

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

Nike Dattani^{1,2,*} and Sichao Xiong^{1,†}

¹HPQC Labs, Waterloo, Canada

²HPQC College, Waterloo, Canada

(Dated: 18th October 2023)

For each axial point group, we determine all possible numbers of atoms and chemical formulas. We also 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 chemical formula. For triatomic and tetra-atomic molecules, we provide examples of molecules of each possible chemical formula if such molecules are known, and hence we determine the types of molecules with under five atoms that are still yet-to-be discovered (in terms of point groups). 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 [1], 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 [2–21] if they have ever come across a list of all possible point groups for a 4-atom (tetra-atomic) 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 now determined all possible point groups for systems with four atoms and also for systems with *any* number of atoms, 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 [2–21] 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 [7]. Nothing similar was found in any of [2–21], and the table of “shapes” in [7] 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 [2–21] was a fair survey of the known literature on the topic. The 2007 paper [21] reported a thorough survey of known published works that provided character tables for the S_8 and D_{8h} point groups, and our list [2–21] includes all of those works, in addition to [21] itself, the papers that cited [21], and also thoroughly selected additional works.

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} and 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 , and 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 [2–21]), 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 [22]. 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 [22].

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 in the next section. 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 summarized below, and all results appear to be novel or unusually difficult to find in existing places.

- 1) In Section III, we provide a table that presents visual examples of structures for all aforementioned axial point groups with $2 \leq n \leq 6$ and $n = \infty$,
 - 1.1) in a more clear way than in [7],
 - 1.2) with inconsistencies in [7] resolved,
 - 1.3) with more of the $n = \infty$ shapes presented than

* nike@hpqc.org

† sichao@hpqc.org

- in [7],
- 1.4) with structures requiring fewer atoms than in [7] for the C_n and D_n series,
 - 1.5) with an extra column showing additional structures for the $n = 2$ cases (in some cases these are more realistic structures than in [7]), and
 - 1.6) with a more logical order, especially in terms of pedagogical considerations, than in [7].
- 2) In Section IV, we describe some properties of various point groups and point group elements. We have found that formulating these properties was very helpful for obtaining or explaining new results, and very helpful for pedagogical purposes.
- 3) In Section V, for each axial point group, we determine:
- 3.1) the minimum number of atoms (the minimum cardinality) required;
 - 3.2) all possible cardinalities allowed (the minimum cardinality results are a special case of this); and
 - 3.3) all possible chemical formulas allowed (the list of all possible cardinalities is a special case of this).
- 4) In Section VI, as a consequence of the results in Section V,
- 4.1) 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 also be generated with the results from Section V but $m = 8$ was the largest number of atoms that we could include for the table to stay within the paper margins without making significant changes to the font, column spacing, or orientation of the table);
 - 4.2) 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 tetra-atomic systems (analogous lists for *arbitrary* chemical formulas can also be generated with the results from Section V);
 - 4.3) we attempted to provide examples of known molecules with each possible point group and for each triatomic or tetra-atomic chemical 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 (for structures with more atoms it would not be a surprise if certain shapes of molecules have not yet been discovered, but we determined that the minimum number of atoms required for the point group D_2 is 4 and yet 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.
- Throughout the paper, we use "Schönflies notation" for the point groups and point group elements, and to distinguish point *groups* such as C_2 from point group *elements* such as \hat{C}_2 , we will place a circumflex symbol on top of point group elements. An axis associated with a point group element will be named after that point group element, for example the axis associated with a \hat{C}_2 operation will be named a \hat{C}_2 axis; likewise a plane associated with a point group element will be named after that point group element, for example the plane associated with a $\hat{\sigma}_h$ operation will be named a $\hat{\sigma}_h$ plane. Also, all of the diagrams that we present for systems of atoms, were made with the most recent version of Avogadro 2 [23].

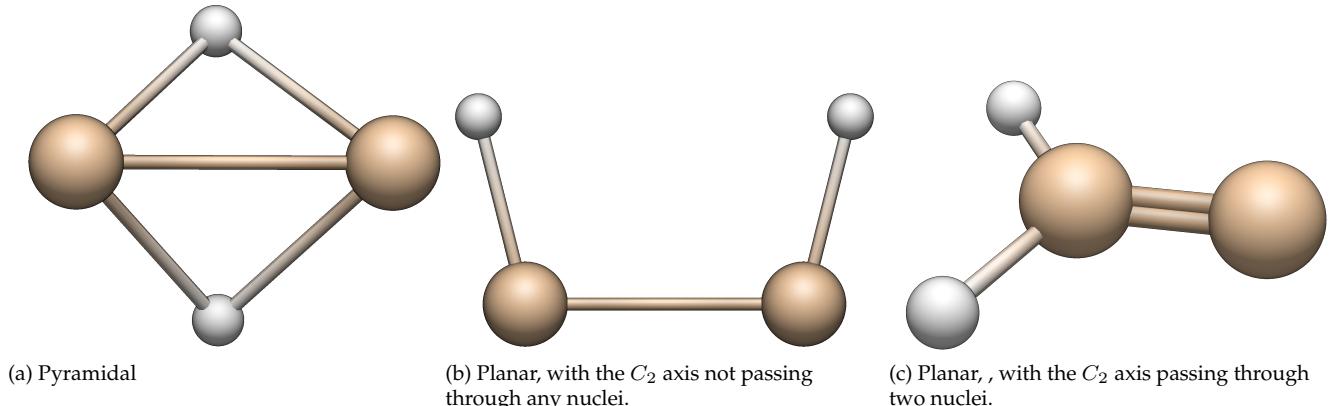
II. MOTIVATION FOR THIS WORK

Our need for answers to the question in the title of the present paper, arose during other work. GW100 is a dataset of 100 systems (95 molecules and 5 atoms) which was introduced in a 2015 paper [24] 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 [25] and 2023 [26] to improve the most accurate calculations that had been done on the system until then [27], 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, the GW100 community has been doing calculations using geometries that were optimized using density functional theory, which means that the results are dependent on the basis set and density functional used for the geometry optimization. This has an effect on (among other things) the comparisons that have been made between calculations that were done for the GW100 dataset and data from laboratory experiments. 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–Börnstein volumes and related books [28–40] and also in the original papers in which the geometric data in those compilations were determined.

The books [28–40] provide only the *point group*, a 3D diagram, and usually a "minimal" set of geometric data (bond lengths, bond angles and dihedrals). An example is the entry for H_2Si_2 in [30] (this molecule is not in GW100 but it is useful for illustrating this point) which provides the point group (C_{2v}), a 3D

Figure 1: Three molecules that form very different shapes despite having the same chemical formula (H_2Si_2) and the same point group (C_{2v}).



shape that is redrawn in Fig. 2a of this paper, and *only* two bond lengths and a dihedral angle: $r_{\text{Si-Si}} = (2.2154 \pm 0.0020)\text{\AA}$, $r_{\text{Si-H}} = (1.6680 + 0.0030)\text{\AA}$, $\varphi = (104.22 + 0.30)^\circ$ (the original experimental paper from which this data came [41] and the previous paper on the same molecule [42], 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 entire system of atoms, 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 φ , 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 applications 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.* [43]) 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 the data that we had available for H_2Si_2), would require specific code blocks for each molecular shape. It is possible for multiple molecules with the same point group and formula to have very different shapes (for example, see Fig. 1, in which three isomers of H_2Si_2 are shown, all with the C_{2v} point group, but one of them being pyramidal, one of them being planar with its C_2 axis not going through any nuclei, and one of them being planar with its C_2 axis going through two nuclei), 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 chemical formulas and point groups (the Landolt-Börnstein volumes and related books [28–40], the CCCBDB [44], 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

Table I: An improved version of the table presented on Page 37 of the 2008 version of [7], 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 D_{nh} and C_{nh} series are the only series in which all systems are planar (although the molecule depicted in the second column of the C_{2v} row is also planar). In the D_{nd} and D_n series, bonded atoms are always on the same plane as each other, and non-bonded atoms are always on different planes from each other. In the S_{2n} and C_n series, the red atoms form planes based on the plane of gray atoms to which the red atoms are bonded: if a group of red atoms are bonded to gray atoms that are on the same plane, then those red atoms will form a plane of their own, and will not include any atoms that are not bonded to gray atoms of the respective gray-atom plane.

The colors of the atoms are not intended to have any meaning apart from showing that the gray atoms have to all be of the same periodic table element as each other, and the red atoms have to all be of the same periodic table element as each other (the red atoms in all of these diagrams can also be of the same periodic table 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.

	$n = 2$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = \infty$
D_{nh}							
D_{nd}							
D_n							
S_{2n}							
C_{nh}							
C_n							
C_{nv}							

XYZ-construction code for each case, adding sub-cases when necessary (e.g. to distinguish between the three types of C_{2v} shapes for A_2B_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 title of the present paper, 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 $AB_2C_2D_2$ (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 C_3H_3N 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 [7] contains a table of “shapes” on Page 37, which appears to be the only table available of its kind (in all of the literature, see Section I). For each of the axial point group series (C_n , D_n , C_{nv} , C_{nh} , D_{nh} , D_{nd} , and S_{2n}), schematic diagrams are presented for n from 2 to 6 in attempt to provide an impression of how a molecule with the specified 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 and compared to analogous lines in the C_{4h} , C_{5h} and C_{6h} diagrams). Although every 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 are not the same size as others, and why some extra purple lines exist, most readers will be comfortable to ignore these inconsistencies and to simply attribute them to human error.

However, the aforementioned table in [7] has some other consistencies which are much more major, and which we would like to address in this section of the paper. For example, the table in [7] 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 [7] 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 cir-

cles 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 [7] 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 [7] that we addressed was the readability and interpretability of their diagrams. This problem is perhaps most severe in their diagrams for S_{2n} , in which the atoms in the first and third planes appear to be of one element of the periodic table (depicted in a beige color in their diagram), and the atoms in the second and fourth planes are of another periodic table element (depicted in a purple or maroon color in their diagram); 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 of one periodic table element, and the atoms in the second and third planes being of one periodic table element (as depicted in our Table III). This issue is also a problem for the D_n and D_{nd} series in their table: the diagrams for the D_{nd} series have two different colors, even though all atoms actually need to be the *same* type (element-wise), and the diagrams in the D_{nd} series do not make it clear that the two polygons shown for each value of n , are on different planes (this can be difficult to show graphically, but we have explained it in the caption to our Table III, whereas no caption or further explanations were given for the analogous table in [7]). 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 [7] 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, nor explained anywhere), 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 depict more than one additional atom each? In fact, both the straight lines and the small triangles probably 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 and other researchers that we asked, 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). Since our diagrams in the $n = 2$ to $n = 6$ columns of Table III included spheres to depict each atom, and thick lines to depict “bonding” between

the atoms, it is *much* easier for our readers to associate our diagrams with the shapes of actual molecules or systems of atoms with some bonds between them.

We also note that the “appendages” in the C_{3h} , C_{4h} , C_{5h} , C_{6h} , C_4 , C_5 and C_6 diagrams in [7] are parallel (in fact co-linear) with the edges of their respective polygons, 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 vertices of the polygons, and as long as the “appendage” atoms are not in the centers of the edges of the polygons, 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.

We also attempted to resolve a confusion associated with the $n = 2$ column in the table in [7]. For instance, the example structure shown for C_2 , is obscure and a far less common example for the C_2 point group than the structure of, for example, H_2O_2 (hydrogen peroxide). Likewise, H_2O (water) 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 [7] 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 un-interpretable “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 [7] 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 of a system to be G (for any chosen point group G)?”, and after answering that question we found that the table in [7] correctly depicts the minimum number of atoms required for all point groups except for the $n = 2$ ones (of every point group series!) and except for the C_n and D_n series (for all values of n). In our Table I, we have shown “*minimal*” structures (*i.e.* structures that have the smallest possible number of atoms for a specific point group) that we found for all C_n and D_n point groups shown, 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 diagrams that are presented in our Table I and the diagrams that the authors of [7] presented in their original table (when correct), 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 criterion, which was to only include for each point group, the structure that requires the minimum number of atoms (except for the first $n = 2$ col-

umn for each point group series, in which we decided that it would be interesting to show a structure requiring the minimum number of atoms when using a “rectangular base”).

The final difference that we will note between the table in [7] 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, all possible 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 our arguments 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 [2–21], although parts of the first statement are mentioned in other places, such as [45].

Ability to orient the coordinate system according to a \hat{C}_n axis and $\hat{\sigma}_h$ plane:

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

Existence of a \hat{C}_n axis:

The existence of a \hat{C}_n axis is only possible if all atoms are on the \hat{C}_n axis or if the system contains a subsystem 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 somewhere in the molecule. Therefore, a minimum of n atoms is needed in order to have a \hat{C}_n axis.

A \hat{C}_n axis would be perpendicular to this polygon and would go through the center of it (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. If a \hat{C}_n axis is

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

the z -axis, then after a \hat{C}_n rotation, an atom that originally had coordinates (x_1, y_1, z_1) , will have coordinates (x_2, y_2, z_1) in which the z -coordinate does not change, and the new coordinates (x_2, y_2) can be different from the old ones (x_1, y_1) but will coincide with the original (x, y) coordinates of another identical atom in the system.

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 a 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 \hat{S}_{2n} operation to be valid when $n \geq 3$:

The \hat{S}_{2n} operation is valid on a regular $2n$ -gon with uniform vertices, because the \hat{S}_{2n} operation is the combination of a \hat{C}_{2n} operation (which is valid on such a system) followed by a $\hat{\sigma}_h$ operation (which is also valid on such a system, since it is planar); and any non-linear object in which the \hat{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 \hat{S}_{2n} operation to be valid in a system in which the $\hat{\sigma}_h$ operation is not valid (as in the D_{nd} , S_{2n} , and T_d point groups), we cannot just have a $2n$ -gon since the $2n$ -gon itself forms a plane that can act as a $\hat{\sigma}_h$ plane. It turns out that for reasons that will be explained in the next paragraph, instead of one regular $2n$ -gon, *two* regular n -gons with one of them rotated by an angle of π/n from the other one, and both of them being on different parallel planes with a \hat{C}_n axis going through the center of each of the n -gons, is a sufficient property and a necessary part of any system in which there is a valid \hat{S}_{2n} operation and no valid $\hat{\sigma}_h$ operations (when $n \geq 3$). Each D_{nd} point group contains \hat{S}_{2n} elements and no $\hat{\sigma}_h$ elements, and the examples for D_{nd} systems that we presented for $n \geq 3$ in Table I have exactly the form described in the previous sentence.

The $\hat{\sigma}_h$ plane that is part of the \hat{S}_{2n} operation, cannot contain all atoms of the system, because then this plane would not only be the $\hat{\sigma}_h$ part of the \hat{S}_{2n} operation, but it would also be a valid $\hat{\sigma}_h$ operation for the entire system. At least one atom needs to not be on the $\hat{\sigma}_h$ plane of the \hat{S}_{2n} operation, and if it is on the \hat{C}_{2n} axis, then since a \hat{C}_{2n} operation does nothing to atoms that are on the \hat{C}_{2n} axis, a valid \hat{S}_{2n} operation would require the $\hat{\sigma}_h$ plane of the \hat{S}_{2n} operation to be a valid $\hat{\sigma}_h$ plane for the atom too; so at least one atom needs to be neither on the \hat{S}_{2n} operation's \hat{C}_{2n} axis nor on its $\hat{\sigma}_h$ plane. Since we can make the \hat{C}_{2n} axis of the \hat{S}_{2n} operation the z -axis, and the $\hat{\sigma}_h$ plane of the \hat{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 \hat{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 \hat{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 π/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 π/n with respect to corresponding vertices on the other n -gon, and the \hat{C}_{2n} axis of the \hat{S}_{2n} operation (also called the \hat{S}_{2n} axis) going through the centers of each n -gon. So not only is this shape sufficient for the existence of a valid \hat{S}_{2n} operation in a system that does not have a valid $\hat{\sigma}_h$ operation on its own, but it is a necessary sub-unit of any system whose point group has an \hat{S}_{2n} element and no $\hat{\sigma}_h$ 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 π/n from each other, so we will again have a $2n$ -gon, as described earlier.

Action of the inversion operation, i:

If there is a center of inversion, we can choose for it 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 homo-nuclear diatomic molecule into an identical molecule, the point group is considered to be $D_{\infty h}$ which is a super-group of D_{2h} that contains *all* of the elements that transform a homo-nuclear diatomic molecule into an identical molecule. We will use this convention throughout this paper, so for example, we will say that the point group of H_2 is $D_{\infty h}$ and not any other point group such as $C_{\infty v}$ or D_{2h} .

A. Axial point groups without axial rotational symmetry

The point groups considered in this subsection do not contain any \hat{C}_n axes in which $n > 1$.

C_s

Any two atoms that are at different positions form a line, and if all atoms of a system are on the same line, then the point group will be either $C_{\infty v}$ or $D_{\infty h}$. Likewise, a single atom has the point group K_h , so it is not possible for the point group of a system with fewer than three atoms to be C_s . Three atoms is enough for the point group of a system to be C_s though, as in the example HNO (azanone). Therefore **three is the minimum number of atoms required to form a structure whose point group is C_s , and such a structure can be formed with any triatomic chemical formula.**

We can add as many more atoms as we wish without changing the point group of the system from C_s : for example, the point group will not change if all of these additional atoms remain on the same plane as the first three atoms and they are not of the same periodic table element (having atoms with different elements guarantees that no *new* symmetry operations become valid on the system, and if we run out of periodic table elements, we just need to ensure that additional atoms are not matched symmetrically with other atoms of the same element). The point group C_s is also possible for non-planar structures (e.g. Cl_2OS which is depicted in our Table V), but we were already able to prove in the previous sentence that even if we restrict ourselves to planar systems, the point group C_s is possible for any number of atoms that is larger than or equal to three. Therefore **any number of atoms that is greater than or equal to three can form a structure whose point group is C_s , and the most general chemical formula for a C_s structure is $\text{ABCD}\cdots$** , but any letter can be identical to

one of the previous ones (e.g. B, C and D can be chosen to be equal to A if desired).

C_1

Any three atoms that are at different positions form a plane or a line, and if all atoms of a molecule are on the same plane, then that plane is a mirror plane. Therefore if a molecule has three or fewer atoms, it will at least have one mirror 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 H_2OS (whose structure is depicted in Table V). Therefore **four is the minimum number of atoms required to form a structure whose point group is C_1 , and such a structure can be formed with any tetra-atomic chemical formula.**

We can add as many more atoms as we wish without changing the point group of the system from C_1 : for example, the point group will not change if all of these additional atoms are not of the same periodic table element (having atoms with different elements guarantees that no *new* symmetry operations become valid on the system, and if we run out of periodic table elements, we just need to ensure that additional atoms are not matched symmetrically with other atoms of the same element). Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is C_1 , and the most general chemical formula for a C_1 structure is $\text{ABCDE}\cdots$** , but any letter can be identical to one of the previous ones (e.g. B, C, D and E can be chosen to be equal to 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 *both* an inversion center *and* a mirror plane, must be larger than C_i . If we have three atoms, then they are confined to a plane (or line), and if all atoms are on the same plane, then that plane is a mirror plane, which is an element that the C_i group does not contain. Therefore we need more than three atoms for the point group of the system to be C_i .

If we have four atoms, and we make the inversion center the origin, the atoms must come in pairs according to the inversion operation \hat{i} : an atom at the point $A = (x_1, y_1, z_1)$ will be paired with an atom at the point $-A = (-x_1, -y_1, -z_1)$ and an atom at the point $B = (x_2, y_2, z_2)$ will be paired with an atom at the point $-B = (-x_2, -y_2, -z_2)$. Assuming that no two atoms will have the same xyz -coordinates, none of these four atoms can be on the origin. Therefore the points A and $-A$ form a line, and the points B and $-B$ form another line, and these two lines both intersect at the origin, but two lines that intersect are always co-planar, so the inversion center guarantees that our four atoms lie on a mirror plane, which is an element that the C_i group does not contain. Therefore we need more than four atoms for the point group of the system to be C_i .

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

Point groups	Number of atoms																Possible chemical formula		
	Minimum	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···) ₂	A(BCDE···) ₂
$C_{\infty v}$	—	2	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	ABCDE···	
D_{nh}	$n \geq 3$	n	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	A_n (BCDE···) ₂	AB_n (CDE···) ₂
D_{nh}	$n = 2$	4	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	(ABCDE···) ₂	A(BCDE···) ₂
D_{nd}	$n \geq 2$	$2n$	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	(A _n BCDE···) ₂	A(B _n CDE···) ₂
D_n	$n \geq 2$	$2n$	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	(A _n BCDE···) ₂	A(B _n CDE···) ₂
S_{2n}	$n \geq 2$	$4n$	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	(A _n B _n CDE···) ₂	A(B _n C _n DE···) ₂
C_{nv}	$n \geq 2$	$n + 1$	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	AB _n CDE···	
C_{nh}	$n \geq 2$	$2n$	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	A_nB_n (CDE···) ₂	AB _n C _n (CDE···) ₂
C_n	$n \geq 2$	$2n$	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	A_nB_n CDE···	
C_s	—	3	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	ABCDE···	
C_i	—	6	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	(ABCDE···) ₂	A(BCDE···) ₂
C_1	—	4	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	ABCDE···	

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 again we will have a mirror plane that the C_i group does not contain. Therefore we need more than five atoms for the point group of the system to be 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. Therefore **six is the minimum number of atoms required to form a structure whose point group is C_i , and the most general chemical formula allowed for a structure with the minimum number of atoms is $\text{A}_2\text{B}_2\text{C}_2$** , and if desired, B and C can be chosen to be equal to A without changing the point group of the structure.

Any even number of atoms can be added to the minimal 6-atom structure while maintaining the point group C_i , as long as they are added in pairs of identical periodic table elements symmetrically around the inversion center. One additional atom of any periodic table element can be added to the inversion center without introducing any new point group elements and without breaking any of the existing ones (the existing ones being only the identity element and the inversion element). Therefore **any number of atoms that is greater than or equal to six can form a structure whose point group is C_i , and the most general chemical formula for a C_i structure with an even number of atoms is $\text{A}_2\text{B}_2\text{C}_2\text{D}_2\cdots$, and for an odd number of atoms the most general chemical formula is $\text{AB}_2\text{C}_2\text{D}_2\cdots$, but**

in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

B. Axial 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 \hat{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 position $(-x_1, -y_1, -z_1)$ 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 Br_2Na_2 (whose structure is depicted in Table V). Therefore **four is the minimum number of atoms required to form a structure whose point group is D_{2h} , and the most general chemical formula for a D_{2h} structure with the minimum number of atoms is A_2B_2** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

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 be paired with an atom of the *same element* symmetrically around the inversion center in order for the inversion operation \hat{i} to remain valid, and such pairs of atoms can be added to the primary \hat{C}_2 axis in the D_{2h} point group without breaking any of the other symmetry elements of the point group. Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is D_{2h} , and the most general chemical formula for a D_{2h} structure with an even number of atoms is $A_2B_2C_2D_2\cdots$, and for an odd number of atoms the most general chemical formula is $AB_2C_2D_2\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to 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 $n \geq 3$, and the point group of such a system is D_{nh} (see Section IV), we can conclude that **n is the minimum number of atoms required to form a structure whose point group is D_{nh} with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is A_n .**

An even number of additional atoms can be added on the \hat{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 *homo-nuclear 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 **any number of atoms that is greater than or equal to n can form a structure whose point group is D_{nh} with $n \geq 3$, and the most general chemical formula for such a structure with n atoms plus an even number of additional atoms is $A_nB_2C_2D_2\cdots$, and for an odd number of additional atoms the most general chemical formula is $A_nBC_2D_2\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

D_{nd}

$$\underline{n = 2}$$

The D_{2d} point 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 at least 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 the second $n = 2$ column in Table III, in which *we have two identical homo-nuclear diatomic molecules that are perpendicular to each other, but on different planes*. Therefore **four is the minimum number of atoms required to form a structure whose point group is D_{2d} , and the most**

general chemical formula for a D_{2d} structure with the minimum number of atoms is A_4 .

An even number of additional atoms can be added on the \hat{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 *homo-nuclear pairs* around the plane containing the two \hat{C}'_2 axes (since we can label the \hat{C}_2 axis as the z -axis, we can make this plane the xy -plane). One additional atom can also be added at the center of the entire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is D_{2d} , and the most general chemical formula for a D_{2d} structure with four atoms plus an even number of additional atoms is $A_4B_2C_2D_2\cdots$, and for an odd number of additional atoms the most general chemical formula is $AB_4C_2D_2\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

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

For any specific integer $n \geq 3$, the D_{nd} point group contains an \hat{S}_{2n} operation and no $\hat{\sigma}_h$ operation on its own, and in Section IV we explained that any system with a valid \hat{S}_{2n} operation with $n \geq 3$, but no $\hat{\sigma}_h$ operation on its own, must at least have a subset of atoms that form the shape shown for the appropriate n value in the D_{nd} row of Table III. Since for each integer $n \geq 3$, the corresponding D_{nd} structure that is depicted in Table III allows all elements of D_{nd} and no elements that are not in D_{nd} , the point group of these structures are D_{nd} . Therefore **$2n$ is the minimum number of atoms required to form a structure whose point group is D_{nd} with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is A_{2n} .**

An even number of additional atoms can be added on the primary \hat{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 *homo-nuclear pairs* around the plane that contains all n of the \hat{C}'_2 axes (since we can choose the z -axis to be the \hat{C}_n axis, this plane can be chosen to be the xy -plane). One additional atom can also be added at the center of the entire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to $2n$ can form a structure whose point group is D_{nd} with $n \geq 3$, and the most general chemical formula for such a structure with $2n$ atoms plus an even number of additional atoms is $A_{2n}B_2C_2D_2\cdots$, and for an odd number of**

additional atoms the most general chemical formula is $\text{AB}_{2n}\text{C}_2\text{D}_2 \dots$, but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

D_n

$n = 2$

The D_2 point group does not contain 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 homo-nuclear diatomic molecules that are not perpendicular to each other, and are on different planes*. Therefore **four is the minimum number of atoms required to form a structure whose point group is D_2 , and the most general chemical formula for a D_2 structure with the minium number of atoms is A_4 .**

An even number of additional atoms can be added along a \hat{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 *homo-nuclear pairs* around the plane containing the two other \hat{C}_2 axes (since we can label the first \hat{C}_2 axis as the z -axis, the other two \hat{C}_2 axes can be labeled as $\hat{C}_2(x)$ and $\hat{C}_2(y)$ axes, and therefore the plane containing them would be the xy -plane). One additional atom can also be added at the center of the entire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is D_2 , and the most general chemical formula for a D_2 structure with four atoms plus an even number of additional atoms is $\text{A}_4\text{B}_2\text{C}_2\text{D}_2 \dots$, and for an odd number of additional atoms the most general chemical formula is $\text{AB}_4\text{C}_2\text{D}_2 \dots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

$n \geq 3$

The D_n point groups can be generated by a \hat{C}_n operation and a perpendicular \hat{C}'_2 operation, whereas the D_{nd} point groups can be generated by an \hat{S}_n operation and a perpendicular \hat{C}'_2 operation (the only difference being the \hat{S}_{2n} generator for D_{nd} instead of the \hat{C}_n generator for D_n). If we start with the structure that we described in the previous subsection for D_{nd} with $n \geq 3$, one way that we can “demote” the \hat{S}_{2n} axis to a \hat{C}_n axis without changing any of the struc-

ture’s 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 π/n (as would be required for an \hat{S}_{2n} axis by the logic that was presented in Section IV). This minor rotation will break the validity of the \hat{S}_{2n} operation while maintaining the validity of the \hat{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 the corresponding D_{nd} point groups, as are the chemical formulas for structures with the minimum number of atoms. Likewise, the possible numbers of atoms for structures with D_n point groups are the same as for the corresponding D_{nd} point groups, as are the possible chemical formulas.**

Table IV: Possible point groups and example systems for all possible chemical formula types for triatomic systems.

Point group	Chemical formula type		
	A_3	A_2B	ABC
$D_{\infty h}$	C_3	CO_2	✗
$C_{\infty v}$	$\text{H} \cdots \text{H}_2$	N_2O	HOS
D_{3h}	H_3^+	✗	✗
C_{2v}	O_3	H_2O	✗
C_s	$\text{Li} \cdots \text{Li}_2$	HO_2	HNO

S_{2n}

$n = 2$

From Section IV we know that within one coordinate system, we can simultaneously choose a \hat{C}_n axis to be the z -axis and a $\hat{\sigma}_h$ plane to be the xy -plane. With that convention, we can use the well-known $R_z(\pi)$ rotation matrix to see that a \hat{C}_4 operation will transform the point (x, y, z) to $(-y, x, z)$. Then since a reflection across the xy -plane will not change the x or y coordinates and will just reflect the z coordinate, the $\hat{\sigma}_h$ operation will transform the point $(-y, x, z)$ to $(-y, x, -z)$; so the \hat{S}_4 operation which is the \hat{C}_4 operation followed by the $\hat{\sigma}_h$ 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 \hat{S}_4 transformation will move it to the position $(-y_1, x_1, -z_1)$, and an identical atom must be there for the \hat{S}_4 operation to be valid. Also after the same \hat{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 yet replaced our first atom which was at the position (x_1, y_1, z_1) , we need to keep applying \hat{S}_4 transformations and to keep 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 \hat{S}_4 transformation and recover the same system:

$$(x_1, y_1, z_1) \rightarrow (-y_1, x_1, -z_1) \rightarrow (-x_1, -y_1, z_1) \rightarrow \quad (2) \\ (y_1, -x_1, -z_1) \rightarrow (x_1, y_1, z_1).$$

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 homo-nuclear in comparison to the first set of four atoms) in order to break the $\hat{\sigma}_d$ planes that are in D_{2d} , but adding any number of atoms on the \hat{S}_4 axis, or on the $\hat{\sigma}_h$ plane of the \hat{S}_4 operation, or on either one of the $z = \pm z_1$ planes will not break the $\hat{\sigma}_d$ planes. Since we have chosen the convention in which the \hat{S}_4 axis is the z -axis, adding atoms on this axis would mean adding atoms with $(x, y) = (0, 0)$. Also, since we have chosen the convention in which the $\hat{\sigma}_h$ plane associated with the \hat{S}_4 operation is the xy -plane, adding atoms on this plane would mean adding atoms with $z = 0$. Furthermore, regardless of the conventions that were mentioned in the previous two sentences, any atom on one of the $\pm z$ planes will have a z -coordinate of z or $-z$. To break the $\hat{\sigma}_d$ 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 (which would be the fifth atom), the pattern in (2) shows that we will not recover the position of the new atom by applying \hat{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 relationships that were described in the previous paragraph, among the coordinates of the atoms. Since the S_4 point group is generated by the \hat{S}_4 operation only, and an \hat{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 . Therefore **eight is the minimum number of atoms required to form a structure whose point group is S_4 , and the most general chemical formula for an S_4 structure with the minimum number of atoms is A_4B_4** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

An even number of additional atoms can be added on the \hat{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 *homo-nuclear pairs* around the $\hat{\sigma}_h$ plane of the \hat{S}_4 operation. One additional atom can also be added at the center of the en-

tire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to eight can form a structure whose point group is S_4 , and the most general chemical formula for an S_4 structure with eight atoms plus an even number of additional atoms is $A_4B_4C_2D_2\cdots$, and for an odd number of additional atoms the most general chemical formula is $AB_4C_4D_2\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

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

For any specific integer $n \geq 3$, in section IV we described structures that have the minimum number of atoms required for an \hat{S}_{2n} operation to be valid with $n \geq 3$ and for no $\hat{\sigma}_h$ operations to be valid on their own. We also described earlier that for the specified value of n , the point group of such a structure by itself is D_{nd} . In order to make a structure in which the point group is S_{2n} with as few atoms as possible, we can start with the D_{nd} structure for the same value of n (since the D_{nd} structure is a required subsystem of any S_{2n} structure) and we can 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 \hat{S}_{2n} axis or the $\hat{\sigma}_h$ plane that forms part of the \hat{S}_{2n} operation, it will not break the $\hat{\sigma}_d$ planes that are elements of D_{nd} , so we need to add at least one atom to a location that is neither on the \hat{S}_{2n} axis nor on the $\hat{\sigma}_h$ plane associated with it. This atom does not need to be homo-nuclear 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 with uniform vertices, such that these two n -gons are identical to each other and parallel to each other, but not co-planar and not on the same planes as the two previously existing n -gons in the structure, and the \hat{S}_{2n} axis must pass perpendicularly 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 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). Therefore **$4n$ is the minimum number of atoms required to form a structure whose point group is S_{2n} with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is $A_{2n}B_{2n}$** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

An even number of additional atoms can be added on the \hat{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 *homo-nuclear*

pairs symmetrically on each side of the $\hat{\sigma}_h$ plane associated with the \hat{S}_{2n} operation. One additional atom can also be added at the center of the overall system without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to $4n$ can form a structure whose point group is S_{2n} with $n \geq 3$, and the most general chemical formula for such a structure with $4n$ atoms plus an even number of additional atoms is $A_{2n}B_{2n}C_2D_2\cdots$, and for an odd number of additional atoms the most general chemical formula is $AB_{2n}C_{2n}D_2\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

C_{nv}

$$\underline{n = 2}$$

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 (water). Therefore **three is the minimum number of atoms required to form a structure whose point group is C_{2v} , and the most general chemical formula for a C_{2v} structure with the minimum number of atoms is A_2B** , since at least two atoms need to be the same in order for the \hat{C}_2 operation to be valid (but if desired, B can be chosen to be equal to A without changing the point group of the structure).

Since extra atoms can be added on the \hat{C}_2 axis without losing any valid symmetry operations and without introducing any new ones, **any number of atoms that is greater than or equal to three can form a structure whose point group is C_{2v} , and the most general chemical formula for a C_{2v} structure is $A_2BCD\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

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

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 \hat{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 such atom is sufficient, since if this atom is placed on a \hat{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 $\hat{\sigma}_h$ plane and any \hat{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 \hat{C}_n axis and n vertical mirror planes, without having any horizontal mirror planes or \hat{S}_n axes, is C_{nv} , so the point group of the pyramid that we have formed in this way must be C_{nv} . Therefore **$n + 1$ is the minimum number of atoms required to form a structure whose point group is C_{nv} with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is A_nB** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

Since any number of additional atoms can be added on the \hat{C}_n axis without losing any valid symmetry operations and without introducing any new ones, therefore **any number of atoms that is greater than or equal to $n + 1$ can form a structure whose point group is C_{nv} with $n \geq 3$, and the most general chemical formula for such a structure is $A_nBCD\cdots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

C_{nh}

$$\underline{n = 2}$$

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 \hat{i} , and we explained that this 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_2N_2 , whose structure is depicted in Figure V). Therefore **four is the minimum number of atoms required to form a structure whose point group is C_{2h} , and the most general chemical formula for a C_{2h} structure with the minimum number of atoms is A_2B_2** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

An even number of additional atoms can be added on the \hat{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* that are placed symmetrically around the $\hat{\sigma}_h$ plane. One additional atom can also be added at the center of the entire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is C_{2h} , and the most general chemical formula for a C_{2h} structure with four atoms plus an even number of additional atoms is $A_2B_2C_2D_2\cdots$** ,

and for an odd number of additional atoms the most general chemical formula is $\text{AB}_2\text{C}_2\text{D}_2 \dots$, but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

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

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 \hat{C}_n axis (which we will label as our z -axis) will not break the $\hat{\sigma}_v$ planes, so we are forced to add another regular n -gon comprised of n identical atoms at its vertices, but these atoms do not need to be the same as the n identical atoms 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 vertices of the first n -gon, and as long as the vertices of the second n -gon are not on the edges of the first n -gon (both of these placements of the second n -gon's vertices would result in a failure to break the $\hat{\sigma}_v$ planes and would not successfully demote the point group from D_{nh} to C_{nh}). Therefore **$2n$ is the minimum number of atoms required to form a structure whose point group is C_{nh} with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is A_nB_n** , and if desired, B can be chosen to be equal to A without changing the point group of the structure.

An even number of additional atoms can be added on the \hat{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* that are placed symmetrically around the $\hat{\sigma}_h$ plane. One additional atom can also be added at the center of the entire structure without introducing any new symmetry operations and without invalidating any existing ones. Therefore **any number of atoms that is greater than or equal to $2n$ can form a structure whose point group is C_{nh} with $n \geq 3$, and the most general chemical formula for such a structure with $2n$ atoms plus an even number of additional atoms is $\text{A}_n\text{B}_n\text{C}_2\text{D}_2 \dots$, and for an odd number of additional atoms the most general chemical formula is $\text{AB}_n\text{C}_n\text{D}_2 \dots$** , but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B, C, and D can be chosen to be equal to A if desired).

$$\underline{C_n}$$

Table V: Possible point groups and example systems for all possible chemical formula types for tetra-atomic systems.

Group	Chemical formula type				
	A_4	A_3B	A_2B_2	A_2BC	ABCD
T_d					
$D_{\infty h}$					
$C_{\infty v}$					
D_{4h}					
D_{3h}					
D_{2h}					
D_{2d}					
D_2					
C_{3v}					
C_{2v}					
C_{2h}					
C_2					
C_s					
C_1					

$n = 2$

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 can act as 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_2O_2 (hydrogen peroxide, whose structure is depicted in Fig. V). Therefore **four is the minimum number of atoms required to form a structure whose point group is C_2 .** In terms of the possible chemical formulas, if we have two atoms at positions A and B , then after rotating them by an angle of $\pi/2$ 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 \hat{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' . Therefore **the most general chemical formula for a C_2 structure with the minimum number of atoms is A_2B_2 ,** and if desired, B can be chosen to be equal to A without changing the point group of the structure.

Any number of additional atoms can be added on the \hat{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 \hat{C}_2 element and the identity element in the case of the C_2 point group) are invalidated. Therefore **any number of atoms that is greater than or equal to four can form a structure whose point group is C_2 , and the most general chemical formula for a C_2 structure is $\text{A}_2\text{B}_2\text{CD}\cdots$,** but in either chemical formula, any letter can be identical to one of the previous ones (e.g. B , C , and D can be chosen to be equal to A if desired).

$n \geq 3$

Since the point group of a regular n -gon with identical atoms at each vertex is required in order to have a \hat{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 the point group D_{nh} that are not present in the point group C_n . In order for the \hat{C}_n operation to remain valid, we can either add any number of atoms (including just one) on the \hat{C}_n axis, or a full set of n identical atoms (but not necessarily the same as the n identical atoms that formed the original n -gon) that form a new regular n -gon with the \hat{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 $\hat{\sigma}_h$ element that the C_{nh} point groups have;

and if the second n -gon is not rotated by exactly an angle of π/n radians in comparison to the first n -gon, then we will not have the \hat{S}_n operation that the D_{nd} point groups have. The only difference between the D_n and C_n point groups are the extra \hat{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. Therefore **2n is the minimum number of atoms required to form a structure whose point group is C_n with $n \geq 3$, and the most general chemical formula for such a structure with the minimum number of atoms is A_nB_n ,** and if desired, B can be chosen to be equal to A without changing the point group of the structure.

Any number of additional atoms can be added on the \hat{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. Therefore **any number of atoms that is greater than or equal to $2n$ can form a structure whose point group is C_n with $n \geq 3$, and the most general chemical formula for such a structure is $\text{A}_n\text{B}_n\text{CD}\cdots$,** and if desired, any letter can be made identical to one of the previous ones without changing the point group of the structure (e.g. if desired B , C , and/or D can be chosen to be equal to A without changing the point group of the structure).

VI. SUMMARY OF OUR RESULTS AND OUTLOOK

The results of the previous section (Section V) are summarized in Tables II and III. Table II only shows the possible point groups for systems 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 Section V 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 to include an arbitrary number of atoms, with each column after 16 atoms 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. Also, better flowcharts for students, and non-axial point groups, and bigger version of Table 1 that contains other possible shapes.

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.

-
- [1] H. Bethe, *Annalen der Physik* **395**, 133 (1929).
 - [2] E. B. Wilson, *Molecular vibrations : the theory of infrared and Raman vibrational spectra* (McGraw-Hill, New York ;, 1955).
 - [3] F. A. Cotton, *Chemical Applications of Group Theory* (Interscience, New York, 1963).
 - [4] R. M. Hochstrasser, *Molecular Aspects of Symmetry* (W. A. Benjamin, New York, 1966).
 - [5] L. Hall, *Group Theory and Symmetry in Chemistry* (McGraw-Hill, 1969).
 - [6] J. Ferraro and J. Ziomek, *Introduction to Group Theory and Its Application to Molecular Structure* (Plenum Press, 1969).
 - [7] P. W. Atkins, *Tables for Group Theory* (London, Oxford University Press, 1970).
 - [8] D. S. Urch, *Orbitals and Symmetry*, Penguin library of physical sciences. Chemistry. (Penguin Books, Harmondsworth, 1970).
 - [9] J. M. Hollas, *Symmetry in Molecules*, Chapman and Hall chemistry textbook series (Chapman and Hall, London, 1972).
 - [10] D. M. Bishop, *Group Theory and Chemistry* (Clarendon Press, Oxford, 1973).
 - [11] A. Vincent, *Molecular Symmetry and Group Theory : a Programmed Introduction to Chemical Applications* (Wiley, Chichester, 1977).
 - [12] R. S. Drago, *Physical Methods in Chemistry*, Saunders golden sunburst series (Saunders, Philadelphia, 1977).
 - [13] J. P. Lowe, *Quantum Chemistry* (Academic Press, New York, 1978).
 - [14] D. Harris and M. Bertolucci, *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy* (Oxford University Press, 1978).
 - [15] G. Davidson, *Group Theory for Chemists*, Macmillan physical science (Macmillan, Basingstoke, 1991).
 - [16] M. Diem, *Introduction to Modern Vibrational Spectroscopy* (Wiley, New York, 1993).
 - [17] J. P. Lowe, *Quantum Chemistry*, 2nd ed. (Academic Press, Boston, 1993).
 - [18] R. L. Carter, *Molecular Symmetry and Group Theory* (J. Wiley, New York, 1998).
 - [19] Y. Ohrn, *Elements of Molecular Symmetry* (Wiley Interscience, New York, 2000).
 - [20] P. W. M. Jacobs, *Group Theory with Applications in Chemical Physics* (Cambridge University Press, Cambridge, 2005).
 - [21] R. B. Shirts, *Journal of Chemical Education* **84**, 1882 (2007).
 - [22] Private conversation between Nike Dattani and Gabe Cunningham (2023).
 - [23] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison, *Journal of Cheminformatics* **4**, 17 (2012).
 - [24] 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.
 - [25] F. Bruneval, N. Dattani, and M. Setten, *Frontiers in Chemistry* **9** (2021), 10.3389/fchem.2021.749779.
 - [26] L. Monzel, C. Holzer, and W. Klopper, *The Journal of Chemical Physics* **158**, 144102 (2023).
 - [27] K. Krause, M. E. Harding, and W. Klopper, *Molecular Physics* **113**, 1952 (2015).
 - [28] L. E. Sutton and H. J. M. Bowen, *Tables of Interatomic Distances and Configuration in Molecules and Ions* (Chemical Society London, London, 1958).
 - [29] L. E. Sutton, *Tables of Interatomic Distances and Configuration in Molecules and Ions* (Chemical Society London, London, 1965).
 - [30] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Inorganic Molecules* (Springer, 1998).
 - [31] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules containing One or Two Carbon Atoms* (Springer, 1999).
 - [32] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules containing Three or Four Carbon Atoms* (Springer, 2001).
 - [33] K. Kuchitsu, *Structure Data of Free Polyatomic Molecules · Molecules Containing Five or More Carbon Atoms* (Springer, 2003).
 - [34] N. V. K. Kuchitsu, M. Tanimoto, *Structure Data of Free Polyatomic Molecules · Inorganic Molecules Supplement to Volume II/25* (Springer, 2006).
 - [35] 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).
 - [36] 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).
 - [37] 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).
 - [38] 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).
 - [39] 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).
 - [40] N. Vogt, *Structure Data of Free Polyatomic Molecules* (Springer, Cham, Switzerland, 2019).
 - [41] M. Bogey, H. Bolvin, M. Cordonnier, C. Demuynck, J. L. Destombes, and A. G. Császár, *The Journal of Chemical Physics* **100**, 8614 (1994).
 - [42] M. Bogey, H. Bolvin, C. Demuynck, and J. L. Destombes, *Phys. Rev. Lett.* **66**, 413 (1991).
 - [43] A. Stuke, C. Kunkel, D. Golze, M. Todorović, J. T. Margraf, K. Reuter, P. Rinke, and H. Oberhofer, *Scientific*

- Data **7**, 58 (2020).
- [44] R. D. Johnson, “NIST Computational Chemistry Comparison and Benchmark Database,” (2022), NIST Standard Reference Database Number 101.
- [45] V. Bakken and T. Helgaker, The Journal of Chemical Physics **117**, 9160 (2002).