INFINITE VECTOR SPACES

INFINITE VECTOR SPACES
PHYS3591

David G. Cerdeño
(November 18, 2019)

Durham University

CONTENTS

Disclaimer: These notes may (and most likely will) contain typographical erorrs and must be used with care. They are solely meant as a guideline of the materials that will be covered in the class but by no means can substitute the basic references.

CHAPTER 1

VECTOR SPACES

We will start these lectures with a warmup, reviewing vector and coordinate transformations and applying them to the three dimensional case (with the known example of Euler angles).

1.1 Vectors and coordinate transformations

1.1.1 Vector spaces

Let us start by recalling the definition of a vector space.

A vector space is a set of elements $E = \{\vec{x}\}$ with a series of properties, with respect to a scalar Field, F, (for example the real or the complex numbers).

There are two operations: sum ($\vec{x} + \vec{y}$) and product with a scalar ($\alpha \cdot \vec{x}$).

Properties of the sum:

- commutative
- associative
- neutral element
- inverse element

Properties of the product with a scalar

- associative
- neutral element in F
- distributive

Examples of vector spaces

- $E = \{C\}$; $F = R$ with $+$: usual complex sum and $\cdot$: usual product by a scalar
- $E = \{R^n\} = \left\{x^1, x^2 \ldots x^n\right\}$; $F = R$ with $+$: usual sum and $\cdot$: usual product
- $E = f(x)$ with $f(x)$ all the real functions defined in (a, b) and $F = R$. $+$: usual sum of functions and $\cdot$: usual product. This is an infinite dimension space

1.1.2 Basis and eigenvectors

We can represent a vector, $\vec{x}$, in a n-dimensional space in terms of its decomposition in the orthonormal basis $\vec{e}_i$, with $i = 1, \ldots, n$:

$$\vec{x} = \sum_{i=1}^{n} \vec{e}_i x^i = \vec{e}_1 x^1 + \vec{e}_2 x^2 + \ldots + \vec{e}_n x^n \ . \tag{1.1}$$

The quantities $(x^1,\ x^2,\ \ldots, x^n)$, or x^i for short, are the coordinates and $\vec{e}_i$ the vector basis. The expression above can be written as a matrix product as follows

$$\vec{x} = [\vec{e}_1,\ \vec{e}_2,\ \ldots \vec{e}_n] \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} \ , \tag{1.2}$$

where we use row vectors for the basis eigenvectors and column vectors for the coordinates.

Vectors are objects which are *independent* of the coordinate systems, but their coordinates are not. If we now choose a transformed *new* basis of orthonormal eigenvectors, $\vec{e}'_i$, the new coordinates x'^i satisfy

$$\vec{x} = \sum_{i=1}^{n} \vec{e}'_i x'^i = \vec{e}'_1 x'^1 + \vec{e}'_2 x'^2 + \ldots + \vec{e}'_n x'^n \ . \tag{1.3}$$

The relation between the *new basis* and the *old basis* can be written in terms of a set of coefficients S_1^1 as follows

$$\begin{aligned} \vec{e}'_1 &= \vec{e}_1 S_1^1 + \vec{e}_2 S_1^2 + \ldots + \vec{e}_n S_1^n = \sum_i \vec{e}_i S_1^i \\ \vec{e}'_2 &= \vec{e}_1 S_2^1 + \vec{e}_2 S_2^2 + \ldots + \vec{e}_n S_2^n = \sum_i \vec{e}_i S_2^i \\ &\vdots \qquad (1.4) \\ \vec{e}'_n &= \vec{e}_1 S_n^1 + \vec{e}_2 S_n^2 + \ldots + \vec{e}_n S_n^n = \sum_i \vec{e}_i S_n^i \qquad (1.5) \end{aligned}$$

This can also be written in matrix form

$$\vec{e}\,' = [\vec{e}\,'_1,\ \vec{e}\,'_2,\ \ldots \vec{e}\,'_n] = [\vec{e}_1,\ \vec{e}_2,\ \ldots \vec{e}_n] \begin{bmatrix} S^1_1 & S^1_2 & \ldots & S^1_n \\ S^2_1 & S^2_2 & \ldots & S^2_n \\ \vdots & & & \\ S^n_1 & S^n_2 & \ldots & S^n_n \end{bmatrix} = \vec{e}\, S \tag{1.6}$$

The matrix S^j_i is used to compute the transformed coordinates for a given vector

$$\vec{x} = \sum_{i=1}^{n} \vec{e}\,'_i x'^i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \vec{e}_j S^j_i \right) x'^i = \sum_{j=1}^{n} \vec{e}_j x^j \tag{1.7}$$

from where

$$x^j = \sum_{i=1}^{n} S^j_i x'^i \tag{1.8}$$

or, in matrix notation,

$$\begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} = \begin{bmatrix} S^1_1 & S^1_2 & \ldots & S^1_n \\ S^2_1 & S^2_2 & \ldots & S^2_n \\ \vdots & & & \\ S^n_1 & S^n_2 & \ldots & S^n_n \end{bmatrix} \begin{bmatrix} x'^1 \\ x'^2 \\ \vdots \\ x'^n \end{bmatrix} \tag{1.9}$$

Notice at this point that we can also use the inverse matrix $(S^{-1})^i_j \equiv T^i_j$ to relate x'^i with x^i as follows.

$$\begin{bmatrix} x'^1 \\ x'^2 \\ \vdots \\ x'^n \end{bmatrix} = [T] \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{bmatrix} . \tag{1.10}$$

1.1.3 Einstein's Notation (Implicit Summation)

The student might have already noticed that summation symbols are ubiquitous (and often quite cumbersome). When the range of summation can be inferred from the context, summation symbols are in fact redundant. Einstein's notation proposes to get rid of the summation symbol, thus, for example

$$m^2 = |\vec{p}\,|^2 = \sum_{i=1}^{4} p_i p^i = p_i p^i \; . \tag{1.11}$$

The basic set of rules of Einstein's notation are as follows

- Whenever the same index appears twice in an expression, *once as a superscript and once as a subscript*, summation over the range of that index is implied.
- The range of summation can be inferred from the context (in case of ambiguity, the summation is written explicitly).

- The index used in implicit summation is a *dummy* index and may be replaced by any other that is not already in use, e.g., $p_i p^i = p_j p^j$.
- *Free* indices are not summed on (e.g., p^i, $T^i_j x^j$, $x_i x^j$).
- It is conventional to place the symbol carrying the summation index as subscript on the left of the symbol carrying the summation index as superscript, e.g., $p_i p^i$, or $g_{ij} x^i x^j$.

At this point, it is useful to re-introduce the *Kronecker* delta symbol, δ^j_i, which is defined as follows

$$\delta^j_i = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{1.12}$$

which can be interpreted as a representation of the identity matrix

$$x^i = \delta^i_j x^j \tag{1.13}$$

Using Einstein notation, we can rewrite the expressions from the previous subsection that apply to a change of basis in a much more compact form:

$$\begin{aligned} \vec{x} &= \vec{e}_i x^i = \vec{e}\,'_i x'^i \\ \vec{e}\,'_j &= \vec{e}_i S^i_j \\ x^j &= S^j_i x'^i \\ x'^j &= T^j_i x^i \end{aligned} \tag{1.14}$$

The product of two matrices $A^i{}_j$ and $B^j{}_k$ is also written in a trivial way

$$(AB)^i{}_j = A^i{}_k B^k{}_j \tag{1.15}$$

Some identities are very easy to prove using Einstein's notation. For example,

$$\text{Tr}(AB) = (AB)^i{}_i = A^i{}_j B^j{}_i = B^j{}_i A^i{}_j = (BA)^j j = \text{Tr}(BA) \tag{1.16}$$

EXAMPLE 1.1

3D Rotations as a change of basis: Let us consider vectors in three dimensions and a given choice of orthonormal basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Let us assume a new coordinate system $\{\vec{e}\,'_1, \vec{e}\,'_2, \vec{e}\,'_3\}$ defined by a rotation of angle φ around the $3-$axis.

The new eigenvectors can be expressed in terms of the old basis as follows

$$\begin{aligned} \vec{e}\,'_1 &= \vec{e}_1 \cos\varphi + \vec{e}_2 \sin\varphi \\ \vec{e}\,'_2 &= \vec{e}_1(-\sin\varphi) + \vec{e}_2 \cos\varphi \\ \vec{e}\,'_3 &= \vec{e}_3 \end{aligned} \tag{1.17}$$

In matrix language, the change of basis matrix reads

$$S(\vec{e}_3, \varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.18}$$

Thus, a given vector $\vec{p}$, with coordinates p^1, p^2, p^3 in the old basis, would have the following coordinates in the rotated basis:

$$\begin{bmatrix} p'^1 \\ p'^2 \\ p'^3 \end{bmatrix} = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p^1 \\ p^2 \\ p^3 \end{bmatrix} \equiv R(\vec{e}_3, \varphi) \begin{bmatrix} p^1 \\ p^2 \\ p^3 \end{bmatrix} \tag{1.19}$$

such that

$$\begin{aligned} p'_1 &= p_1 \cos\varphi + p_2 \sin\varphi \\ p'_2 &= p_1(-\sin\varphi) + p_2 \cos\varphi \\ p'_3 &= p_3 \end{aligned} \tag{1.20}$$

One can do the same exercise around the $1-$ and $2-$axis as well, remembering that rotation of a positive angle is counterclockwise. The rotation matrices around the $1-$, $2-$, and $3-$ axis would read, respectively

$$R(\vec{e}_1, \varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{bmatrix} \tag{1.21}$$

$$R(\vec{e}_2, \varphi) = \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix} \tag{1.22}$$

$$R(\vec{e}_3, \varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.23}$$

It is now easy to understand that the consecutive rotation around the $3-$axis by angles φ_1 and φ_2 is described by the following matrix:

$$R(\vec{e}_3, \varphi_1 + \varphi_2) = \begin{bmatrix} \cos(\varphi_1 + \varphi_2) & \sin(\varphi_1 + \varphi_2) & 0 \\ -\sin(\varphi_1 + \varphi_2) & \cos(\varphi_1 + \varphi_2) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.24}$$

Note that we have found a matrix representation of $e^{i\varphi_1} e^{i\varphi_2} = e^{i(\varphi_1 + \varphi_2)}$

We should note an interesting property of these rotation matrices, $R^T = R^{-1}$ or equivalently, $R^T R = 1$, which can also be written as $(R^T R)_i^j = \delta_i^j$. A matrix with these properties is called **orthogonal matrix**.

Any orthogonal 3×3 matrix, $\mathcal{O}$, can be written as a product of the three $R(\vec{e}_i, \varphi_i)$ and the reflexion matrix

$$S = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{1.25}$$

This is, $\mathcal{O} = R(\vec{e}_1, \varphi_1) R(\vec{e}_2, \varphi_2) R(\vec{e}_3, \varphi_3)$. This is equivalent to Euler's theorem.

PROBLEMS

1.1 First problem

$$x^i \tag{1.26}$$

First solution

1.2 Linear Forms

A linear form (or functional) is an application F that transforms elements of the vector space E into those of the scalar set K.

$$\begin{array}{rcl} F:E & \longrightarrow & K \\ \vec{x} & \longrightarrow & F(\vec{x}) \end{array} \tag{1.27}$$

Theorem 1.1 *All linear form F is fully defined if the following numbers are specified.*

$$F(\vec{e}_i) = f_i \ , \tag{1.28}$$

where $\{\vec{e}_i\}$ is a basis in E.

Proof:

$$F(\vec{x}) = F\left(\sum_{i=1}^{n} \vec{e}_i x^i\right) = \sum_{i=1}^{n} F\left(\vec{e}_i\right) x^i = \sum_{i=1}^{n} f_i x^i \tag{1.29}$$

■

1.3 Dual Space

E^*, dual space with respect to E, is the set of all the linear forms (functionals) that can be defined in E.

It is easy to show that E^* is also a vector space: the sum of two linear functionals is a linear functional; the sums of linear functionals and products by scalars also obey the properties of a vector space; we can define a neutral element (a zero functional that maps every vector in E to the number 0) and an additive inverse element.

From now on, we will denote as $\vec{x}^{\,*}$ the elements (linear forms) of the dual vector space.

1.3.1 Canonical dual basis (in E^*)

Given a basis $\{\vec{e}_i\}$ in E, we can define a canonical dual basis in E^*, $\left\{\vec{e}^{\,*i}\right\}$, which is made out of linear forms which, acting on the basis of E we obtain the Kronecker delta. This is,

$$\vec{e}^{\,*i}(\vec{e}_j) = \delta^i_j \tag{1.30}$$

It can be shown that these linear forms are linearly independent and that every linear form can be expressed as a linear combination of $\left\{\vec{e}^{\,*i}\right\}$. Thus, they form a basis. From now on, we will omit the $*$ on the dual basis, thus $\vec{e}^{\,*i} = \vec{e}^{\,i}$.

From here, it also follows that the components of $\vec{y}$, y^i, can be expressed as $y^i = \vec{e}^{\,i}(\vec{y})$.

Given a linear form (or functional) $\vec{x}^{\,*}$, we can study how it acts on an element of E, $\vec{y}$:

$$\vec{x}^{\,*}(\vec{y}) = \vec{x}^{\,*}\left(\vec{e}_i y^i\right) = \vec{x}^{\,*}\left(\vec{e}_i\right) y^i \equiv x^*_i y^i = x^*_i \vec{e}^{\,i}(\vec{y}) \tag{1.31}$$

In this expression, $\vec{x}^{\,*}_i$ are the components of the linear form in the dual basis (they are just numbers). From this expression,

$$\vec{x}^{\,*} = x^{\,*}_{\,i} \vec{e}^{\,i} \tag{1.32}$$

1.3.2 Covectors

We will refer to the elements of the dual space E^* as *covectors*. As we saw above, a covector $\vec{x}^*$ can be expressed as a linear combination of the dual basis members. $\vec{x}^* = x_i^* \vec{e}^i$.

It can be shown that, upon a change of basis, the components x_i^* transform as

$$x_i^{*\prime} = x_j^* S_i^j \tag{1.33}$$

which can be compared to the rest of the transformations in Eq. 1.69. As we can see, covectors behave differently than vectors under change of basis. Vector bases are transformed using S and coordinates using T. Covectors bases are transformed using T and coordinates using S. Covectors are said to be *covariant* and vectors *contravariant*.

If the basis vectors are the same i.e. we had orthonormal bases then the contravariant and covariant components are IDENTICAL. But of course in general they are not.

1.4 Tensors

We will now generalise the objects that we can define in a vector space, introducing the concept of tensors.

A tensor of type $(r,\, s)$ over an $n-$dimensional vector space is an object, $A^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ consisting on n^{r+s} components that obeys the following transformation rule under change of basis

$$A'^{\,i_1 \ldots i_r}_{j_1 \ldots j_s} = T^{i_1}_{m_1} \ldots T^{i_s}_{m_s} A^{m_1 \ldots m_r}_{n_1 \ldots n_s} S^{n_1}_{j_1} \ldots S^{n_s}_{j_s} \tag{1.34}$$

The coordinates $i_1, \ldots i_r$ are called *contravariant coordinates* and $j_1, \ldots j_s$ are the *covariant coordinates*, generalising the concept that we introduced above for vectors and covectors.

Remember that contravariant coordinates appear above and are transformed using T, whereas covariant coordinates appear below and are transformed using S.

1.4.1 Examples, Rank ordering, and Transformations

We can classify tensors of different ranks, depending on how they transform under rotations (orthogonal coordinate transformations) $\mathbf{R}$ and coordinate inversions $\mathbf{S}$.

- **Scalars:** for example, the mass

$$\begin{aligned} \mathbf{R} : \phi \quad &\longrightarrow \quad \phi' = \phi \\ \mathbf{S} : \phi \quad &\longrightarrow \quad \phi' = \phi \end{aligned}$$

- **Pseudoscalars:** for example, the scalar triple product $\rho = a \cdot (\vec{x} \times \vec{y})$

$$\begin{aligned} \mathbf{R} : \rho \quad &\longrightarrow \quad \rho' = \rho \\ \mathbf{S} : \rho \quad &\longrightarrow \quad \rho' = -\rho \end{aligned}$$

- **Vectors:** (or tensors of rank 1) for example, the velocity

$$\begin{aligned} \mathbf{R} : v^i \quad &\longrightarrow \quad v'^i = v^j \mathbf{R}^i_j \\ \mathbf{S} : v^i \quad &\longrightarrow \quad v'^i = v^j \mathbf{S}^i_j = -v^i \end{aligned}$$

- **Axial vector (pseudovector):** a quantity that transforms as a vector under a proper rotation $\mathbf{R}$ but that gains a sign flip under reflections $\mathbf{S}$.

$$\begin{aligned} \mathbf{R} : v^i \qquad &\longrightarrow \qquad v'^i = (\det \mathbf{R}) v^j \mathbf{R}^i_j = v^j \mathbf{R}^i_j \\ \mathbf{S} : v^i \qquad &\longrightarrow \qquad v'^i = (\det \mathbf{S}) v^j \mathbf{S}^i_j = -v^j \mathbf{S}^i_j \end{aligned}$$

Consider for example the vector product of two vectors $\vec{x} \times \vec{y}$, which under

- **Moment of inertia** In tensorial notation, the moment of inertia of a rigid body (which relates the angular momentum and the angular velocity) corresponds to a rank 2 tensor, which, in terms of the position vectors $\vec{r_n}$ can be expressed as

$$I^{ij} = \sum_n m_n \left[r_n^2 \delta^{ij} - r_n^i r_n^j \right] \tag{1.35}$$

The moment of inertia relates the angular momentum $\vec{L}$ and the angular velocity $\vec{w}$,

$$L^i = w_j I^{ji} = \delta_{jk} w^k I^{ji} \tag{1.36}$$

Notice that, by definition, it is a symmetric tensor, and therefore it can always be diagonalised to obtain the *principal moments of inertia*

$$I' = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \tag{1.37}$$

- **Isotropic tensors** are invariant under coordinate rotations.

All rank 0 tensors (scalars and pseudoscalars) are isotropic. No rank 1 tensor is isotropic (except for the null vector). The unique rank-2 isotropic tensor is the Kronecker delta, δ_{ij}, and the unique rank-3 isotropic tensor is the permutation symbol, ϵ_{ijk}

- **Kronecker delta**

It is easy to prove that the Kronecker delta is an isotropic tensor. We only need to prove that it is invariant under rotations:

$$\delta'_{mn} = \delta_{ij} R^i_m R^j_n = R_{jm} R^j_n = (R^T)_{mj} R^j_n = (R^T R)_{mn} = \delta_{mn} \tag{1.38}$$

- **Levi Civita tensor** The rank-3 Levi-Civita tensor is defined as follows

$$\epsilon_{ijk} = \begin{cases} 1 & (i,j,k) = (1,2,3),\ (3,1,2),\ (2,3,1) \\ -1 & (i,j,k) = (1,3,2),\ (2,1,3),\ (3,2,1) \\ 0 & \text{if any two indices are equal} \end{cases} \tag{1.39}$$

1.5 Scalar and vector products

The scalar and vector products can be expressed in terms of tensorial contractions as follows:

$$\vec{a} \cdot \vec{b} = \delta_{ij} a^i b^j = a_i b^i \tag{1.40}$$

$$(\vec{a} \times \vec{b})_i = \epsilon_{ijk} a^j b^k \tag{1.41}$$

PROBLEMS

1.1 First problem

$$x^i \tag{1.42}$$

First solution

1.6 Groups and Fields

Let us now introduce some concepts which we are going to need when describing infinite vector spaces.

1.6.1 Groups

A group G is a set of elements together with an operation, $*$, that assigns every *ordered* pair of elements $f, g \in G$ another element, $h = f * g$ of group G. We have the following set of properties

G1 If $f, g \in G$ then $h = f * g \in G$

G2 For any $f, g, h \in G$ we have $f * (g * h) = (f * g) * h$

G3 There exists a *unique* identity element, e, such that $f * e = e * f = f$

G4 $\forall f \in G$ there exists an inverse element f^{-1} such that $f * f^{-1} = f^{-1} * f = e$. (It can also be shown that the inverse of $f * g$ is $g^{-1} * f^{-1}$).

It is customary to define a *multiplication table* that show how the given operator acts on an ordered pair of elements of the group. For example, for a group G with elements $\{e, f_1, f_2, f_3 \ldots\}$

$(G, *)$	e	f_1	f_2	f_3	$\ldots$
e	$e * e$	$e * f_1$	$e * f_2$	$e * f_3$	$\ldots$
f_1	$f_1 * e$	$f_1 * f_1$	$f_1 * f_2$	$f_1 * f_3$	$\ldots$
f_2	$f_2 * e$	$f_2 * f_1$	$f_2 * f_2$	$f_2 * f_3$	$\ldots$
f_3	$f_3 * e$	$f_3 * f_1$	$f_3 * f_2$	$f_3 * f_3$	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$

1.6.2 Representation of a group

A representation of a group G is a mapping D of the elements of G onto a set of linear operators. This mapping is not unique. It satisfies the following properties:

- For the identity operator $D(e) = 1$
- The group multiplication law $*$ is mapped onto the multiplication of objects in the representation: $D(g1)D(g2) = D(g_1 * g_2)$

EXAMPLE 1.2

Consider the cyclic group Z_3. This is an example of a *finite group*, with three elements. Thus the dimension of G is three. We label the three elements as $\{emamb\}$ We can see that the group is *abelian*, this is $f*g = g*f$ (leading to a symmetric multiplication table).

We can choose different representations. The dimension of the representation

$(Z_3, *)$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Some representations of Z_3 in different dimensions

- 1 Dimensional Rep. $D(e) = 1$, $D(a) = e^{i2\pi/3}$, $D(b) = e^{i4\pi/3}$
- 2 Dimensional Rep. $D(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $D(a) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$, $D(b) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$
- Regular Representationx can be constructed from the multiplication table. In particular we take the elements of the group and consider that they form an orthonormal basis. In this case of Z_3, using the bra and ket notation, we have the elements $|e\rangle$, $|a\rangle$, $|b\rangle$ and we define the multiplication rule of the group as follows:

$$g_1 * g_1 \equiv D(g_1) |g_1\rangle = |g_1 g_1\rangle \tag{1.43}$$

In such a way that we can construct the matrix

$$[D(g)]_{ij} = \langle g_i| D(g) |g_j\rangle \tag{1.44}$$

that corresponds to the matrix for the element g in the regular representation. In the case of Z_3, we are left with

$$\begin{aligned}
D(e) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D(a) &= \begin{pmatrix} \langle e| D(a) |e\rangle & \langle e| D(a) |a\rangle & \langle e| D(a) |b\rangle \\ \langle a| D(a) |e\rangle & \langle a| D(a) |a\rangle & \langle a| D(a) |b\rangle \\ \langle b| D(a) |b\rangle & \langle b| D(a) |a\rangle & \langle b| D(a) |b\rangle \end{pmatrix} \\
&= \begin{pmatrix} \langle e| |a\rangle & \langle e| |b\rangle & \langle e| |e\rangle \\ \langle a| |a\rangle & \langle a| |b\rangle & \langle a| |e\rangle \\ \langle b| |a\rangle & \langle b| |b\rangle & \langle b| |e\rangle \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
D(e) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}
\end{aligned} \tag{1.45}$$

1.6.3 Fields

We introduce two operations ($+$ and $\cdot$) on a set of elements F. That satisfy the following properties

- $(F, +)$ is an abelian group (with identity element 0)

- $(F, \cdot)$ is an abelian group (with neutral element 1)

Examples of fields:

- $(\mathbb{R}, +, \cdot)$ the real numbers with the usual definitions of sum and multiplication
- $(\mathbb{C}, +, \cdot)$ the complex numbers with the usual definitions of sum and multiplication

1.6.4 Vector spaces

We define a vector space over a field F as a set V with the following operations

$+$ given $v, w \in V$ then $v + w = z \in V$

$\cdot$ given $\lambda \in F$ and $v \in V$ then $\lambda \cdot v \in V$

with the following properties. Given $v, w, z \in V$ and $\alpha, \beta \in F$ then the following properties are fulfiled

(V1) $(v + w) + z = v + (w + z)$

(V2) Neutral element: $\exists 0 \in V$ sich that $v + 0 = 0 + v = v$

(V3) Incense element: $\forall v \in V$, $\exists - v$ such that $v + (-v) = 0$

(V4) V is abelian under the $+$ operation, $v + w = w + v$.

V(5) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

V(6) $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$

V(7) $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \ \cdot v$

V(8) $1 \cdot v = v$

Let us now end up with some examples of vector spaces

EXAMPLE 1.3

- Consider the set of $n-$tuples, i.e., $\mathbb{R}^n$, where we can define elements such as $P = (p_1, p_2, \dots p_n)$
- the set of $n-$tuples with complex numbers $\mathbb{C}^n$, where we can define elements such as $P = (p_1, p_2, \dots p_n)$ where $p_i \in \mathbb{C}$.
- Set of all complex polynomials of degree $\leq n$, $P^n(t)$, where elements are of the form $a_0 + a_1 t + a_2 t^2 + \dots$
- Continuous functions (real valued) in the interval $[a, b]$, $C_{\mathbb{R}}[a, b]$
- Complex valued functions of a single real variable which are square-integrable in the interval $[a, b]$, which we refer to as ${\mathbb{L}_{\mathbb{C}}}^2[a, b]$.

PROBLEMS

1.1 First problem

$$x^i \tag{1.46}$$

First solution

1.7 Linear Dependence, Basis and dimension of a group

Consider a vector space V with elements $\vec{v}$ and $\vec{w}$ defined over a field F which can be $\mathbb{R}$ or $\mathbb{C}$. A finite set of vectors $\vec{v}_i$ with $i = 1, 2, 3 \ldots k$ is linearly dependent if there exists a linear combination $\sum_{i=1}^{k} \alpha_i \vec{v}_i = 0$ (where α_i are elements of F) with at least one $\alpha_i \neq 0$.

If the only way to satisfy $\sum_{i=1}^{k} \alpha_i \vec{v}_i = 0$ is with all $\alpha_i = 0$ then the set of vectors are said to be linearly independent.

EXAMPLE 1.4

Consider the vector space $\mathbb{C}^3$ defined over $\mathbb{C}$. We can check explicity that the set of vectors

$$\vec{v}_1 = \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}, \tag{1.47}$$

are linearly dependent, as we can construct $i\vec{v}_1 + (-1)\vec{v}_2 + (1/k)\vec{v}_3 = 0$

A vector space is said to be $n-$dimensional if there exists a subset of n linearly independent vectors but there is no subset with $n + 1$ linearly independent vectors.

The *dimension* of a vector space is the maximum number of linearly independent vectors. For example, $\mathbb{R}^3$ over $\mathbb{R}$ is $3-$dimensional.

1.7.1 Basis of a vector space

A subset of n linearly independent vectors $\{\vec{e}_i\}$ is a basis iff they are linearly independent and every vector $\vec{p} \in V$ can be constructed as a linear combination of the elements of the basis $\vec{p} = \vec{e}_i p^i$.

If a vector space is n dimensional then the basis consists of n linearly independent vectors.

EXAMPLE 1.5

Consider $\mathbb{C}^3$ as a vector space over $\mathbb{R}$. We can define the following basis

$$\begin{aligned} \vec{v}_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\ \vec{v}_4 &= \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_5 = \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix}, \quad \vec{v}_6 = \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix}. \end{aligned} \tag{1.48}$$

Notice that $\mathbb{C}^3$ over $\mathbb{R}$ is 6 dimensional whereas $\mathbb{C}^3$ over $\mathbb{C}$ is 3 dimensional (in particular, $\vec{v}_1$ and $\vec{v}_4$ are linearly independent because $i \notin \mathbb{R}$).

1.7.2 Scalar product and orthonormal basis

We can define a *scalar product* (or inner product). Consider $\vec{p}, \vec{q}, \vec{r} \in V$ and $\alpha, \beta \in F$. The scalar product assigns an ordered pair of vectors $(\vec{p}, \vec{q})$ an element in F that satisfies

- $(\vec{p}, \alpha\vec{q} + \beta\vec{r}) = \alpha(\vec{p}, \vec{q}) + \beta(\vec{p}, \vec{r})$
- $(\vec{p}, \vec{q}) = (\vec{q}, \vec{p})^*$
- $(\vec{p}, \vec{p}) \geq 0$ with $(\vec{p}, \vec{p}) = 0$ iff $\vec{p} = 0$

Note that as a consequence of these properties, the scalar product is antilinear in the first element: $(\alpha\vec{q} + \beta\vec{r}, \vec{p}) = \alpha^*(\vec{q}, \vec{p}) + \beta^*(\vec{r}, \vec{p})$.

In what follows, we are going to concentrate on Real Vector Spaces. The scalar product can be defined in terms of the elements of the basis $B = \{\vec{e}_i\}$. Consider two vectors $\vec{x} = \vec{e}_i x^i$ and $\vec{y} = \vec{e}_i y^i$.

$$(\vec{x}, \vec{y}) = (\vec{e}_i, \vec{e}_j)x^i y^j = g_{ij} x^i y^j \tag{1.49}$$

where the metric tensor are n^2 numbers, defined as $(\vec{e}_i, \vec{e}_j) = g_{ij}$. (notice that there would be complex conjugation in the case of complex vector spaces).

1.7.3 Orthonormal basis

Orthonormal basis satisfy $g_{ij} = \delta_{ij}$.

- $g_{ij} = g_{ji}$ in a real vector space ($G = G^T$)
- $\det G \neq 0$

We can now define the concept of norm

$$||\vec{v}|| = \sqrt{(\vec{v}, \vec{v})} \tag{1.50}$$

which satisfies the following properties

- positivity $||\alpha\vec{v}|| = |\alpha|\ ||\vec{v}|| \geq 0$
- symmetry $||\vec{v} - \vec{w}|| = ||\vec{w} - \vec{v}||$
- triangular inequality $||\vec{v} + \vec{w}|| \leq ||\vec{w}|| + ||\vec{v}||$

1.7.4 Gram Schmidt orthogonalisation

Assume that we have a set of n linearly independent vectors $\{\vec{v}_i\}$ in an n dimensional vector space. The Gram Schmidt method allows us to rotate these vectors to form an orthogonal basis (and ultimately, by dividing by their norm, to construct an orthonormal basis).

We start constructing the vector of the orthogonal basis $\{\vec{u}_i\}$ as follows

$$\begin{aligned}
\vec{u}_1 &= \vec{v}_1 \\
\vec{u}_2 &= \vec{v}_2 - \mathrm{proj}_{\vec{u}_1}\vec{v}_2 \\
\vec{u}_3 &= \vec{v}_3 - \mathrm{proj}_{\vec{u}_1}\vec{v}_3 - \mathrm{proj}_{\vec{u}_2}\vec{v}_3 \\
\ldots &
\end{aligned} \tag{1.51}$$

where $\text{proj}_{\vec{u}_i}\vec{v}_j$ is the projection of vector $\vec{v}_j$ onto the vector $\vec{u}_i$, which can be computed as follows

$$\text{proj}_{\vec{u}_i}\vec{v}_j = \frac{(\vec{v}_j, \vec{u}_i)}{(\vec{u}_i, \vec{u}_i)}\vec{u}_i \tag{1.52}$$

This leads to an orthogonal basis, which obviously can be transformed into an orthonormal basis $\{\vec{e}_i\}$ simply by dividing each vector by its norm.

$$\vec{e}_i = \frac{\vec{u}_i}{||\vec{u}_i||} \tag{1.53}$$

1.7.5 Isomorphisms

Assume that we have two vector spaces V_1 and V_2. If we can construct a liner map L that relates the elements between the two vector spaces then the vector spaces are said to be isomorphic.

$$\begin{array}{ccc} V_1 & \longrightarrow & V_2 \\ \alpha\vec{p}_1 + \beta\vec{p}_2 & \longrightarrow & L(\alpha\vec{p}_1 + \beta\vec{p}_2) = \alpha L(\vec{p}_1) + \beta L(\vec{p}_2) \end{array} \tag{1.54}$$

The map has to be bijective, essentially meaning that we are mapping all the objects in V_1 and all the objects on V_2 and viceversa.

EXAMPLE 1.6

The complex vector space $\mathbb{C}^n$ is isomorphic to the complex vector space of polynomials of order $n-1$ with complex coefficients $\mathcal{P}^{n-1}$.

We only need to find a bijective mapping between the elements of both vector spaces (there is not just one!). For example, given a vector $\vec{p}$ in $\mathbb{C}^n$:

$$\vec{p} = (p^1, p^2, \ldots, p^n)\ , \tag{1.55}$$

I can choose a linear map L such that

$$L(\vec{p}) = p^1 + p^2 x + \ldots + p^n x^{n-1} \tag{1.56}$$

which is a vector of $\mathcal{P}^{n-1}$. We can easily convince ourselves that the map is bijective. The mapping is not unique, of course!

All n dimensional complex vector spaces are isomorphic to $\mathbb{C}^n$ (all n dimensional real vector spaces are isomorphic to $\mathbb{R}^n$).

1.7.6 Tensor product

Consider two vector spaces: V (which is n-dimensional) and W (which is m-dimensional). I can find a basis in each of these vector spaces:

$$\begin{array}{ll} V : E = \{\vec{e}_i\} & i = 1\ldots n \\ V : G = \{\vec{g}_j\} & j = 1\ldots m \end{array}$$

I can now define a new vector space X, which we will call the *tensor* product of V and W, in the following way: $X = V \otimes W$.

A basis in X will be made out of pairs of elements from each vector space

$$E \otimes G = \{\vec{e}_i \otimes \vec{g}_j\} \tag{1.57}$$

where indices i and j can run over different dimensions (n and m, respectively)

The vectors in X can thus be understood as tensor products of vectors $\vec{v}$ and $\vec{w}$ in V and W, respectively, and expanded in terms of these pairs

$$\vec{x} = \vec{v} \otimes \vec{w} = \vec{e}_i \otimes \vec{g}_j v^i w^j \tag{1.58}$$

Notice that we can simplify to the case where $V = W = \mathbb{R}^n$

PROBLEMS

1.1 First problem

$$x^i \tag{1.59}$$

First solution

1.8 Linear Operators

A linear operator A maps the vector $\vec{x}$ from a given domain in vector space X into its image $\vec{y} = A(\vec{x})$ of a target (or codomain) in vector space Y.

$$\begin{array}{ccc} X & \xrightarrow{A} & Y \\ \vec{x} & \longrightarrow & \vec{y} = A(\vec{x}) \end{array} \tag{1.60}$$

We will often consider the case $X = Y$ (for example, operators that act as $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ or $\mathbb{C}^n \longrightarrow \mathbb{C}^n$). This map is linear and satisfies

$$\begin{aligned} A(\vec{x}_1 + \vec{x}_2) &= A(\vec{x}_1) + A(\vec{x}_2) \qquad \forall \vec{x}_1, \vec{x}_2 \in X \\ A(\alpha \vec{x}_1) &= \alpha A(\vec{x}) \qquad \forall \vec{x} \in X, \forall \alpha \in F \end{aligned}$$

Notice that X and Y can in general have different dimensionality and we can choose different basis on each.

EXAMPLE 1.7

Identity operator

$$O(\vec{x}) = \vec{x} \qquad \forall \vec{x} \in X \tag{1.61}$$

Null operator

$$E(\vec{x}) = \vec{0}_Y \qquad \forall \vec{x} \in X \tag{1.62}$$

1.8.1 Matrix associated to a linear operator

Every linear operator A in $\mathbb{C}^n$ (or $\mathbb{R}^n$) can be represented by an $n \times n$ complex (real) matrix.

Consider the linear operator A, such that

$$\begin{array}{ccc} X & \xrightarrow{A} & Y \\ \vec{x} & \longrightarrow & \vec{y} = A(\vec{x}) \end{array} \tag{1.63}$$

where X is an $n-$dimensional vector space with basis $B_X = \vec{e}_i$ and Y is an $m-$dimensional vector space with basis $B_Y = \vec{\epsilon}_j$.

Thus, a vector $\vec{x} \in X$ can be expressed in terms of its components as

$$\vec{x} = \vec{e}_i x^i \tag{1.64}$$

and a vector $\vec{y} \in Y$ can be expressed in the basis B_Y as

$$\vec{y} = \vec{\epsilon}_i y^i \tag{1.65}$$

Thus, we can do the same for the result of the linear operator A acting on $\vec{x}$

$$A(\vec{x}) = A\left(\vec{e}_j x^j\right) = A\left(\vec{e}_j\right) x^j = \vec{\epsilon}_i A^i_j x^j \tag{1.66}$$

Thus $A(\vec{x})$ has components $A^i_j x^j$ in the basis B_Y with $i = 1, \ldots m$ and $j = 1, \ldots n$. Notice that we can now write this as a matrix equation.

$$\vec{y} = A\vec{x} \tag{1.67}$$

Notice that the matrix A is $m \times n$, it is not a square matrix unless X and Y have the same dimension. Also, the matrix representation depends on the choice of basis B_X and B_Y.

As mentioned above, we will often consider applications within the same vector space. In that case, we would identify $X = Y$, $B_X = B_Y$ and the matrix A would be square $n \times n$.

1.8.2 Change of basis

For simplicity, we are going to consider here the case where the domain and codomain coincide $X = Y$, and we can choose the same original basis $B = B_X = B_Y = \vec{e_i}$. Consider an operator A, which in basis B has the components A^i_j. Consider now the change of basis, defined by

$$\begin{aligned} \vec{e}\,'_j &= \vec{e}_i S^i_j \\ x^j &= S^j_i x\,'^i \\ x'^j &= (S^{-1})^j_i x^{\,i} = T^j_i x^{\,i} \end{aligned} \tag{1.68}$$

Then, given $\vec{y} = A\vec{x}$, in the old basis we have $y^j = A^j_i x^i$ and in the new basis $y\,'^j = A'^j_i x\,'^i$. Starting from the first expression and transforming the coordinates x^i and y^j, we have

$$\begin{aligned} y^j &= A^j_i x^i \\ S^j_k y\,'^k &= A^j_i S^i_l x\,'^l \\ y\,'^k &= (S^-1)^k_j A^j_i S^i_l x\,'^l \end{aligned} \tag{1.69}$$

which allows us to identify $A'^k_l = (S^-1)^k_j A^j_i S^i_l$

$$A' = S^{-1} A S \tag{1.70}$$

or, equivalently,

$$A' = T A T^{-1} \tag{1.71}$$

1.8.3 Adjoint operators

Given a linear operator A, we define the adjoint operator $A^\dagger$ in terms of the scalar product as follows

$$(\vec{y}, A(\vec{x})) = (A^\dagger(\vec{y}), \vec{x}) \tag{1.72}$$

For all linear operators, A, there exists an adjoint operator $A^\dagger$, which is linear and unique. This can be shown as follows. Notice that, given a vector $\vec{x}$ we can express it as

$$\vec{x} = \vec{e}_i x^i = \vec{e}_i(\vec{e_i}, \vec{x}) \tag{1.73}$$

Thus, we can use the same expression for $A^\dagger(\vec{y})$ as follows

$$\begin{aligned} A^\dagger(\vec{y}) &= \vec{e}_i(\vec{e_i}, A^\dagger(\vec{y})) \\ &= \vec{e}_i\overline{(A^\dagger(\vec{y}), \vec{e_i})} \\ &= \vec{e}_i\overline{(\vec{y}, A^\dagger(\vec{e_i}))} \\ &= \vec{e}_i(A(\vec{e_i}), \vec{y}) \end{aligned} \tag{1.74}$$

We can also find a matrix representation for $A^\dagger$ in terms of those of A. Using

$$\begin{aligned} A^\dagger(\vec{e_i}) &= \vec{e}_j(A(\vec{e_i}), \vec{e_j}) \equiv \vec{e}_j B^j_i \\ A^\dagger(\vec{e_i}) &= \vec{e}_j(\vec{e_j}, A(\vec{e_i})) \equiv \vec{e}_j A^i_i \end{aligned} \tag{1.75}$$

then,

$$B^i{}_j = (A(\vec{e_i}), \vec{e_j}) = \overline{(\vec{e_j}, A(\vec{e_i}))} = \overline{A_j^i} \tag{1.76}$$

This is, the matrix $A^\dagger$ is the result of transposing and taking the complex conjugate on A.

$$A^\dagger = \overline{A^T} \tag{1.77}$$

Some properties of the adjoint operator

$$\begin{aligned} (A^\dagger)^\dagger &= A \\ (A+B)^\dagger &= A^\dagger + B^\dagger \\ (AB)^\dagger &= B^\dagger B^\dagger \\ (\alpha A)^\dagger &= \overline{\alpha} A^\dagger \end{aligned} \tag{1.78}$$

1.8.4 Hermitian operators

Hermitian (or self-adjoint) operators satisfy $A^\dagger = A$.

1.8.5 Unitary operators

An operator A defined in a real vector space $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ is **orthogonal** if it preserves the scalar product. This is,

$$(A(\vec{x}), A(\vec{y})) = (\vec{x}, \vec{y}) \tag{1.79}$$

An operator A defined in a complex vector space $\mathbb{C}^n \longrightarrow \mathbb{C}^n$ is **unitary** if it preserves the scalar product. This is,

$$(A(\vec{x}), A(\vec{y})) = (\vec{x}, \vec{y}) \tag{1.80}$$

It can be shown that if A is unitary, then $A^\dagger A = \mathbb{1}$, or equivalently, $A^\dagger = A^{-1}$.
Notice that unitary operators preserve the norm of a vector.

1.8.6 Determinants and inverses

The determinant of a matrix can be expressed in terms of the Levi-Civita tensor as follows

$$\det(A) = \epsilon_{i_1,i_2,\ldots i_n} A_1^{i_1} A_2^{i_2} \ldots A_n^{i_n} \tag{1.81}$$

PROBLEMS

1.1 First problem

$$x^i \tag{1.82}$$

First solution

1.9 Systems of linear equations and inverse of a matrix

Consider a system of n linear equations with n unknowns, $\{x^i\}$.

$$\begin{aligned}
a^1_1 x^1 + a^1_2 x^2 + \dots a^1_n x^n &= y^1 \\
a^2_1 x^1 + a^2_2 x^2 + \dots a^2_n x^n &= y^2 \\
\vdots &\quad \vdots \\
a^n_1 x^1 + a^n_2 x^2 + \dots a^n_n x^n &= y^n
\end{aligned} \tag{1.83}$$

This can be written in matrix form as

$$A\vec{x} = \vec{y} \tag{1.84}$$

The solution implies inverting A, such that

$$\vec{x} = A^{-1}\vec{y} \tag{1.85}$$

with $A^{-1}A = \mathbb{1}$, or $(a^{-1})^i_j a^j_k = \delta^i_k$. The question is, of course, does A^{-1} exist? Or equivalently, is there a solution to this system of equations?

Theorem 1.2 *Let A be a linear operator in $\mathbb{C}^n$. The following statements are all equivalent*

- *A is invertible*
- *A is injective (one-to-one)*
- *A is surjective*
- $\det A \neq 0$

1.10 Eigenvalues and Eigenvectors

If $A(\vec{v}) = \alpha\vec{v}$ for a non-null vector $\vec{v} \neq \vec{0}$, then $\vec{v}$ is an eigenvector of A with eigenvalue $\alpha \in \mathbb{C}$.

If $\vec{v}$ is an eigenvector of A, so is $\beta\vec{v}, \forall\beta \in \mathbb{C}$, with the same eigenvalue. This simply follows from the linearity of A.

Theorem 1.3 $\alpha \in \mathbb{C}$ *is an eigenvalue of A* $\iff \det(A - \alpha\mathbb{1}) = 0$.

Proof:

$\Rightarrow$ Suppose that α is an eigenvalue of A. This implies that there exists a non-vanishing eigenvector, $\exists\vec{v} \neq 0$, such that $A\vec{v} = \alpha\vec{v}$. Thus, $(A - \alpha\mathbb{1})\vec{v} = 0$. However, we also have $(A - \alpha\mathbb{1})\vec{0} = 0$. There are therefore two different vactors $\vec{v} \neq \vec{0}$ with the same image under $(A - \alpha\mathbb{1})$, which necessarily means that $(A - \alpha\mathbb{1})$ is not injective and therefore $\det(A - \alpha\mathbb{1}) = 0$

$\Leftarrow$ If $\det(A - \alpha\mathbb{1}) = 0$, then it is not injective. This means that there are at least two different vectors $\vec{v} \neq \vec{y}$ that satisfy $(A - \alpha\mathbb{1})\vec{v} = (A - \alpha\mathbb{1})\vec{y}$. Thus, choosing $\vec{z} = \vec{v} - \vec{y}$, we have $(A - \alpha\mathbb{1})\vec{z} = 0$. Thus $A\vec{z} = \alpha\vec{z}$, and $\vec{z}$ is an eigenvector of A with eigenvalue α.

■

1.10.1 Computing Eigenvalues

The expression $\det(A - \alpha\mathbb{1}) = 0$ is a polynomial in α called **characteristic polynomial**. To compute the eigenvalues of a linear operator A one can solve the characteristic polynomial. This is a polynomial equation of degree n which therefore has n complex solutions

- if A is real (in a real vector space), and not all the solutions of the characteristic polynomial are real, then A is not diagonalisable.
- if A is complex, there are always n complex solutions α_i with $i = 1, \ldots, n$ (n eigenvalues). The corresponding eigenvectors are found solving $A\vec{v}_i = \alpha_i \vec{v}_i$.

A certain eigenvalue α_i may appear more than once, for example j times. We say that the eigenvalue os $j-$fold degenerate.

We can choose the n eigenvectors $\{\vec{x}_i\}$ as a new basis. Then, for a given vector $\vec{p}$,

$$A(\vec{p}) = A(\vec{v}_i p^i) = A(\vec{v}_i)p^i = \alpha_i \vec{v}_i p^i \tag{1.86}$$

The matrix representing the operator A in the basis $\{\vec{x}_i\}$ is therefore diagonal, with the eigenvalues α_i along the diagonal elements.

If T is the coordinate transformation from the old (i.e., standard) to the new (eigenvectors) basis, such that $\vec{x}\,' = T\vec{x}$, then the operator is transformed as $A_T = TAT^{-1}$ and is diagonal.

The new basis $\{\vec{v}_i\}$ is in general not an orthonormal basis (i.e., T is not necessarily unitary). However, if A is either unitary ($A^\dagger = A^{-1}$) or Hermitian ($A^\dagger = A$), then the new basis is orthonormal (and T is unitary).

The characteristic polynomial is independent of the basis

$$P(\alpha) = \det(A - \alpha\mathbb{1}) = \det(A' - \alpha\mathbb{1}) \tag{1.87}$$

This is easy to prove, using $A' = T^{-1}AT$

$$\begin{aligned} P'(\alpha) &= \det(A' - \alpha\mathbb{1}) \\ &= \det(T^{-1}AT - \alpha\mathbb{1}) \\ &= \det(T^{-1}AT - \alpha T^{-1}T) \\ &= \det(T^{-1}(A - \alpha\mathbb{1})T) \\ &= \det(T^{-1})\det(A - \alpha\mathbb{1})\det(T) \\ &= \det(A - \alpha\mathbb{1}) \end{aligned} \tag{1.88}$$

1.11 The Spectral Theorem

Lemma: the eigenvalues of a unitary operator are all of the form $e^{i\varphi}$. Eigenvectors to different eigenvalues are orthogonal.

Let $\vec{v}$ be an eigenvector of U with eigenvalue α

$$(\vec{v}, \vec{v}) = (U(\vec{v}), U(\vec{v})) = |\alpha|^2 (\vec{v}, \vec{v}) \Rightarrow \alpha = e^{i\varphi} \tag{1.89}$$

Thus,

$$(\vec{v_1}, \vec{v_2}) = (U(\vec{v_1}), U(\vec{v_2})) = \alpha_1^* \alpha_2 (\vec{v_1}, \vec{v_2}) = \frac{\alpha_2}{\alpha_1}(\vec{v_1}, \vec{v_2}) \tag{1.90}$$

since $\alpha_1 \neq \alpha_2$ then $\frac{\alpha_2}{\alpha_1} \neq 1$ and then $(\vec{v_1}, \vec{v_2}) = 0$.

Lemma: the eigenvalues of a Hermitian operator are real. The eigenvectors to different eigenvalues are orthogonal.

With these premises, the spectral theorem states the following.

If A is unitary or Hermitian, we can find an orthonormal basis $\{\vec{a}_i\}$ where all the basis vectors are eigenvectors of A. The matrix representing the operator in this basis is diagonal (and the entries are the eigenvalues).

Thus, for all $\vec{p}$ we have $\vec{p} = (\vec{a}_i, \vec{p})\vec{a}_i$ and $A(\vec{a}_i) = \alpha_i \vec{a}_i$ and $A(\vec{p}) = (\vec{a}_i, \vec{p})\alpha_i \vec{a}_i$

PROBLEMS

1.1 First problem

$$x^i \tag{1.91}$$

First solution

1.12 Infinite Vector Spaces

So far we have considered the case of vector spaces with finite dimensionality, in which a vector could be expressed in terms of a linear combination of the elements of the $n-$dimensional basis

$$\vec{p} \in V \longrightarrow \vec{p} = \vec{e}_i p^i \tag{1.92}$$

This seems to imply that I can do the same if the dimension of the vector space is infinite, if I find an infinite set of linearly independent vectors $\{\vec{e}_i\}$, $i = 1, 2....N$ with $N \to \infty$.

$$\vec{p} \in H \longrightarrow \vec{p} = \vec{e}_i p^i \tag{1.93}$$

Given that this is an infinite series, there are two questions that are not completely trivial: Does the series converge? Does it converge to an element of H?

We need to define a concept of distance to test convergence. Thus, Hilbert spaces are *metric spaces*.

1.13 Hilbert Spaces

We will define a Hilbert Space, H, as an Euclidean (where we include the concept of distance or "norm") vector space which can have infinite dimensionality.

H is a vector space (typically with dimension $n = \infty$) where we define a scalar product.

$$\forall f,\, g \in H \longrightarrow (f, g) \in F \tag{1.94}$$

which satisfies the following properties

- $(f, g) = \overline{(g, f)}$
- $(f, \alpha g) = \alpha(f, g)$
  $(\alpha f, g) = \alpha^*(f, g)$
- $(f, g + h) = (f, g) + (f, h)$
- $(f, f) > 0, \forall f \neq 0$

We define the norm of f as the (positively defined) square root of the scalar product

$$||f|| = \sqrt{(f, f)} \tag{1.95}$$

The scalar product satisfies the Schwartz inequality:

$$|(f, g)| \leq ||f||\, ||g|| \tag{1.96}$$

Proof: If $g = 0$ the identity is trivially true.
If $g \neq 0$ then, we consider

$$(f + \alpha g, f + \alpha g) = ||f||^2 + \overline{\alpha}(g, f) + \alpha(f, g) + |\alpha|^2 ||g||^2 \tag{1.97}$$

Taking $\alpha = -(g, f)/||g||^2$ in the expression above,

$$= ||f||^2 - \frac{|(g,f)|^2}{||g||^2} - \frac{|(g,f)|^2}{||g||^2} + \frac{|(g,f)|^2}{||g||^2} \geq 0 \tag{1.98}$$

thus,

$$|(f,g)| \leq ||f||\,||g|| \tag{1.99}$$

■

Triangular inequality

$$||f+g|| \leq ||f|| + ||g|| \tag{1.100}$$

Hilbert spaces are metric spaces, we can define the concept of distance between two elements. We do it using the previously defined norm as follows.

$$d(f,g) \equiv ||f-g|| = \sqrt{(f-g, f-g)} \tag{1.101}$$

which fulfils the three properties for a distance in a metric space

1. $d(f,g) = d(g,f)$
2. $d(f,g) > 0$ and it only vanishes for $f = g$
3. $d(f,g) \leq d(f,h) + d(h,g)$

1.13.1 Convergence of sequences

When working on a Hilbert space, the dimensionality can be infinity and this raises important questions about the convergence of sequences. In particular, we are forced to work with an infinite basis, which in itself is an interesting concept. It means that one can choose n linearly independent vectors, with $n \to \infty$. If such a basis exists, $\{\vec{e}_i\}$, this means that a given vector $\vec{p}$ can be expressed as

$$\vec{p} = \sum_{i=1}^{\infty} \vec{e}_i p^i \tag{1.102}$$

Strong convergence criterion : If we have a sequence $\left\{\vec{f}_1, \vec{f}_2, \dots \vec{f}_n\right\}$, with $\vec{f}_k = \sum_{i=1}^{k} \vec{e}_i p^i \in H$ this sequence converges strongly to $\vec{f}$ if $||\vec{f} - \vec{f}_n|| \to 0$ when $N \to \infty$.

Weak convergence criterion : The sequence $\left\{\vec{f}_i\right\}$ converges weakly to $\vec{f}$ if the scalar product converges $(\vec{g}, \vec{f}_n) \to (\vec{g}, \vec{f})$ when $N \to \infty$.

Strong convergence implies weak convergence but not otherwise.

A Hilbert space is *complete* if all Cauchy sequencies converge in H. Cauchy sequences imply $||\vec{f}_n - \vec{f}_m|| \to 0$ when $n, m \to \infty$. In other words, H contains the limits of all sequencies.

1.13.2 Examples of Hilbert Spaces

- Consider an infinite column vector, with elements a^1, a^2...., we can define the space $l^2(\mathbb{C}) = \left\{ \vec{f} = \left\{ a^1, a^2, a^3, \ldots \right\}, a_i \in \mathbb{C} / \sum_i^\infty |a_i|^2 < \infty \right\}$

  All complex Hilbert spaces are isomorphic to $l^2(\mathbb{C})$

  The scalar product can be defined in the usual way $(f, g) = \sum_i^\infty |f_i g^i|$ and the norm is $||f|| = \sqrt{\sum_i^\infty |f_i|^2} < \infty$

- $L^2[a, b] = \left\{ f(x) / \int_a^b |f(x)|^2 dx < \infty \right\}$ where $f(x)$ are complex functions of a real variable.

  The scalar product is defined as $(f, g) = \int_a^b f^*(x) g(x) dx$

1.13.3 Basis in a Hilbert space

It is a complete set in a vector space, meaning that it contains all the linear combinations and all elements in sequencies. In other words, it is a set of objects that allow to obtain all the elements in the Hilbert space.

A separable Hilbert space is separable if there exists a sequence $\left\{ \vec{f_n} \right\}$ with a generic term $\vec{f_n} = c_1 \vec{f_1} + c_2 \vec{f_2} + \ldots$ such that $\vec{f_n} \rightarrow$
$vecf$ when $n \rightarrow \infty$. Then all elements in H can be expressed as a linera combination

$$\vec{g} = \sum_{i=1}^{\infty} \vec{f_i} c^i \tag{1.103}$$

I can now define an orthonormal basis in such a way that the scalar product of the elements of the basis are the Kronecker delta function $(\vec{f_i}, \vec{f_j}) = \delta_{ij}$. For two general vectors, the scalar product is also expressed as a product of the elements of the basis

$$(\vec{p}, \vec{q}) = \left(\sum_{i=1}^{\infty} \vec{f_i} p^i, \sum_{i=1}^{\infty} \vec{f_i} q^i \right) = \delta_{ij} p^i q^j = p_i q^i \tag{1.104}$$

1.14 Formulation of Quantum Mechanics

We are now ready to address the formulation of Quantum Mechanics.

- Rule 1: The states in a quantum mechanical system are represented by unit vectors $|v\rangle$ in a Hilbert Space.

$$||v|| = \langle v|v\rangle^{1/2} = 1$$

- Rule 2: Physical (measurable) quantities (which we will refer to as observables) are represented by Hermitian operators, A. Whenever a measurement of this quantity is made, the result is one of the eigenvalues $\alpha_m \in \mathbb{R}$ of A.

If the physical state $|\psi\rangle$ is an eigenvector of A then $A|m\rangle = \alpha_m |m\rangle$. If the physical state is not an eigenvector then it can be expressed in the basis of eigenvector $|m\rangle$.

The probability of measuring α_m is given by the projection of the physical state onto the corresponding eigenvector.

$$\mathrm{prob}(\alpha_m)_\psi = \sum_{\alpha_m} |\langle m|\Psi\rangle|^2 \tag{1.105}$$

where we su to all eigenvectors with eigenvalue α_m.

For example,

$$|\Psi\rangle = \Psi_m |m\rangle \tag{1.106}$$

The probability of measuring *any* eigenvalue

$$\sum_m |\langle m|\Psi\rangle|^2 = \sum_m \langle\Psi|m\rangle\langle m|\Psi\rangle = \langle\Psi|\Psi\rangle = 1 \tag{1.107}$$

since $|m\rangle\langle m| = \mathbb{1}$.

1.14.1 Expectation value of an observable

Given the observable A, the average result of a measurement of a given state $|\Psi\rangle$ is given by the expectation value of A as follows

$$\begin{aligned}
\langle A\rangle_\Psi &= \sum_i \alpha_i \mathrm{prob}(\alpha_i)_\Psi \\
&= \sum_i \alpha_i |\langle i|\Psi\rangle|^2 \\
&= \sum_i \alpha_i \langle\Psi|i\rangle\langle i|\Psi\rangle \\
&= \sum_i \langle\Psi| A |i\rangle\langle i|\Psi\rangle \\
&= \langle\Psi| A |\Psi\rangle
\end{aligned} \tag{1.108}$$

1.14.2 Example: spin 1/2 particles

- What is the corresponding Hilbert space?

  We will choose physical states as vectors in $\mathbb{C}^2$ (spinors).

- Choose a basis

  We will represent spins up and down as follows

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad ; \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.109}$$

  An arbitrary state is a linear combination of the form $|\Psi\rangle = v_1 |\uparrow\rangle + v_2 |\downarrow\rangle$ with $v_1, v_2 \in \mathbb{C}$ and a normalisation as $||\Psi|| = 1$ which implies $|v_1|^2 + |v_2|^2 = 1$.

- What are the physical quantities that we want to describe and define operators on the Hilbert space

  We will represent rotations and measure the components of the spin along each axis. Thus s_x, s_y, s_z are the operators that represent measurements along each of the axis. In matrix form

$$\begin{aligned} s_z &= \frac{1}{2}\sigma_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ s_x &= \frac{1}{2}\sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ s_y &= \frac{1}{2}\sigma_z = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{aligned} \tag{1.110}$$

  It can be checked explicitly that $s_z |\uparrow\rangle = \frac{1}{2} |\uparrow\rangle$, and $s_z |\downarrow\rangle = -\frac{1}{2} |\downarrow\rangle$ and also that $\langle s_z \rangle = \frac{1}{2}(|v_1|^2 + |v_2|^2)$

1.14.3 Example: Particle in a 1-dimensional potential well

Consider a particle in a 1-dimensional potential well between $x = [-L/2, L/2]$.

- What is the corresponding Hilbert space?

  We will consider $L^2_{\mathbb{C}}$. States are vectors in this Hilbert space, this is functions $\Psi(x)$ such that $\int_{-L/2}^{L/2} |\Psi|^2 = 1$

- Basis in the Hilbert space

  We will label the elements with an integer index $n = -\infty, \ldots, \infty$ such that

$$|n\rangle = \frac{1}{\sqrt{L}} e^{i\frac{2\pi n}{L}x} \equiv \Psi_n(x) \tag{1.111}$$

- Observables

  Position X such that $X\,|Psi\rangle = x\,|Psi\rangle$

  Momentum P such that $P\,|Psi\rangle = -i\frac{\partial}{\partial x}\,|Psi\rangle$

  Energy $\frac{P^2}{2m} - V(x)$ where $V(x)$ is the potential

  Note that

$$P\,|n\rangle = \frac{2\pi n}{L}\,|n\rangle \tag{1.112}$$

1.14.4 Time evolution

Time evolution is determined by Schroedinger equation

$$-i\hbar\frac{d}{dt}\,|\Psi,t\rangle = H\,|\Psi,t\rangle \tag{1.113}$$

where $|\Psi,t\rangle$ represents the physical state at time t. This can be expressed as the evolved state from $t=0$ as

$$|\Psi,t\rangle = U(t,t_0)\,|\Psi,t_0\rangle \tag{1.114}$$

where $U(t,t_0) = e^{-iH(t-t_0)}$ is the time evolution operator. If we expand the state on a basis of eigenvectors of the Hamiltonian

$$H\,|m\rangle = E_m\,|m\rangle \tag{1.115}$$

then

$$|\Psi,t\rangle = \sum_m e^{-iE_m(t-t_0)}\,|m\rangle\,\langle m|\,|\Psi,t\rangle \tag{1.116}$$

REFERENCES

[1] Mathematical Methods for Physicists, G. B. Arken and H. J. Webber