

MATH 210B. WHY STUDY REPRESENTATION THEORY?

1. MOTIVATION

Books and courses on group theory often introduce groups as purely abstract algebraic objects, but in practice groups G tend to arise through their actions on other things: a manifold, a molecule, solutions to a differential equation, solutions to a polynomial equation, and so on. There are often vector spaces V naturally attached to such data (as we will see below), and that in turn gives rise to a linear action ρ of G on V . The purpose of representation theory is to understand the ways in which G can act on vector spaces (subject to various appropriate hypotheses), and especially the following two basic questions:

- (i) Does V have nonzero proper subspaces stable by the G -action, and if so then how do we detect their presence?
- (ii) Can we classify the cases in which there is no such nonzero proper G -stable subspace (the “irreducible case”)?

It is a remarkable fact that (ii) above often has a reasonable answer: typically a given G only has a handful of irreducible representations in a given dimension (a priori it might seem like it could have a tremendous number), and one can say a lot about their properties. This imposes seriously non-trivial constraints on how G can act even in the setting of (i). (When G is a non-compact Lie group then the V of interest tend to be infinite-dimensional.)

The consequences of these constraints can only be fully appreciated with experience, and we will get some glimpse of this in our study of the basics of representation theory. What we will see in this course is just the tip of the iceberg. The purpose of this handout is to provide a wide range of contexts in which linear actions of groups on vector spaces arise in mathematics, and ways in which answers to (i) and (ii) are relevant to interesting issues.

Example 1.1. Suppose V is a k -vector space with $\text{char}(k) \neq 2$. The space $V^{\otimes 2}$ has a natural action by the symmetric group \mathfrak{S}_2 on 2 letters under which $\iota := (12)$ acts by swapping tensor factors: $\rho(\iota) : v \otimes v' \mapsto v' \otimes v$ for all $v, v' \in V$. The space $V^{\otimes 2}$ naturally decomposes into a direct sum of so-called *symmetric* and *anti-symmetric* tensors. That is, if $(V^{\otimes 2})^+$ denotes the space of ι -fixed vectors in $V^{\otimes 2}$ and $(V^{\otimes 2})^-$ denotes the -1 -eigenspace for ι in $V^{\otimes 2}$ then we have

$$V^{\otimes 2} = (V^{\otimes 2})^+ \oplus (V^{\otimes 2})^-$$

via the formula $t = (1/2)(t + \rho(\iota)(t)) + (1/2)(t - \rho(\iota)(t))$. This is a decomposition into \mathfrak{S}_2 -stable subspaces.

For $m > 2$, does $V^{\otimes m}$ admit an analogous decomposition? There is a natural action on $V^{\otimes m}$ by the symmetric group \mathfrak{S}_m on m letters via

$$\rho_m(\sigma) : v_1 \otimes \cdots \otimes v_m \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(m)}$$

for $\sigma \in \mathfrak{S}_m$ and $v_1, \dots, v_m \in V$, and the useful generalization of the decomposition of $V^{\otimes 2}$ involves an advanced understanding of the representation theory of \mathfrak{S}_m (when $\#\mathfrak{S}_m = m! \in k^\times$, especially $\text{char}(k) = 0$). The case $m = 2$ was able to be worked out by bare hands without awareness of representation theory simply because the group \mathfrak{S}_2 is abelian and we'll see that the representation theory of finite abelian groups is very explicit and elementary

(related to Fourier analysis for finite abelian groups). For $m > 2$ the group \mathfrak{S}_m is non-abelian, due to which one really cannot proceed without some awareness of the answer to (i) and (ii) for $G = \mathfrak{S}_m$.

The study of the \mathfrak{S}_m -action on $V^{\otimes m}$ (at least when $m! \in k^\times$) amounts to the theory of “Schur functors” that arises in algebraic combinatorics and a variety of other areas of math.

Although representation theory of finite groups G is literally the study of homomorphisms $\rho : G \rightarrow \mathrm{GL}(V)$ for finite-dimensional vector spaces V (often over \mathbf{C} , but also over \mathbf{R} or \mathbf{Q} or even finite fields), that gives a completely misguided impression about the purpose of the subject (akin to saying that Galois theory is the study of roots of polynomials entirely misses the deeper aspects such as the structure of field extensions and as a tool to explore phenomena connected with the effect of ground field extension).

In particular, *the purpose of representation theory is not to “represent G in terms of matrices”!* It is a fact of experience (to be illustrated below) that groups acting on vector spaces arise in many situations, and representation theory aims to study the structure of such actions and how they behave under operations of linear algebra (see §3 for tensor products). We say ρ is *faithful* if $\ker \rho = 1$, but one is often studying non-faithful ρ (even the trivial action plays a role in the general theory).

Also, many questions about general finite groups can often be reduced to the case of simple groups, and it is a remarkable aspect of the classification of finite simple groups that nearly all of them arise as “matrix groups over finite fields”. For this reason, it really is a natural and useful task to deeply study the \mathbf{C} -linear representation theory of finite groups such as $G = \mathrm{GL}_n(\mathbf{F}_q)$; the fact that this group happens to be literally defined via “matrices” (over \mathbf{F}_q) in no way diminishes the importance of understanding its \mathbf{C} -linear representation theory (and in fact its concrete matrix description provides a useful grip on the subgroup structure that in turn is very helpful in attempts to answer question (ii) above for this G).

2. SAMPLE CONTEXTS

Example 2.1. Many problems in physics concern systems with rotational symmetry, and consequently the space of solutions to the system of PDE’s describing the physics inherits a natural action by the connected compact special orthogonal group $\mathrm{SO}(3) \subset \mathrm{GL}_3(\mathbf{R})$ of orientation-preserving rigid motions of \mathbf{R}^3 preserving the origin. A general understanding of the structure and properties of representations of compact Lie groups explains conceptually the existence and relations among many “special function” solutions to such PDE’s.

A prominent instance of this is given by the so-called spherical harmonics $Y_m^\ell(\theta, \varphi) \in L^2(S^2)$ ($\ell \geq 0, -\ell \leq m \leq \ell$) that are smooth eigenfunctions for the spherical Laplacian Δ_{S^2} . The significance of these functions is completely explained by studying the action of $\mathrm{SO}(3)$ on $L^2(S^2)$ induced by the transitive rotation action of $\mathrm{SO}(3)$ on S^2 .

More specifically, general principles in representation theory for compact Lie groups explain why the irreducible $\mathrm{SO}(3)$ -stable closed subspaces of $L^2(S^2)$ must be finite-dimensional, labeled in this case as $\{H_\ell\}_{\ell \geq 0}$ with $\dim H_\ell = 2\ell + 1$, and the subgroup structure of $\mathrm{SO}(3)$ (especially the action of specific “circle subgroups”) picks out preferred lines that orthogonally span each H_ℓ . The functions $\{Y_m^\ell\}_{-\ell \leq m \leq \ell}$ are unit vectors in H_ℓ spanning those distinguished

lines, and considerations with Lie algebras explain why Δ_{S^2} acts on each H_ℓ by the scalar $\ell(\ell + 1)$ (and in fact H_ℓ is the full eigenspace for that eigenvalue).

The representation theory of compact Lie groups also explains why there is a Hilbert direct sum decomposition $L^2(S^2) = \widehat{\bigoplus} H_\ell$. This is a “Fourier decomposition” much as the circle $S^1 = \mathbf{R}/\mathbf{Z} = \text{SO}(2)$ with its natural rotation action on itself has the lines $\mathbf{C}e^{2\pi i\theta} \subset L^2(S^1)$ ($n \in \mathbf{Z}$) as the irreducible S^1 -stable closed subspaces of $L^2(S^1)$ that provide the classical Fourier decomposition $\widehat{\bigoplus} \mathbf{C}e^{2\pi i\theta} = L^2(S^1)$. Representation theory also explains conceptually why S^1 has its “irreducible” Fourier components of dimension 1 in its L^2 -space whereas S^2 has Fourier components of higher dimension: this is ultimately because $\text{SO}(2)$ is commutative whereas $\text{SO}(3)$ is not. We will see an incarnation of this dichotomy between abelian and non-abelian groups when we explore the representation theory of finite groups.

There is a real zoo of identities satisfied by the spherical harmonics Y_m^ℓ (“addition formulas”, formulas for products of spherical harmonics, explicit differential equations that they satisfy, etc.), and these are all explained in a systematic way via representation theory. As an illustration, for $k \leq \ell$ we have $Y_m^k Y_{m'}^\ell = \sum_{\ell-k \leq L \leq \ell+k} C(L, k, \ell, m, m') Y_{m+m'}^L$ with explicit coefficients $C(L, k, \ell, m, m')$ (“Clebsch-Gordan coefficients”), and this is ultimately due to a general isomorphism of $\text{SO}(3)$ -representations

$$H_k \otimes H_\ell \simeq \bigoplus_{j=0}^k H_{k+\ell-j}$$

and the realization of $Y_n^r(\theta, \varphi)$ as the matrix entry for the action of a certain $R(\theta, \varphi) \in \text{SO}(3)$ in a specific model for the irreducible representation H_r of $\text{SO}(3)$. The “addition formulas” for spherical harmonics express the homomorphism property of certain representations, generalizing that irreducible 2-dimensional representations of $S^1 = \mathbf{R}/2\pi\mathbf{Z}$ over \mathbf{R} are precisely $\rho_n : x \mapsto \begin{pmatrix} \cos nx & -\sin nx \\ \sin nx & \cos nx \end{pmatrix}$ for $n \neq 0$ and the equality $\rho_1(x+y) = \rho_1(x)\rho_1(y)$ expresses addition formulas for sin and cos. These matters are discussed in Chapter III (§3, §4.1–§4.4, §8) of the amazing book *Special Functions and the Theory of Group Representations* by N. Vilenkin; later chapters in this book use the representation theory of other physically significant Lie groups (the Lorentz group, the group of Euclidean motions, etc.) to explain a vast array of properties of many special functions of mathematical physics.

In general, theoretical physicists are very experienced at finding symmetry information lurking within PDE’s and using it in conjunction with the representation theory of compact Lie groups to find good coordinate systems in which to rewrite PDE’s so that they can dramatically simplify the appearance and thereby find explicit solutions.

Example 2.2. In the study of molecular structure in chemistry, unlike in many physics problems, the relevant symmetry groups tend to be finite groups rather than “continuous groups” (i.e., positive-dimensional Lie groups) due to the more discrete nature of molecules. Consequently, being able to classify and describe properties of irreducible representations of specific finite groups turns out to be the key to explaining many features of the chemical properties of molecules with not too many atoms. Many books on physical chemistry have a discussion of representation theory of finite groups for this reason, though it is usually presented there in a language that is difficult to read for mathematicians (and the books on representation theory written by mathematicians tend to be unreadable to chemists!).

Example 2.3. If a finite group G acts on a manifold M , then G naturally acts on the (often finite-dimensional) cohomology spaces $H^i(M, \mathbf{R})$ respecting natural operations (such as cup product, and Poincaré duality when G is orientation-preserving). Thus, to the extent we have a good understanding of the representation theory of G over \mathbf{R} , this can impose non-trivial constraints on the cohomology ring of M .

Many interesting manifolds M have a high degree of symmetry in the sense that some Lie group G has a transitive action $G \times M \rightarrow M$ that is C^∞ . For example, via rotations $\mathrm{SO}(n)$ acts transitively on S^{n-1} for $n \geq 2$, via its natural action on \mathbf{R}^n the group $\mathrm{GL}_n(\mathbf{R})$ acts transitively on the so-called Grassmann manifold $\mathrm{Gr}(k, n)$ that parameterizes k -dimensional subspaces of \mathbf{R}^n ($1 \leq k \leq n-1$), and for a non-degenerate quadratic form q on \mathbf{R}^n with possibly mixed signature (i.e., $\sum_{i=1}^r x_i^2 - \sum_{j=r+1}^n x_j^2$) the associated orthogonal group $\mathrm{O}(q)$ of linear automorphisms of \mathbf{R}^n preserving q naturally acts on each (smooth) level set $\{q = c\}$ for $c \in \mathbf{R}^\times$ and this is a transitive action (due to a general theorem of Witt).

It is a general fact that closed subgroups H of Lie groups G are always submanifolds and that the coset space G/H always has a unique manifold structure making $G \rightarrow G/H$ a submersion. Moreover, when G has a smooth transitive action on M then for the stabilizer $H = \mathrm{Stab}_G(m)$ at a point $m \in M$ one can show that the bijective orbit map $G/H \rightarrow M$ defined by $gH \mapsto g(m)$ is a diffeomorphism. Thus, in all situations as above with a transitive G -action, the manifold is in fact a coset space for G equipped with its natural manifold structure. The structure of the cohomology $H^i(G/H, \mathbf{R})$ is often very much informed by knowledge of the representation theory of both G and H .

Many manifolds M that do not at first sight appear to have any symmetry have universal covering spaces that are highly symmetric (for compact oriented 3-manifolds this is one of the lessons of Thurston’s geometrization conjecture proved by Perelman), and due to this the original manifold can often be described as a double coset space $\Gamma \backslash G/K$ for a Lie group G , compact subgroup K (typically with large dimension), and *discrete* subgroup $\Gamma \subset G$. (All compact connected Riemann surfaces arise in this way, by the Uniformization Theorem.) For example, the “space” of lattices in \mathbf{R}^n taken up to rigid motion is naturally parameterized by the points of a manifold given by

$$\mathrm{GL}_n(\mathbf{Z}) \backslash \mathrm{GL}_n(\mathbf{R}) / \mathrm{O}(\sum x_i^2).$$

In such cases, “Matsushima’s formula” describes $H^i(\Gamma \backslash G/K, \mathbf{R})$ in terms of the action of K on the space of smooth functions $C^\infty(\Gamma \backslash G)$ through the right-multiplication action of K on $\Gamma \backslash G$, and this involves much information from the representation theory of K (and the structure of G).

Example 2.4. Many questions on the algebraic side of combinatorics related to symmetric functions of various sorts and positivity properties of various identities among polynomials (“Littlewood-Richardson coefficients” and so on) involve the representation theory of the symmetric group \mathfrak{S}_n . Remarkably, the representation theory of this *finite* group is very closely related to the representation theory of the *connected compact* unitary group $\mathrm{U}(n)$ (linear automorphisms of \mathbf{C}^n preserving the standard hermitian form $h(z, w) = \sum z_j \bar{w}_j$, or in matrix terms $\{M \in \mathrm{GL}_n(\mathbf{C}) \mid M \bar{M}^\top = 1\}$). This link is called *Schur-Weyl duality*, and is a rather special feature of \mathfrak{S}_n not available for typical finite groups.

Other topics in discrete probability, such as random walks on groups and average length of the longest subsequences of a permutation, involve the use of ideas from representation theory; ask Diaconis about this.

Example 2.5. There are significant problems in pure group theory which can only be understood by bringing in techniques from representation theory (or at least whose most meaningful explanation is given in such terms). For example, one of the first results pointing towards the possibility of trying to classify finite simple groups was Burnside’s so-called “ $p^a q^b$ theorem”: every finite group whose size is divisible by exactly 2 primes is solvable. (Of course, the solvability of p -groups is a much more elementary fact in finite group theory.) The original proof of this result in 1904 made essential use of techniques from representation theory, and around 6 decades later a (much more complicated) proof was finally found that did not use representation theory.

Example 2.6. Harmonic analysis is concerned with the study of the G -action on $L^2(G)$ induced by the translation action of G on itself for Lie groups G . The case $G = \mathbf{R}^n$ constitutes much of classical Fourier analysis, and extending the scope of those techniques to the non-abelian setting, especially for the most interesting case of “semisimple” G (such as $\mathrm{SL}_n(\mathbf{R})$ or compact Lie groups such as $\mathrm{SO}(n)$), is the subject of non-abelian harmonic analysis. The focus on $L^2(G)$ might appear to be unusually specific among Hilbert spaces equipped with an isometric G -action, but this particular representation space turns out to play a prominent role even for the study of G -actions on quite general Hilbert spaces V .

Example 2.7. In number theory, representations of Galois groups encode deep information in subtle ways, and employing techniques of representation theory in such situations is a powerful tool.

For example, if $f \in \mathbf{Q}[x]$ is a cubic polynomial with no repeated root then the plane curve E defined by $y^2 = f(x)$ (really its Zariski closure in the projective plane) is called an *elliptic curve* (over \mathbf{Q}), and algebraic geometry provides a natural commutative group structure on the set of points $E(\overline{\mathbf{Q}})$ with coordinates in $\overline{\mathbf{Q}}$. The structure of the torsion in this group is quite concrete: the group $E(\overline{\mathbf{Q}})[n]$ of points killed by n in the group law turns out to be a free module of rank 2 over $\mathbf{Z}/(n)$ (seen by working with $E(\mathbf{C})$ as a genus-1 compact Riemann surface).

By design, the group law on $E(\overline{\mathbf{Q}})$ is given by expressions in rational functions of the coordinates with coefficients in \mathbf{Q} , so it commutes with the action on points by the so-called absolute Galois group $\Gamma_{\mathbf{Q}} := \mathrm{Aut}(\overline{\mathbf{Q}}/\mathbf{Q})$ acting on coordinates. Hence, the $\Gamma_{\mathbf{Q}}$ -action on points yields an action of $\Gamma_{\mathbf{Q}}$ on $E(\overline{\mathbf{Q}})[n] \simeq (\mathbf{Z}/(n))^2$ respecting its additive structure, so this is a representation $\rho_{E,n} : \Gamma_{\mathbf{Q}} \rightarrow \mathrm{GL}_2(\mathbf{Z}/(n))$. Remarkably, such $\rho_{E,n}$ ’s considered for many n at once encode a lot of the important number-theoretic information about E , and it was through the study of these representations that Wiles proved Fermat’s Last Theorem.

3. TENSOR DECOMPOSITION

Let G be a group (no hypotheses on its structure!) and let

$$\rho : G \rightarrow \mathrm{GL}(V), \quad \rho' : G \rightarrow \mathrm{GL}(V')$$

be two representations of G on finite-dimensional nonzero vector spaces V and V' over a field k . The tensor product $V \otimes V'$ inherits a natural G -action via

$$g(v \otimes v') = \rho(g)(v) \otimes \rho'(g)(v')$$

(i.e., the right side is visibly bilinear in (v, v') and so there really is a unique linear endomorphism of $V \otimes V'$ satisfying $v \otimes v' \mapsto \rho(g)(v) \otimes \rho'(g)(v')$ on elementary tensors). It is easy to check that this is indeed a representation of G on $V \otimes V'$, called the *tensor product* of ρ and ρ' , often denoted $\rho \otimes \rho'$.

Assume that ρ and ρ' are irreducible (i.e., neither V nor V' has a nonzero proper subspace stable under the G -action). Typically $\rho \otimes \rho'$ is not irreducible (we will see many instances of this later on), so one is led to ask if it is at least a direct sum of irreducible representations.

Example 3.1. The decomposition of $V \otimes V'$ into a direct sum of irreducible representations is of interest to physicists because for certain compact connected Lie groups G the irreducible continuous representations over $k = \mathbf{C}$ are related to elementary particles, and decomposing tensor product representations expresses physical information related to particle interactions.

Example 3.2. Many applications of representation theory in pure mathematics involve decomposing tensor product representations into a direct sum of irreducible representations. Many identities for special functions are also explained in this way, and calculations in cohomology rings of some important manifolds rest on it too.

When $\text{char}(k) = p > 0$ then it can happen that $V \otimes V'$ is *not* a direct sum of irreducible representations (i.e., it has a nonzero proper G -stable subspace with no G -stable linear complement). For example, if $G = \text{SL}_2(k)$ and $V = k^2$ is the standard 2-dimensional representation of G , $V^{\otimes p}$ equipped with its natural G -action is not a direct sum of irreducible representations. In characteristic 0 the situation is much better:

Theorem 3.3 (Chevalley). *If $\text{char}(k) = 0$ then $V \otimes V'$ is a direct sum of irreducibles.*

When G is finite (or if G is compact and $k = \mathbf{C}$ with the representations continuous) this theorem is not so interesting because we will see that *every* finite-dimensional representation of a finite group in characteristic 0 is a direct sum of irreducible representations.

But in characteristic 0 many infinite groups not arising as finite groups or compact groups admit a plethora of representations that are *not* a direct sum of irreducible representations. (For instance, if $G \subset \text{GL}_n(k)$ is the subgroup of upper-triangular matrices then its standard representation on k^n has no such direct-sum decomposition.) This makes Chevalley's theorem remarkable: no matter what G at all we begin with, when $\text{char}(k) = 0$ the representation $V \otimes V'$ is a direct sum of irreducible representations when V and V' are irreducible.

More remarkable than the result is the proof: we have no information about G , but techniques from affine algebraic geometry and the (non-obvious) general structure theory of Zariski-closed subgroups of GL_n reduce the task to a problem in the representation theory of finite-dimensional Lie algebras in characteristic 0, with which the problem can be solved.

One of the essential features of the case when G is finite with $k = \bar{k}$ (or compact with $k = \mathbf{C}$) is that for irreducible V and V' the decomposition of $V \otimes V'$ into a direct sum of irreducibles can be completely determined: exactly *which* irreducibles occur inside $V \otimes V'$, and with what “multiplicity”.