

# What are the possible axial point groups for a molecule of a certain size or formula?

Nike Dattani<sup>1,2,\*</sup> and Sichao Xiong<sup>1,†</sup>

<sup>1</sup>HPQC Labs, Waterloo, Canada

<sup>2</sup>HPQC College, Waterloo, Canada

(Dated: 3rd October 2023)

For each axial point group, we determine all possible structures in terms of the number of atoms and the chemical formula. We provide examples of molecules of each size and formula if they are known, and hence we determine the yet-to-be discovered molecules with under five atoms for each point group. We present a set of principles that can be used for proofs about the possibility or impossibility for a certain point group to apply to a molecule of a certain size and/or formula. We also resolve some inconsistencies in a table by P. Atkins, M. Child and C. Phillips, add missing entries to it, and improve it further in a number of ways.

## I. INTRODUCTION

One might wonder why, despite group theory being applied to molecules for about 100 years [19], the questions that are being addressed in this paper do not appear to have been answered yet. Indeed we have asked all living authors of [1–18, 36, 40] if they have ever come across a list of all possible point groups for a 4-atom (tetraatomic) system, and in all cases we were explicitly told that they were unaware of such a list, or we received no response at all. We have determined all possible point groups for systems with *any* number of atoms (not just four), and we have gone a step even further by determining all possible point groups for structures with a given chemical formula.

After searching through all of [1–18, 36, 40] thoroughly, the closest work that we found to the results in this paper, is presented in the table of “shapes” found in the 2008 version of the famous “Tables for Group Theory” document by P. Atkins, M. Child and C. Phillips [1]. Nothing similar was found in any of [1–18, 36, 40], and the table of “shapes” in [1] has several shortcomings which we will address and fix in Section III, before we begin presenting the main results of our work in the following sections.

For investigating the novelty of our work, we believe that our thorough examination through all of [1–18, 36, 40] was a fair survey of the known literature on the topic. The 2007 paper [40] reported a thorough survey of known published works that provided character tables for the  $S_8$  and  $D_{8h}$  point groups, and our list [1–18, 36, 40] includes all of those works, in addition to [40] itself, the papers that cited [40], and also thoroughly selected work on the topic that did not provide  $S_8$  and  $D_{8h}$  character tables.

Our paper focuses on *axial* point groups, which are those belonging to the series  $C_n$ ,  $C_{nv}$ ,  $C_{nh}$ ,  $S_{2n}D_n$ ,  $D_{nh}$ ,  $D_{nd}$  (including  $C_s$  which

is equivalent to  $C_{1h}$  and  $C_{1v}$ , and  $C_i$  which is equivalent to  $S_2$ ). For the *polyhedral* point groups  $T$ ,  $T_d$ ,  $T_h$ ,  $O$ ,  $O_h$ ,  $I$ ,  $I_h$ , we will present even more novel results in a forthcoming paper. For example, in all of the literature that we have seen (including [1–18, 36, 40]), the shape with the smallest number of vertices (or the system containing the smallest number of atoms) for which the point group is  $O$ , is the snub cube which has 24 vertices (or atoms), but within the last month we have discovered that it is possible to construct an object with only 20 vertices (or for the atoms in a 20-atom system to be arranged) such that the point group is  $O$  [44]. Likewise, in all of the literature that we have seen, the shape with the smallest number of vertices for which the point group is  $I$ , is the snub dodecahedron which has 60 vertices, but within the last month we have discovered that it is possible to construct an object with only 50 vertices such that the point group is  $I$  [44].

The next section of this paper will explain the context which made *us* feel compelled to answer the questions that this paper addresses; for any readers that wish to skip the motivation section, the rest of the paper is self-contained and does not require knowledge of any of the material presented there. The second paragraph of this introduction already described the contents of Section III. The diagrams in Section III will make it easier to follow Section IV, which contains the foundation for obtaining our main results which are provided in Section V and summarized in Section VI. Overall our results are as follows (all results appear to be novel or unusually difficult to find in existing places):

- In Section III: A table that presents visual examples of structures for all aforementioned axial point groups for  $2 \geq n \geq 6$  and  $n = \infty$ ,
  - in a more clear way than in [1],
  - with inconsistencies in [1] resolved,
  - with more of the  $n = \infty$  shapes presented than in [1],
  - with structures requiring fewer atoms than in [1] for the  $C_n$  and  $D_n$  series,

\* [nike@hpqc.org](mailto:nike@hpqc.org)

† [sichao@hpqc.org](mailto:sichao@hpqc.org)

- with an extra column showing additional structures (in some cases these are more realistic structures than in [1]) for the  $n = 2$  cases, and
- with a more logical order, especially for pedagogical reasons, than in [1].

• In Section IV: Properties of some point groups that are very helpful for obtaining or explaining new results, and for pedagogical purposes.

• In Section V: For each axial point group, we determine

- the minimum number of atoms (cardinality) required;
- all possible cardinalities allowed (the minimum cardinality is a special case of this); and
- all possible chemical formulas allowed (the list of all possible cardinalities is a special case of this).

• In Section VI: As a consequence of the results in Section V,

- we were able to list in Table III all possible point groups allowed for systems with  $m$  atoms for  $m = 1 - 8$  (and a table with *arbitrary* values of  $m$  can be generated with the results from V but would take up too much space to be presented here);
- we also explicitly listed the specific point groups (and diagrams of example structures) that are possible for systems with the specific formulas  $A_3$ ,  $A_2B$ , and  $ABC$  for triatomic systems and  $A_4$ ,  $A_3B$ ,  $A_2B_2$ ,  $A_2BC$ , and  $ABCD$  for tetraatomic systems (analogous lists for *arbitrary* chemical formulas can also be generated with the results from V but would take up too much space to be presented here);
- we attempted to provide examples of known molecules with each possible point group for each formula  $A_3$ ,  $A_2B$ ,  $ABC$ ,  $A_4$ ,  $A_3B$ ,  $A_2B_2$ ,  $A_2BC$ , and  $ABCD$ , and this led us to the discovery of shapes for which no molecule/complex has yet been synthesized or found, despite the existence of such molecules/complexes being proven to be possible (this was not a surprise for larger structures, but we determined that the minimum number of atoms required for the point group  $D_2$  is 4 but we could not find any example of a 4-atom molecule with the  $D_2$  point group despite the apparent simplicity that would come with such a small molecular size).

Section VIA briefly mentions how the results of this work can be used for forthcoming projects.

## II. MOTIVATION FOR THIS WORK

Our need for answers to the question in the present paper's title, arose during other work. GW100 is a dataset of 100 systems (95 molecules and 5 atoms)

which was introduced in a 2015 paper [38] and by the end of 2020, already more than 100 electronic structure methods/codes were used to calculate energies for all systems in the GW100 dataset. It has therefore become a monumental dataset for benchmarking, and efforts were made in 2021 [20] and 2023 [21] to improve the most accurate calculations that had been done on the system until then [22], and therefore to create one of the best known sets of benchmark data in existence for electronic structure calculations.

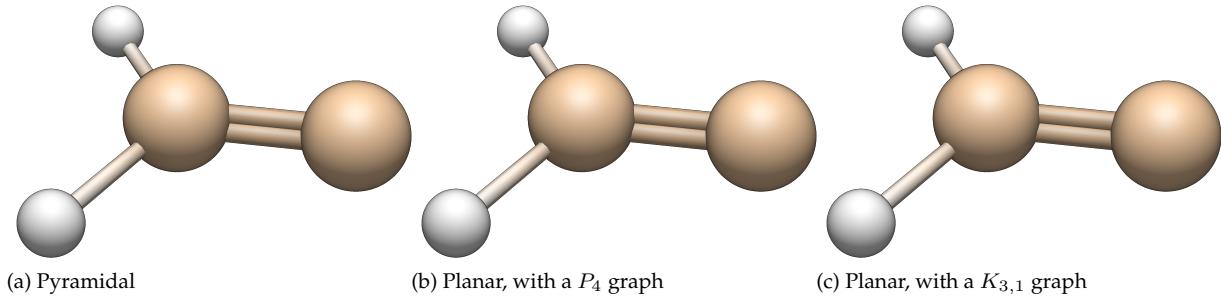
However, for many of the molecules in the GW100 dataset, calculations are done at geometries that were optimized using density functional theory, which means that the results are dependent on the basis set and density functional used; and it also had an effect on the comparisons that have been made between calculations that were done for the GW100 dataset and experimental data on the same dataset. For many of those molecules, very accurate bond lengths and angles that were determined from high-resolution spectroscopic experiments, are available in places such as the Landolt–Bornstein volumes and related books [23–35] and also in the original papers in which the geometric data was determined.

The books [23–35] provide only the *point group*, a 3D diagram, and usually the minimum number of bond lengths, bond angles and dihedrals necessary to generate all XYZ coordinates or a z-matrix. An example is the entry for  $H_2Si_2$  in [26] (this molecule is not in GW100 but it is useful for illustrating this point) which provides the point group ( $C_{2v}$ ), a 3D shape that is redrawn in Fig. 1 of this paper, and *only* two bond lengths and a dihedral angle:  $r_{Si-Si} = (2.2154 \pm 0.0020)\text{\AA}$ ,  $r_{Si-H} = (1.6680 + 0.0030)\text{\AA}$ ,  $\varphi = (104.22 + 0.30)^\circ$  (the original experimental paper from which this data came [42] and the previous paper on the same molecule [41], also provided no additional bond lengths or angles). But to do electronic structure calculations we need either the XYZ coordinates of all atoms, or a z-matrix for the molecule, and there is no software for obtaining either of these from just the limited information given (point group, 3D diagram, two bond lengths and a dihedral angle). A software for generating a z-matrix from this data would be significantly simpler to write than a software for generating XYZ coordinates of all atoms, but the z-matrix for this system would need at least one bond angle, and the reader might find it surprising that the simplest formula for the H-Si-H angle in  $H_2Si_2$  is (in terms of the provided dihedral angle  $\varphi$ , the provided Si-Si bond length  $r_1$  and the provided Si-H bond length  $r_2$ ):

$$\theta = \arcsin \left( \frac{\sin \left( \frac{\varphi}{2} \right) \sqrt{4r_2^2 - r_1^2}}{r_2} \right). \quad (1)$$

Such a formula is not expected to be found in any compilation of formulas. Determining this formula can be accomplished with some high school level trigonometry (applications of the ordinary cosine law), and ap-

Figure 1: Three molecules that have different geometries despite having the same chemical formula ( $\text{H}_2\text{Si}_2$ ) and the same point group ( $C_{2v}$ ).



plications of the *dihedral cosine law* which is unlikely to be taught in high school or university-level curricula, but can be learned and applied to such small molecules within a day by most researchers; however with the pace of research today, the derivation of such a formula *would ideally be done by software rather than by hand*. Unfortunately, not even the most advanced symbolic computation or computer algebra software can provide a user with Eq. 1 based on the information given.

Furthermore, the interest today in machine learning and big data analysis has lead to the introduction of much larger datasets than GW100, such as GW5000 and OE62 (both from Stuke *et al.* [39]) which involve more than 5000 and 61000 molecules respectively, but the geometries for the molecules in these datasets have issues similar to the ones for the GW100 dataset. This reinforces our motivation for constructing a software that could, for example, convert the thousands of geometries from the Landolt–Börnstein series and related compilations of spectroscopically determined geometry data, into z-matrices or XYZ coordinates.

We believe that as machine learning and big data analysis become more popular over the next century, there will be a rise in interest in being able to categorize molecules (e.g. for training neural networks) and the question of which point groups are possible for molecules of a certain size or formula, will be asked more and more. Anyone seeking the answers to these questions will be able to find them in this paper.

**A software to generate z-matrices (or even XYZ coordinates) for molecules with the data presented in experimental papers, such as for  $\text{H}_2\text{Si}_2$ , would require specific code blocks for each molecular shape.** It is possible for two molecules with the same point group and formula to have very different shapes (for example see Fig. X, in which two isomers of  $\text{H}_2\text{Si}_2$  are shown, both with the  $C_{2v}$  point group, but one of them being pyramidal and the other one being planar), so a general-purpose “z-matrix constructor” or “XYZ-coordinates constructor” that uses only a “minimal” set of bond lengths, bond angles and dihedral angles, would need specific code blocks based on their point groups and chemical formulas, in addition to other characteristics. However, a natural first step towards constructing such a software would be to categorize  $n$

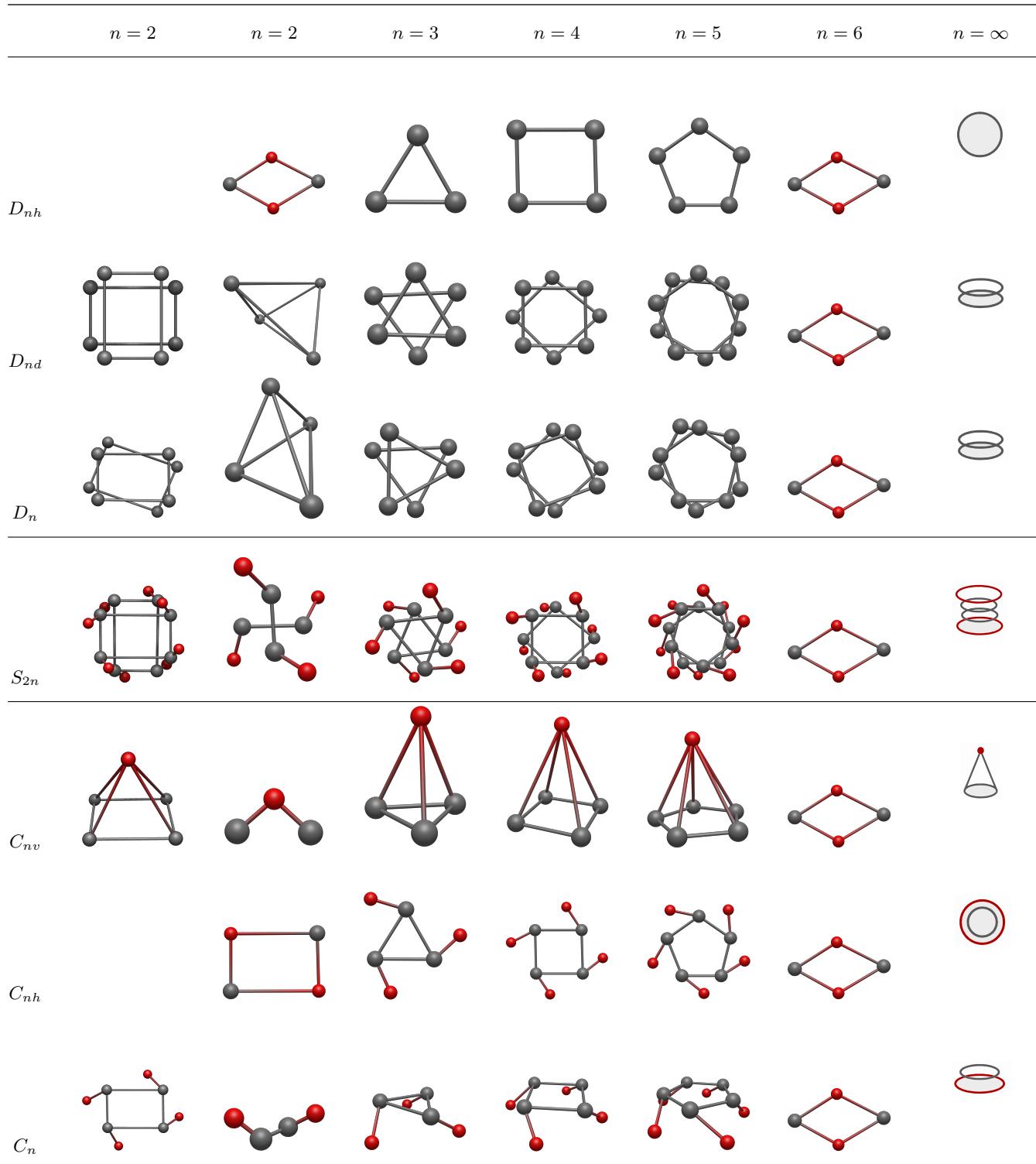
atom molecules based on their molecular formulas and point groups (the Landolt–Börnstein volumes and related books [23–35], the CCCBDB [43], and the CRC Handbook of Chemistry and Physics are some examples in which point groups are already used to help describe molecules), and then to write the relevant z-matrix-construction or XYZ-construction code for each case, adding sub-cases when necessary (e.g. to distinguish between the two types of  $C_{2v}$  shapes for  $\text{A}_2\text{B}_2$  molecules as depicted in Fig. 1).

**For  $n$  atoms, how many cases do we need to consider in order to cover each point group at least once in the software?** That is the essence of the question in the present paper’s title, and as mentioned in the opening paragraph of the Introduction section, it appears that this question has not been answered before, or at least the answer is not easily available. We also go even one step further, by determining all possible *chemical formulas* for a particular point group of an  $n$ -atom system. For example, in order for a 7-atom molecule to have the point group  $C_i$  its formula needs to be  $\text{AB}_2\text{C}_2\text{D}_2$  (with any of the variables B, C, or D being allowed to be a duplicate of one of the previous letters) but a 7-atom molecule whose chemical formula does not have that structure cannot have the point group  $C_i$  (e.g. for reasons that will be explained in this work, there is no arrangement of the atoms of  $\text{C}_3\text{H}_3\text{N}$  such that the point group of the molecule can be  $C_i$ ).

### III. TABLE OF EXAMPLE SHAPES FOR ALL AXIAL POINT GROUP SERIES

The 2008 version of [1] contains a table of “shapes” on Page 37, which appears to be the only table available (in all of the scientific literature) of its kind (see Section I). For the point groups  $C_n$ ,  $D_n$ ,  $C_{nv}$ ,  $C_{nh}$ ,  $D_{nh}$ ,  $D_{nd}$ , and  $S_{2n}$ , with  $n$  from 2 to 6, schematic diagrams are presented in attempt to provide an impression of how a molecule with the point group would look. This table has some minor inconsistencies (e.g. one of the lines in the  $C_4$  diagram is black instead of purple, the  $C_6$  diagram has some extra unintended purple lines, and the  $C_4$ ,  $C_5$  and  $C_6$  diagrams have lines with inconsistent lengths compared to other lines in the same diagrams

Table I: An improved version of the table presented on Page 37 of [1], depicting example structures for each point group provided in the table. For  $n = 2$ , two families of structures are presented (one with a rectangular “base” and one with a diatomic “base”). For each point group series and for each value of  $n$ , more families of structures also exist, so the examples that are presented are not exhaustive. All objects that are depicted for the  $n = \infty$  point groups satisfy the definition of either  $D_{\infty h}$  or  $C_{\infty v}$ , so they are labeled as such. The colors of the atoms are not intended to have any meaning apart from showing that the grey/black atoms have to all be the same element, and the red atoms have to all be the same element (the red atoms in all of these diagrams can also be the same element as the black atoms without affecting the point group, but it is not necessary for them to be the same). No structure is intended to depict a real molecule.



and compared to analogous lines in the  $C_{4h}, C_{5h}$  and  $C_{6h}$  diagrams). Although any inconsistency in such a table can hamper the experience of its readers, and although in these cases some readers are likely to lose a bit of time wondering why some lines aren't the same size and why there's some extra lines, most readers will be comfortable to ignore these ones and to simply attribute them to human error.

However, the aforementioned table in [1] has more major consistencies which we would like to address in this section of the paper. For example, the table in in [1] shows the  $C_{nv}$  series converging to a cone in the  $n = \infty$  case, but does not show the  $D_{nh}$  series converging to a circle in the  $n = \infty$  case (the reader may compare with our Table I for context). The table in [1] showed the  $D_{nd}$  series converging to what appears to be a single circle in the  $n = \infty$  case, but it would be more logical for the  $D_{nd}$  series to converge to two equivalent and co-axial circles on different parallel planes (as shown in our Table I) and to label the  $n = \infty$  object as having the point group  $D_{\infty h}$  since whether we have one circle (as in their table) or two parallel co-axial circles (as in our table), the point group is in fact  $D_{\infty h}$  if we are dealing with physical objects (the distinction between  $D_{\infty d}$  and  $D_{\infty h}$  arises only when considering vector fields rather than physical objects). The table in [1] also does not show the objects to which the point group series  $D_n, S_{2n}, C_{nh}$ , and  $C_n$  converge in the  $n = \infty$  case, so we have shown these objects (and whether the point group series converges to  $C_{\infty v}$  or  $D_{\infty h}$ ) in our Table I.

The next issue with the table in [1] that we addressed was the readability and interprability of their figures. This problem is perhaps most severe in their figures for  $S_{2n}$ , in which the atoms in the first and third planes are of one elemental type (depicted in a beige color in their figure), and the atoms in the second and fourth planes are of another elemental type (depicted in a purple or maroon color in their figure); whereas for reasons that will be explained in Section IV, it would be more correct to show the atoms in the first and fourth planes being one elemental, and the atoms in the second and third planes being one elemental type (as depicted in our Table III). This issue is also a problem for the  $D_n$  and  $D_{nd}$  series in their table: the figures for the  $D_{nd}$  series have two different colors, even though all atoms actually need to be the *same* type, and the figures in the  $D_{nd}$  series do not make it clear that the two polygons shown for each value of  $n$ , are on different planes. For the  $D_n$  series, again the colors appear to be incorrect. The chemical formulas and object shapes are made much more clear in our Table III and are explained in detail in subsequent sections of this paper.

On a more fundamental level, the table in [1] does not make it clear what precisely is being shown: most readers will likely assume that each polygon vertex represents an atom (although this is not made obvious), but what about the thin lines and small triangles that appear at some vertices? Do the thin lines represent one additional atom each, while the triangles de-

pict more than one additional atom each? In fact, both the straight lines and the small triangles represent just one additional atom in each case, but the triangles depict atoms that are outside of the plane of the shown polygon, and the straight lines depict atoms that are in the same plane as the shown polygon (but many readers, like us, may at first think that the use of straight lines versus small triangles was just a "typo" akin to the various "minor inconsistencies" that we mentioned earlier).

We also note that the "appendages" in the  $C_{3h}, C_{4h}, C_{5h}, C_{6h}, C_4, C_5, C_6$  figures in [1] are parallel (in fact co-linear) with the edges of the triangle, but for  $C_3$  they are *not*. This is likely another "typo" but we will note that the "appendages" do not need to be parallel, as long as they do not bisect the exterior angles at the polygons' vertices, and as long as the atoms are not in the center of the polygons' edges, since if either of these situations were to occur, the point group of the system would be promoted to  $D_{nh}$ . In our Table I, we kept the angles of the "appendages" consistent throughout the table, and chose to make them all non-parallel with the "bonds" in the structures.

We also attempted to resolve a confusion associated with the  $n = 2$  column in the table in [1]. The example structure shown for  $C_2$ , for example is obscure and a far less common example for the  $C_2$  point group than the structure of  $\text{H}_2\text{O}_2$  for example. Likewise,  $\text{H}_2\text{O}$  is the first system that usually comes to mind when we consider the point group  $C_{2v}$ , but the structure shown in the table for  $C_{2v}$  is very different. Also to add to the confusion, the table in [1] uses a rectangular "base" for the example systems for the  $D_2, C_{2h}, D_{2h}, D_{2d}$  and  $S_4$  point groups, but an ambiguous or uninterpretable "base" for  $C_2$  and  $C_{2v}$ . We have resolved this confusion by including two  $n = 2$  columns in our Table I: one with a diatomic "base" and one with a rectangular "base". so that readers can see example shapes for both cases.

For the  $C_n$  and  $D_n$  series, perhaps the authors of [1] had specific molecules in mind, but one of the early questions that arose for us before we began writing this paper, was "what is the minimum number of atoms required for the point group so a system to be  $G$  [with  $G$  being the point group]", and after answering that question we found that the table in [1] correctly depicts the minimum number of atoms required for all point groups except for the  $n = 2$  ones (all of them!) and for  $C_n$  and  $D_n$ . In our Table I, we have shown the "minimal" structures that we found for  $C_n$  and  $D_n$  as well as for the  $n = 2$  cases for all other point group series. It is important to note that for each point group, multiple structures can be valid, even when the number of atoms and the chemical formula are the same (we already mentioned this in the previous section and illustrated it in Figure 1), so the figures that we presented in I and that the authors of [1] presented in their original table, are not the only correct examples that can be provided for each point group; but ours were at least

chosen based on a non-arbitrary criterium, which was to only include for each point group, the structure that require the minimum number of atoms (except for the first  $n = 2$  column for each point group series, in which we decided that it would be interesting to show a structure requiring the minimum number of atoms that uses a “rectangular base”).

The final difference that we will note between the table in [1] and our Table I is the order in which the point group series are presented. The order in which we present them is far better from a pedagogical standpoint, because with the  $D_{nh}$  series presented first (these are the simplest shapes in the entire table), all other shapes in the table are built on top of these  $D_{nh}$  shapes and contain the  $D_{nh}$  shapes as a subsystem of a larger shape. The order that we presented the point group series in Table I is the same as the order in which we present the arguments for the minimum numbers of atoms and the possible chemical formulas for each point group in Section V.

#### IV. GENERAL PRINCIPLES FOR POSSIBLE MOLECULES

In this section we will state some general principles that we will use in the following section when formulating proofs for the required numbers of atoms and required chemical formulas for each point group. Some of the statements made in this section may already be familiar or intuitive to the reader, but we have not seen most of them stated in textbooks about group theory for chemistry [1–18, 36, 40], although parts of the first statement are mentioned in other places, such as [37].

##### Ability to orient the coordinate system according to a $C_n$ axis and $\sigma_h$ plane:

Any molecule with a rotation axis, called a  $C_n$ -axis, can be oriented in an xyz (Cartesian) coordinate system such that for the largest possible  $n$  value in  $C_n$ , the  $z$ -axis is aligned with a  $C_n$  axis, and a  $\sigma_h$  plane (if one exists) is a mirror plane coinciding with the  $xy$ -plane in the coordinate system.

##### Existence of a $C_n$ axis:

The existence of a  $C_n$  axis is only possible if all atoms are on the  $C_n$  axis or if the system contains a subset of identical atoms that are vertices of an  $n$ -sided polygon with equal side-lengths (a regular  $n$ -gon), or if  $n < 3$  we simply need to have  $n$  identical atoms in the molecule. A  $C_n$  axis would go through the center of this polygon (or in the case of  $n = 2$ , it would go through the centre of the line joining the two identical atoms), and since we can call this axis the  $z$ -axis, all atoms that form vertices of the polygon (or the two identical atoms in the case of  $n = 2$ ), would have to have the same  $z$ -coordinates. Therefore, a minimum of  $n$  atoms is needed in order to have a  $C_n$  axis.

If a  $C_n$  axis is the  $z$ -axis, then after a  $C_n$  rotation, an atom that originally had coordinates  $(x, y, z)$ , will have

coordinates  $(x', y', z)$  in which the  $z$ -coordinate does not change, and the new coordinates  $(x', y')$  can be different from the old ones  $(x, y)$  but will coincide with the original  $(x, y)$  coordinates of another identical atom in the molecule.

##### Necessary and sufficient conditions for the $D_{nh}$ point group when $n \geq 3$ :

A regular  $n$ -gon with uniform vertices (in a molecule, this would mean the presence of identical atoms at the vertices) has a point group of  $D_{nh}$  (this was already depicted visually in Table I), and any non-linear object with the  $D_{nh}$  point group must contain a subset of identical vertices that form a regular  $n$ -gon if  $n \geq 3$ .

##### Necessary and sufficient conditions for the $S_{2n}$ operation to be valid when $n \geq 3$ :

The  $S_{2n}$  operation is valid on a regular  $2n$ -gon with uniform vertices, because the  $S_{2n}$  operation is the combination of a  $C_{2n}$  operation (which is valid on such a system) followed by a  $\sigma_h$  operation (which is also valid on such a system, since it is planar); and any non-linear object in which the  $S_{2n}$  operation is valid, must contain a subset of identical vertices that form a regular  $n$ -gon if  $n \geq 3$ .

However, for the  $S_{2n}$  operation to be valid in a system in which the  $\sigma_h$  operation is not valid (as in the  $D_{nd}$ ,  $S_{2n}$ , and  $T_d$  point groups), we can not just have a  $2n$ -gon since the  $2n$ -gon itself forms a plane that can act as a  $\sigma_h$  plane. It turns out that instead of one regular  $2n$ -gon, two regular  $n$ -gons with one of them rotated by an angle of  $\pi/n$  from the other one, and both of them being on different parallel planes with a  $C_n$  axis going through the center of each of the  $n$ -gons, is a sufficient property and a necessary part of any system for there to be a valid  $S_{2n}$  operation and no valid  $\sigma_h$  operations (when  $n \geq 3$ ). Each  $D_{nd}$  point group contains  $S_{2n}$  elements and no  $\sigma_h$  elements, and the examples for  $D_{nd}$  that we presented for  $n \geq 3$  in Table I have exactly the form described in the previous sentence.

The  $\sigma_h$  plane that is part of the  $S_{2n}$  operation, cannot contain all atoms of the system, because then this plane would not only be the  $\sigma_h$  part of the  $S_{2n}$  operation, but it would also be a valid  $\sigma_h$  operation for the entire system. At least one atom needs to not be on the  $\sigma_h$  plane of the  $S_{2n}$  operation, and if it is on the  $C_{2n}$  axis, then since a  $C_{2n}$  operation does nothing to atoms that are on the  $C_{2n}$  axis, a valid  $S_{2n}$  operation would require the  $\sigma_h$  plane of the  $S_{2n}$  operation to be a valid  $\sigma_h$  plane for the atom too; so at least one atom needs to be neither on the  $S_{2n}$  operation's  $C_{2n}$  axis nor on its  $\sigma_h$  plane. Since we can make the  $C_{2n}$  axis of the  $S_{2n}$  operation the  $z$ -axis, and the  $\sigma_h$  plane of the  $S_{2n}$  operation the  $xy$ -plane, we can summarize the previous sentence by saying that one point exists at a point  $(x_1, y_1, z_1)$  in which  $(x_1, y_1) \neq (0, 0)$  and  $z_1 \neq 0$ . One  $S_{2n}$  operation will rotate the  $(x, y)$  coordinates of the point by an angle of  $\frac{2\pi}{2n} = \pi/n$  and the  $z$ -coordinate of

that point to  $-z_1$ , and an identical atom needs to be at this new point:  $(x_2, y_2, -z_1)$ . The  $S_{2n}$  operation on the atom at  $(x_2, y_2, -z_1)$  will rotate the  $(x, y)$  coordinates of the point by an angle of  $\pi/n$  and the  $z$ -coordinate of that point to  $+z_1$ , and an identical atom needs to be at this new point. By repeating this procedure, we will be forced to have a regular  $n$ -gon with uniform vertices on the plane  $z = +z_1$  and another one on the plane  $z = -z_1$  but with all vertices rotated by  $\pi/n$  with respect to corresponding vertices on the other  $n$ -gon, and the  $C_{2n}$  axis of the  $S_{2n}$  operation (also called the  $S_{2n}$  axis) going through the centers of each  $n$ -gon. So not only is this shape sufficient for the existence of a valid  $S_{2n}$  operation in a system that does not have a valid  $\sigma_h$  operation on its own, but it is a necessary subunit of any system whose point group has an  $S_{2n}$  element and no  $\sigma_h$  elements.

When  $z_1 = 0$  in the shape described in the previous paragraph, both  $n$ -gons will be on the same plane, and rotated by an angle of  $\pi/n$  from each other, so we will again have a  $2n$ -gon, as described earlier.

#### Action of the inversion operation, $i$ :

We can always choose center of inversion to be at the origin  $(0, 0, 0)$ , so if an atom is located at  $(x, y, z)$ , the inversion operator will move it to  $(-x, -y, -z)$ .

## V. POSSIBLE NUMBER OF ATOMS AND CHEMICAL FORMULAS FOR EACH AXIAL POINT GROUP

In all of the literature that we know, “the point group” of a molecule is only considered to be the *largest* point group such that all elements (symmetry operations) transform the molecule into an identical one. For example, even though all elements of  $D_{2h}$  transform a homonuclear diatomic molecule into an identical molecule, the point group is considered to be  $D_{\infty h}$  which is a supergroup of  $D_{2h}$  that contains *all* of the elements that transform a homonuclear diatomic molecule into an identical molecule. We will use this convention throughout this paper.

### A. Point groups without axial rotational symmetry

The point groups considered in this subsection do not contain any  $C_n$  axes in which  $n > 1$ .

#### $C_s$

Any two atoms form a line, and if all atoms of a molecule are on the same line, then the point group will be promoted to either  $C_{\infty v}$  or  $D_{\infty h}$ . Likewise, a single atom has the point group  $K_h$ , so it is not possible for a molecule with fewer than three atoms to have the point group  $C_s$ . Three atoms is enough for the point group of a molecule to be  $C_s$  though, as in the example HNO (azanone, whose structure is depicted in Table IV). **The minimum number of atoms required is three, and any chemical formula is allowed.**

We can add as many more atoms as we wish without changing the point group of the system from  $C_s$ , as long as these atoms all remain on the same plane as the first three atoms, and as long as they are not from the same element (having atoms with different elements guarantees that no *new* symmetry operations become valid on the system). The point group  $C_s$  is also possible for non-planar structures, and structures containing more than one atom from the same element, are also possible (e.g.  $\text{Cl}_2\text{OS}$  which is both non-planar and both has two atoms of the element Cl), but we were already able to prove in the previous sentence that the point group  $C_s$  is possible for any number of atoms that is larger than or equal to three. **The general chemical formula is  $\text{ABCD}\cdots$ ,** in which any letter can be identical to one of the previous ones (e.g. B, C and D can be chosen to equal A if desired).

#### $C_1$

Any three atoms form a plane or a line, and if all atoms of a molecule are on the same plane, then that plane is a reflection plane (also known as a mirror plane). Therefore if a molecule has three or fewer atoms, it will at least have one reflection plane, and would be promoted from  $C_1$  to at least  $C_s$ . Four atoms is enough for the point group of a molecule to be  $C_1$  though, as in the example  $\text{H}_2\text{OS}$  (whose structure is depicted in Table V). **The minimum number of atoms required is four, and any chemical formula is allowed.** We can add as many more atoms as we wish without changing the point group of the system from  $C_1$ , as long as none of these additional atoms are from the same element (having atoms with different elements guarantees that no symmetry operations become valid on the molecule). If we run out of elements, we can still add atoms such that no new symmetry operations are valid on the system. **The general formula is  $\text{ABCDE}\cdots$ ,** in which any letter can be identical to one of the previous ones (e.g. B, C, D and E can be chosen to equal A if desired).

#### $C_i$

Since the  $C_i$  point group does not have any reflection elements, the point group of any system that contains a reflection plane and an inversion center must be larger than  $C_i$ .

If we have three or fewer atoms, then there does not even need to be an inversion center for the atoms to be confined to a plane (or line). If we have four atoms, and we make the inversion center the origin, the atoms must come in pairs according to the inversion operation  $i$ :  $A = (x_1, y_1, z_1)$  will be paired with  $-A = (-x_1, -y_1, -z_1)$  and  $B = (x_2, y_2, z_2)$  will be paired with  $-B = (-x_2, -y_2, -z_2)$ . Assuming that no two atoms will have the same  $xyz$ -coordinates, none of these four atoms can be on the origin.  $A$  and  $-A$  form a line, and  $B$  and  $-B$  form another line, and these two

Table II: Possible point groups for systems with  $m = 1$  to  $m = 8$  atoms.

$m$  Possible groups (not including subgroups)

1											$K_h$
2											$D_{\infty h} C_{\infty v}$
3	$C_s$										$D_{\infty h} C_{\infty v}$
4	$C_s$ $C_1$ $C_2$										$D_{\infty h} C_{\infty v}$
5	$C_s$ $C_1$ $C_2$										$D_{\infty h} C_{\infty v}$
6	$C_s$ $C_i$ $C_1$ $C_2$ $C_3$										$O_h D_{\infty h} C_{\infty v}$
7	$C_s$ $C_i$ $C_1$ $C_2$ $C_3$										$O_h D_{\infty h} C_{\infty v}$
8	$C_s$ $C_i$ $C_1$ $C_2$ $C_3$ $C_4$ $C_{2h}$ $C_{3h}$ $C_{4h}$ $C_{2v}$ $C_{3v}$ $C_{4v}$ $C_{5v}$ $C_{6v}$ $C_{6v}$ $C_{7v}$ $S_4$										$O_h D_{\infty h} C_{\infty v}$

lines both intersect at the origin, but two lines that cross are always co-planar. If we have five atoms, the fifth one must be at the origin, because atoms need to either be paired or at the origin in order for the inversion operation to be valid on the system of atoms. However, the other four atoms in the five-atom system will form a plane that contains the origin (because two lines that intersect at the origin form a plane that contains the origin). Therefore the fifth atom will also be on the same plane as the other four atoms, and the point group will be promoted beyond  $C_i$ .

Six atoms is enough for the point group of a system to be  $C_i$ , for example  $\text{C}_2\text{H}_2\text{Br}_2\text{Cl}_2$  with the two carbons removed. Any number of atoms can be added while maintaining the point group  $C_i$ , but an even number of atoms would need to be added in pairs of identical elements, and/or one additional atom of any element can be added to the inversion center. **The minimum number of atoms required is six, and the chemical formula must be of the form  $A_2B_2C_2D_2\dots$  or  $AB_2C_2D_2\dots$  depending on if the number of atoms is even or odd**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

### B. Point groups with axial rotational symmetry

#### $D_{nh}$

$$n = 2$$

If we have any fewer than three atoms, the point group of the system will be promoted to  $K_h$ ,  $D_{\infty h}$  or  $C_{\infty v}$ , however all  $D_{nh}$  point groups with even values of  $n$  (including  $D_{2h}$ ) have an inversion element  $i$ , and for a reason that is explained in the next sentence, the only way for a three atom system to have an inversion element is for it to be linear, which would promote it either to  $D_{\infty h}$  or  $C_{\infty v}$ . If we choose the origin  $(0, 0, 0)$  to be the center of inversion, and one atom is at the position  $(x_1, y_1, z_1)$ , then an identical atom would need to be at the positon  $(-x_1, -y_1, -z_1)$  and the third atom would need to be at the origin. We therefore would have two identical atoms on a *line* going through the origin, and a third atom in the center of that line. We

therefore would need a minimum of four atoms for the  $D_{2h}$  point group, and indeed four atoms is enough as in the case of  $\text{Br}_2\text{Na}_2$  (whose structure is depicted in Table V). **The minimum number of atoms required is four, and the most general chemical formula for a structure with the minimum number of atoms is  $A_2B_2$** , and B can be chosen to equal A if desired. Any atom can be added to the inversion center of the system without breaking any of the  $D_{2h}$  symmetry elements. If an atom is to be added anywhere else, it needs to *paired* with an atom of the *same element* in order for the inversion operation  $i$  to remain valid, and such pairs of atoms can be added to the primary  $C_2$  axis in the  $D_{2h}$  point group without breaking any of the other symmetry elements of the group. **The chemical formula must be of the form  $A_2B_2C_2D_2\dots$  or  $AB_2C_2D_2\dots$  depending on if the number of atoms is even or odd**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$$\underline{n \geq 3}$$

Since a regular  $n$ -gon with identical atoms at each vertex is a *necessary* component of any system with a point group of  $D_{nh}$  for  $n \geq 3$ , we can conclude the **the minimum number of atoms required is  $n$ , and the formula for this minimal structure would be  $A_n$** . An even number of additional atoms can be added along the  $C_n$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* on each side of the plane formed by the  $n$ -gon. One additional atom can also be added at the center of the  $n$ -gon. Therefore **the most general formula for a structure with  $n$  atoms plus an even number of additional atoms would be  $A_nB_2C_2D_2\dots$ , and for an odd number of additional atoms the most general formula would be  $A_nBC_2D_2\dots$** , and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

#### $D_{nd}$

$$\underline{n = 2}$$

Table III: Possible number of atoms and possible chemical formulas for each point group or point group series.

Point groups	Minimum	Number of atoms															Possible chemical formula		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Even # of atoms	Odd # of atoms
$K_h$	—	1	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	A
$D_{\infty h}$	—	2	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$(ABCDE \dots)_2$	$A(BCDE \dots)_2$
$C_{\infty v}$	—	2	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$ABCDE \dots$	
$D_{nh}$	$n \geq 3$	$n$	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$A_n(BCDE \dots)_2$	$AB_n(CDE \dots)_2$
$D_{nd}$	$n = 2$	4	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$(ABCDE \dots)_2$	$A(BCDE \dots)_2$
$D_n$	$n \geq 2$	$2n$	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$(A_n BCDE \dots)_2$	$A(B_n CDE \dots)_2$
$S_{2n}$	$n \geq 2$	$4n$	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	$(A_n B_n CDE \dots)_2$	$A(B_n C_n DE \dots)_2$
$C_{nv}$	$n \geq 2$	$n + 1$	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$AB_n CDE \dots$	
$C_{nh}$	$n \geq 2$	$2n$	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$A_n B_n (CDE \dots)_2$	$AB_n C_n (CDE \dots)_2$
$C_n$	$n \geq 2$	$2n$	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$A_n B_n CDE \dots$	
$C_s$	—	3	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$ABCDE \dots$	
$C_i$	—	6	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$(ABCDE \dots)_2$	$A(BCDE \dots)_2$
$C_1$	—	4	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	$ABCDE \dots$	

The  $D_{2d}$  group does not include any horizontal mirror planes, and since any three atoms can form a plane which can serve as a horizontal mirror plane, we need more than four atoms for the point group of the system to be  $D_{2d}$ . Four atoms is indeed enough for the point group of the system to be  $D_{2d}$  as depicted in Table V, in which we have two diatomic molecules that are perpendicular to each other, but on different planes. **The minimum number of atoms required is 4, and the formula for this minimal structure would be  $A_4$ .** An even number of additional atoms can be added along the  $C_2$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that (since we can label the  $C_2$  axis as the  $z$ -axis) these atoms are added in *homonuclear pairs* around the plane containing the two  $C_2(x)$  and  $C_2(y)$  axes (*i.e.* the  $xy$ -plane). One additional atom can also be added at the center of the entire structure. Therefore **the most general formula for a structure with 4 atoms plus an even number of additional atoms would be  $A_4 B_2 C_2 D_2 \dots$ , and for an odd number of additional atoms the most general formula would be  $AB_4 C_2 D_2 \dots$** , and any letter can be identical to one of the previous ones (*e.g.* B, C, and D can be chosen to equal A if desired).

$n \geq 3$

Each  $D_{nd}$  point group contains an  $S_{2n}$  operation and no  $\sigma_h$  operation on its own, and in Section IV we explained that any system with a valid  $S_{2n}$  operation (with  $n \geq 3$ ) but no  $\sigma_h$  operation on its own, must at least have a subset of atoms that form the shape (depending on the value of  $n$ ) shown for  $D_{nd}$  in Table

III. Since all other elements in  $D_{nd}$  are valid symmetry operations of these shapes, and no other symmetry operations that are not in  $D_{nd}$  are valid on these shapes, the point group of these structures are  $D_{nd}$ . Since these structures constitute “minimal” structures with a valid  $S_{2n}$  operation and no valid  $\sigma_h$  operations, we know that **the minimum number of atoms required is  $2n$ , and the most general formula for the minimal structure would be  $A_{2n} B_{2n}$** . An even number of additional atoms can be added along the primary  $C_n$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* around the plane that contains all  $n$  of the  $C'_2$  axes (since we can choose the  $z$ -axis to be the  $C_n$  axis, this plane would be the  $xy$ -plane). One additional atom can also be added at the center of the entire structure. Therefore **the most general formula for a structure with  $2n$  atoms plus an even number of additional atoms would be  $A_{2n} B_{2n} C_2 D_2 \dots$ , and for an odd number of additional atoms the most general formula would be  $AB_{2n} C_2 D_2 \dots$** , and any letter can be identical to one of the previous ones (*e.g.* B, C, and D can be chosen to equal A if desired).

$D_n$

$n = 2$

The  $D_2$  group does not include any horizontal mirror planes, and since any three atoms can form a plane which can serve as a horizontal mirror plane, we need more than four atoms for the point group of the system to be  $D_{2d}$ . Four atoms is indeed enough for the

point group of the system to be  $D_2$  as depicted in the second  $n = 2$  column in Table III, in which we have two identical homonuclear diatomic molecules that are *not* perpendicular to each other, and are also on different planes. **The minimum number of atoms required is 4, and the most general formula for the minimal structure would be  $A_4$ .** An even number of additional atoms can be added along the  $C_2$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that (since we can label the  $C_2$  axis as the  $z$ -axis) these atoms are added in *homonuclear pairs* around the plane containing the two  $C_2(x)$  and  $C_2(y)$  axes (*i.e.* the  $xy$ -plane). One additional atom can also be added at the center of the entire structure. Therefore **the most general formula for a structure with 4 atoms plus an even number of additional atoms would be  $A_4B_2C_2D_2\cdots$ , and for an odd number of additional atoms the most general formula would be  $AB_4C_2D_2\cdots$** , and any letter can be identical to one of the previous ones (*e.g.* B, C, and D can be chosen to equal A if desired).

$n \geq 3$

The  $D_n$  point groups can be generated by a  $C_n$  operation and a perpendicular  $C'_2$  operation, whereas the  $D_{nd}$  point groups can be generated by an  $S_n$  operation and a perpendicular  $C'_2$  operation (the only difference being the  $S_{2n}$  generator for  $D_{nd}$  instead of the  $C_n$  generator for  $D_n$ ). If we start with the structure that we described in the previous subsection for  $D_{nd}$ , one way that we can demote the  $S_{2n}$  axis to a  $C_n$  axis without changing any of the other valid symmetry operations and without introducing any new ones, would be to rotate one of the two  $n$ -gons in the  $D_{nd}$  shape so that the two  $n$ -gons are no longer perfectly rotated by an angle of  $\pi/n$  (as would be required for an  $S_{2n}$  axis by the logic that was presented in Section IV). This minor rotation will break the validity of the  $S_{2n}$  operation while maintaining the validity of the  $C_n$  operation. The resulting structures are depicted in Table III, and the **minimum number of atoms required for each  $D_n$  point group are the same as for  $D_{nd}$ , as are the chemical formulas for the minimal structures.** Likewise, the **possible numbers of atoms for structures with  $D_n$  point groups are the same as for  $D_{nd}$  point groups, as are the possible chemical formulas**, and any letter can be identical to one of the previous ones (*e.g.* B, C, and D can be chosen to equal A if desired).

$S_{2n}$

$n = 2$

From Section IV we know that a  $C_n$  axis can simultaneously be chosen to be along the  $z$ -axis and a  $\sigma_h$  plane can be chosen to be the  $xy$ -plane in the same

coordinate system. With that convention, a  $C_4$  operation will transform the point  $(x, y, z)$  to  $(-y, x, z)$  and the  $\sigma_h$  operation will transform the point  $(-y, x, z)$  to  $(-y, x, -z)$ ; so the  $S_4$  operation which is the  $C_4$  operation followed by the  $\sigma_h$  operation, will transform  $(x, y, z)$  to  $(-y, x, -z)$ . If our first atom is at the position  $(x_1, y_1, z_1)$  then the  $S_4$  transformation will move it to the position  $(-y_1, x_1, -z_1)$ , and an identical atom must be there for the  $S_4$  operation to be valid. Also after the same  $S_4$  transformation, the atom at the position  $(-y_1, x_1, -z_1)$  will move to the position  $(-x_1, -y_1, z_1)$  and an identical atom must be there too. Since no atom has replaced our first atom which was at the position  $(x_1, y_1, z_1)$ , we need to keep applying  $S_4$  transformations and keeping track of the atoms that are necessary to include in our system, until we do successfully replace the first atom. We show below that  $m = 4$  identical atoms are required in order for an  $m$ -atom system to undergo an  $S_4$  transformation and recover the same system:

$$\begin{aligned} (x_1, y_1, z_1) &\rightarrow (-y_1, x_1, -z_1) \rightarrow (-x_1, -y_1, z_1) \rightarrow \quad (2) \\ (y_1, -x_1, -z_1) &\rightarrow (x_1, y_1, z_1). \end{aligned}$$

Unfortunately the coordinates in (2) form the shape that we presented in the second  $n = 2$  column in our  $D_{nd}$  examples in Table III, because  $(x_1, y_1, z_1)$  and  $(-x_1, -y_1, z_1)$  form a line on the  $z = z_1$  plane, whereas  $(-y_1, x_1, -z_1)$  and  $(y_1, -x_1, -z_1)$  form a perpendicular line of the same size but on the  $z = -z_1$  plane. We can add at least one more atom (it does not necessarily need to be homonuclear in comparison to the first set of four atoms) in order to break the  $\sigma_d$  planes that are in  $D_{2d}$ , but adding any number of atoms on the  $C_4$  axis of the  $S_4$  operation, or on the  $\sigma_h$  plane of the  $S_4$  operation, or on either one of the  $z = \pm z_1$  planes will not break the  $\sigma_d$  planes. Since we have chosen the convention in which the  $S_4$  axis is the  $z$ -axis, adding atoms on this axis would mean adding atoms with  $(x, y) = 0$ ; since we have chosen the convention in which the  $\sigma_h$  plane associated with the  $S_4$  operation is the  $xy$ -plane, adding atoms on this plane would mean adding atoms with  $z = 0$ ; and regardless of those conventions, any atom on one of the  $\pm z$  planes will have a  $z$ -coordinate of  $z$  or  $-z$ . To break the  $\sigma_d$  planes, we therefore need an atom at a different  $z$ -value such as  $z_2$  and with non-zero  $x$ - and  $y$ -values. Regardless of the  $x$ - and  $y$ -values of this new atom (the fifth atom), the pattern in (2) shows that we will not recover the new atom's position by applying  $S_4$  operations unless we have another three atoms that are identical to the fifth atom.

The example that we have provided in the second  $n = 2$  column for  $S_{2n}$  in Table III is able to satisfy the above relationships among the coordinates of the atoms; since the  $S_4$  group is generated by the  $S_4$  operation only, and an  $S_4$  operation can successfully be applied to this example, and there are no other symmetry operations valid on this structure which are not in the  $S_4$  point group, the point group of this 8-atom

structure is  $S_4$ . **The minimum number of atoms required is 8 with the chemical formula  $A_4B_4$ .** An even number of additional atoms can be added along the  $S_4$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* around the  $\sigma_h$  plane of the  $S_4$  operation. One additional atom can also be added at the center of the entire structure. Therefore **the most general formula for a structure with 8 atoms plus an even number of additional atoms would be  $A_4B_4C_2D_2\cdots$ , and for an odd number of additional atoms the most general formula would be  $AB_4C_4D_2\cdots$** , and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$n \geq 3$

In section IV we described structures that have the minimum number of atoms required for an  $S_{2n}$  operation to be valid with  $n \geq 3$  and for no  $\sigma_h$  operations to be valid on their own. We also described earlier that the point group of such structures is  $D_{nd}$ : in order to make a minimal structure in which the point group is  $S_{2n}$ , we can start with the  $D_{nd}$  structure and add atoms to break the extra symmetry elements that are present in  $D_{nd}$  but not in  $S_{2n}$ . If we add any number of atoms to the  $S_{2n}$  axis or the  $\sigma_h$  plane that forms part of the  $S_{2n}$  operation, it will not break the  $\sigma_d$  planes that are elements of  $D_{nd}$ , so we need to add at least one atom to a location that is neither on the  $S_{2n}$  axis nor on the  $\sigma_h$  plane associated with it. This atom does not need to be homonuclear with the existing atoms in the structure, but as we described in Section IV, we would have no choice but to add not just this one atom but two entire regular  $n$ -gons which are identical to each other and parallel to each other but not on the same planes as the two previously existing  $n$ -gons in the structure, and the  $S_{2n}$  axis must pass through the centre of all four  $n$ -gons.

We have shown examples of such structures in Table I (our choice of chemical “bonds” shown in these examples, was made to resemble some known real  $S_{2n}$  structures, but we could also have chosen them differently, for example we could have had four polygons rather than two polygons with appendages). **The minimum number of atoms required is  $4n$ , and the formula for this minimal structure would be  $A_{2n}B_{2n}$ .** An even number of additional atoms can be added along the  $S_{2n}$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* symmetrically on each side of the  $\sigma_h$  plane associated with the  $S_{2n}$  operation. One additional atom can also be added at the center of the overall system. Therefore **the most general formula for a structure with  $4n$  atoms plus an even number of additional atoms would be  $A_{2n}B_{2n}C_2D_2\cdots$ , and for an odd number**

**of additional atoms the most general formula would be  $AB_{2n}C_{2n}D_2E_2\cdots$** , and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$C_{nv}$

$n = 2$

If we have any fewer than three atoms, the point group of the system will be promoted to  $K_h$ ,  $D_{\infty h}$  or  $C_{\infty v}$ , however three atoms is indeed enough for the point group of a molecule to be  $C_{2v}$ , as in the case of  $H_2O$ . **The minimum number of atoms required is three, and the most general chemical formula for structure with the minimum number of atoms, is  $A_2B$**  since at least two atoms need to be the same in order for the  $C_2$  operation to be valid. Since extra atoms can be added on the  $C_2$  axis without losing any valid symmetry operations and without introducing any new ones, the **the most general chemical formula is  $A_2BCD\cdots$** , and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$n \geq 3$

Since we know from Section IV that the point group of a regular  $n$ -gon with identical atoms at each vertex is required in order to have the  $C_n$  axis that forms part of the point group  $C_{nv}$ , and we also know from Section IV that the point group of such a structure would be  $D_{nh}$ , we would need at least one additional atom to break any symmetries in  $D_{nh}$  that are not present in  $C_{nv}$ . It turns out that only one atom is sufficient, since if this atom is placed on a  $C_n$  axis that goes through the center of the  $n$ -gon, but not on the plane of the  $n$ -gon, it will break the  $\sigma_h$  plane and any  $S_n$  axes of the  $D_{nh}$  point group, while maintaining the  $n$  vertical mirror planes that are in the  $C_{nv}$  point group. The only point group with a  $C_n$  axis and  $n$  vertical mirror planes, without having any horizontal mirror planes or  $S_n$  axes, is  $C_{nv}$ , so the point group of the pyramid that we have formed must be  $C_{nv}$ . **The minimum number of atoms required is  $n+1$ , and the most general formula for a structure with the minimum number of atoms is  $A_nB$ .** Since extra atoms can be added on the  $C_n$  axis without losing any valid symmetry operations and without introducing any new ones, the **most general chemical formula is  $A_nBCD\cdots$** , and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$C_{nh}$

$n = 2$

We need a minimum of four atoms for the  $C_{2h}$  point group, for the same reason that we gave when we made this statement for the  $D_{2h}$  point group (because the  $C_{2h}$  point group has an inversion element  $i$ , which is not possible for non-linear molecules with fewer than four atoms). Indeed four atoms is enough for the point group of a molecule to be  $C_{2h}$ , as in the case of trans-H<sub>2</sub>N<sub>2</sub> whose structure is depicted in Figure V). **The minimum number of atoms required is four, and the most general chemical formula for a minimal structure is A<sub>2</sub>B<sub>2</sub>.** An even number of additional atoms can be added along the  $C_2$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* around the  $\sigma_h$  plane. One additional atom can also be added at the center of the entire structure. Therefore **the most general formula for a structure with 4 atoms plus an even number of additional atoms would be A<sub>2</sub>B<sub>2</sub>C<sub>2</sub>D<sub>2</sub> ..., and for an odd number of additional atoms the most general formula would be AB<sub>2</sub>C<sub>2</sub>D<sub>2</sub> ...**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$n \geq 3$

From Section IV we know that the system needs to contain a subset of atoms that form a regular  $n$ -gon, in order to have a  $C_n$  axis, but if we only have one regular  $n$ -gon, the point group would be  $D_{nh}$ , so we need to add at least one more atom in order to break the symmetry elements that  $D_{nh}$  has which are not in  $C_{nh}$ . Adding any number of atoms to the  $C_n$  axis (which we will label as our  $z$ -axis) will not break the  $\sigma_v$  planes, so we are forced to add another regular  $n$ -gon with comprised of  $n$  identical atoms at its vertices, but these atoms do not need to be the same as the  $n$  identical atom that comprised our first  $n$ -gon. We can add this second regular  $n$ -gon on the same plane as the first one, as long as its vertices do not bisect the exterior angles at the first  $n$ -gon's vertices, and as long as the second  $n$ -gon's vertices are not on the edges of the first  $n$ -gon (both of these placements of the vertices would result in a failure to break the  $\sigma_v$  planes and would not successfully demote the point group from  $D_{nh}$  to  $C_{nh}$ ). Therefore, **the minimum number of atoms required is  $2n$ , and the most general chemical formula for a minimal structure is A<sub>n</sub>B<sub>n</sub>.** An even number of additional atoms can be added along the  $C_n$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations are invalidated, provided that these atoms are added in *homonuclear pairs* around the  $\sigma_h$  plane. One additional atom can also

be added at the center of the entire structure. Therefore **the most general formula for a structure with  $2n$  atoms plus an even number of additional atoms would be A<sub>n</sub>B<sub>n</sub>C<sub>2</sub>D<sub>2</sub> ..., and for an odd number of additional atoms the most general formula would be AB<sub>n</sub>C<sub>n</sub>D<sub>2</sub> ...**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

Table IV: Possible point groups and example systems for all possible chemical formulas of triatomic systems.

Point group	Chemical formula type		
	A <sub>3</sub>	A <sub>2</sub> B	ABC
$D_{\infty h}$	C <sub>3</sub>	CO <sub>2</sub>	✗
$C_{\infty v}$		N <sub>2</sub> O	HOS
$D_{3h}$	H <sub>3</sub> <sup>+</sup>	✗	✗
$C_{2v}$	O <sub>3</sub>	H <sub>2</sub> O	✗
$C_s$	H...H <sub>2</sub>	HO <sub>2</sub>	HNO

$C_n$

$n = 2$

If we have any fewer than three atoms, the point group of the system will be promoted to  $K_h$ ,  $D_{\infty h}$  or  $C_{\infty v}$ ; and if we have only three atoms, they form a plane which will be a mirror plane, but the  $C_2$  point group does not include any mirror planes, so three atoms would not be enough for the point group of a system to be  $C_2$ . Four atoms would be enough, as in the case of H<sub>2</sub>O<sub>2</sub> (whose structure is depicted in Fig. V). **The minimum number of atoms required is four.** In terms of the possible formulas, if we have two atoms at positions  $A$  and  $B$ , then after rotation around the  $C_2$  axis (which we will make the  $z$ -axis), these atoms will be at positions  $A' = (-x_1 - y_1, z_1)$  and  $B' = (-x_2 - y_2, z_2)$  respectively, and for the  $C_2$  operation to be valid, the atoms that were originally at positions  $A$  and  $A'$  need to be identical, and likewise for the atoms that were originally at  $B$  and  $B'$ . **The most general formula for the minimal structure is therefore A<sub>2</sub>B<sub>2</sub>.** Any number of additional atoms can be added along the  $C_2$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations (only the  $C_2$  element and the identity element in the case of the  $C_2$  point group) are invalidated. Therefore any number of atoms equal to four or more can

Table V: Possible point groups and example systems for all possible chemical formulas of tetratomic systems.

Group	Chemical formula type				
	A <sub>4</sub>	A <sub>3</sub> B	A <sub>2</sub> B <sub>2</sub>	A <sub>2</sub> BC	ABCD
T <sub>d</sub>		X	X	X	X
D <sub>∞h</sub>		X		X	X
C <sub>∞v</sub>		C <sub>3</sub> H		C <sub>2</sub> HF	HCNO
D <sub>4h</sub>		X	X	X	X
D <sub>3h</sub>			X	X	X
D <sub>2h</sub>		X		X	X
D <sub>2d</sub>		X	X	X	X
D <sub>2</sub>		X	X	X	X
C <sub>3v</sub>			X	X	X
C <sub>2v</sub>					X
C <sub>2h</sub>		X		X	X
C <sub>2</sub>		X		X	X
C <sub>s</sub>					
C <sub>1</sub>					

form a structure with the point group  $C_2$  and **the most general formula for such a structure is A<sub>2</sub>B<sub>2</sub>CD...**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

$$\underline{n \geq 3}$$

Since the point group of a regular  $n$ -gon with identical atoms at each vertex is required in order to have a  $C_n$  axis, and we have already mentioned in Section IV that the point group of such a structure would be  $D_{nh}$ , we would need at least one additional atom to break any symmetry elements in  $D_{nh}$  that are not present in  $C_n$ . In order for the  $C_n$  operation to remain valid, we can either add any number of atoms (including just one) on the  $C_n$  axis, or a full set of  $n$  identical (but not necessarily the same as the  $n$  identical atoms that formed the first  $n$ -gon) atoms that form a regular  $n$ -gon with the  $C_n$  axis going through its center and perpendicular to the plane formed by the  $n$ -gon. If the second  $n$ -gon is not on the same plane as the the first one, then we will not have the  $\sigma_h$  element that the  $C_{nh}$  point groups have; and if the second  $n$ -gon is not rotated by exactly an angle of  $\pi/n$  radians in comparison to the first  $n$ -gon, then we will not have the  $S_n$  operation that is present in the  $D_{nd}$  point groups have. The only difference between the  $D_n$  and  $C_n$  point groups are the extra  $C'_2$  elements that the  $D_n$  groups have, and these symmetry elements can be removed (without introducing any new symmetry elements) simply by choosing the bonds between atoms as we have done in the  $C_n$  row in Table III. **The minimum number of atoms required is  $2n$ , and the most general formula for this minimal structure is A<sub>n</sub>B<sub>n</sub>.** Any number of additional atoms can be added along the  $C_n$  axis in such a way that no new valid symmetry operations are made possible, and such that none of the existing valid symmetry operations (only the  $C_n$  element and the identity element in the case of the  $C_n$  point groups) are invalidated. Therefore any number of atoms equal to  $n$  or more can form a structure with the point group  $n$  and **the most general formula for such a structure is A<sub>n</sub>B<sub>n</sub>CD...**, and any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to equal A if desired).

## VI. SUMMARY OF OUR RESULTS AND OUTLOOK

The results of the previous section are summarized in Tables II and III. Table II only shows the possible point groups for system with up to 8 atoms because a smaller font or a change in other aspects of the formatting would be needed to extend the table to include systems with 9 atoms, and such an extension would not be particularly enriching, especially since the general chemical formulas presented in the previous section and summarized in Table III can be used to generate a table that includes systems with any *arbitrary*

number of atoms. Likewise the reason for Table III having columns up to 16 atoms is similar, but such a table can be constructed with up to an arbitrary number of atoms, with each column after 8 atoms containing indicating that the point group of such a system cannot be  $K_h$  and it can be any of the other point groups that are listed in the table.

### A. Outlook

Apart from finding molecules to fill in the tables, z-matrix generator in OpenMolcas.

- 
- [1] P. W. Atkins, *Tables for Group Theory* (London, Oxford University Press, 1970).
- [2] D. M. Bishop, *Group Theory and Chemistry* (Clarendon Press, Oxford, 1973).
- [3] R. L. Carter, *Molecular Symmetry and Group Theory* (J. Wiley, New York, 1998).
- [4] F. A. Cotton, *Chemical Applications of Group Theory* (Interscience, New York, 1963).
- [5] G. Davidson, *Group Theory for Chemists*, Macmillan physical science (Macmillan, Basingstoke, 1991).
- [6] M. Diem, *Introduction to Modern Vibrational Spectroscopy* (Wiley, New York, 1993).
- [7] R. S. Drago, *Physical Methods in Chemistry*, Saunders golden sunburst series (Saunders, Philadelphia, 1977).
- [8] J. Ferraro and J. Ziomek, *Introduction to Group Theory and Its Application to Molecular Structure* (Plenum Press, 1969).
- [9] L. Hall, *Group Theory and Symmetry in Chemistry* (McGraw-Hill, 1969).
- [10] D. Harris and M. Bertolucci, *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy* (Oxford University Press, 1978).
- [11] R. M. Hochstrasser, *Molecular Aspects of Symmetry* (W. A. Benjamin, New York, 1966).
- [12] J. M. Hollas, *Symmetry in Molecules*, Chapman and Hall chemistry textbook series (Chapman and Hall, London, 1972).
- [13] P. W. M. Jacobs, *Group Theory with Applications in Chemical Physics* (Cambridge University Press, Cambridge, 2005).
- [14] Y. Ohrn, *Elements of Molecular Symmetry* (Wiley Interscience, New York, 2000).
- [15] D. S. Urch, *Orbitals and Symmetry*, Penguin library of physical sciences. Chemistry. (Penguin Books, Harmondsworth, 1970).
- [16] J. P. Lowe, *Quantum Chemistry* (Academic Press, New York, 1978).
- [17] J. P. Lowe, *Quantum Chemistry*, 2nd ed. (Academic Press, Boston, 1993).
- [18] A. Vincent, *Molecular Symmetry and Group Theory : a Programmed Introduction to Chemical Applications* (Wiley, Chichester, 1977).
- [19] H. Bethe, *Annalen der Physik* **395**, 133 (1929).
- [20] F. Bruneval, N. Dattani, and M. Setten, *Frontiers in Chemistry* **9** (2021), 10.3389/fchem.2021.749779.
- [21] L. Monzel, C. Holzer, and W. Klopper, *The Journal of Chemical Physics* **158**, 144102 (2023).
- [22] K. Krause, M. E. Harding, and W. Klopper, *Molecular Physics* **113**, 1952 (2015).
- [23] L. E. Sutton and H. J. M. Bowen, *Tables of Interatomic Distances and Configuration in Molecules and Ions* (Chemical Society London, London, 1958).
- [24] L. E. Sutton, *Tables of Interatomic Distances and Configuration in Molecules and Ions* (Chemical Society London, London, 1965).
- [25] N. Vogt, *Structure Data of Free Polyatomic Molecules* (Springer, Cham, Switzerland, 2019).
- [26] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Inorganic Molecules* (Springer, 1998).
- [27] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules containing One or Two Carbon Atoms* (Springer, 1999).
- [28] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules containing Three or Four Carbon Atoms* (Springer, 2001).
- [29] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules Containing Five or More Carbon Atoms* (Springer, 2003).
- [30] N. V. K. Kuchitsu, M. Tanimoto, *Structure Data of Free Polyatomic Molecules · Inorganic Molecules Supplement to Volume II/25* (Springer, 2006).
- [31] N. V. K. Kuchitsu, M. Tanimoto, *Structure Data of Free Polyatomic Molecules · Molecules containing One or Two Carbon Atoms Supplement to Volume II/25* (Springer, 2006).
- [32] N. V. K. Kuchitsu, M. Tanimoto, *Structure Data of Free Polyatomic Molecules · Molecules containing Three or Four Carbon Atoms Supplement to Volume II/25* (Springer, 2007).
- [33] N. V. K. Kuchitsu, M. Tanimoto, *Structure Data of Free Polyatomic Molecules · Molecules containing Five or More Carbon Atoms Supplement to Volume II/25* (Springer, 2007).
- [34] M. T. K. Kuchitsu, N. Vogt, *Structure Data of Free Polyatomic Molecules · Molecules Containing No Carbon Atoms and Molecules Containing One or Two Carbon Atoms* (Springer, 2014).
- [35] M. T. K. Kuchitsu, N. Vogt, *Structure Data of Free Polyatomic Molecules · Molecules Containing Three or Four Carbon Atoms and Molecules Containing Five or More Carbon Atoms* (Springer, 2014).
- [36] E. B. Wilson, *Molecular vibrations : the theory of infrared and Raman vibrational spectra* (McGraw-Hill, New York ;, 1955).

### VII. ACKNOWLEDGMENTS

We thank Hemanth Haridas and Henri Abensour of HPQC Labs for helping us with this project. We thank Hemanth for discussions about point groups from around 15 July to 13 September 2023. We thank Henri for assistance in finding literature references (particularly the ones that were books rather than journal articles) and for assistance with typesetting the horizontal lines in Table II.

- [37] V. Bakken and T. Helgaker, *The Journal of Chemical Physics* **117**, 9160 (2002).
- [38] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, *Journal of Chemical Theory and Computation* **11**, 5665 (2015), pMID: 26642984.
- [39] A. Stuke, C. Kunkel, D. Golze, M. Todorović, J. T. Margraf, K. Reuter, P. Rinke, and H. Oberhofer, *Scientific Data* **7**, 58 (2020).
- [40] R. B. Shirts, *Journal of Chemical Education* **84**, 1882 (2007).
- [41] M. Bogey, H. Bolvin, C. Demuynck, and J. L. Destombes, *Phys. Rev. Lett.* **66**, 413 (1991).
- [42] M. Bogey, H. Bolvin, M. Cordonnier, C. Demuynck, J. L. Destombes, and A. G. Császár, *The Journal of Chemical Physics* **100**, 8614 (1994).
- [43] R. D. Johnson, “NIST Computational Chemistry Comparison and Benchmark Database,” (2022), NIST Standard Reference Database Number 101.
- [44] Private conversation between Nike Dattani and Gabe Cunningham (2023).