

## 2.

I motivate representation theory from scratch, focusing on the Maschke–Peter–Weyl theorem. All vector spaces will be over the complex numbers  $\mathbb{C}$ .

Throughout mathematics, we find the motif of “atoms” or “building blocks.” For instance, every whole number can be decomposed as a product of prime numbers. A fancier example: If a linear operator  $T$  acts on a vector space  $V$ , then to *diagonalize*  $T$  means to decompose  $V$  into a direct sum of eigenlines for  $T$ . Doing so clarifies the geometry of  $T$ , since it scales each eigenline by the associated eigenvalue.

Representation theory can be viewed as a generalization of the problem of finding eigenspaces for linear operators. But, instead of starting with a single linear operator on a fixed vector space, we start with an abstract *group*, i.e., a set  $G$  equipped with a composition law  $\star : G \times G \rightarrow G$  that satisfies certain axioms (associativity, identity, inverses). A *representation* of  $G$  is a choice of vector space  $V$ , together with a linear operator  $\rho(g) : V \rightarrow V$  for each  $g \in G$ , such that

$$\rho(g \star h) = \rho(g) \circ \rho(h)$$

for all  $g, h \in G$ . That is,  $\rho$  transforms the composition law of  $G$  into the composition of operators. If  $G$  is equipped with a topology that isn’t discrete, then  $\rho$  is required to satisfy a continuity property that we won’t discuss here.

If  $G$  is abelian, meaning  $g \star h = h \star g$  for all  $g, h$ , then we can simultaneously diagonalize all the operators  $\rho(g)$ . In the language of representation theory, we’ve decomposed  $(V, \rho)$  into *subrepresentations*: namely, the simultaneous eigenlines. In general, a subrepresentation is defined as a subspace  $V' \subseteq V$  such that  $\rho(g)V' \subseteq V'$  for all  $g$ , which needn’t be a simultaneous eigenspace of the  $\rho(g)$ . A (nonzero) representation  $(V, \rho)$  is *irreducible*, or *simple*, iff its only subrepresentations are the zero subspace and itself.

If  $G$  is not abelian, then it may not be possible to diagonalize the  $\rho(g)$  simultaneously. Nonetheless, there’s a more general statement where irreducible subrepresentations take the place of eigenlines.

**Theorem 2.1** (Semisimplicity). *Suppose that  $G$  is compact (e.g., finite).*

- (1) (Maschke). *If  $V$  is finite-dimensional, then  $(V, \rho)$  is a direct sum of irreducible subrepresentations.*
- (2) (Peter–Weyl). *More generally, if  $V$  is a Hilbert space and  $\rho(g)$  is unitary for all  $g \in G$ , then there is a dense subspace of  $(V, \rho)$  that is a direct sum of finite-dimensional irreducible subrepresentations.*

Below, the first two examples illustrate Maschke, while the last two illustrate Peter–Weyl. The first and third are abelian, while the second and fourth are nonabelian.

**Example 2.2.** *If  $T : V \rightarrow V$  is an operator on a finite-dimensional vector space, and  $T^m$  is the identity map for some integer  $m > 0$ , then  $T$  is diagonalizable.*

Here, we set  $G = \mathbf{Z}/m\mathbf{Z}$ , the additive group of integers modulo  $m$ . We can turn  $V$  into a representation of  $G$  by setting  $\rho(n) = T^n$ . Using a famous result called *Schur's lemma*, we can check that any irreducible representation of an abelian group must be 1-dimensional. Since  $G$  is, moreover, finite, Theorem 2.1 now shows that  $(V, \rho)$  decomposes into a direct sum of 1-dimensional subrepresentations, i.e., lines that are stable under the operation of  $G$ . Each such line is an eigenline of  $T$ .

**Example 2.3.** Let  $X \subseteq \mathbf{C}^d$  be a set of vectors that is stable under any permutation  $\sigma$  of the coordinates, i.e.,  $(x_1, \dots, x_d) \in X \implies (x_{\sigma(1)}, \dots, x_{\sigma(d)}) \in X$ . If the span of  $X$  is nonzero, then it must be one of the following:

- (1) The subspace where  $x_1 = \dots = x_d$ .
- (2) The subspace where  $x_1 + \dots + x_d = 0$ .
- (3)  $\mathbf{C}^d$  itself.

Here, we take  $G$  to be the group of permutations of the set  $\{1, \dots, d\}$ , also known as the  $d$ th symmetric group, and  $V = \mathbf{C}^d$ . For every permutation  $\sigma$ , we define the linear operator  $\rho(\sigma)$  by setting

$$\rho(\sigma)(x_1, \dots, x_d) = (x_{\sigma(1)}, \dots, x_{\sigma(d)}).$$

Let  $V'$  and  $V''$  be the subspaces in items (1) and (2), respectively. Then  $V = V' \oplus V''$  as vector spaces, and in fact,  $V'$  and  $V''$  are both stable under  $\rho(\sigma)$  for all  $\sigma$ , so they form subrepresentations of  $(V, \rho)$ . Since  $V'$  is 1-dimensional, it must be irreducible. It is harder, but still possible, to show that  $V''$  is also irreducible. These results amount to the statement about  $X$ . They also imply that  $V = V' \oplus V''$  is the decomposition predicted by Theorem 2.1.

**Example 2.4 (Fourier Series).** If  $f : [0, 1) \rightarrow \mathbf{C}$  is square-integrable in the sense that  $\|f\|_2 = \int_0^1 |f|^2 dx < \infty$ , then there are constants  $a_n \in \mathbf{C}$  such that

$$\sum_{-N \leq n \leq N} a_n e_n \xrightarrow{\|\cdot\|_2} f \quad \text{as} \quad N \rightarrow \infty,$$

where  $e_n(x) = e^{2\pi i n x}$ .

Here, we set  $G = \mathbf{R}/\mathbf{Z}$ , the additive group of real numbers modulo 1, and  $V = L^2(\mathbf{R}/\mathbf{Z})$ , the Hilbert space of square-integrable functions on  $G$ . (Note that  $[0, 1)$  is a fundamental domain for  $\mathbf{R}/\mathbf{Z}$ .) We define  $\rho$  by

$$(\rho(y)f)(x) = f(x - y)$$

for all  $x, y \in G$  and  $f \in L^2(G)$ . Then the 1-dimensional subspace  $\mathbf{C}e_n \subseteq V$  forms a subrepresentation of  $(V, \rho)$  because  $\rho(y)e_n = e^{-2\pi i n y}e_n$  for all  $y$ . It turns out that the direct sum  $V^\circ = \bigoplus_n \mathbf{C}e_n$  is dense in  $V$  with respect to the norm  $\|\cdot\|_2$ , so the function  $f : [0, 1) \rightarrow \mathbf{C}$ , viewed as a vector in  $V$ , is the limit of a sequence of vectors  $f_N \in V^\circ$ . The scalar  $a_n$  is the projection of  $f_N$  onto  $\mathbf{C}e_n$  when  $N \gg n$ .

To state the next example, we need more terminology. In coordinates  $x, y, z$  on  $\mathbf{R}^3$ , the *Laplacian* is the differential operator  $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ . A twice-differentiable function  $f : \mathbf{R}^3 \rightarrow \mathbf{C}$  is *harmonic* iff  $\Delta f = 0$ . We let  $A_\ell$  be the vector space of harmonic polynomials in  $x, y, z$  of homogeneous degree  $\ell$ :

$$A_0 = \mathbf{C}, \quad A_1 = \mathbf{C}\langle x, y, z \rangle, \quad A_2 = \mathbf{C}\langle xy, xz, yz, x^2 - y^2, x^2 - z^2 \rangle, \quad \text{etc.}$$

Next, we let  $H_\ell$  be the vector space of functions on the 2-dimensional sphere  $S^2 \subseteq \mathbf{R}^3$  that arise from elements of  $A_\ell$  by restriction of domain. The elements of  $H_\ell$  are known as the *spherical harmonics of degree  $\ell$* .

It turns out that  $\dim A_\ell = \dim H_\ell = 2\ell + 1$ . Laplace introduced explicit functions  $Y_\ell^m : S^2 \rightarrow \mathbf{C}$ , with  $-\ell \leq m \leq \ell$ , that together form a basis of  $H_\ell$ .

**Example 2.5 (Spherical Harmonics).** (1) For any choice of  $\ell, m$  with  $|m| \leq \ell$ , the vector space  $H_\ell$  is spanned by the functions we get by rotating  $Y_\ell^m$ . Conversely, any such function belongs to  $H_\ell$ .

(2) If  $f : S^2 \rightarrow \mathbf{C}$  is square-integrable, then there are constants  $a_\ell^m \in \mathbf{C}$  such that

$$\sum_{0 \leq \ell \leq N} \sum_{-\ell \leq m \leq \ell} a_\ell^m Y_\ell^m \xrightarrow{\|\cdot\|_2} f \quad \text{as} \quad N \rightarrow \infty,$$

where  $\|f\|_2 = \int_{S^2} |f|^2 dA$ .

Here, we set  $G = \text{SO}(3)$ , the group of  $3 \times 3$  rotation matrices. Explicitly,

$$\text{SO}(3) = \{g \in \text{Mat}^{3 \times 3}(\mathbf{R}) : g^t g = I\},$$

where  $I$  is the identity matrix and  $g^t$  is the transpose of  $g$ . We take  $V = L^2(S^2)$  and

$$(\rho(g)f)(x, y, z) = f(g^{-1}(x, y, z)),$$

where  $g^{-1}(x, y, z)$  means we apply the matrix  $g^{-1}$  to the vector  $(x, y, z)$ . As every continuous function on  $S^2$  is square-integrable,  $H_\ell$  is a subspace of  $V$  for all  $\ell \geq 0$ . Item (1) says that, more strongly,  $H_\ell$  is an irreducible subrepresentation of  $(V, \rho)$ . Item (2) says that the direct sum  $\bigoplus_\ell H_\ell$  is dense in  $V$ .

Here are two examples where  $G$  is not compact and the conclusion of Theorem 2.1 fails, but a weaker conclusion does hold.

- (1) Take  $G = \mathbf{Z}$ . If  $\rho(1)$  is not diagonalizable, then  $V$  won't be a direct sum of irreducible subrepresentations. But the *Jordan normal form* of  $\rho(1)$  essentially encodes how the generalized eigenspaces of  $\rho(1)$  are *filtered* by irreducibles.
- (2) Take  $G = \mathbf{R}$  and  $V = L^2(\mathbf{R})$ . Then  $V$  is not a direct sum, but a *direct integral* of 1-dimensional subrepresentations. This integral is essentially described by the *Fourier transform* on  $L^2(\mathbf{R})$ .