Lecture 4

II. Quantum Spin Systems

2.4 On Dirac Notation

The Dirac "bra" and "ket" notation is used often in quantum mechanics as well as in quantum information theory. It is a convenient way to express certain calculations that arise frequently in the Hilbert space setting. We will only discuss/use "Dirac Notation" in the context of finite dimensional Hilbert spaces.

Let $\mathcal{H} = \mathbb{C}^d$ be a finite dimensional Hilbert Space. To any vector $\phi \in \mathcal{H}$, we can associate 2 linear maps:

i) $|\phi\rangle : \mathbb{C} \to \mathcal{H}$, $\quad z \mapsto z\phi$

and

ii) $\langle\phi| : \mathcal{H} \to \mathbb{C}$, $\quad \psi \mapsto \langle \phi, \psi \rangle$.

Again, with $\phi \in \mathcal{H}$ fixed, each of these maps is linear.

In their dependence on $\phi \in \mathcal{H}$, however, the 1^{st} is linear and the second is anti-linear.

Recall: For any complex vector spaces $\mathfrak{X}$ and $\mathfrak{Y}$, we denote by $\mathcal{L}(\mathfrak{X}, \mathfrak{Y})$ the collection of linear maps T from $\mathfrak{X}$ to $\mathfrak{Y}$.

We now define:

1) $\quad | \, \rangle : \mathcal{H} \to \mathcal{L}(\mathbb{C}, \mathcal{H}), \qquad \phi \mapsto |\phi\rangle$

and

2) $\quad \langle \, | : \mathcal{H} \to \mathcal{L}(\mathcal{H}, \mathbb{C}), \qquad \phi \mapsto \langle\phi|$

The linear map in 1) is called "ket".
The anti-linear map in 2) is called "bra".

Note: The map in 2), i.e. bra, is the map which identifies $\mathcal{H}$ with its dual space $\mathcal{H}^* = \mathcal{L}(\mathcal{H}, \mathbb{C})$. The existence of such a map is guaranteed by the (Hilbert Space version of the) Riesz Representation Theorem.

I will have more information on this theorem in the notes on basic linear algebra.

Let us now fix an orthonormal basis in $\mathcal{H}$.

For each $\phi \in \mathcal{H}$, $|\phi\rangle \in \mathcal{L}(\mathbb{C}, \mathcal{H})$ can be identified with a column vector.

For each $\phi \in \mathcal{H}$, $\langle\phi| \in \mathcal{L}(\mathcal{H}, \mathbb{C})$ can be identified with a row vector.

These identifications are "matrix representations" of these linear maps. I will say more on this in the notes.

For any pair of vectors, $\phi_1, \phi_2 \in \mathcal{H}$, one can define a map $|\phi_1\rangle\langle\phi_2| : \mathcal{H} \to \mathcal{H}$ by setting.

$$\big(|\phi_1\rangle\langle\phi_2|\big)(\psi) = \langle \phi_2, \psi\rangle \phi_1$$

This rank one map can be understood as the composition of $|\phi_1\rangle$ and $\langle\phi_2|$.

We have seen a special case of this before.

If $\phi_1 = \phi_2 = \phi \in \mathcal{H}$, then

$$(|\phi\rangle\langle\phi|)(\psi) = \langle \phi, \psi\rangle \phi = P_\phi(\psi)$$

where P_ϕ is the orthogonal projection onto the span of $\phi \in \mathcal{H}$

"Dirac Notation" is the convention that vectors $\psi \in \mathcal{H}$ and kets $|\psi\rangle$ are used interchangeably.

Viewing the inner product as a "braket"

$$\underbrace{\langle \phi , \psi \rangle}_{\text{the braket of } \phi \text{ and } \psi} \sim \underbrace{\langle \phi |}_{\text{bra}} \; \underbrace{| \psi \rangle}_{\text{ket}}$$

one sees a relation to the semantics.

Here is a common useage of Dirac notation.

Let $\mathcal{H} = \mathbb{C}^d$.

Let $\{v_1, v_2, \dots, v_d\}$ be an orthonormal basis of $\mathbb{C}^d$.

In Dirac's notion, the vectors are simply labeled:

$$\{v_1, v_2, \dots, v_d\} \longrightarrow \{ |1\rangle, |2\rangle, \dots, |d\rangle \}.$$

orthonormality is replaced by

$$\langle v_i , v_j \rangle = \delta_{ij} \longrightarrow \langle i | j \rangle = \delta_{ij}$$

and completeness is replaced by:

$$v = \sum_{i=1}^{d} \langle v_i, v \rangle v_i \longrightarrow \sum_{i=1}^{d} |i\rangle\langle i| = \mathbb{1}$$

The latter means, for any $v \in \mathbb{C}^d$

$$v = \mathbb{1} v = \sum_{i=1}^{d} |i\rangle\langle i| v = \sum_{i=1}^{d} \langle i | v\rangle |i\rangle$$

$$= \sum_{i=1}^{d} \langle v_i, v\rangle v_i$$

in vector notation

Here is a calculation that occurs often:

Let $\{v_1, v_2, \ldots, v_d\}$ be an orthonormal basis on $\mathcal{H}$.

Let $x, y \in \mathcal{H}$. Use completeness to express $\langle x, y\rangle$.

vectors

$$\langle x, y\rangle = \left\langle \sum_{i=1}^{d} \langle v_i, x\rangle v_i, \sum_{j=1}^{d} \langle v_j, y\rangle v_j \right\rangle$$

$$= \sum_{i=1}^{d} \overline{\langle v_i, x\rangle} \sum_{j=1}^{d} \langle v_j, y\rangle \underbrace{\langle v_i, v_j\rangle}_{\delta_{ij}}$$

$$= \sum_{i=1}^{d} \overline{\langle v_i, x\rangle} \langle v_i, y\rangle$$

$$= \sum_{i=1}^{d} \langle x, v_i\rangle \langle v_i, y\rangle$$

Dirac

$$\langle x, y\rangle = \langle x, \mathbb{1} y\rangle$$

$$= \sum_{i=1}^{d} \langle x | i\rangle\langle i | y\rangle.$$

This is more efficient!

Let us now briefly review some important results in Linear Algebra. You will find these statements in the notes aswell.

Appendix Some important theorems in linear algebra.

A.1 The Spectral Theorem and Functional Calculus

Let V be a complex vector space.

By $\mathcal{L}(V)$ we denote the collection of all linear maps T from V to V.

Def. Let V be a complex inner product space. $T \in \mathcal{L}(V)$ is said to be normal if and only if $TT^* = T^*T$.

In words, T is normal if and only if T commutes with its adjoint.

Recall: If $S, T \in \mathcal{L}(V)$, then

$[S,T] = ST - TS$ is called the commutator of S and T

S and T are said to commute if $[S,T] = 0$.

In this case, $[S,T] = 0 \Rightarrow ST - TS = 0 \Rightarrow ST = TS$.

Theorem (The Spectral Theorem)

Let V be a finite dimensional, complex inner product space.

$T \in \mathcal{L}(V)$ is normal $\iff$ $\exists$ an orthonormal basis of V which consists entirely of eigenvectors of T.

We will not prove this. It is Theorem 11.3.1 in reference (5) from syllabus.

Let us focus on matrices.

Two common examples of normal matrices:

- $A \in M_d$ is self-adjoint $\Rightarrow$ A is normal
- $A \in M_d$ is unitary $\Rightarrow$ A is normal

Recall: Unitary means $A^*A = \mathbb{1}$.

Let $A \in M_d$ be self-adjoint.

By the spectral theorem, there is an orthonormal basis $\{u_1, u_2, \ldots, u_d\}$ of $\mathbb{C}^d$ for which

$$Au_j = \lambda_j u_j \quad \text{for all } 1 \le j \le d.$$

The numbers λ_j are the eigenvalues of A corresponding to the eigenvectors. One easily checks that, since A is self-adjoint, each $\lambda_j \in \mathbb{R}$.

Using the fact that these vectors form a basis, one sees that:

For any $v \in \mathbb{C}^d$, $$v = \sum_{j=1}^{d} \langle u_j, v\rangle u_j$$

and moreover, A acts trivially with respect to this basis:

$$Av = \sum_{j=1}^{d} \langle u_j, v\rangle A u_j = \sum_{j=1}^{d} \lambda_j \underbrace{\langle u_j, v\rangle u_j}_{P_{u_j}(v)}$$

The orthogonal projection onto span(u_j)!

We conclude that for any self-adjoint $A \in M_d$

$$A = \sum_{j=1}^{d} \lambda_j P_{u_j} \quad \left(= \sum_{j=1}^{d} \lambda_j |u_j\rangle\langle u_j| \right)$$

in Dirac notation.

More is true.

Let $U \in M_d$ be the matrix whose columns are given by this orthonormal basis, i.e.

$$U = (u_1 \,\vdots\, u_2 \,\vdots\, \cdots \,\vdots\, u_d)$$

Then it is easy to check that U is unitary $(U^*U = \mathbb{1})$

and moreover,

$$U^* A U = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{pmatrix} = \mathrm{diag}(\lambda_j)$$

i.e. A is unitarily diagonalizable.

<u>Note</u>: One easily checks that: If $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis of $\mathbb{C}^n$, then

$$V = \left(v_1 \,\vdots\, v_2 \,\vdots\, \cdots \,\vdots\, v_n \right) \quad \text{is unitary}$$

and

$$\hat{V} = \begin{pmatrix} \text{---}\; v_1 \;\text{---} \\ \text{---}\; v_2 \;\text{---} \\ \vdots \\ \text{---}\; v_n \;\text{---} \end{pmatrix} \quad \text{is unitary}$$

This diagonalization is the key insight into defining functions of matrices. In general, this is called functional calculus.

Functional Calculus

Let $f : \mathbb{R} \to \mathbb{C}$ be any function.

Let $A \in M_d$ be self-adjoint.

Let $U \in M_d$ be the unitary that diagonalizes A:

$$U^* A U = \mathrm{diag}(\lambda_j) \iff A = U \,\mathrm{diag}(\lambda_j) U^*.$$

We define $f(A) \in M_d$ by setting

$$f(A) = U \,\mathrm{diag}(f(\lambda_j)) U^*.$$

Note: One readily checks that if f is real valued then $f(A)$ is self-adjoint.

A short calculation shows that

standard basis vector in $\mathbb{C}^d$

$$f(A) u_K = U \,\mathrm{diag}(f(\lambda_j)) U^* u_K = U \,\mathrm{diag}(f(\lambda_j))\, e_K$$

$$= f(\lambda_K) U e_K = f(\lambda_K) u_K$$

$$\Rightarrow \quad f(A) = \sum_{j=1}^{d} f(\lambda_j) P_{u_j} \quad \left(= \sum_{j=1}^{d} f(\lambda_j) |u_j\rangle\langle u_j| \right)$$

in Dirac notation.

Ex An important example of this functional calculus is as follows.

Let $H \in M_d$ be self-adjoint.

For any $t \in \mathbb{R}$, consider the function $f_t : \mathbb{R} \to \mathbb{C}$ given by $f_t(x) = e^{-itx}$.

The matrix $f_t(H) = e^{-itH} \in M_d$ is defined using the functional calculus above.

One checks that $f_t(H)$ is unitary. (In fact, $f_t(H)^* = e^{itH}$.)

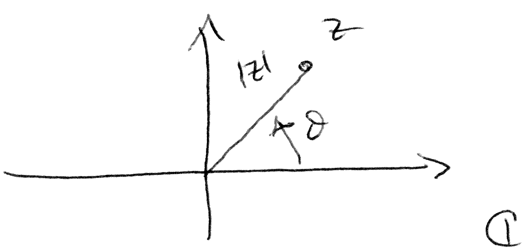
This matrix is crucial in solving the Schrödinger equation.

The Singular Value decomposition

Another important result from linear algebra is the singular value decomposition. We state two forms of this result below.

Version 1 (on square matrices)

<u>Recall</u>: If $z \in \mathbb{C}$, then $z = e^{i\theta}|z|$ where $|z|$ is the modulus of z and $0 \leq \theta < 2\pi$. This is called the polar form of z. It corresponds to:



There is an analogous result for matrices $A \in M_d$. Before we state it, let us recall some notation.

For any $A \in M_d$, the matrix $B = A^*A$ is clearly non-negative. As such, it is self-adjoint with non-negative eigenvalues. By the spectral theorem, we may write:

$$B = A^*A = \sum_{j=1}^{d} (\sigma_j(A))^2 P_j$$

where P_j is the orthogonal projection onto the eigenvector associated to $(\sigma_j(A))^2$ and we have written the non-negative eigenvalues as "squares" for convenience.

Using functional calculus, we may define $|A| \in M_d$

$$|A| = \sqrt{A^*A} = \sum_{j=1}^{d} \sigma_j(A) P_j$$

The numbers $\{\sigma_j(A)\}_{j=1}^{d}$ are called the singular values of A.

Theorem (Polar decomposition)

Let $A \in M_d$. There is a unitary $U \in M_d$ for which

$$A = U|A|.$$

Writing A, as above, is called the polar decomposition of A.

We will not prove this result. It is theorem 11.6.1 in reference (5) from the syllabus.

Using the polar decomposition, our first result on the Singular Value Decomposition follows easily.

<u>Theorem</u> (Singular Value Decomposition version 1)

Let $A \in M_d$. There are 2 orthonormal bases of $\mathbb{C}^d$, labeled $\{e_1, e_2, \ldots, e_d\}$ and $\{f_1, f_2, \ldots, f_d\}$, for which

$$Ax = \sum_{j=1}^{d} \sigma_j(A) \langle e_j, x \rangle f_j \quad \text{for all } x \in \mathbb{C}^d.$$

Here the numbers $\{\sigma_j(A)\}_{j=1}^{d}$ are the singular values of A.

<u>Proof:</u>

Let $\{e_1, e_2, \ldots, e_d\}$ be the orthonormal basis of eigenvectors of $|A|$. (The existence of this is guaranteed by the spectral theorem since $|A|$ non-neg. $\Rightarrow$ $|A|$ self-adjoint.)

In this case,
$$|A|x = \sum_{j=1}^{d} \langle e_j, x \rangle |A| e_j = \sum_{j=1}^{d} \sigma_j(A) \langle e_j, x \rangle e_j$$

$$\Rightarrow \quad Ax = U|A|x = \sum_{j=1}^{d} \sigma_j(A) \langle e_j, x \rangle U e_j.$$

One checks that $\{f_1, f_2, \ldots, f_d\}$ with $f_j = U e_j$ for $1 \le j \le d$ is an orthonormal basis of $\mathbb{C}^d$. This completes the proof.

Version 2 (on non-square matrices)

Let $A \in \mathbb{C}^{m\times n}$. Write $A = \{a_{ij}\}$ where $a_{ij} \in \mathbb{C}$ for $1 \le i \le m$ and $1 \le j \le n$.

Define $A^* \in \mathbb{C}^{n\times m}$ as follows.

Write $A^* = \{a^*_{ji}\}$ and set $a^*_{ji} = \overline{a_{ij}}$ for all $1\le i\le m$, $1 \le j \le n$

We refer to A^* as the adjoint of A.

A short calculation shows that:

$$(*) \quad \langle Ax, y\rangle_{\mathbb{C}^m} = \langle x, A^*y\rangle_{\mathbb{C}^n} \quad \text{for all } x\in\mathbb{C}^n \text{ and } y \in \mathbb{C}^m.$$

This relation "justifies" the label A^* is the adjoint of A.

Now, for any $A \in \mathbb{C}^{m\times n}$, the matrix $B \in M_n$ given by

$$B = A^*A \qquad \text{is clearly non-negative.}$$

In fact, for any $x \in \mathbb{C}^n$,

use $(*)$

$$\langle x, Bx\rangle_{\mathbb{C}^n} = \langle x, A^*Ax\rangle_{\mathbb{C}^n} = \langle Ax, Ax\rangle_{\mathbb{C}^m} = \|Ax\|^2_{\mathbb{C}^m} \ge 0.$$

Since B is non-negative, the spectral theorem applies. We can write

$$B = A^* A = \sum_{j=1}^{n} (\sigma_j(A))^2 P_j$$

and we again declare that the numbers $\{\sigma_j(A)\}_{j=1}^{n}$ <u>are</u> the singular values of A.

<u>Theorem</u> (Singular Value Decomposition Version 2)

For any $A \in \mathbb{C}^{m \times n}$, there are unitary matrices $U_1 \in M_m$ and $U_2 \in M_n$ for which

$$A = U_1 D U_2^* .$$

Here the matrix $D \in \mathbb{C}^{m \times n}$ is diagonal with non-negative entries. In fact, the diagonal entries are the singular values of A.

<u>Note:</u>

1) $D \in \mathbb{C}^{m \times n}$ is diagonal means that if $D = \{d_{ij}\}$ then $d_{ij} = 0$ if $i \neq j$.

2) The final statement should read: The non-zero entries in D are the non-zero singular values of A counted according to multiplicity.

Proof:

Let us 1st construct a singular value decomposition for $A \in \mathbb{C}^{m \times n}$ when $n \leq m$. We discuss other cases later.

As discussed before the statement of this result, $B = A^* A \in M_n$ is non-negative; hence self-adjoint. In this case the spectral theorem guarantees the existence of an orthonormal basis of eigenvectors, label these by $\{u_1, u_2, \dots, u_n\} \subset \mathbb{C}^n$, and

$$A^* A u_j = B u_j = \sigma_j(A)^2 u_j \qquad \text{for } 1 \leq j \leq n.$$

By reordering, let us choose a basis so that

$$\sigma_1(A) \geq \sigma_2(A) \geq \dots \sigma_j(A) \geq 0 \qquad \text{for all } 1 \leq j \leq n.$$

In this case, take $U_2 \in M_n$ by setting

$$U_2 = \left(u_1 \,\middle|\, u_2 \,\middle|\, \cdots \,\middle|\, u_n \right).$$

By previous arguments, $U_2 \in M_n$ is unitary.

Note also that

use (*)

$$\langle A u_j, A u_k \rangle_{\mathbb{C}^m} = \langle A^* A u_j, u_k \rangle_{\mathbb{C}^n} = \sigma_j(A)^2 \delta_{jk}$$

We conclude that the vectors

$$\{Au_1, Au_2, \ldots, Au_n\} \subset \mathbb{C}^m$$

are pairwise orthogonal and normalized so that:

$$\|Au_j\|_{\mathbb{C}^m}^2 = \sigma_j(A)^2.$$

It is, of course, possible that $\sigma_j(A) = 0$ for some j.

In this case, $Au_j = 0$ as well.

If $A = 0$, take $D = 0$.

If $A \neq 0$, then $\sigma_1(A) \neq 0$.

(In fact, if $\sigma_1(A) = 0$, then all singular values are zero. In this case $A^*A = 0$. As a result, for any x, $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = 0$, and thus $A = 0$.)

Let $j_0 \geq 1$ be the maximum integer for which $\sigma_j(A) > 0$.

For any $1 \leq j \leq j_0$, take $w_j \in \mathbb{C}^m$ by setting

$$w_j = \frac{1}{\sigma_j(A)} Au_j.$$

This collection of vectors is normalized ($\|w_j\| = 1$) and pairwise orthogonal. As such it may be extended to an orthonormal basis of $\mathbb{C}^m$. Let us label these vectors as

$$\{w_1, w_2, \ldots, w_m\}.$$

The matrix $U_1 \in M_m$ given by

$$U_1 = \left(w_1 \,\vdots\, w_2 \,\vdots\, \cdots \,\vdots\, w_m \right)$$

is clearly unitary. Take $D \in \mathbb{C}^{m \times n}$ to be the diagonal matrix with entries:

$$d_{jj} = \begin{cases} \sigma_j(A) & 1 \le j \le j_0 \\ 0 & j > j_0 \end{cases}.$$

We claim that $A = U_1 D U_2^*$.

To see this, we check that this equality is true on the basis $\{u_1, u_2, \dots, u_n\}$.

Let us take $\{e_1, e_2, \dots, e_n\} \subset \mathbb{C}^n$ to be the standard basis.
Similarly take $\{e_1', e_2', \dots, e_m'\} \subset \mathbb{C}^m$ to be the standard basis.

Note that: i) $U_2^* u_j = e_j$ for $1 \le j \le n$

ii) $U_1 e_j' = w_j$ for $1 \le j \le m$.

Now for any $1 \le j \le n$

$$(U_1 D U_2^*)(u_j) = (U_1 D)(e_j) = d_{jj} U_1 e_j' = d_{jj} \cdot w_j$$

if $1 \le j \le j_0 \Rightarrow d_{jj} w_j = \sigma_j(A) \cdot \frac{1}{\sigma_j(A)} A u_j = A u_j$

if $j > j_0 \Rightarrow d_{jj} w_j = 0 = A u_j$ ✓

Now, let us say that $A \in \mathbb{C}^{m\times n}$ has a singular value decomposition if there are unitaries $U_1 \in M_m$ and $U_2 \in M_n$ for which

$$A = U_1 D U_2^*$$

for some diagonal matrix $D \in \mathbb{C}^{m\times n}$ with non-negative entries.

Observe: If A has a singular value decomposition, then

$$A^* A = U_2 D^* U_1^* U_1 D U_2^*$$
$$= U_2 D^* D U_2^*$$

and $D^* D \in M_n$ is diagonal. By definition, the diagonal entries of D (at least those that are non-zero) must be the singular values of A; counting multiplicity.

Note also that: If A has a singular value decomposition, then

$$AA^* = U_1 D U_2^* U_2 D^* U_1^* = U_1 D D^* U_1^*$$

and $DD^* \in M_m$ is diagonal. Thus the non-zero entries of D must also be the non-zero singlar values of A^*.

We conclude that the non-zero singular values of A coincide (counting multiplicity) with the non-zero singular values of A^*.

To finish, we proved that any $A \in \mathbb{C}^{m\times n}$ with $n \leq m$ has a singular value decomposition. If $A \in \mathbb{C}^{m\times n}$ satisfies $m < n$, then $A^* \in \mathbb{C}^{n\times m}$ is a matrix to which we can apply the previous result. Since the non-zero singular values of A and A^* agree, this proves the result in general.