

Further Review

Wayne Dam

Physics Cafe

November 15, 2020

1 Projection Operators

2 The Trace

3 Matrix Functions

Projection Operators

- Recall from the Spectral Decomposition Theorem that we are able to diagonalize Hermitian operators into a weighted sum of rank one outer products,

$$A = \sum_i \lambda_i |i\rangle\langle i|, \quad (1)$$

- The set of eigenvectors, $|i\rangle$, are orthonormal and span the entire vector space.
They form a convenient coordinate system for the vector space.

- So for any vector $|\psi\rangle$ we can write it as,

$$|\psi\rangle = \sum_i \psi_i |i\rangle \quad \text{Usual basis expansion.} \quad (2)$$

$$= \sum_i |i\rangle \cdot \underbrace{\langle i|\psi\rangle}_{\psi_i} \quad \text{Writing components as inner product projections.} \quad (3)$$

$$= \sum_i |i\rangle\langle i| \cdot |\psi\rangle \quad \text{Shifted the order of multiplication by associativity.} \quad (4)$$

- The quantity on rhs of the first line is a ket multiplied by a scalar (it's an inner product).

- The quantity on the last line is a linear operator multiplying a ket.
- But these two quantities are equal: $\psi_i|i\rangle = |i\rangle\langle i|\cdot|\psi\rangle$
- The operators $|i\rangle\langle i|$ are special operators — they are **rank-one projection operators** that project onto the one-dimensional subspaces spanned by the $|i\rangle$.
- If we take a collection of these projectors, the first k , for example,

$$P \equiv \sum_{i=1}^k |i\rangle\langle i|, \quad (5)$$

it will be the projector onto the k -dimensional linear subspace spanned by the first k eigenvectors, $|1\rangle, \dots, |k\rangle$.

- The **orthogonal complement** of P is $Q \equiv I - P$. It is a projector onto the vector space spanned by $|k+1\rangle, \dots, |n\rangle$. **Why?**
- Also, $|v\rangle\langle v| = (|v\rangle\langle v|)^\dagger$. **Why?** This implies P is Hermitian.
- Note, too, that $P^2 = P$. **Why?**

The Trace

- Previously the trace of a matrix was defined as the sum of its diagonal elements,

$$\text{tr}(A) = \sum_i a_{ii}. \quad (6)$$

- It's important to understand **we are really evaluating a property**, that we call the trace, **of the linear operator** that the matrix represents.
- It's **intrinsic to the operator** and **does not depend on the particular orthonormal coordinate system** we are using to represent the operator as as a matrix.
- This is the same as the norm (or length) of a geometric vector. The numbers in the column matrix we use to represent that vector will change with the coordinate system chosen. But it's *norm* is invariant. It's intrinsic to that vector.

Example (1. The length (Euclidean norm) of a vector is invariant.)

- Recall that the norm squared of a vector equals its inner product with itself,

$$|v|^2 = \langle v | v \rangle = v^\dagger v. \quad (7)$$

- Recall the change of basis on slide 13 from the previous talk.

Given a matrix U with orthonormal columns containing the basis vectors of the new coordinate system, the new vector is given by,

$$v' = U^\dagger v. \quad (8)$$

- But the norm squared in the new coordinate system is just

$$v'^\dagger v' = v^\dagger U U^\dagger v \quad (9)$$

$$= v^\dagger I v \quad (10)$$

$$= v^\dagger v. \quad (11)$$

- Recall that the (i, j) element of a matrix representation of a linear operator is given by,

$$a_{ij} = \langle i | A | j \rangle. \quad (12)$$

- We can then write the trace as,

$$\text{tr } A = \sum_i \langle i | A | i \rangle. \quad (13)$$

- Note that it is linear. Why?
- Next we show two useful and properties of the trace:

Properties of the trace

- The trace is independent of the coordinate system used.
- The trace equals the sum of the operator's eigenvalues.

(Proof) 1. The Trace is Independent of the Coordinate System.

- The trick is using the completion of unity from last slides,

$$I = \sum_i |i\rangle\langle i| \quad (14)$$

$$\text{tr } A = \sum_i \langle i | A | i \rangle \quad (15)$$

$$= \sum_i \langle i | A \left(\sum_j |\phi_j\rangle\langle\phi_j| \right) | i \rangle \quad (16)$$

$$= \sum_i \sum_j \langle i | A | \phi_j \rangle \langle \phi_j | i \rangle \quad (17)$$

$$= \sum_i \sum_j \langle \phi_j | i \rangle \langle i | A | \phi_j \rangle \quad (18)$$

$$= \sum_j \langle \phi_j | \left(\sum_i |i\rangle\langle i| \right) A | \phi_j \rangle \quad (19)$$

$$= \sum_j \langle \phi_j | A | \phi_j \rangle \quad (20)$$

(Proof) 2. The Trace is the Sum of the Eigenvalues.

- Recall the trace of A is given by the sum of its diagonal elements,

$$\text{tr } A = \sum_i a_{ii} = \sum_i \langle i | A | i \rangle \quad (21)$$

- But in the same basis (just using a different index variable, j) we can expand A as,

$$A = \sum_j \lambda_j |j\rangle \langle j| \quad (22)$$

- Substituting this expansion into the trace definition,

$$\text{tr } A = \sum_i \langle i | \cdot \sum_j \lambda_j |j\rangle \langle j| \cdot | i \rangle \quad (23)$$

$$= \sum_i \sum_j \lambda_j \langle i | j \rangle \langle j | i \rangle \quad (24)$$

$$= \sum_{ij} \lambda_j \delta_{ij} \delta_{ji} \quad (25)$$

$$\implies \text{tr } A = \sum_i \lambda_i \quad (26)$$



Matrix Functions

- In quantum mechanics **scalar quantities** that we measure classically, such as position, energy, etc., **become linear operators** represented by matrices in a particular coordinate system.
- So in the classical case where we might have a function of position or energy, we now have to consider **functions with matrix arguments**.
- In reality these are functions whose arguments are linear operators and return new linear operators.
- We are used to functions of scalar values, such as $f(x) = \sin(x)$ or $f(x) = e^x$.
- From first year we know functions can generally be expanded in a Taylor Series,

$$f(x) = \sum_{i=0}^{\infty} a_i x^i. \quad (27)$$

- For example, the exponential function,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad (28)$$

- So we simply *define* **matrix functions** based on their scalar Taylor series.

- So we now have,

$$f(X) = \sum_i a_i X^i, \quad (29)$$

- For example,

$$e^X = I + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots \quad (30)$$

- In principle we may have **hundreds of terms**, requiring raising a matrix to a hundredth power by repeated multiplication. These matrices may be of **dimensions of thousands or millions**, again in principle.
 \implies See the **Cayley-Hamilton Theorem** for more details. E.g., wikipedia.
- There is a remarkable fact that we can write this matrix function strictly in terms of the eigenvalues of X !

Theorem (1. Cayley-Hamilton Theorem (sort of))

- A function of an $n \times n$ Hermitian matrix can be written in terms of n scalar functions of its eigenvalues.
- If X has the eigendecomposition $X = U\Lambda U^\dagger$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, then $f(X)$ can be written as,

$$f(X) = U \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & f(\lambda_n) \end{pmatrix} U^\dagger \quad (31)$$

- Why is this such a big deal?

For example, what is X^{100} ?

$$X^{100} = U \begin{pmatrix} \lambda_1^{100} & 0 & \cdots & 0 \\ 0 & \lambda_2^{100} & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & \lambda_n^{100} \end{pmatrix} U^\dagger \quad (32)$$

Example (1. Examples of Matrix Functions)

- Some matrix functions,

$$e^X = U \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & e^{\lambda_n} \end{pmatrix} U^\dagger \quad (33)$$

$$\sin(X) = U \begin{pmatrix} \sin(\lambda_1) & 0 & \cdots & 0 \\ 0 & \sin(\lambda_2) & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & \sin(\lambda_n) \end{pmatrix} U^\dagger \quad (34)$$

$$X^{-1} = U \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & \frac{1}{\lambda_n} \end{pmatrix} U^\dagger \quad (35)$$

1. Proof (Cayley-Hamilton Theorem (sort of)).

- Define the function $f(x) = \sum_i a_i X^i$ where $X = U\Lambda U^\dagger$
- The i 'th power of X is then,

$$X^i = (U\Lambda U^\dagger)^i \quad (36)$$

$$= U\Lambda U^\dagger \cdot U\Lambda U^\dagger \dots U\Lambda U^\dagger \cdot U\Lambda U^\dagger \quad (37)$$

$$= U\Lambda^i U^\dagger \quad (38)$$

- So carrying the a_i coefficients into the diagonal eigenvalue matrix,

$$f(x) = \sum_i a_i U\Lambda^i U^\dagger \quad (39)$$

$$= \sum_i U \begin{pmatrix} a_i \lambda_1^i & & \\ & a_i \lambda_2^i & \\ & & \ddots \end{pmatrix} U^\dagger \quad (40)$$

$$= U \begin{pmatrix} \sum_i a_i \lambda_1^i & & \\ & \sum_i a_i \lambda_2^i & \\ & & \ddots \end{pmatrix} \quad (41)$$

1. Proof (Cayley-Hamilton Theorem (sort of), cont'd.

- And so we have,

$$f(X) = U \begin{pmatrix} f(\lambda_1) & & \\ & f(\lambda_2) & \\ & & \ddots \end{pmatrix} U^\dagger \quad (43)$$

$$= U f(\Lambda) U^\dagger. \quad (44)$$

- So the function is applied element-wise to the *eigenvalues* of X . Not the elements of X .



Example (2. Examples of Matrix Functions)

- There are interesting analytical results from this.
- Suppose we exponentiate the antisymmetric (not Hermitian here!) matrix multiplied by a parameter θ ,

$$A = \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (45)$$

So we have e^A .

- You can show that A has eigenvalues $\pm i\theta$ with corresponding eigenvectors $(\pm i, 1)$.
- So the eigen-decomposition of A is,

$$\Lambda = \begin{pmatrix} i\theta & 0 \\ 0 & -i\theta \end{pmatrix}, U = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, U^{-1} = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \quad (46)$$

Example (2. Examples of Matrix Functions, cont'd)

- And we have,

$$e^A = U e^\Lambda U^{-1} \quad (47)$$

$$= \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \quad (48)$$

$$= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad (49)$$

- This is the two dimensional rotation matrix and is a linear operator in its own right.

Example (3. Application: Solving Schrodinger's Equation)

- Suppose we have a system with a Hamiltonian $H = \hbar\omega\sigma_z$, where σ_z is the Pauli spin matrix $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

- Recall Schrodinger's equation, which gives the time evolution of a system,

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle. \quad (50)$$

- For a time independent Hamiltonian, Mike has shown it has the solution,

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(t=0)\rangle. \quad (51)$$

- This looks complicated but it's a straightforward generalization of Mike's scalar case.

Recall the d.e. $\frac{dx}{dt} = Hx$ where x and H are just scalar numbers.

Positing a solution $x = e^{Ht}$ we see it satisfies the d.e. by taking derivatives of both sides of the equation.

We have the same case here with a few constants (i and \hbar) thrown in and **we think of the variables being vectors and matrices now.**

- So we need to solve to the exponential of a two by two matrix!

Example (3. Application: Solving Schrodinger's Equation, cont'd)

- Easy. Follow the steps from the previous example, modifying as appropriate.
- After a bit of algebra we find

$$e^{-iHt/\hbar} \equiv U(t) = \begin{pmatrix} \cos(\omega t) & -i \sin(\omega t) \\ -i \sin(\omega t) & \cos(\omega t) \end{pmatrix} \quad (52)$$

where $U(t)$ is a unitary matrix.

- Suppose our initial quantum state at time $t = 0$ is $|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- Then the state as it evolves in time under the given Hamiltonian is,

$$|\psi(t)\rangle = U(t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (53)$$

$$= \begin{pmatrix} \cos(\omega t) \\ -i \sin(\omega t) \end{pmatrix} \quad (54)$$

- What is the probability of finding the particle in its original state if we make a measurement at time t ?
- This will be the magnitude squared of the projection between the two vectors.

- This can be seen as a prototype example of a particle with spin precessing in constant magnetic field (and so constant Hamiltonian).
- The Hamiltonian is constant in time *but* the particle state varies in time. Interesting!

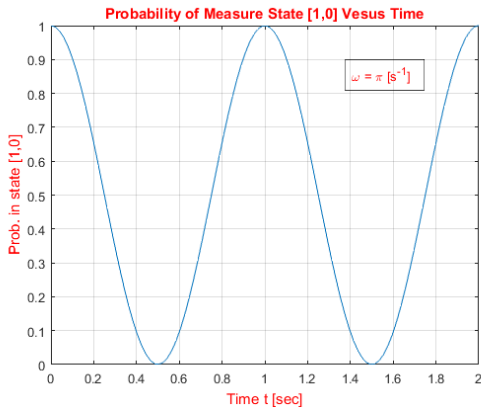


Figure: Probability of find state in $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ versus time for $\omega = \pi \text{ s}^{-1}$.

End

END.