Functions of Matrices

\[ \exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \]
\[ \det(\exp(AB)) = \exp\left( \text{Tr}(A) \right) \]

\[ \sin(A) = \frac{\exp(iA) - \exp(-iA)}{2i} \]
\[ \cos(A) = \frac{\exp(iA) + \exp(-iA)}{2} \]

Baker-Hausdorff formula

\[ \exp(iG)H \exp(-iG) = H + [G,H] + \frac{1}{2} [G, [G, H]] + \ldots \]

Pauli and Dirac Matrices

\[ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \]

were introduced by W. Pauli to describe a particle of spin \( \frac{1}{2} \) in non-relativistic quantum mechanics. It can be shown they satisfy:

\[ \sigma_i \sigma_j + \sigma_j \sigma_i = 2 \delta_{ij} I \]

\[ \sigma_i \sigma_j = i \delta_{ij} \epsilon_{ijk} \text{ cyclic permutation } 1, 2, 3 \]

\[ \sigma_i^2 = I, \quad e^{i \theta} \sigma_k = \cos \theta + i \sin \theta \]

\[ [\sigma_i, \sigma_j] = \sigma_i \sigma_j - \sigma_j \sigma_i = 2i \epsilon_{ijk} \sigma_k \]

Anti-symmetric tensor

Levi-Civita symbol

+1 for even permutations and -1 for odd permutations zero repeated indices
as the orbital angular momentum \( \mathbf{L} \) \((L \times L = iL)\) and the SO(3) and SU(2) groups.

The three Pauli matrices and the unit matrix form a complete set; any Hermitian matrix can be expanded as

\[
\mathbf{M} = m_0 \mathbb{I} + \mathbf{m} \cdot \mathbf{\sigma} = m_0 + \mathbf{m} \cdot \mathbf{\sigma} \quad \mathbf{m} = \text{vector}
\]

\[
\text{Tr} (\mathbf{\sigma}_i) = 0
\]

\[
\text{Tr} (\mathbf{M}) = 2m_0 \quad \text{Tr} (\mathbf{M} \mathbf{\sigma}_i) = 2m_i \quad (i = 1, 2, 3)
\]

In 1927 Paul Dirac extended this formalism to moving relativistic particles with spin \( \frac{1}{2} \) such as electrons (or neutrinos).

To be consistent with special relativity, he started from Einstein energy Eq

\[
E^2 = \mathbf{p}^2 c^2 + m^2 c^4 \quad \text{instead of} \quad E = \frac{\mathbf{p}^2}{2m}
\]

The key to Dirac Eq is to factorize

\[
E^2 - \mathbf{p}^2 c^2 \quad \text{using} \quad (\mathbf{\sigma} \cdot \mathbf{a}) (\mathbf{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} \mathbf{\Gamma} + \mathbf{\sigma} \cdot (\mathbf{\sigma} \times \mathbf{\sigma})
\]

\[
E^2 - \mathbf{p}^2 c^2 = E^2 - (c \mathbf{\sigma} \cdot \mathbf{p})^2 = (E - c \mathbf{\sigma} \cdot \mathbf{p}) (E + c \mathbf{\sigma} \cdot \mathbf{p}) = m^2 c^4
\]

Equivalently, we can introduce the matrices \( \mathbf{\Gamma} \) and \( \mathbf{\chi} \).
To factorize we can introduce new $4 \times 4$ matrices - so called Dirac matrices.

1. $\gamma^0 = \sigma_3 \otimes I = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

2. $\gamma^{41} = i \sigma_2 \otimes \sigma_1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$

3. $\gamma^2 = \gamma^1 \gamma^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

4. $\gamma^{12} = \gamma \otimes \sigma_2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}$

5. $\gamma^3 = \gamma \otimes \sigma_3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

$\vec{\gamma} = \gamma \otimes \sigma$ is a vector with three components each of a $4 \times 4$ matrix. A generalization of Pauli matrices to $4 \times 4$ matrices.
\[ x^\mu = (x_0, x_1, x_2, x_3) \]

If we recognize
\[ \left( (E x^\mu \otimes I - c (r x \bar{\sigma})) \cdot \bar{p} \right)^2 \]
\[ = E^2 r^2 \otimes I + c^2 r^2 \otimes (\bar{\sigma} \cdot \bar{p})^2 - E c (r^1 \nu + r^2 \nu') \otimes \bar{\sigma} \cdot \bar{p} \]
\[ = E^2 - p^2 c^2 \]

Using \( r^2 = 1 = -x^2 \), \( r^1 \nu + r^2 \nu' = 0 \)

Then \( E r^\mu \otimes I - c (r x \bar{\sigma}) \bar{p} = x^\mu p_\mu = (x^0, x^1) \cdot (E, c\bar{p}) \)

scalar product of two vectors \( r^\mu \) and \( p^\mu \)

Then \( p^2 = p^\mu p_\mu = E^2 - p^2 c^2 \)
\[ (x^\mu p_\mu)^2 = p^\mu p_\mu \]

for vector generalization of \( (\bar{\sigma} \cdot \bar{\sigma})^2 = \bar{\sigma}^\mu \bar{\sigma}_\mu \)

Summarizing:

"The relativistic treatment of spin \( \frac{1}{2} \) particles led to 4x4 matrices. Non-relativistic particles are described by 2x2 \( \sigma \) matrices."
Rank of a matrix:

Maximum number of linearly independent rows (or columns)

A matrix has a rank $r$ if:

1. At least one of its $r 	imes r$ minor determinants $\neq 0$
2. Each of its $(r+1) \times (r+1)$ minor determinants $= 0$

A square $N \times N$ matrix $A$ is called non-singular if $r = N$, i.e., $\det A \neq 0$ and called singular if $r < N$ ($\det A = 0$)

- An inhomogeneous system of Eqs

$$AX = \vec{y} \neq 0$$

with $A$ and $\vec{y}$ known

has a unique solution if $\det A \neq 0$

Cramer's rule: Only of formal interest, impractical

Call $A_i$ the $N \times N$ matrix obtained by replacing the $i$th column of $A$ by the vector $\vec{y}$

° The solution $AX = \vec{y}$ is given by

$$x_i = \frac{\det A_i}{\det A}$$

E.g.,

$$
\begin{bmatrix}
0 & 1 \\
4 & 1
\end{bmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} =
\begin{pmatrix}
1 \\
2
\end{pmatrix}
$$

$\det A = -a$

Solve for $x_1$
\[ A_1 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \rightarrow \det A_1 = -1 \]

\[ x_1 = \frac{1}{a}, \quad x_2 = \frac{\det \begin{pmatrix} 0 & 1 \\ -a & 1 \end{pmatrix}}{-a} = 1 \rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{a} \\ 1 \end{pmatrix} \]

**Homogeneous System**

1) This is a system that has the trivial solution \( \vec{x} = 0 \)

2) There is at least one non-trivial solution if \( \det A = 0 \)

3) If the rank of \( A \) is \( r < N \), then the system has exactly \( N - r \) linearly independent solutions

**Eigenvalue Problems**  \( AX = \lambda X \Rightarrow (A - \lambda I)X = 0 \)

with \( A \) = real, symmetric matrix (Hermitian if complex)

\( \lambda \): unknown eigenvalue

\( X \): unknown eigenvector corresponding to \( \lambda \)

This system has a non-trivial solution if

\[ \det (A - \lambda I) = 0 \]

When you expand this out you get a polynomial \( P \) of order \( N \) in \( \lambda \), whose \( N \) solutions of \( P(\lambda) = 0 \) are the desired \( \lambda \).
There will be $N$ roots $\lambda_k$ to $p(\lambda) = 0$, $\lambda_k = \text{Eigenvalues}$

For Hermitian matrices the $\lambda_k$ are all real

Note that if $A\mathbf{x} = \lambda\mathbf{x}$

$$\mathbf{x}^* A \mathbf{x} = \lambda \mathbf{x}^* \mathbf{x}$$

$$\lambda = \frac{\mathbf{x}^* A \mathbf{x}}{\mathbf{x}^* \mathbf{x}}$$

$$\frac{\sum_{i=1}^{N} x_i^* A x_i}{\sum_{i=1}^{N} x_i^* x_i} = \lambda$$

To each $\lambda_k$, there is a corresponding eigenvector $\mathbf{x}_k$.

If we collect the $\lambda_k$ into the diagonal elements of the matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

Then we can write the eigenvalue equation as

$$A \mathbf{x} = \lambda \mathbf{x}$$ i.e. each column vector of $\mathbf{X}$ is a separate eigenvector.
Orthogonality of eigenvectors

(i) If the eigenvalues of a symmetric matrix $A$ are distinct, then their corresponding eigenvectors are orthogonal.

e.g. consider two distinct eigenvalues $\lambda_1, \lambda_2$

$\lambda_1 \neq \lambda_2$ and their respective eigenvectors $\vec{x}_1$ and $\vec{x}_2$

If $A$ is symmetric then

(1) $\vec{x}_1 \cdot A \vec{x}_2 = \vec{x}_1 \cdot \lambda_2 \vec{x}_2$

(2) $\vec{x}_2 \cdot A \vec{x}_1 = \vec{x}_2 \cdot \lambda_1 \vec{x}_1$

$A - \lambda_1 I = 0$ since $A$ is symmetric

$(\lambda_1 - \lambda_2) \vec{x}_1 \cdot \vec{x}_2 = 0 \implies \vec{x}_1 \cdot \vec{x}_2 = 0$ orthogonal

(2) If $\lambda_1 = \lambda_2$ = Degenerate eigenvalues

If there are $r$ eigenvalues (out of the $N$ total) which are all equal to $\lambda_1$, then we say $\lambda_1$ is $r$-fold degenerate.

The corresponding eigenvectors at the degenerate eigenvalues are not automatically guaranteed to be orthogonal.

However, there are $r$ linearly-independent
eigenvectors which can be orthogonalized using the Gram-Schmidt process.

Once your eigenvectors are made orthogonal, one can construct a matrix \( X \) that is orthogonal (unitary if \( A \) is Hermitian):

\[
X^+ X = I \quad \text{or} \quad X = X^{-1}
\]

\( AX = \lambda X \) can be also written as:

\[(\lambda) \quad X^+ A X = \lambda \quad \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}
\]

The process of determining \( \lambda \) and \( X \) is often called 'diagonalization' of \( A \) even if the step is not performed.

\[
X = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nn} \end{pmatrix}
\]

And the elements of \((X^+ X)_{ij}\)

are just \( x_i \cdot x_j \)

\[
(X_{11}, x_{12}, \ldots, x_{1n}), \quad (x_{11}) = \delta_{1i}
\]

Since eigenvectors are orthogonal:

\[
\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}
\]