb{C}^N$ which will also have applications in quantum physics. A general vector in the vector space $\mathbb{C}^N$ can be written as

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_N \end{pmatrix} \tag{1.8}$$

1.3 Function spaces as vector spaces

Polynomial functions of degree up to D have the general form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_D x^D \tag{1.9}$$

They are closed under addition, can be scaled. They form a vector space of dimension $D + 1$.

1.4 The lesson so far :

The same mathematical structure can be recognized in different physical situations, be it the geometry of displacements or quantum physics. The key properties shared is the existence of a property of addition, under which the set of objects under consideration is closed ; and the existence of a scaling operation by real or complex numbers. So we are recognizing some analogies between geometry and quantum physics and the common properties lead to the definitions of a vector space, which we will get to shortly. Once we set down the key definitions we will see that the same definitions allow us to deal with function spaces as examples of vector spaces.

But before that let us continue to explore the analogies. It turns out that there are analogies between angles in geometry and probability amplitudes in quantum mechanics.

1.5 The dot product of vectors in geometry

Given two vectors $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ in $\mathbb{R}^2$, we recall the dot products :

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= v_1 w_1 + v_2 w_2 \\ \mathbf{v} \cdot \mathbf{v} &= v_1^2 + v_2^2 \\ \mathbf{w} \cdot \mathbf{w} &= w_1^2 + w_2^2 \end{aligned} \tag{1.10}$$

Exercise : Prove, using the trigometric identity,

$$\cos(\theta_1 - \theta_2) = (\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2)$$

the familiar result that

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}|\cos\theta \tag{1.11}$$

where

$$\begin{aligned} |\mathbf{v}| &\equiv \sqrt{\mathbf{v} \cdot \mathbf{v}} \\ |\mathbf{w}| &\equiv \sqrt{\mathbf{w} \cdot \mathbf{w}} \end{aligned} \tag{1.12}$$

are the lengths of the respective vectors and θ is the angle between the vectors.

NOTE : The symbol $\equiv$ means "*is defined as*"

When $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^N$, we define the standard inner product as

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_N w_N \tag{1.13}$$

It is a general result true for $\mathbb{R}^N$ that the dot product of two vectors is the product of their lengths (also called magnitudes) times the cosine of the angle between them.

Dot products are also called scalar products or inner products. Sometimes you will see slightly different notation :

$$(\mathbf{v}, \mathbf{w}) = v_1 w_1 + v_2 w_2 + \cdots + v_N w_N \tag{1.14}$$

1.6 Inner products in quantum physics

Inner products are also extensively used in quantum physics. They are used to calculate amplitudes. The absolute values of amplitudes give probabilities. As you know from QP, because of the uncertainty principle, probabilities are very important in Quantum physics.

Given $\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_N \end{pmatrix}$ and $\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_N \end{pmatrix}$, which are two columns of complex numbers, elements of $\mathbb{C}^N$, the standard inner product is

$$(\Phi, \Psi) = \phi_1^* \psi_1 + \phi_2 \psi_2^* + \cdots + \phi_N^* \psi_N \tag{1.15}$$

Here ϕ^* is the complex conjugate of the complex number ϕ. YOu will recall, for real numbers a, b, r, θ and the imaginary unit $i = \sqrt{-1}$.

$$\begin{aligned} (a + ib)^* &= a - ib \\ (re^{i\theta})^* &= re^{-i\theta} \end{aligned} \tag{1.16}$$

1.7 The lesson

The common property between inner products in $\mathbb{R}^N$ and in quantum physics is that they are both ways of getting a number from a pair of vectors. We will give formal definitions of inner products for vector spaces later.

2 Vector Spaces : The definitions

THese notes are meant to be a summary of key points. It is useful to consult the books e.g Riley-Hobson-Bence Chapter 8

Having reviewed examples of vector spaces from geometry and from quantum mechanics, we give the precise definitions of Vector spaces.

A vector space V is a set of elements called vectors, $\mathbf{u}, \mathbf{v}, \mathbf{w}, \cdots$. For any pair of elements $\mathbf{u}, \mathbf{v}$ there is an operation of **addition** which combines two vectors to give a third, denoted as $\mathbf{u} + \mathbf{v}$, and an operation of **scaling** which multiplies any vector with numbers (e.g λ, μ below).

The definition of a vector space specifies an allowed set of numbers for the scaling. In all cases of interest to us the numbers will belong to the set $\mathbb{R}$ of real numbers or the set $\mathbb{C}$ of complex numbers.

When the scalars are in $\mathbb{R}$, we say that we have a **vector space over** $\mathbb{R}$. When they are in $\mathbb{C}$, we say that we have **a vector space over** $\mathbb{C}$. More loosely, we sometimes speak of a real or a complex vector space respectively.

2.1 Properties of Vector Spaces

Commutativity of vector addition : The addition of two elements $v, w \in V$ is commutative.

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \tag{2.1}$$

Associativity of vector addition : For any three elements of V

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \tag{2.2}$$

Multiplication by scalars :

$$\begin{aligned} &\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w} \\ &(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v} \\ &\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v} \\ &1(\mathbf{v}) = \mathbf{v} \end{aligned} \tag{2.3}$$

Existence of Null vector $\mathbf{0}$: For all $\mathbf{v}$

$$\mathbf{v} + \mathbf{0} = \mathbf{v} \tag{2.4}$$

The null vector is also called the *identity for the addition operation.*
Existence of negative vectors :

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0} \tag{2.5}$$

The negative is also called the *the inverse for the addition operation.*
Remarks :

- $\mathbb{C}^N$ is a vector space of dimension N over $\mathbb{C}$, or more loosely a complex vector space of dimension N. $\mathbb{R}^N$ is a vector space of dimension N over $\mathbb{R}$, or more loosely a real vector space of dimension N.

- **Subtle point :** The complex numbers $\mathbb{C}$ can be viewed as a two-dimensional vector space over $\mathbb{R}$, since any complex number z can be written as $x + iy$.

- *Exercise :* Verify the above statement, by checking that the definitions of a vector space above are satisfied.

2.2 Inner products

Sometimes one is interested only in the above properties. When a vector space is equipped with the additional structure of an inner product, it is called an **inner product space**.

The inner product is a rule for assiging a number to any pair of vectors. I will denote the inner product as $(\mathbf{v}, \mathbf{w})$. RHB denotes is as $\langle bv, bw \rangle$. Don't be alarmed, just different symbols for the same thing, like different names. But in any calculation, pick a convention and stick with it.

For Real vector spaces, where the above rules for scalar multiplication hold with the scalars being real numbers, we have the following properties of inner products :

$$\begin{aligned} &(\mathbf{v}, \mathbf{w}) = (\mathbf{w}, \mathbf{v}) \\ &(\mathbf{v}, \lambda\mathbf{w} + \mu\mathbf{x}) = \lambda(\mathbf{v}, \mathbf{w}) + \mu(\mathbf{v}, \mathbf{x}) \end{aligned} \tag{2.6}$$

For complex vector spaces, the above rules become

$$\begin{aligned} &(\mathbf{v}, \mathbf{w}) = (\mathbf{w}, \mathbf{v})^* \\ &(\mathbf{v}, \lambda\mathbf{w} + \mu\mathbf{x}) = \lambda(\mathbf{v}, \mathbf{w}) + \mu(\mathbf{v}, \mathbf{x}) \end{aligned} \tag{2.7}$$

The complex conjugations above are very important.

Note that if you remember (2.7) then you can recover (2.6) simply by dropping the complex conjugations.

Exercise : Show that the definitions (2.7) imply

$$\begin{aligned} (\lambda\mathbf{v} + \mu\mathbf{w}, \mathbf{x}) &= \lambda^*(\mathbf{v}, \mathbf{x}) + \mu^*(\mathbf{w}, \mathbf{x}) \\ (\lambda\mathbf{v}, \mu\mathbf{w}) &= \lambda^*\mu(\mathbf{v}, \mathbf{w}) \end{aligned} \tag{2.8}$$

3 Index Notation

Mathematics can be viewed as a set of tools for disciplined thinking. One of the tools Physicists find immensely powerful is index notation.

We will introduce index notation provides a powerful technical tool for understanding the properties of Matrices and vectors. When you study Einstein's relativity in "SPace, Time and Gravity," "Physical Dynamcs" and advanced Quantum physics courses, you will find this very useful.

Later we will also use it to manipulate Div, Grad and Curl, which you have encountered before.

What is an index ?

A symbol, such as $i, j, k \cdots$ which can take a set of values. They can be used to write equations in more compact form. Let me illustrate.

Here is a sequence of sums :

$$\begin{aligned} 1 &= 1 \\ 1+2 &= 3 \\ 1+2+3 &= 6 \\ &\vdots \end{aligned} \tag{3.1}$$

In general you may know that

$$1+2+3+\cdots+n = \frac{n(n+1)}{2} \tag{3.2}$$

A more elegant way to write the same equation, without the $\cdots$ is

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \tag{3.3}$$

The symbol $\sum$ is an *instruction to sum.* The symbol i is an **index** . It takes values from 1 to n. Read the left hand side(LHS) of equation (3.3) as *Sum over i from* 1 *to n* .

In Mathematica you recover the above formula as

$$\begin{aligned} &\text{In[1]: Sum}[i, \{i, 1, n\}] \\ &\text{Out[1]: } \frac{1}{2} n(n+1) \end{aligned} \tag{3.4}$$

As a slight detour to the main storyline here about index notation, I recommend the following review exercise.
Exercise : Prove the equation (3.3) by induction.

Index as a label for components of a vector

Suppose I have a column vector

$$\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \tag{3.5}$$

I can write

$$\mathbf{v} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{3.6}$$

Basis vectors are often denoted as

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{3.7}$$

The above expansion of $\mathbf{v}$ can be written as

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 \tag{3.8}$$

Using an index i to label the different terms in the sum

$$\mathbf{v} = \sum_{i=1}^{3} v_i\mathbf{e}_i \tag{3.9}$$

You can now start to see the power of Index Notation. A vector of length 10000 has an expansion

$$\mathbf{v} = \sum_{i=1}^{10000} v_i\mathbf{e}_i \tag{3.10}$$

In fact for any length n we can write

$$\mathbf{v} = \sum_{i=1}^{n} v_i\mathbf{e}_i \tag{3.11}$$

which is equivalent to but clearly more compact than writing

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + v_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{3.12}$$

The **dot product** of two three-component vectors v, w in space, with v as in (3.6), (3.8) and w given by

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \sum_{i=1}^{3} w_i\mathbf{e}_i \tag{3.13}$$

is

$$(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{3} v_i w_i \tag{3.14}$$

For two n-component vectors, the dot product is

$$(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{n} v_i w_i \tag{3.15}$$

4 Working with Vector Spaces

We will follow RHB Chapter 8 :
Linear Independence
Basis Vectors
Orthonormal Basis

5 Linear Operators

A linear operator associates with every vector $\mathbf{x}$ a vector $\mathbf{y}$

$$\mathbf{y} = \mathcal{A}\mathbf{x} \tag{5.1}$$

in such a way that

$$\mathcal{A}(\lambda\mathbf{v} + \mu\mathbf{w}) = \lambda\mathcal{A}\mathbf{v} + \mu\mathcal{A}\mathbf{w} \tag{5.2}$$

$\mathcal{A}$ operates on or transforms a vector. Defined independently of basis. We will see later that, after choosing a basis, $\mathcal{A}$ will be related to matrices.
Properties

$$(\mathcal{A} + \mathcal{B})\mathbf{x} = \mathcal{A}\mathbf{x} + \mathcal{B}\mathbf{x} \tag{5.3}$$

Definitions

$$\begin{aligned} (\lambda\mathcal{A}) &\equiv \lambda(\mathcal{A}\mathbf{x}) \\ (\mathcal{A}\mathcal{B})\mathbf{x} &\equiv \mathcal{A}(\mathcal{B}\mathbf{x}) \end{aligned} \tag{5.4}$$

5.1 Linear Operators, Matrices and more on Indices.

By choosing a basis $\{\mathbf{e}_i\}$, we can write

$$\mathcal{A}\mathbf{e}_i = \sum_j A_{ji}\mathbf{e}_j \tag{5.5}$$

The array of numbers A_{ji} are called the *the matrix elements of the linear operator* $\mathcal{A}$. The linear operator $\mathcal{A}$ is descibed by a matrix (array of numbers) which we can call A. Sometimes we uses the same symbol for the operator and the matrix, but you will see when discussing base change that keeping different symbols is useful. In the following we won't make this distinction between $\mathcal{A}$ and A.

Using (5.5) for the linear operator $(\mathcal{AB})$, we can write

$$(\mathcal{AB})\mathbf{e}_i \quad = \sum_j (\mathcal{AB})_{ji}\mathbf{e}_j \tag{5.6}$$

Using the definition (5.4)

$$\begin{aligned}(\mathcal{AB})\mathbf{e}_i \quad &= \mathcal{A}(\sum_k \mathcal{B}_{ki}\mathbf{e}_k) \\ &= \sum_k \mathcal{B}_{ki} \sum_j \mathcal{A}_{jk}\mathbf{e}_j \\ &= \sum_j (\mathcal{A}_{jk}\mathcal{B}_{ki})\mathbf{e}_j \end{aligned} \tag{5.7}$$

Comparing (5.6) and (5.7) we see that

$$(\mathcal{AB})_{ji} = \sum_k (\mathcal{A}_{jk}\mathcal{B}_{ki}) \tag{5.8}$$

Remark Note that summed indices are repeated indices. Sometimes we adopt the convention that we will not explicitly write out the summation symbol and just take it for granted that any repeated index is summed. This is called the summation convention. So using the summation convention, we can write :

$$(\mathcal{AB})_{ji} = \mathcal{A}_{jk}\mathcal{B}_{ki} \tag{5.9}$$

Exercise : Prove that

$$(\mathcal{ABC})_{ij} = \sum_k \sum_l \mathcal{A}_{ik}\mathcal{B}_{kl}\mathcal{C}_{lj} \tag{5.10}$$

or, using the summation convention :

$$(\mathcal{ABC})_{ij} = \mathcal{A}_{ik}\mathcal{B}_{kl}\mathcal{C}_{lj} \tag{5.11}$$

Remarks :

- Note that this equation contains two dummy indices and two free indices. The free indices must always match on both sides of a meaningful equation. The free indices here are i, j. The dummy indices are k, l.

- The dummy indices can be renamed, e.g the same equation can be written

$$(\mathcal{ABC})_{ij} = \sum_p \sum_q \mathcal{A}_{ip}\mathcal{B}_{pq}\mathcal{C}_{qj}$$

Note that $\sum_p \sum_q = \sum_q \sum_p$. Convince yourself this is true by writing out for some arbitrary function $F(i, j)$

$$\begin{aligned}
\sum_{j=1}^{2}\sum_{i=1}^{2} F(i,j) &= \sum_{j=1}^{2}(F(1,j) + F(2,j)) \\
&= (F(1,1) + F(2,1)) + (F(1,2) + F(2,2)) \\
&= F(1,1) + F(2,1) + F(2,1) + F(2,2)
\end{aligned}$$

And also

$$\begin{aligned}
\sum_{i=1}^{2}\sum_{j=1}^{2} F(i,j) &= \sum_{i=1}^{2}(F(i,1) + F(i,2)) \\
&= (F(1,1) + F(1,2)) + (F(2,1) + F(2,2)) \\
&= F(1,1) + F(2,1) + F(2,1) + F(2,2) \\
&= F(1,1) + F(2,1) + F(2,1) + F(2,2)
\end{aligned}$$

Clearly you can freely exchange the order of the summation symbols, so

$$(\mathcal{ABC})_{ij} = \sum_p \sum_q \mathcal{A}_{ip}\mathcal{B}_{pq}\mathcal{C}_{qj} = \sum_q \sum_p \mathcal{A}_{ip}\mathcal{B}_{pq}\mathcal{C}_{qj}$$

For this reason, sometimes we write

$$(\mathcal{ABC})_{ij} = \sum_{p,q} \mathcal{A}_{ip}\mathcal{B}_{pq}\mathcal{C}_{qj}$$

In fact,in most formulae of matrix algebra you will find that dummy indices always appear twice. So one adopts the summation convention which says that any index appearing twice is automatically understood to be summed, so we drop the summation sign altogether

$$(\mathcal{ABC})_{ij} = \mathcal{A}_{ip}\mathcal{B}_{pq}\mathcal{C}_{qj}$$

- When more than one dummy index appears in the same equation, do not use the same symbol for both. If you do, you will get confused, especially if you are using the summation convention.

5.2 Linear operators, Base change, Similarity transformation

Bases are not unique. The same vector can be expressed in two different bases. Coonsider the basis $\{\mathbf{e}_i\}$ and a new basis $\{\mathbf{e}'_i\}$, where $i = 1, 2, \cdots, N$. A vector can be written as a sum of the vectors in either basis.

$$\mathbf{u} = \sum_{i=1}^{N} u_i \mathbf{e}_i = \sum_{i=1}^{N} u'_i \mathbf{e}'_i \tag{5.12}$$

The numbers u_i are the components of the vector $\mathbf{u}$ in the basis $\{\mathbf{e}_i\}$. The numbers u'_i are the components of the vector $\mathbf{u}$ with respect to the basis $\{\mathbf{e}'_i\}$. In fact we can write

$$\mathbf{e}'_i = \sum_j S_{ji} \mathbf{e}_j \tag{5.13}$$

since it is possible to expand the basis vectors of the primed-basis set in terms of the unprimes basis set. The numbers S_{ji} are the numbers appearing in this expansion. COn-versely, we can write

$$\mathbf{e}_i = \sum_j T_{ji} \mathbf{e}_j \tag{5.14}$$

Combining (5.13) and (5.14) we have

$$\mathbf{e}'_i = \sum_j S_{ji} \sum_k T_{kj} \mathbf{e}'_k = \sum_j \sum_k T_{kj} S_{ji} \mathbf{e}'_k \tag{5.15}$$

On the other hand we have

$$\mathbf{e}'_i = \sum_k \delta_{ki} \mathbf{e}'_k \tag{5.16}$$

Comparing (5.16) and (5.15) we find

$$(ST)_{ik} = \delta_{ik} \tag{5.17}$$

which is the index form of the matrix equation

$$ST = \mathbf{1} \tag{5.18}$$

We can also show in the same way $TS = \mathbf{1}$.

Going back to (5.19)

$$\mathbf{u} = \sum_{i=1}^{N} u_i \mathbf{e}_i = \sum_i u'_i \sum_j S_{ji} \mathbf{e}_j = \sum_i \sum_j S_{ij} u'_j \mathbf{e}_i \tag{5.19}$$

so we learn

$$u_i = S_{ij} u'_j$$

Likewise, an operator $\mathcal{A}$ can be expressed as

$$\begin{aligned}\mathcal{A}\mathbf{e}_i &= \sum_j A_{ji}\mathbf{e}_j \\ \mathcal{A}\mathbf{e}'_i &= \sum_j A'_{ji}\mathbf{e}'_j\end{aligned} \tag{5.20}$$

The numbers A_{ji} can be collected in the matrix A and give the matrix elements of the operator $\mathcal{A}$ in the basis $\{\mathbf{e}_i\}$. The numbers A'_{ji} can be collected in the matrix A' and give the matrix elements of the operator $\mathcal{A}$ in the basis $\{\mathbf{e}_i\}$. NOtice that, we have taken care to distinguish the symbol $\mathcal{A}$ for the linear operator from the symbol A for the matrix with respect to a basis. This is especially important we are working with more than one basis.

From the two expansions (5.20) , we can obtain the relation between the matrices A and A' in terms of S. This was done in class to obtain

$$A' = S^{-1} A S \tag{5.21}$$

You should find the steps in your hand-written notes from class. This is called a **similarity transformation** .

Exercise Show that

$$\begin{aligned}trA' &= trA \\ tr(A')^k &= trA^k \quad \text{for any positive integer } k \\ Det(A') &= Det(A) \\ Det(A-\lambda\mathbf{1}) &= Det(A'-\lambda\mathbf{1})\end{aligned}$$

For the last step, you may like to use the identity $Det(AB) = Det(A)Det(B)$.

5.3 Rotations and orthogonal matrices

The rotation matrix for an anti-clockwise rotation by an angle θ in the x, y plane can be worked out by drawing pictures. The unit vector $\mathbf{e}_1$ along the x-axis goes to $\cos\theta\ \mathbf{e}_1 + \sin\theta\ \mathbf{e}_2$, where $\mathbf{e}_2$ is a unit vector along the y-axis. The unit vector $\mathbf{e}_2$ goes to $-\sin\theta\ \mathbf{e}_1 + \cos\theta \mathbf{e}_2$.

We express this as

$$\begin{aligned}R\mathbf{e}_x &= \cos\theta\ \mathbf{e}_x + \sin\theta\ \mathbf{e}_y \\ R\mathbf{e}_y &= -\sin\theta\ \mathbf{e}_x + \cos\theta\mathbf{e}_y\end{aligned} \tag{5.22}$$

Compare this to the equation

$$Re_i = \sum_j R_{ji}\mathbf{e}_j \tag{5.23}$$

which can be expanded as

$$\begin{aligned} R\mathbf{e}_1 &= R_{11}\mathbf{e}_1 + R_{21}\mathbf{e}_2 \\ R\mathbf{e}_2 &= R_{12}\mathbf{e}_1 + R_{22}\mathbf{e}_2 \end{aligned} \tag{5.24}$$

to read off

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{5.25}$$

Exercise Check that $RR^T = \mathbf{1}$
In index notation, this reads

$$\sum_j R_{ij}(R^T)_{jk} = \sum_j R_{ij}(R)_{kj} = \delta_{ik} \tag{5.26}$$

Matrices, in any dimension N, obeying $RR^T = \mathbf{1}$ are called orthogonal matrices.
For a vector $\mathbf{x} = \sum_i x_i\mathbf{e}_i$

$$R\mathbf{x} = \sum_i Rx_i\mathbf{e}_i = \sum_i \sum_j R_{ji}x_i\mathbf{e}_j \tag{5.27}$$

Using the standard inner product on $\mathbb{R}^N$, the square of the norm of $R\mathbf{x}$ is

$$\begin{aligned} (R\mathbf{x}, R\mathbf{x}) &= \sum_i \sum_j R_{ji}x_j \sum_k R_{ki}x_k \\ &= \sum_j \sum_k (\sum_i R_{ji}R_{ki})x_i x_k \\ &= \sum_j \sum_k \delta_{jk}x_i x_k \\ &= (\mathbf{x}, \mathbf{x}) \end{aligned} \tag{5.28}$$

This shows that, in any dimension N, it makes sense to define rotation operators using orthogonal matrices. These operators preserve lengths (norms). Similar steps, done in class, show that

$$(R\mathbf{x}, R\mathbf{y}) = (\mathbf{x}, \mathbf{y}) \tag{5.29}$$

This shows that angles are preserved since $(\mathbf{x}, \mathbf{y})$ can shown in any dimension to be equal to $|\mathbf{x}||\mathbf{y}|\cos\theta$ where θ is the angle between the two vectors.

6 Hermitian Operators

Definition of hermitian conjugate operator $A^\dagger$ of an operator A is given by teh property that

$$(\mathbf{v}, \mathcal{A}\mathbf{w}) = (\mathcal{A}^\dagger \mathbf{v}, \mathbf{w}) \tag{6.1}$$

for any vectors $\mathbf{v}, \mathbf{w}$.

In an orthonormal basis, the matrix elements A_{ji} in

$$\mathcal{A}\mathbf{e}_i = \sum_j A_{ji}\mathbf{e}_j$$

can be written as

$$A_{ji} = (\mathbf{e}_j, \mathcal{A}\mathbf{e}_i) \tag{6.2}$$

A Hermtian operator is defined by $\mathcal{H} = \mathcal{H}^\dagger$, so that

$$(\mathbf{v}, \mathcal{H}\mathbf{w}) = (\mathcal{H}\mathbf{v}, \mathbf{w}) \tag{6.3}$$

for any vectors $\mathbf{v}, \mathbf{w}$. Applying this equation in the orthonormal basis leads to

$$H_{ij} = H_{ji}^* \tag{6.4}$$

which is the familiar equation defining hermitian matrices.

Exercise Complete the steps in proving (6.4) from (6.3). Consult class notes if necessary.

6.1 Real eigenvalues

An eigenvector of the hermitian operator $\mathcal{H}$ is a non-zero vector which obeys $\mathcal{H}\mathbf{v} = \lambda\mathbf{v}$. COnsider

$$(\mathbf{v}, \mathcal{H}\mathbf{v}) = (\mathcal{H}\mathbf{v}, \mathbf{v}) \tag{6.5}$$

from the definiton of hermitian operator. Write out each side of above equation

$$\begin{aligned} (\mathbf{v}, \mathcal{H}\mathbf{v}) &= (\mathbf{v}, \lambda\mathbf{v}) = \lambda(\mathbf{v}, \mathbf{v}) \\ (\mathcal{H}\mathbf{v}, \mathbf{v}) &= (\lambda\mathbf{v}, \mathbf{v}) = \lambda^*(\mathbf{v}, \mathbf{v}) \end{aligned} \tag{6.6}$$

Compare the two sides to get

$$(\lambda - \lambda^*)(\mathbf{v}, \mathbf{v}) = 0 \tag{6.7}$$

For a non-zero vector $(\mathbf{v}, \mathbf{v}) \neq 0$. So we conclude

$$\lambda = \lambda^* \tag{6.8}$$

which is equivalent to sayign that λ is real. To summarise
Eigenvalues of hermitian operators are real

Another result, proved by similar steps in class,
If two eigenvectors of $\mathcal{H}$ correspond to distinct eigenvectors they must be orthogonal

In a basis,which we will take to be orthonormal, the equation $\mathcal{H}\mathbf{v} = \lambda\mathbf{v}$ is expressed as

$$(H\mathbf{v})_i = \lambda \mathbf{v}_i \tag{6.9}$$

We have seen before that $(\mathcal{H}\mathbf{v})_i = \sum_j H_{ij}\mathbf{v}_j$. So

$$H_{ij}\mathbf{v}_j = \lambda \mathbf{v}_i \tag{6.10}$$

This is the eigenvalue equation for the matrix H of matrix elements of the operator $\mathcal{H}$. The above (index-free) proof of reality of eigenvalues proves that the eigenvalues of Hermitian matrices are real.

7 Further Notes

Some remarks on how the Linear algebra we have covered (and some extensions) are useful in subsequent Physics courses.

More on Vectors in quantum physics :

1. State vectors in Quantum mechanics. The quantum state

$$|\uparrow_z\rangle$$

gives the mathematical description of an electron with spin up in the z-direction.

The quantum state

$$|\downarrow_z\rangle$$

corresponds to an electron with spin down in the z-direction.

The state

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(\,|\uparrow_z\rangle + |\downarrow_z\rangle\,)$$

corresponds to spin up in the x-direction. The state

$$| \downarrow_x\rangle = \frac{1}{\sqrt{2}}(\, | \uparrow_z\rangle - | \downarrow_z\rangle \,)$$

corresponds to spin up in the x-direction.

The state for spin up in the y-direction is

$$| \uparrow_y\rangle = \frac{1}{\sqrt{2}}(| \uparrow_z\rangle + i| \downarrow_z\rangle)$$

Notice the imaginary unit i. The state for spin down in the y-direction is

$$| \downarrow_y\rangle = \frac{1}{\sqrt{2}}(| \uparrow_z\rangle - i| \downarrow_z\rangle)$$

The derivation of these formulae uses Eigenvalues and Eigenvectors which we will turn to in Weel 2-3.

A general spin-state of an electron can be written as a

$$|\psi\rangle = \psi_1| \uparrow_z\rangle + \psi_2| \downarrow_z\rangle$$

where ψ_1, ψ_2 are complex numbers. We can arrange these numbers in a column

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

Any vector in this vector space of quantum states is specified by two numbers. The vector space is called $\mathbb{C}^2$. This space is **closed under addition** . It is a **complex vector space**. But it shares with the real vector spaces of geometry the property of admitting addition.

2. For a particle of spin J we have $\mathbb{C}^N$ where $N = 2J+1$ A general vector in the vector space $\mathbb{C}^N$ can be written as

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_n \end{pmatrix}$$

Unitary matrices

- $UU^\dagger = \mathbf{1}$
- Matrices relating two orthonormal bases are unitary.
- After finding eigenvalues and eigenvectors of H, we can write $H = UDU^\dagger$, where U is constructed from the orthonormalized eigenvectors. If eigenvalues not zero, $H^{-1} = UD^{-1}U^\dagger$.
- Unitary and hermitian matrices related by exponentials of matrices. If H is hermitian, then e^{itH} is unitary. (Useful in Quantum Physics-QMS)

Generalized Inner products

The condition that the only vector with zero norm is the zero-vector is given up. Whereas the usual inner product is given by δ_{ij}, in special relativity one uses $\eta_{\mu\nu}$ which is zero when $\nu \neq \nu$ and

$$\begin{aligned}\eta_{00} &= -1\\ \eta_{11} &= 1\\ \eta_{22} &= 1\\ \eta_{33} &= 1\end{aligned}$$

LOrentz transformations obey a generalization of "orthogonality" given by $L^T \eta L = \eta$.