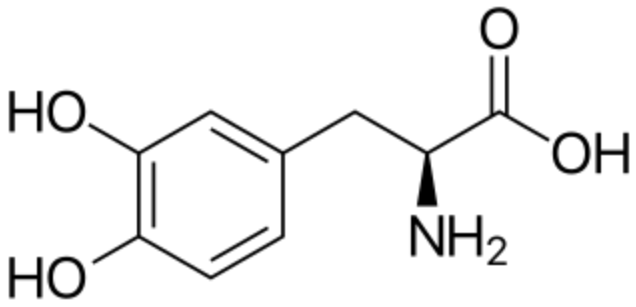
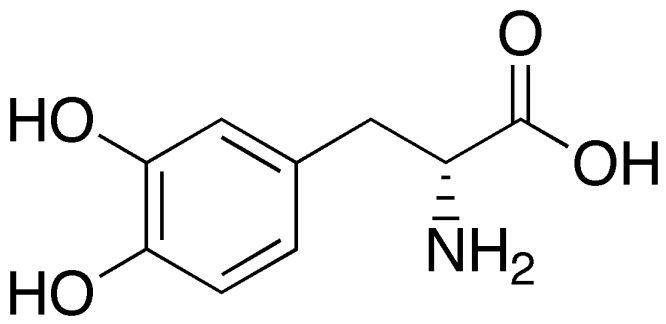
Survey on Representation Theory and Molecular Structure
May Cai

I first came across the use of representation theory in chemistry when I was searching up the character table for the cyclic group on three elements. I found a tool that gave me the characters that I was looking for, but there was a lot more information that I cared about, and about something called the "point group C_3." Furthermore, lower down the list were a couple properties that left me dumbfounded. I saw a mention to chirality, as well as a list of molecules. Why would chemical properties be placed next to a character table? What does representation theory have to do with molecules? I hope to give an answer to that statement in this survey.



L-DOPA (taken from https://en.wikipedia.org/wiki/L-DOPA)

First, the concept of symmetry is important in chemistry. Consider a molecule known as L-DOPA. It is an important amino acid, produced by the human body and used as a drug to treat Parksinson's disease. It has a molecular structure depicted above, and this molecular structure has the important property that it doesn't have mirror symmetry. If we view most of the molecule as lying flat on a plane, then the NH_2 amino group would be "coming out" of the plane, and reflection across the plane would give a different molecule, known as D-DOPA.



D-DOPA (taken from https://en.wikipedia.org/wiki/D-DOPA)

The only difference between these two molecules is that the NH_2 amino group is on the other side of the plane. D-DOPA and L-DOPA are mirror images of one another, but D-DOPA is completely ignored by the human body, and thus "is basically useless." [J20] Thus, whether or not a molecule has such mirror images can greatly change its behaviors in chemical processes, and so it is useful to know these things when we want to predict a molecule's chemical properties. Likewise, rotational symmetry can help us determine other properties of molecules, such as their magnetic dipoles.

Thus, given a molecule, we care about the various symmetries that act on it, and of course the natural thing to do is to consider the group formed by the symmetries that act on it. In addition to the rotational and reflectional symmetries mentioned earlier, chemists also consider inversion, or the map $(x,y,z) \mapsto (-x,-y,-z)$, the composition of rotation around an axis and then reflection across a plane normal to that axis, and of course the identity as the 5 main types of symmetries. As an example, H_2O, for instance, has a rotation by π, and 2 reflections, one lying along the plane of the molecule, and one perpendicular, forming the dihedral group with 4 elements, D_4. Chemists call this the point group of the molecule, as each symmetry operation leaves at least one point fixed, and so D_4 is the point group of H_2O. [P15]

Of course, as these groups arise from operations on three-dimensional space, it is the most natural thing in the world to consider their representations onto Euclidean three-dimensional space. Here is the character table for D_4, as a chemist would see it.

C_{2v}	E	$C_2(z)$	$\sigma_v(xz)$	$\sigma_v(yz)$	linear/rotation	quadratic	cubic
A_1	1	1	1	1	z	x^2, y^2, z^2	z^3, zx^2, zy^2
A_2	1	1	-1	-1	R_z	xy	xyz
B_1	1	-1	1	-1	x, R_y	xz	xz^2, x^3, xy^2
B_2	1	-1	-1	1	y, R_z	yz	yz^2, x^2y, y^3

Character table for D_4 (reproduced from [Gel])

The symbol C_{2v} refers to the group D_4 in Schoenflies notation, a point group notation used primarily by chemists. Specifically, it refers to the group corresponding to a 2-fold rotation with additionally 2 planes of reflection containing the axis of rotation. The identity element is denoted E, $C_2(z)$ is the symbol for rotation by π (specifically along the principal z-axis), and σ_v referrs to reflection across a plane containing the principal axis (z). The parentheticals (xz) and (yz) indicate exactly which axes the plane lies on.

The labels of each character, known as Mulliken symbols, also have a conventional meaning. A and B labels are used when they correspond to 1-dimensional representations, and E and T for 2- and 3-dimensional representations respectively. Furthermore, A is used when said one-dimensional representation is symmetric to the principal axis, and B otherwise, and the subscripts are used to distinguish between representations with otherwise identical Mulliken symbols.

The linear/rotation column is for indicating common sets of basis vectors that produce each representation. For instance, a unit vector along the z-axis, when acted upon by C_{2v}, is

fixed by each group operation. Thus, it gives rise to the identity character A_1. A unit vector along the x-axis is fixed by E and $\sigma_v(xz)$, and reversed by C_2 and $\sigma_v(yz)$, giving rise to the character B_1, and likewise a y-axis unit vector gives rise to the character B_2. And finally, the symbols R_x, R_y, and R_z correspond to special "rotational" vectors. The symbol R_z is the rotational vector collinear with the z-axis but with an orientation of counterclockwise rotation around it. A rotation operation will fix R_z, but any reflection will reverse the orientation of reflection, and thus that the rotational vector R_z gives rise to the character A_2.

The last two columns are for the degree-2 and degree-3 monomials of x, y, and z, respectively. Their entry comes from direct products of characters; the pointwise product of B_1B_2 is A_2, and likewise you'll see that xy is in the row corresponding to the irrep A_2. These are useful for the symmetry properties of electron orbitals (probability distributions for the locations of electrons of the molecule, given certain information about the energy and angular momentum of the molecule.) For instance, the 5 *d* orbitals d_{xy}, d_{xz}, d_{yz}, d_{z^2}, and $d_{x^2-y^2}$ of H_2O have physical shapes with symmetries corresponding to the "quadratic" column of the character table. For instance, the d_{xy} orbital has representation A_2. The orbital $d_{x^2-y^2}$ has as representation the difference of the x^2 and y^2 representations, and likewise the action of C_{2v} on the 7 *f* orbitals of water correspond to linear combinations of the "cubic" column of the character tables.

Of course, these are not the only representations that we might come across. In general, given a representation inspired by a molecule (say, the representation from degrees of freedom of a molecule, where we assign 3 unit vectors to each of the molecules, which gives a 9-dimensional representation for water) we can of course find its component irreps with Schur's orthogonality formula:

$$\frac{1}{|G|}\sum_{g\in G}\chi_{irr}(g)\chi(g^{-1}) = n$$

This formula says that, given a starting representation χ and a target irreducible representation χ_{irr} we can find *n* the number of copies of the irreducible representation that appear in our starting representation, by evaluating that sum over the elements of the group. Of course, since we're working with real numbers (and therefore the field of complex numbers), we can use that

$$\chi(g^{-1}) = \overline{\chi(g)}$$

and the fact that we're working exclusively with the reals to simplify the orthogonality formula to

$$\frac{1}{|G|}\sum_{g\in G}\chi_{irr}(g)\chi(g) = n$$

giving us a relatively simple formula to decompose any representation of a point group (once we have its character table, of course).

This approach does elide over some issues, the most blatant being an assumption that all our representations are real. [P15] In fact, even relatively simple groups like C_3 (the cyclic group on 3 elements, in Schoenflies notation) has as character table the trivial representation and then two complex representations. Since we are working with structures grounded in real space, it is common to simply use the real representations, yielding

C_3	E	C_3	$(C_3)^2$
A	1	1	1
E	1 1	ε ε^2	ε^2 ε

or

C_3	E	C_3	$(C_3)^2$
A	1	1	1
E	2	-1	-1

Character tables for C_3 (reproduced from [Gel])

where $\varepsilon = e^{2\pi i/3}$. To be clear, C_3 in the above tables is overloaded as both the cyclic group on three elements, as well as a generator for that group, and E is both a character (the first column and the third row) as well as the identity element (in the second column of the first row). The

character A is the identity character, and E is a 2-dimensional real character, since we're primarily concerned with real representations. When looking for irreducible components of characters on a character with complex irreps, we can use the simplified formula with the second representation in the second table, and instead divide by 2.

References

[Gel] Gelessus, Achim, *Character tables for chemically important point groups,* Available at http://symmetry.jacobs-university.de/

[J20] Fu, Jieming, private communication, March. 2020.

[P15] Pfennig, Brian William, *Principles of Inorganic Chemistry*, Wiley, 2015.

[SS94] Sakurai, J. J. and San Fu Tuan, *Modern Quantum Mechanics Revised Edition*, Reading, MA: Addison-Wesley Pub., 1994.

School of Mathematics, Georgia Tech, Atlanta GA 30332, USA

Email address: mcai@gatech.edu