Physics 307

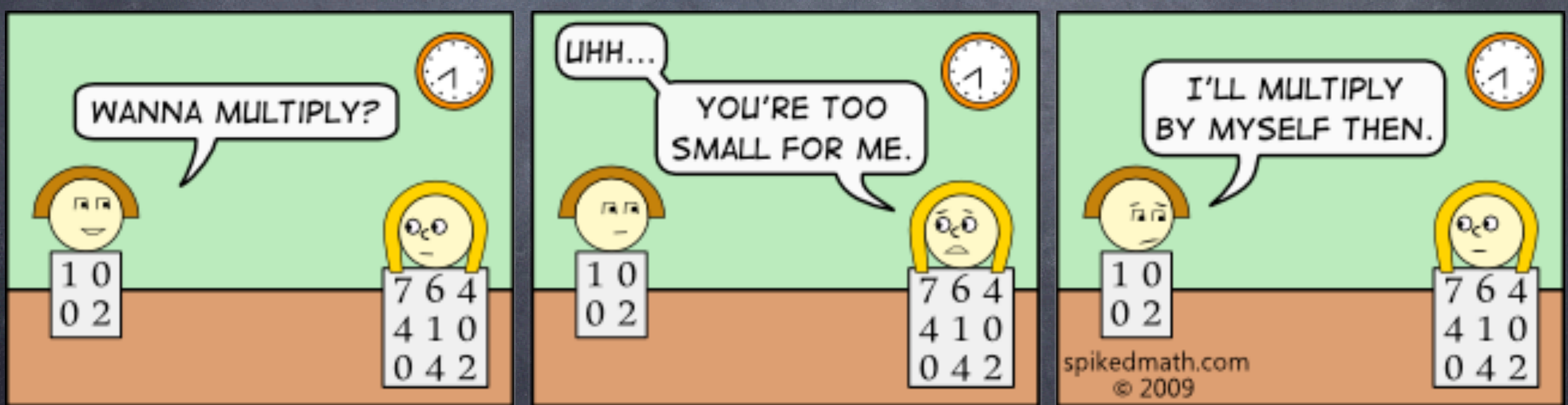
Mathematical Physics

Luis Anchordoqui

Elements of Linear Algebra

1.1 Linear Spaces

1.2 Matrices and Linear Transformations



LINEAR SPACES

Definition 2.1.1.

A *field* is a set F together with two operations $+$ and $\cdot$ for which all axioms below hold $\forall\ \lambda,\ \mu,\ \nu \in F$:

$(i) - closure-$ sum $\lambda + \mu$ and product $\lambda \cdot \mu$ again belong to F

$(ii) - associative\ law - \lambda + (\mu + \nu) = (\lambda + \mu) + \nu$ & $\lambda \cdot (\mu \cdot \nu) = (\lambda \cdot \mu) \cdot \nu$

$(iii) - commutative\ law - \lambda + \nu = \nu + \lambda$ & $\lambda \cdot \mu = \mu \cdot \lambda$

$(iv) - distributive\ laws - \lambda \cdot (\mu + \nu) = \lambda \cdot \mu + \lambda \cdot \nu$
and $(\lambda + \mu) \cdot \nu = \lambda \cdot \nu + \mu \cdot \nu$

$(v) - existence\ of\ an\ additive\ identity-$ there exists an element $0 \in F$ for which $\lambda + 0 = \lambda$

$(vi) - existence\ of\ a\ multiplicative\ identity-$ there exists an element $1 \in F$ with $1 \neq 0$ for which $1 \cdot \lambda = \lambda$

$(vii) - existence\ of\ additive\ inverse-$ to every $\lambda \in F$ there corresponds an additive inverse $-\lambda$ such that $-\lambda + \lambda = 0$

$(viii) - existence\ of\ multiplicative\ inverse-$ to every $\lambda \in F$ there corresponds a multiplicative inverse λ^{-1} such that $\lambda^{-1} \cdot \lambda = 1$

Example 2.1.1.

Underlying every linear space is a field F

examples are $\mathbb{R}$ and $\mathbb{C}$

Definition 2.1.2.

A linear space V is a collection of objects with a (vector) addition and scalar multiplication defined which is closed under both operations

Such a vector space satisfies following axioms:

- commutative law of vector addition

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}, \ \forall \mathbf{x}, \mathbf{y} \in V$$

- associative law of vector addition

$$\mathbf{x} + (\mathbf{y} + \mathbf{w}) = (\mathbf{x} + \mathbf{y}) + \mathbf{w}, \ \forall \mathbf{x}, \mathbf{y}, \mathbf{w} \in V$$

- There exists a zero vector $\mathbf{0}$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}, \forall \mathbf{x} \in V$

- To every element $\mathbf{x} \in V$

  there corresponds an inverse element $-\mathbf{x}$

  such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$

- associative law of scalar multiplication

$$(\lambda\,\mu)\,\mathbf{x} = \lambda\,(\mu\,\mathbf{x}),\ \forall\,\mathbf{x} \in V \text{ and } \lambda, \mu \in F$$

- distributive laws of scalar multiplication

$$(\lambda + \mu)\,\mathbf{x} = \lambda\,\mathbf{x} + \mu\,\mathbf{x},\ \forall\,\mathbf{x} \in V \text{ and } \lambda, \mu \in F$$

$$\lambda\,(\mathbf{x} + \mathbf{y}) = \lambda\,\mathbf{x} + \lambda\,\mathbf{y},\ \forall\,\mathbf{x}, \mathbf{y} \in V \text{ and } \lambda \in F$$

- $1 \cdot \mathbf{x} = \mathbf{x},\ \forall\,\mathbf{x} \in V$

Example 2.1.2.

Cartesian space $\mathbb{R}^n$ is prototypical example of real n-dimensional vector space

Let $\mathbf{x} = (x_1, \ldots, x_n)$ be an ordered n tuple of real numbers x_i

to which there corresponds a point $\mathbf{x}$ with these Cartesian

coordinates and a vector $\mathbf{x}$ with these components

We define addition of vectors by component addition

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n) \qquad (2.1.1.)$$

and scalar multiplication by component multiplication

$$\lambda\,\mathbf{x} = (\lambda x_1, \ldots, \lambda x_n) \qquad (2.1.2.)$$

Definition 2.1.3.

Given a vector space V over a field F

a subset W of V is called subspace

if W is vector space over F under operations already defined on V

Corollary 2.1.1

A subset W of a vector space V is a subspace of $V \Leftrightarrow$

$(i)\, W$ is nonempty (ii) if $\mathbf{x}, \mathbf{y} \in W$ ☛ then $\mathbf{x} + \mathbf{y} \in W$

(iii) $\mathbf{x} \in W$ and $\lambda \in F$ ☛ then $\lambda \cdot \mathbf{x} \in W$

After defining notions of vector spaces and subspaces

next step is to identify functions that can be used

to relate one vector space to another

Functions should respect algebraic structure of vector spaces

so we require they preserve addition and scalar multiplication

Definition 2.1.4.

Let V and W be vector spaces over field F

A linear transformation from V to W is a function $T : V \to W$

such that $$T(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda T(\mathbf{x}) + \mu T(\mathbf{y}) \quad (1.1.3.)$$

for all vectors $\mathbf{x}, \mathbf{y} \in V$ and all scalars $\lambda, \mu \in F$

If a linear transformation is one-to-one and onto

it is called vector space isomorphism ☛ or simply isomorphism

Definition 2.1.5.

Let $S = \mathrm{x_1}, \cdots, \mathrm{x_n}$ be a set of vectors in vector space V over field F

Any vector of form $$\mathbf{y} = \sum_{i=1}^{n} \lambda_i \mathrm{x_i} \quad \text{for } \lambda_i \in F$$

is called linear combination of vectors in S

Set S is said to span V if each element of V

can be expressed as linear combination of vectors in S

Definition 2.1.6.

Let $\mathbf{x_1}, \ldots, \mathbf{x_m}$ be m given vectors
and $\lambda_1, \ldots \lambda_m$ an equal number of scalars

We can form a linear combination or sum

$$\lambda_1 \mathbf{x_1} + \cdots + \lambda_k \mathbf{x_k} + \cdots + \lambda_m \mathbf{x_m} \quad (2.1.4.)$$

which is also an element of the vector space

Suppose there exist values $\lambda_1 \ldots \lambda_n$ which are not all zero
such that above vector sum is the zero vector

Then the vectors $\mathbf{x_1}, \ldots, \mathbf{x_m}$ are said to be linearly dependent

Contrarily vectors $\mathbf{x_1}, \ldots, \mathbf{x_m}$ are called linearly independent

if $\lambda_1 \mathbf{x_1} + \cdots + \lambda_k \mathbf{x_k} + \cdots + \lambda_m \mathbf{x_m} = \mathbf{0} \quad (2.1.5.)$

demands scalars λ_k must all be zero

Definition 2.1.7

Dimension of V

☛ maximal number of linearly independent vectors of V

Definition 2.1.8.

Let V be an n dimensional vector space

and $S = \mathbf{x_1}, \ldots, \mathbf{x_n} \subset V$ (2.1.6.)

a linearly independent spanning set for V

➪ S is called a basis of V

Definition 2.1.9.

Let S be a nonempty subset of vector space V

➪ S is a basis for V if and only if each vector in V can be written uniquely as a linear combination of vectors in S

Definition 2.1.10.

An inner product $\langle , \rangle : V \times V \to F$ is a function that takes each ordered pair $(\mathbf{x}, \mathbf{y})$ of elements of V to a number $\langle \mathbf{x}, \mathbf{y} \rangle \in F$ and has following properties:

- conjugate symmetry or Hermiticity $\langle \mathbf{x}, \mathbf{y} \rangle = (\langle \mathbf{y}, \mathbf{x} \rangle)^*$
- linearity in second argument
$$\langle \mathbf{x}, \mathbf{y} + \mathbf{w} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{w} \rangle \text{ and } \langle \mathbf{x}, \lambda\, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$$
- definiteness $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

Corollary 2.1.2.

Conjugate symmetry and linearity in second variable gives
$$\langle \lambda\, \mathbf{x}, \mathbf{y} \rangle = (\langle \mathbf{y}, \lambda \mathbf{x} \rangle)^* = \lambda^* (\langle \mathbf{y}, \mathbf{x} \rangle)^* = \lambda^* (\langle \mathbf{x}, \mathbf{y} \rangle)$$
$$\langle \mathbf{y} + \mathbf{w}, \mathbf{x} \rangle = (\langle \mathbf{x}, \mathbf{y} + \mathbf{w} \rangle)^* = (\langle \mathbf{x}, \mathbf{y} \rangle)^* + (\langle \mathbf{x}, \mathbf{w} \rangle)^* = \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{w}, \mathbf{x} \rangle$$

Remark 2.1.1.

In $\mathbb{R}$ inner product is symmetric whereas in $\mathbb{C}$ is a sesquilinear form (i.e. is linear in one argument and conjugate-linear in other)

Definition 2.1.11.

An inner product $\langle \, , \rangle$ is said to be positive definite $\Leftrightarrow$

for all non-zero x in $V, \langle \mathrm{x}, \mathrm{x} \rangle \geq 0$

A positive definite inner product is usually referred to as genuine inner product

Definition 2.1.12.

An inner product space is a vector space V over field F equipped with an inner product $\langle \, , \rangle : V \times V \to F$

Definition 2.1.13.

Vector space V on F endowed with a positive definite inner product (a.k.a. scalar product) defines Euclidean space $\mathcal{E}$

Example 2.1.3.

For $x, y \in \mathbb{R}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k \qquad (2.1.7.)$$

Example 2.1.4.

For $x, y \in \mathbb{C}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k^* y_k \qquad (2.1.8.)$$

Example 2.1.5.

Let $\mathcal{C}([a,b])$ denote set of continuous functions $x(t)$ defined on closed interval $-\infty < a \leq t \leq b < \infty$

This set is structured as vector space with respect to usual operations of sum of functions and product of functions by numbers whose neutral element is zero function

For $x(t), y(t) \in \mathcal{C}([a,b])$ we can define scalar product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_a^b x^*(t)\; y(t)\, dt, \quad (2.1.9.)$$

which satisfies all necessary axioms

In particular

$$\langle \mathbf{x}, \mathbf{x} \rangle = \int_a^b |x(t)|^2 dt \geq 0 \quad (2.1.10.)$$

and if $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ ☛ $0 = \int_a^b |x(t)|^2\, dt \geq \int_{a_1}^{b_1} |x(t)|^2\, dt \geq 0 \quad (2.1.11.)$

$\forall a \leq a_1 \leq b_1 \leq b$ ➡ $x(t) \equiv 0$

$\mathcal{C}^2([a,b])$ denotes euclidean space of continuous functions on interval $[a,b]$ equipped with scalar product (2.1.9)

Definition 2.1.14.

Axiom of positivity allows one to define a norm or length

For each vector of an euclidean space

$$\|\mathbf{x}\| = +\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \qquad (2.1.12.)$$

In particular $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

Further ☛ if $\lambda \in$ $\quad \|\lambda \mathbf{x}\| = \sqrt{|\lambda|^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda| \|\mathbf{x}\|$ (2.1.13.)

This allows a normalization for any non-zero length vector

Indeed ☛ if $\mathbf{x} \neq \mathbf{0}$ then $\|\mathbf{x}\| > 0$

Thus ☛ we can take $\lambda \in$ such that $|\lambda| = \|x\|^{-1}$ and $\mathbf{y} = \lambda \mathbf{x}$

It follows that $\|\mathbf{y}\| = |\lambda| \|\mathbf{x}\| = 1$

Example 2.1.6.

Length of a vector $\mathbf{x} \in \mathbb{R}^n$ is

$$\|\mathbf{x}\| = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2} \qquad (2.1.14.)$$

Example 2.1.7.

Length of a vector $\mathbf{x} \in \mathcal{C}^2([a,b])$ is

$$\|\mathbf{x}\| = \left\{ \int_a^b |x(t)|^2 \, dt \right\}^{1/2} \qquad (2.1.15.)$$

Definition 2.1.15.

In a real Euclidean space angle between vectors x and y

$$\cos \widehat{xy} = \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \|\mathbf{y}\|} \qquad (2.1.16.)$$

Definition 2.1.16.

Two vectors are orthogonal $\mathrm{x} \perp \mathrm{y}$ if $\langle \mathrm{x}, \mathrm{y} \rangle = 0$

Zero vector is orthogonal to every vector in $\mathcal{E}$

Definition 2.1.17.

In a real Euclidean space

angle between two orthogonal non-zero vectors is $\pi/2$ ↵

i.e. $\cos \widehat{xy} = 0$

Lemma 2.1.1.

If $\{\mathrm{x_1}, \mathrm{x_2}, \cdots, \mathrm{x_k}\}$ is a set of mutually orthogonal non-zero vectors then its elements are linearly independent

Proof.

Assume that vectors are linearly dependent

Then ☛ there exists k numbers λ_i (not all zero) such that

$$\lambda_1 \mathbf{x_1} + \lambda_2 \mathbf{x_2} + \cdots + \lambda_k \mathbf{x_k} = \mathbf{0} \quad (2.1.17.)$$

Further ☛ assume that $\lambda_1 \neq 0$ and consider scalar product of the linear combination (2.1.17) with vector $\mathrm{x_1}$

Since $\mathrm{x_i} \perp \mathrm{x_j}$ for $i \neq j$ ☛ we have

$$\lambda_1 \langle \mathbf{x_1}, \mathbf{x_1} \rangle = \langle \mathbf{x_1}, \mathbf{0} \rangle \quad (2.1.18.)$$

or equivalently

$$\lambda_1 \|\mathbf{x_1}\|^2 = 0 \Rightarrow \mathbf{x_1} = \mathbf{0} \quad (2.1.19.)$$

which contradicts hypothesis

Corollary 2.1.3.

If a sum of mutually orthogonal vectors is $\mathbf{0}$ then each vector must be $\mathbf{0}$

Definition 2.1.18.

A basis $\mathrm{x_1}, \ldots, \mathrm{x_n}$ of V is called orthogonal

if $\langle \mathrm{x_i}, \mathrm{x_j} \rangle = 0$ for all $i \neq j$

☛ basis is called orthonormal

if in addition each vector has unit length

$$\|\mathrm{x_i}\| = 1, \; \forall i = 1, \ldots, n$$

Example 2.1.8.

Simplest example of an orthonormal basis is standard basis

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \qquad \ldots \qquad \mathbf{e_n} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (2.1.20.)$$

Lemma 2.1.2.

If set of vectors $\{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_k}\}$ is orthogonal to $\mathbf{y} \in \mathcal{E}$ then every linear combination of this set of vectors is also orthogonal to $\mathbf{y}$

$$\left\langle \mathbf{y}, \sum_{i=1}^{k} \lambda_i \mathbf{x_i} \right\rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{y}, \mathbf{x_i} \rangle \qquad (2.1.22.)$$

Theorem 2.1.1. (Pythagorean theorem)

If $\mathbf{x} \perp \mathbf{y} \in \mathcal{E}$ then

$$\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \qquad (2.1.23.)$$

In any right triangle ☛ area of square whose side is hypotenuse (side opposite right angle) is equal to sum of areas of squares whose sides are two legs (two sides that meet at a right angle)

Corollary 2.1.4.

If set of vectors $\{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_k}\}$ are mutually orthogonal $\mathbf{x_i} \perp \mathbf{x_j}$ with $i \neq j$ then

$$||\mathbf{x_1} + \cdots + \mathbf{x_k}||^2 = ||\mathbf{x_1}||^2 + \cdots + ||\mathbf{x_k}||^2 \qquad (2.1.24.)$$

Corollary 2.1.5. (Triangle inequality)

For $\mathrm{x}, \mathrm{y} \in \mathcal{E}$ we have

$$\Big| \|\mathrm{x}\| - \|\mathrm{y}\| \Big| \leq \|\mathrm{x} + \mathrm{y}\| \leq \|\mathrm{x}\| + \|\mathrm{y}\| \qquad (2.1.25.)$$

Length of a side of a triangle
does not exceed sum of lengths of other two sides
nor is it less than absolute value of difference of other two sides

Proof.

Consider scalar product

$$\|\mathrm{x} + \mathrm{y}\|^2 = \langle \mathrm{x} + \mathrm{y}, \mathrm{x} + \mathrm{y} \rangle = \|\mathrm{x}\|^2 + 2\Re\mathrm{e}\langle \mathrm{x}, \mathrm{y} \rangle + \|\mathrm{y}\|^2 \qquad (2.1.26.)$$

according to Cauchy-Schwarz inequality

$$|\Re\mathrm{e}\langle \mathrm{x}, \mathrm{y} \rangle| \leq |\langle \mathrm{x}, \mathrm{y} \rangle| \leq \|\mathrm{x}\| \|\mathrm{y}\| \qquad (2.1.27.)$$

therefore

$$(\|\mathrm{x}\| - \|\mathrm{y}\|)^2 \leq \|\mathrm{x} + \mathrm{y}\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|\mathrm{x}\| + \|\mathrm{y}\|)^2 \quad (2.1.28.)$$

Definition 2.1.19.

Let $\mathrm{x} = (x_1, \ldots, x_k, \ldots)$ be an infinite sequence of real numbers such that $\sum_{k=1}^{\infty} x_k^2$ converges

Sequence x defines a point of Hilbert coordinate space $\mathbb{R}^\infty$ with k -th coordinate x_k

It also defines a vector with k-th component x_k which as in $\mathbb{R}^n$ we identify with point

Addition and scalar multiplication are defined analogously to (1.1.1) and (1.1.2)

Norm of Hilbert vector x is Pythagorean expression

$$\|\mathrm{x}\| = \left(\sum_{k=1}^{\infty} x_k^2\right)^{1/2}$$

By hypothesis

this series converges if x is an element of Hilbert space $\mathcal{H} = \mathbb{E}^\infty$

LINEAR OPERATORS ON EUCLIDEAN SPACES

Definition 2.2.0

An operator A on $\mathcal{E}$ is a vector function $A : \mathcal{E} \to \mathcal{E}$

Operator is called linear if

$$A(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha A \mathbf{x} + \beta A \mathbf{y}, \ \forall \mathbf{x}, \mathbf{y} \in \mathcal{E} \text{ and } \forall \alpha, \beta \in \mathbb{C} \,(\text{or}\, \mathbb{R})$$

Definition 2.2.1.

Let $\mathbb{A}$ be an $n \times n$ matrix and $\mathbf{x}$ a vector

➪ the function $T(\mathbf{x}) = \mathbb{A}\mathbf{x}$ is a linear operator

Definition 2.2.2.

A vector $\mathbf{x} \neq \mathbf{0}$ is eigenvector of $\mathbb{A}$ if $\exists \lambda$ satisfying $\mathbb{A}\,\mathbf{x} = \lambda\,\mathbf{x}$

in such a case $(\mathbb{A} - \lambda\,\mathbb{I})\,\mathbf{x} = \mathbf{0}$ ☛ $\mathbb{I}$ is identity matrix

Eigenvalues λ are given by relation $\det\,(\mathbb{A} - \lambda\,\mathbb{I}) = 0$

which has m different roots with $1 \leq m \leq n$

➪ $\det(\mathbb{A} - \lambda\,\mathbb{I})$ is a polynomial of degree n

Eigenvectors associated with eigenvalue λ

are obtained by solving (singular) linear system $(\mathbb{A} - \lambda\,\mathbb{I})\,\mathbf{x} = \mathbf{0}$

Remark 2.2.1.

If x_1 and x_2 are eigenvectors with eigenvalue λ and a, b constants

⇨ $a\mathrm{x}_1 + b\mathrm{x}_2$ is an eigenvector with eigenvalue λ because

$$\mathbb{A}(a\mathrm{x}_1 + b\mathrm{x}_2) = a\mathbb{A}\mathrm{x}_1 + b\mathbb{A}\mathrm{x}_2 = a\lambda\mathrm{x}_1 + b\lambda\mathrm{x}_2 = \lambda(a\mathrm{x}_1 + b\mathrm{x}_2)$$

It is straightforward to show that:

(i) eigenvectors associated to a given eigenvalue form a vector space

(ii) two eigenvectors corresponding to different eigenvalues are lineraly independent

Definition 2.2.3.

A matrix $\mathbb{A}$ is said to be diagonable (or diagonalizable) if the eigenvectors form a base i.e. if any vector v can be written as a linear combination of eigenvectors

A matrix $\mathbb{A}$ is said to be diagonable if there exists n eigenvectors $\mathrm{x}_1, \ldots, \mathrm{x}_\mathrm{n}$ that are linearly independent

In such a case

- ☛ we can form with n eigenvectors an $n \times n$ matrix $\mathbb{U}$ such that k-th column of $\mathbb{U}$ is k-th eigenvector

In this way

- ☛ n relations $\mathbb{A}\mathrm{x_k} = \lambda \mathrm{x_k}$ can be written in a matrix form

$\mathbb{A}\mathbb{U} = \mathbb{U}\mathbb{A}'$ ☛ $\mathbb{A}'$ is a $n \times n$ diagonal matrix such that

$$A'_{ij} = \lambda_i \delta_{ij}$$

The latter can also be written as

$$\mathbb{U}^{-1}\ \mathbb{A}\ \mathbb{U} = \mathbb{A}', \quad \text{or equivalently} \quad \mathbb{A} = \mathbb{U}\ \mathbb{A}'\ \mathbb{U}^{-1}, \quad (2.2.11.)$$

which bind diagonal matrix with original matrix

($\mathbb{U}$ is invertible because eigenvectors are linearly independent)

Definition 2.2.4.

Transformation (2.2.11.) represents a change of base

Note that eigenvalues (and therefore matrix $\mathbb{A}'$) are independent of change of base

if $\mathbb{B} = \mathbb{W}^{-1}\mathbb{A}\mathbb{W}$ with $\mathbb{W}$ an arbitrary (invertible) $n \times n$ matrix

$$\Rightarrow \det\,(\mathbb{B} - \lambda\,\mathbb{I}) = \det\,(\mathbb{W}^{-1}\ \mathbb{A}\ \mathbb{W} - \lambda \mathbb{W}^{-1}\ \mathbb{W}) = \det\,(\mathbb{A} - \lambda\,\mathbb{I}) \quad (2.2.12.)$$

such that it has same eigenvalues

Definition 2.2.5.

If real function $f(x)$ has a Taylor expansion

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \qquad (2.2.13.)$$

matrix function is defined by substituting argument x by $\mathbb{A}$

powers become matrix powers, additions become matrix sums and multiplications become scaling operations

If real series converges for $|x| < r$

corresponding matrix series converges for matrix argument $\mathbb{A}$

if $\|\mathbb{A}\| < r$ for some matrix norm $\|\cdot\|$ which satisfies

$$\|\mathbb{A}\mathbb{B}\| \leq \|\mathbb{A}\| \cdot \|\mathbb{B}\| . \qquad (2.2.14.)$$

It is possible to evaluate an arbitrary matrix function $F(\mathbb{A})$

applying power series definition to decomposition (2.2.11.)

We find that $F(\mathbb{A}) = \mathbb{U} F(\mathbb{A}') \mathbb{U}^{-1}$

with $F(\mathbb{A}')$ given by matrix $[F(A')]_{ij} = F(\lambda_i)\delta_{ij}$

Note that

$$A^n = (UDU^{-1})^n = (UDU^{-1})(UDU^{-1})\cdots(UDU^{-1})$$
$$= UD(U^{-1}U)D(U^{-1}U)D\cdots(U^{-1}U)DU^{-1}$$
$$= UD^nU^{-1}$$

Definition 2.2.6.

A complex square matrix $\mathbb{A}$ is Hermitian if $\mathbb{A} = \mathbb{A}^\dagger$
where $\mathbb{A}^\dagger = (\mathbb{A}^*)^T$ is conjugate transpose of a complex matrix

Remark 2.2.2.

It is easily seen that if $\mathbb{A}$ is Hermitian then:

(i) its eigenvalues are real

(ii) eigenvectors associated to different eingenvalues are orthogonal

(iii) it has a complete set of orthogonal eigenvectors which makes it diagonalizable

Definition 2.2.7.

A partially defined linear operator A on a Hilbert space $\mathcal{H}$
is called symmetric if $\langle A\mathrm{x}, \mathrm{y}\rangle = \langle \mathrm{x}, A\mathrm{y}\rangle, \forall\, \mathrm{x}$ and y in domain of A

A symmetric everywhere defined operator is called self-adjoint or Hermitian

Note that if we take as $\mathcal{H}$ Hilbert space $\mathbb{C}^n$ with standard dot product
and interpret a Hermitian square matrix $\mathbb{A}$ as a linear operator
on this Hilbert space we have ☛ $\langle \mathrm{x}, \mathbb{A}\mathrm{y}\rangle = \langle \mathbb{A}\mathrm{x}, \mathrm{y}\rangle, \forall\, \mathrm{x}, \mathrm{y} \in \mathbb{C}^n$

Example 2.2.1

A convenient basis for traceless Hermitian 2×2 matrices are Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.1.36.)$$

They obey following relations:

$(i)\ \sigma_i^2 = \mathbb{I}$

$(ii)\ \sigma_i \sigma_j = -\sigma_j \sigma_i$

$(iii)\ \sigma_i \sigma_j = i\sigma_k$

☛ (i, j, k) a cyclic permutation of (1,2,3)

These three relations can be summarized as

$$\sigma_i \sigma_j = \mathbb{I}\delta_{ij} + i\epsilon_{ijk}\sigma_k \quad (2.1.37.)$$

☛ ϵ_{ijk} is Levi-Civita symbol


TO BE
CONTINUED....