PHYSICS 491: Symmetry and Quantum Information April 11, 2017

Representation theory, density operators, partial trace

Lecture 3 *Michael Walter, Stanford University*

These lecture notes are not proof-read and are offered for your convenience only. They include additional detail and references to supplementary reading material. I would be grateful if you email me about any mistakes and typos that you find.

In this lecture, we'll revisit some fundamentals: First, we discuss representation theory more systematically and prove the "resolution of the identity" formula (2.6) from last lecture. Then we recall the notion of a density operator and discuss the partial trace, which allows us to define the quantum state of subsystems.

3.1 Representation theory primer

A *(finite-dimensional unitary) representation* of a group G is given by (i) a (finite-dimensional) Hilbert space $\mathcal{H}$, and (ii) unitary operators R_g on $\mathcal{H}$ for every group element $g \in G$ such that the following two laws are satisfied:

$$R_1 = \mathbb{1}, \quad R_{gh} = R_g R_h$$

Every group has a *trivial representation*, given by identity operators $R_g = \mathbb{1}_{\mathcal{H}}$ acting on a one-dimensional space $\mathcal{H}$. We will often simply speak of "the representation H", but we always have associated operators R_g in mind. All representations that we will ever study in this course will be unitary and finite-dimensional.

A useful way of understanding a representation is to decompose it into smaller building blocks. Suppose that $\tilde{\mathcal{H}} \subseteq \mathcal{H}$ is an *invariant subspace*, i.e., a subspace such that $R_g \tilde{\mathcal{H}} \subseteq \tilde{\mathcal{H}}$ for all $g \in G$. Then, the orthogonal complement $\tilde{\mathcal{H}}^\perp$ is also an invariant subspace! Indeed, if $|\phi\rangle \in \tilde{\mathcal{H}}^\perp$ then, for all $|\psi\rangle \in \tilde{\mathcal{H}}$,

$$\langle \psi | R_g | \phi \rangle = \langle R_g^\dagger \psi | \phi \rangle = 0,$$

since $R_g^\dagger |\psi\rangle \in \tilde{\mathcal{H}}$; this shows that $R_g |\phi\rangle \in \tilde{\mathcal{H}}^\perp$. As a consequence, the operators R_g are block diagonal with respect to the decomposition $\mathcal{H} = \tilde{\mathcal{H}} \oplus \tilde{\mathcal{H}}^\perp$, i.e.,

$$R_g = \begin{pmatrix} \tilde{R}_g & 0 \\ 0 & \tilde{R}_g^\perp \end{pmatrix}.$$

Note that the block $\tilde{R}_g$ is a representation on $\tilde{\mathcal{H}}$ and the block $\tilde{R}_g^\perp$ is a representation on $\tilde{\mathcal{H}}^\perp$. Thus we have successfully decomposed the given representation into two "smaller" representations. We can apply the same reasoning separately to $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{H}}^\perp$, and continue until we arrive at a decomposition

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots \oplus \mathcal{H}_m \tag{3.1}$$

that cannot be refined further. That is, the building blocks $\mathcal{H}_j$ have no interesting invariant subspaces (i.e., the only invariant subspaces are $\mathcal{H}_j$ itself and $\{0\}$, neither of which allow us to decompose further). We call such representations $\mathcal{H}_j$ *irreducible representations* – or *"irreps"*.

How can we compare different representations? An *intertwiner* $J: \mathcal{H} \to \mathcal{H}'$ is a map such that

$$JR_g = R'_g J$$

(hence the name). If there exists an *invertible* intertwiner J then we say that the two representations $\mathcal{H}$ and $\mathcal{H}'$ are *equivalent*, and write $\mathcal{H} \cong \mathcal{H}'$. This invertible intertwiner can always be chosen to be a unitary operator, and we will always assume that all invertible intertwiners under consideration are unitary operators. Note that in this case we have

$$JR_g J^{-1} = JR_g J^\dagger = R'_g$$

so the operators $\{R_g\}$ and $\{R'_g\}$ differ only by an overall "base change". We will use the notation $\mathcal{H} \cong \mathcal{H}'$ and $R_g \cong R'_g$.

Example. *An example that you all know well is the group* $\mathrm{SU}(2)$ *of unitary* 2×2*-matrices with unit determinant, which arises in the study of rotational symmetries of quantum systems. Up to equivalence, its irreducible representations are labeled by their* spin

$$j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\}.$$

E.g., V_0 *is the one-dimensional trivial representation (also called the* singlet*),* $V_{1/2} \cong \mathbb{C}^2$*,* V_1 *is the* triplet *representation, etc. We used the decomposition of* $\mathrm{SU}(2)$*-representations into irreducibles briefly in section 1.2 to find a qubit, and will revisit it in greater detail in a later lecture.*

Example 3.1. *The permutation group* S_3 *has three irreducible representations (up to equivalence):*

(i) The trivial representation $W_{\ydiagram{3}} = \mathbb{C}|0\rangle$*, with* $R_\pi|0\rangle = |0\rangle$*.*

(ii) The sign representation $W_{\ydiagram{1,1,1}} = \mathbb{C}|0\rangle$*, with* $R_\pi|0\rangle = \operatorname{sign}\pi|0\rangle$*.*

Here $\operatorname{sign}\pi$ *denotes the* sign *of a permutation* $\pi \in S_n$*, defined to be* -1 *for transpositions ("swaps")* $i \leftrightarrow j$*. It is extended to arbitrary permutations by the requirement that* $\operatorname{sign}\pi\tau = (\operatorname{sign}\pi)(\operatorname{sign}\tau)$*. (This assignment is well-defined, as you may verify, e.g., in the special case* S_3*.)*

Now consider the representation $\mathcal{H} = \mathbb{C}^3$*, with* $R_\pi|i\rangle = |\pi(i)\rangle$*. It is not itself irreducible. However:*

(iii) The invariant subspace

$$W_{\ydiagram{2,1}} = \{\alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle : \alpha + \beta + \gamma = 0\} \subseteq \mathbb{C}^3$$

is a two-dimensional irreducible representation of S_3*.*

Its orthogonal complement is $W_{\ydiagram{2,1}}^\perp = \mathbb{C}(|0\rangle + |1\rangle + |2\rangle) \cong W_{\ydiagram{3}}$*. Hence:*

$$\mathbb{C}^3 \cong W_{\ydiagram{2,1}} \oplus W_{\ydiagram{3}}$$

The curious labeling of the irreps will become more clear when we discuss Schur-Weyl duality (see remark 7.2).

An important tool for us is the following mathematical result, known as *Schur's lemma*.

Lemma 3.2 (Schur). *Let $J: \mathcal{H} \to \mathcal{H}'$ be an intertwiner between* irreducible *representations R_g, R'_g.*

(i) Either J is invertible (and hence $\mathcal{H} \cong \mathcal{H}'$) or $J = 0$.

(ii) If $\mathcal{H} = \mathcal{H}'$ and $R_g = R'_g$ then $J \propto \mathbb{1}_{\mathcal{H}}$ (i.e., any self-intertwiner is necessarily a multiple of the identity operator).

Schur's lemma shows that intertwiners between irreducible representations are rigidly determined. In particular, there are no nonzero intertwiners between inequivalent irreducible representations. We will not prove this result – you are encouraged to look it up in your favorite textbook (e.g., Fulton and Harris, 2013) – but we will profitably use it many times in this class.

Normal forms of representations

Now suppose that someone handed us a list of irreducible representations of a group G. Let us write V_j for the Hilbert space, $R_g^{(j)}$ for the operators, and j runs over some index set J that labels the different irreps. We assume that the list is *complete* (i.e., that any other irreducible representation is equivalent to some V_j) and that it is *irredundant* (i.e., that $V_j \ncong V_{j'}$ if $j \neq j'$). We just saw two such lists for $G = \mathrm{SU}(2)$ and $G = S_3$, respectively.

Then, if $\mathcal{H}$ is an arbitrary representation of G, we can first decompose as in eq. (3.1). Since each $\mathcal{H}_k$ in eq. (3.1) is irreducible, it must be equivalent to some V_j – say $\mathcal{H}_k \cong V_{j_k}$. Thus:

$$\mathcal{H} \cong V_{j_1} \oplus \ldots \oplus V_{j_m} \tag{3.2}$$

Suppose that n_j is the number of times that V_j appears in this list, i.e., $n_j = \#\{k : j_k = j\}$. Let us reorder (3.2) according to the different values of j:

$$\mathcal{H} \cong \bigoplus_{j \in J} \underbrace{V_j \oplus \ldots \oplus V_j}_{n_j \text{ times}} \tag{3.3}$$

The numbers n_j are uniquely determined – as a consequence of Schur's lemma! They fully characterize the representation $\mathcal{H}$, up to equivalence. A useful alternative way of writing down the decomposition (3.3) is as follows:

$$\mathcal{H} \cong \bigoplus_{j \in J} V_j \otimes \mathbb{C}^{n_j}, \tag{3.4}$$

where G acts on the right-hand side by the block-diagonal matrices

$$\bigoplus_k R_g^{(j)} \otimes \mathbb{1}_{n_j}.$$

(We use the notation $\oplus$ to stress that they are block diagonal with respect to the direct sum decomposition of the Hilbert space that they act on, i.e., eq. (3.4).) We may think of eq. (3.3) or eq. (3.4) as a "normal form" of the representation $\mathcal{H}$.

Remark. *The fact that unitary representation $\mathcal{H}$ can be brought into a normal form is completely analogous to how, e.g., a unitary or Hermitian matrix can always be diagonalized.*

Representation theory tells us about the list of irreducible representations for a given group G and how to determine the decomposition (3.3) or (3.4) of a representation into its irreducible pieces (in particular, how to calculate the numbers n_j).

Proof of the resolution of the identity for the symmetric subspace

Schur's lemma allows us to at last deduce eq. (2.6). To see this, we first observe that the space

$$(\mathbb{C}^d)^{\otimes n}$$

is not only a representation of S_n, as discussed in section 2.2, but also of the unitary group $U(d)$. Its elements are the unitary $d \times d$-matrices U, and its representation on $(\mathbb{C}^d)^{\otimes n}$ is defined as follows:

$$T_U = (U \otimes \ldots \otimes U) = U^{\otimes n}$$

Next week, we will learn much more about the way $(\mathbb{C}^d)^{\otimes n}$ decomposes with respect to the groups S_n and $U(d)$. For today, we only note that the two group actions commute:

$$R_\pi T_U = T_U R_\pi, \quad \text{or} \quad [R_\pi, T_U] = 0. \tag{3.5}$$

Let us verify this explicitly:

$$\begin{aligned} &R_\pi T_U(|\psi_1\rangle \otimes \ldots \otimes |\psi_1\rangle) = R_\pi(U|\psi_1\rangle \otimes \ldots \otimes U|\psi_1\rangle) \\ &= U|\psi_{\pi^{-1}(1)}\rangle \otimes \ldots \otimes U|\psi_{\pi^{-1}(n)}\rangle = T_U R_\pi(|\psi_1\rangle \otimes \ldots \otimes |\psi_1\rangle). \end{aligned}$$

Equation (3.5) implies at once that the symmetric subspace $\mathrm{Sym}^n(\mathbb{C}^d)$ is an invariant subspace for $U(d)$. Indeed, if $|\Phi\rangle \in \mathrm{Sym}^n(\mathbb{C}^d)$ then $R_\pi(T_U|\Phi\rangle) = T_U(R_\pi|\Phi\rangle) = T_U|\Phi\rangle$ and so $T_U|\Phi\rangle \in \mathrm{Sym}^n(\mathbb{C}^d)$.

Importantly, the symmetric subspace is in fact an irreducible representation of $U(d)$. You will show this in problem 2.3. It is now easy to see that the operator Π'_n defined in eq. (2.6) is equal to the projector onto the symmetric subspace. First, note that eq. (2.7) asserts precisely that Π'_n is a self-intertwiner, i.e., $T_U\Pi'_n = \Pi'_n T_U$ (this follows from the invariance of the integral under substituting $|\psi\rangle \mapsto U|\psi\rangle$). Second, note that Π'_n is supported only on the symmetric subspace. We may therefore safely think of Π'_n as an operator from $\mathrm{Sym}^n(\mathbb{C}^d)$ to $\mathrm{Sym}^n(\mathbb{C}^d)$. But since the symmetric subspace is irreducible, Schur's lemma tells us that Π'_n must be proportional to the identity operator on $\mathrm{Sym}^n(\mathbb{C}^d)$, i.e., to Π_n. Since moreover

$$\operatorname{tr}\Pi'_n = \binom{n+d-1}{n} \int d\psi \underbrace{\operatorname{tr}\left[|\psi\rangle^{\otimes n}\langle\psi|^{\otimes n}\right]}_{=1} = \binom{n+d-1}{n} = \operatorname{tr}\Pi_n,$$

we conclude that $\Pi_n = \Pi'_n$.

3.2 Density operators and mixed states

Before we proceed with entanglement and symmetries, let us talk a bit about ensembles of quantum states. Many of you know density operators and partial traces, but I hope this might be a good reminder for everyone.

Suppose that $\{p_i, |\psi_i\rangle\}$ is an ensemble of quantum states on some Hilbert space $\mathcal{H}$, i.e., we have the state $|\psi_i\rangle$ with probability p_i. If X is an observable then we can compute its expectation value by

$$\langle X\rangle = \sum_i p_i \langle\psi_i|X|\psi_i\rangle = \sum_i p_i \operatorname{tr}\left[|\psi_i\rangle\langle\psi_i| X\right] = \operatorname{tr}\left[\underbrace{\sum_i p_i |\psi_i\rangle\langle\psi_i|}_{=:\rho} X\right] = \operatorname{tr}[\rho X].$$

The operator ρ is called a *density operator*, or a *density matrix*, or simply a *quantum state* on $\mathcal{H}$. It safisfies $\rho \geq 0$ and $\operatorname{tr}\rho = 1$, and any such operator arises from some ensemble of quantum states (think of the spectral decomposition!). The *Born rule* for density operators reads

$$\Pr(\text{outcome } x) = \operatorname{tr}[\rho Q_x],$$

as follows from our preceding calculation.

If $\rho = |\psi\rangle\langle\psi|$ then we say that it is a *pure state* (and it is not uncommon to simply write $\rho = \psi$ in this case). Otherwise, ρ is called a *mixed state* (but we will often be sloppy and say "mixed state" when we really should say "density operator"). Note that ρ is pure if and only if $\operatorname{rk}\rho = 1$, or if $\rho^2 = \rho$, or if the eigenvalue spectrum is $\{1, 0\}$.

Example 3.3 (Warning!)**.** *In general the ensemble that determines a density operator is* not *unique. E.g.,* $\tau = \mathbb{1}/2$ *can be written in an infinite number of ways:*

$$\tau = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|) = \dots.$$

The states $\tau_{\mathcal{H}} = \mathbb{1}_{\mathcal{H}}/\dim\mathcal{H}$ are known as *maximally mixed states*. They are the analogues of uniform distributions in probability theory.

More generally, if $p(x_1, \dots, x_n)$ is a probability distribution then we may consider the ensemble $\{p(x_1, \dots, x_n), |x_1\rangle \otimes \dots \otimes |x_n\rangle\}$. The corresponding density operator is

$$\rho_{X_1,\dots,X_n} = \sum_{x_1,\dots,x_n} p(x_1, \dots, x_n)\,|x_1\rangle\langle x_1| \otimes \dots \otimes |x_n\rangle\langle x_n| \tag{3.6}$$

and we call such a state a *classical state*. If all probabilities $p(x_1, \dots, x_n)$ are the same then $\rho_{X_1,\dots,X_n}$ is a maximally mixed state, $\rho = \tau$. In a later problem set, you will explore more generally how classical probability theory can be embedded into quantum mechanics.

In quantum physics, density operators arise in a number of places: As statistical ensembles (e.g., Gibbs states in statistical quantum physics), when describing noisy sources, . . . – but importantly, also when describing the state of subsystems, as we will discuss in the following.

Reduced density matrices and partial trace

Suppose that ρ_{AB} is a quantum state on $\mathcal{H}_A \otimes \mathcal{H}_B$ and X_A an observable on $\mathcal{H}_A$. The axioms of quantum mechanics tell us $X_A \otimes \mathbb{1}_B$ is the appropriate observable on the joint system $\mathcal{H}_A \otimes \mathcal{H}_B$. Let's calculate the expectation value of this observable in the state ρ_{AB}:

$$\begin{aligned}
\langle X_A\rangle &= \operatorname{tr}[\rho_{AB}(X_A \otimes \mathbb{1}_B)] = \sum_{a,b}\langle a,b|\rho_{AB}(X_A \otimes \mathbb{1}_B)|a,b\rangle \\
&= \sum_{a,b}(\langle a| \otimes \langle b|)|\rho_{AB}(X_A \otimes \mathbb{1}_B)(|a\rangle \otimes |b\rangle) \\
&= \sum_{a,b}\langle a|\,(\mathbb{1}_A \otimes \langle b|)\rho_{AB}(X_A \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes |b\rangle)\,|a\rangle \\
&= \sum_{a,b}\langle a|\,(\mathbb{1}_A \otimes \langle b|)\,\rho_{AB}\,(\mathbb{1}_A \otimes |b\rangle)\,X_A|a\rangle \\
&= \sum_{a}\langle a|\underbrace{\sum_{b}(\mathbb{1}_A \otimes \langle b|)\,\rho_{AB}\,(\mathbb{1}_A \otimes |b\rangle)}_{=:\operatorname{tr}_B[\rho_{AB}]}X_A|a\rangle
\end{aligned}$$

The operation tr_B just introduced is called the *partial trace* over B. If ρ_{AB} is a quantum state, then $\mathrm{tr}_B[\rho_{AB}]$ is called the *reduced density operator* or the *reduced density matrix* ρ_A of ρ_{AB}. We will often denote it by $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ (even though this can at times seem ambiguous). Dually, ρ_{AB} is said to be an *extension* of ρ_A. By construction,

$$\mathrm{tr}[\rho_{AB}(X_A \otimes \mathbb{1}_B)] = \mathrm{tr}[\rho_A X_A], \tag{3.7}$$

and so the reduced density operator ρ_A contains all information necessary to evaluate observables on A. It therefore faithfully describes the state of the subsystem A.

We can also compute partial traces of operator that are not quantum states: If M_{AB} is an arbitrary operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ then its partial trace over B is defined just as before,

$$\mathrm{tr}_B[M_{AB}] = \sum_b (\mathbb{1}_A \otimes \langle b|)\, M_{AB}\, (\mathbb{1}_A \otimes |b\rangle)\,.$$

(However, if M_{AB} is not a state then we will *never* denote this partial trace by M_A.)

The following useful rule tells us how to compute partial traces of tensor product operators $M_A \otimes N_B$ and justifies the term "partial trace":

$$\mathrm{tr}_B[M_A \otimes N_B] = M_A \,\mathrm{tr}[N_B] \tag{3.8}$$

It follows directly from the definition:

$$\mathrm{tr}_B[M_A \otimes N_B] = \sum_b (\mathbb{1}_A \otimes \langle b|)\,(M_A \otimes N_B)\,(\mathbb{1}_A \otimes |b\rangle) = M_A \sum_b \langle b|N_B|b\rangle = M_A \,\mathrm{tr}[N_B].$$

Other useful properties are

- $\mathrm{tr}_B[(M_A \otimes \mathbb{1}_B)X_{AB}(M'_A \otimes \mathbb{1}_B)] = M_A \,\mathrm{tr}_B[O_{AB}]M'_B$,
- $\mathrm{tr}_B[(\mathbb{1} \otimes M_B)O_{AB}] = \mathrm{tr}_B[O_{AB}(\mathbb{1} \otimes M_B)]$.

Remark. *A useful convention that you will often find in the literature is that tensor products with the identity operator are omitted. E.g., instead of $X_A \otimes \mathbb{1}_B$ we would write X_A, since the subscripts already convey the necessary information. Thus, instead of eqs. (3.7) and (3.8) we would write*

$$\mathrm{tr}[\rho_{AB}X_A] = \mathrm{tr}[\rho_A X_A],$$
$$\mathrm{tr}_B[M_A N_B] = M_A \,\mathrm{tr}[N_B]$$

which is arguably easier to read.

Example (Warning!)**.** *Even if ρ_{AB} is a pure state, ρ_A can be mixed. For example, consider the maximally entangled state $|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Then,*

$$\begin{aligned}\rho_{AB} = |\psi\rangle\langle\psi|_{AB} &= \frac{1}{2}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|) \\ &= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|) \\ &= \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)\,,\end{aligned}$$

and so, using eq. (3.8),

$$\rho_A = \mathrm{tr}_B[|\psi\rangle\langle\psi|_{AB}] = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)\,.$$

Thus ρ_A is a mixed state – in fact, the maximally mixed state τ_A introduced previously in example 3.3.

The preceding example was not an accident. Every pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ has a so-called *Schmidt decomposition*

$$|\psi\rangle_{AB} = \sum_i s_i |e_i\rangle_A \otimes |f_i\rangle_B ,$$

where $s_i > 0$ and the $|e_i\rangle_A$ and $|f_i\rangle_B$ are sets of orthonormal vectors in $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Note:

$$\rho_A = \sum_i s_i^2 |e_i\rangle\langle e_i|_A \quad \text{and} \quad \rho_B = \sum_i s_i^2 |f_i\rangle\langle f_i|_B .$$

Thus the eigenvalues of the reduced density matrices are directly related to the cofficients s_i.

The Schmidt decomposition is a very important tool that we already briefly met in the fine-print of lecture 1. For one, it helps us to understand entanglement in pure states: E.g., if $|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B$ is a product state then the reduced density matrices are pure. Conversely, if the reduced density matrices of a pure state $|\psi\rangle_{AB}$ are mixed then this is a signature of entanglement. You will discuss this in more detail on problem 2.1. (This also justifies why quantities such as *entanglement entropies* that some of you might already know might be good entanglement measures (only) for pure states.)

We mention two last important facts that you will prove in problem 2.2:

(i) Any mixed state ρ_A has a *purification*: That is, there exists a pure state $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, with $\mathcal{H}_B$ an auxiliary Hilbert space, such that

$$\rho_A = \mathrm{tr}_B[|\psi_{AB}\rangle\langle\psi_{AB}|].$$

Remark. *This justifies why in lecture 1 we were allowed to only consider quantum strategies that involved pure states and observables. At the expense of adding an auxiliary Hilbert space, we can always replace mixed states by pure states and generalized measurements by measurements of observables (you proved the latter in problem 1.4).*

(ii) If $\rho_A = |\psi\rangle_A\langle\psi|_A$ is pure then any extension ρ_{AB} is necessarily a product, i.e., $\rho_{AB} = \rho_A \otimes \rho_B$ – whether ρ_{AB} is pure or mixed. We already mentioned this fact when discussing the privacy of random bits in lecture 1.

Bibliography

William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.