



# The determination of point groups from imprecise molecular geometries

Peter J. Knowles<sup>1</sup>

Received: 22 July 2021 / Accepted: 27 October 2021 / Published online: 15 November 2021  
© The Author(s) 2021

## Abstract

We present a new approach for the assignment of a point group to a molecule when the structure conforms only approximately to the symmetry. It proceeds by choosing a coordinate frame that minimises a measure of symmetry breaking that is computed efficiently as a simple function of the molecular coordinates and point group specification.

**Keywords** Symmetry · Point group

## 1 Introduction

In principle, the discovery of the point-group symmetry of an isolated molecule with known atomic coordinates is a straightforward task that can be achieved with simple algorithms that search for particular symmetry elements and eliminate candidate groups based on the outcome. However, except in very simple cases, there are usually complicating factors associated with the provenance of the molecular structure. If it has been obtained from a crystal structure measurement, or from the numerical minimisation of an energy function, there will inevitably be noise in the structure that results in some or all of the likely actual symmetry elements being formally absent. Further complication can arise from the coordinate frame in which the structure is expressed; the task of finding symmetry elements involves varying the position of the coordinate origin and the orientation of the coordinate axes until a match is found. Again, in simple cases, especially if the underlying symmetry is in fact exact, orienting the molecule so that its axes and planes are in the expected places may be straightforward, but in unfavourable cases care is needed.

The imperfect satisfaction of symmetry relations arising from numerical noise is typically desirable to condone, because the structure is taken to stand for an ideal

---

✉ Peter J. Knowles  
KnowlesPJ@Cardiff.ac.uk

<sup>1</sup> School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

exact-symmetry structure. However it may also arise that the molecule has a structure that really is symmetry broken, in the sense that it is close to satisfying all the elements of some particular point group, but would remain distorted even if the precision of data were improved. It then becomes important to have clear criteria that allow one to distinguish unambiguously, against some threshold-based measures, these two possibilities.

The work of Avnir and coworkers [1–10] is largely based on discovering approximate symmetry elements, such as rotational axes, by numerical inspection of approximate motifs, such as regular polygons, in subsets of the atoms. Once an approximate motif has been found, a corresponding exact form is generated by averaging of coordinates; a symmetry measure is then constructed as mean square displacement of the actual structure from the idealised.

Largent, Polik and Schmidt adopt a somewhat simpler approach, again based on the atomic coordinates, but using the actual desired point group operations mapped onto the molecular frame in order to calculate the aggregate deviation from exact symmetry [11]. The coordinate transformation is defined by the principal inertial axes for asymmetric tops, with additional partially-specified criteria for symmetric and spherical tops. An efficient computational implementation is reported. Similar schemes are implemented in other software packages [12–14].

Closely related is the analysis of approximate point group symmetry from the viewpoint of fuzzy sets [15, 16], and using similarity measures based on electron density [17–20], and introducing the syntopy concept [21]. These methods have in common the need to already know the electron density, or even the potential energy surface, for a particular electronic state. The latter [15], in particular, is appealing, since it can give information on whether or not significant symmetry breaking is in play against an energy scale defined by temperature. These methods offer important insights and quantification that can be applied to any molecule. However, in the context of the present work, these approaches cannot be used directly, since we seek an analysis of near symmetry based only on the coordinates of the nuclei, typically to be used before an electronic structure computation is carried out.

Oakley et al. [22] describe use of the symmetry measure of Ref. [1] as a continuous function of the relative orientation of the molecule and point group coordinate frames that can be minimised by reorientation. This approach, which could be superior to simple use of inertial axes when symmetry is not exact, is developed further in the present work.

Grimme [23], and later Casanova and Alemany [24], adopt an approach where the focus is on quantum-mechanical wavefunctions. A symmetry measure can be defined based on the overlap of the wavefunction with its image induced by the action of the operator on the electronic coordinates, which deviates from unity as symmetry is broken.

In this paper, we introduce an alternative scheme that extends and combines these previous approaches, and defines measures of broken symmetry by providing clearly-defined algorithms that are invariant to coordinate system, and which support the automatic evaluation and discovery of point-group symmetry. The approximate symmetry is searched for by optimisation of a symmetry measure with respect to the position and orientation of the molecule, with the target symmetry specified

entirely by the definition of the particular point group being tested, rather than looking for specific features in the atomic coordinates, leading to a modular and robust algorithm. We also discuss the proper refinement of atomic positions in order to satisfy the elements of a chosen point group more precisely. The work is accompanied by a freely-available software implementation.

## 2 Methods

### 2.1 Symmetry measure

The molecule is defined by the coordinates and charges of a set of atoms,  $\{\vec{R}_A, Z_A\}$  for  $A = 1, \dots, N$ . We then consider a set of symmetry operators

$$G = \{\hat{T}_t, t = 1, 2, \dots, g\}. \quad (1)$$

$G$  would normally consist of all of the elements of a particular point group, but need not do so, as none of the closure or other group properties will be used; it could as a special case consist of just a single operator. Each operator acting on each atom gives rise to an image

$$\vec{R}_{At} = \hat{T}_t \vec{R}_A \quad (2)$$

and if the molecule conforms exactly to  $G$ ,  $\vec{R}_{At} \in G$  for all  $A, t$ . Since symmetry operators effect specific geometric transformations, we must be able to specify the location and orientation of any axes, planes, etc., and we will assume initially that the operators are defined in a fixed global coordinate system with cartesian axes  $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$  via real unitary representation matrices that act on the components in this coordinate system of any given position vectors:

$$\vec{r} = \sum_{\alpha} r_{\alpha} \vec{e}_{\alpha} \quad (3)$$

$$\hat{T}_t \vec{r} = \sum_{\alpha\beta} T_{t,\beta\alpha} r_{\alpha} \vec{e}_{\beta} = \mathbf{r}^{\dagger} \mathbf{T}_t^{\dagger} \vec{\mathbf{e}} \quad (4)$$

In order to describe inexact conformity with the group, we define for each image its matcher, the closest atom to the image,

$$B_{At} \leftarrow \min_{B_{At}} \left( d_{At} = |\vec{R}_{At} - \vec{R}_{B_{At}}| \right) \quad (5)$$

in which deviations of  $d_{At}$  from zero indicate symmetry breaking. The deviations can be used to define an overall continuous symmetry measure, for example [1]

$$F = g^{-1} \sum_{At} d_{At}^2 \quad (6)$$

Following instead the signature of symmetry in the electronic wavefunction [23, 24]

$$F = n^{-1} \sum_i^n \sum_t^g \langle \Psi | (1 - \hat{T}_t) | \Psi \rangle \quad (7)$$

where the symmetry operator in wavefunction space is defined as

$$\hat{T}_t \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = \sum_i^N \Psi(\vec{r}_1, \vec{r}_2, \dots, (\hat{T}_t^{-1} \vec{r}_i), \dots, \vec{r}_n) \quad (8)$$

If  $\hat{T}_t$  is an exact symmetry operation, the  $n$ -electron wavefunction is an eigenfunction with unit eigenvalue; otherwise, its expectation value is always less than one.

The coordinate- and wavefunction-based approaches can be unified by adopting a simple model for the wavefunction that reflects most of the features of the symmetry deviation of the true wavefunction induced by displacement of the nuclei from exact symmetry positions. We start with the case of a collection of hydrogen atoms—or more generally, a system where in the vicinity of the nuclei, the wavefunction is dominated by a single orbital, with the overall wavefunction an orbital product. Each orbital has the form  $\phi_A = \pi^{-1/2} Z_A^{3/2} \exp(-Z_A |\vec{r} - \vec{R}_A|)$  multiplied by an appropriate damping factor giving zero value near any other nucleus, and near the nucleus, the overall wavefunction is the product of this orbital and  $n - 1$  slowly-varying normalised orbitals. We adopt (7), but drop the  $n^{-1}$  factor; this gives a formulation that is formally extensive (if all atoms show the same symmetry deviation), and locally intensive (if additional atoms that obey the symmetry exactly are added, the measure does not change). We then obtain

$$F_0 = \sum_A \sum_t \langle \phi_A | (1 - \hat{T}_t) | \phi_A \rangle \quad (9)$$

$$= \sum_A \sum_t \left( 1 - \exp(-Z_A d_{At}) (1 + Z_A d_{At} + \frac{1}{3} Z_A^2 d_{At}^2) \right) \quad (10)$$

$$= \sum_A \sum_t f_0(Z_A d_{At}), \quad \text{where } f_0(x) = 1 - e^{-x} (1 + x + \frac{1}{3} x^2) = \frac{1}{6} x^2 + O(x^4) \quad (11)$$

This agrees with the simple point-charge formula except that near each nucleus, the distance unit is the inverse of the nuclear charge. Additionally, with the closed formula (11), the measure is bounded,  $0 \leq F_0 \leq N g$ , supporting interpretation of the value of  $F_0$ . Note that this formula is additive in both atoms (extensivity) and symmetry operations (descent to subgroups), allowing one to look at a chain of symmetry groups and evaluate the effect of lowering symmetry. As an example, water with one bond infinitely stretched gives  $F_0 = 0$  in  $C_s$ ,  $F_0 = 2.56$  in  $C_{2v}$ , and  $F_0 = 8.26$  in  $D_{2h}$  because of additional contributions from new operators.

For non-hydrogenic atoms, the model can in principle be extended. A simple way to do this is to recognise additional non-zero occupied orbitals near each nucleus that, because of the Coulomb singularity, will have the local shape of the  $2s, 3s, 2p, \dots$

orbitals of the 1-electron ion. The overlap integrals are similar in form to that in (10), varying as  $d_{A_i}^2$  but with a different coefficient [25]. Since the symmetry measure is meant to be only an arbitrary, but well-defined, quantity, we have chosen to ignore the consequences of many-electron wavefunction structure, and use (27) for all atoms without modification. This will tend to underemphasise the importance of positional deviations of heavier elements, but compared to the simple distance-squared criterion, they already have an additional  $Z_A^2$  weight.

## 2.2 Optimisation of symmetry operators

In principle, the measure defined above can be calculated for any suspected point group, and if the value is less than some chosen threshold, the molecule can be taken to be belonging to that group. This opens the way to using the standard point-group decision tree found in many textbooks [26] to systematically discover the highest-order compliant group. However, the molecular coordinates may happen to be expressed in any reference frame that is completely unrelated to the coordinate system used to define  $G$ , and it is then quite unlikely that a small measure will result. One approach to this challenge is to first ensure that the definition of the group is through operators that are in some standard orientation. For example, for an axial group, an obvious choice is for the unique high-order axis to be aligned along  $\vec{e}_z$ , and for some of the other axes or reflection plane normals to be chosen to coincide with  $\vec{e}_x$  or  $\vec{e}_y$ . Then the molecule can be translated so that its centre of charge is at the origin, and rotated so that the axes coincide with appropriate principal axes—for example, the eigenvectors of the second moment tensor. This works well for molecules with exact symmetries that are asymmetric tops, i.e., the three eigenvalues of the second moment tensor are distinct, but is incomplete when there are degeneracies (symmetric tops and spherical tops). If the molecule has only approximate symmetry, it is not obvious that this origin and axis choice is the best one.

Instead, we adopt an approach that is agnostic with respect to the frames in which the point group and molecule are specified, and effect a relative realignment that is varied until the symmetry measure  $F$  is minimum.

We choose to do this by leaving the molecular coordinates untouched, and specifying a rotation and translation of frame holding the symmetry operations.

We need to differentiate  $F$  with respect to the parameters that define the coordinate system (origin  $\vec{o}$ , axes  $\vec{a}_\alpha$ ,  $\alpha = 1, 2, 3$ ) in which the operators are defined. They consist of  $\mathbf{p} = \{o_1, o_2, o_3, q_1, q_2, q_3\}$  defined by

$$\vec{o} = \mathbf{o}^\dagger \vec{\mathbf{e}} = o_x \vec{e}_x + o_y \vec{e}_y + o_z \vec{e}_z \quad (12)$$

$$\vec{\mathbf{a}} = \mathbf{u}^\dagger \vec{\mathbf{e}} \quad \vec{a}_\alpha = \vec{e}_\beta u_{\beta\alpha} \quad (13)$$

$$\mathbf{u} = \mathbf{U}(\mathbf{q}) \quad (14)$$

where  $\vec{\mathbf{e}}$  are the vectors defining the global coordinate system, and  $\mathbf{U}$  is a function that produces a unitary matrix from three unconstrained parameters. Consideration

for the choice of  $\mathbf{U}$  included the matrix logarithm, quaternions and functions of the Euler angles that map to an infinite range; the final choice was the simple use of Euler angles. The position vector of any point can then be represented in either frame,

$$\vec{b} = \mathbf{b}^\dagger \vec{e} = \vec{o} + \bar{\mathbf{b}}^\dagger \vec{a} \quad (15)$$

$$\mathbf{b} = \vec{b} \cdot \vec{e} = \mathbf{o} + \mathbf{u} \bar{\mathbf{b}} \quad \bar{\mathbf{b}} = (\vec{b} - \vec{o}) \cdot \vec{a} = \mathbf{u}^\dagger (\mathbf{b} - \mathbf{o}) \quad (16)$$

and from here on, we assume that the symmetry operations are defined with respect to the frame  $\mathbf{p}$ .

The action of a symmetry operator on a point with position vector  $\vec{b} = \mathbf{b}^\dagger \vec{e}$  can then be represented through its effects on the global-frame coordinates,

$$\hat{T}_t \vec{b} = \vec{o} + (\mathbf{b} - \mathbf{o})^\dagger \mathbf{u} \hat{T}_t \vec{a} \quad (17)$$

$$= \vec{o} + (\mathbf{b} - \mathbf{o})^\dagger \mathbf{u} \bar{\mathbf{T}}_t^\dagger \vec{a} \quad (18)$$

$$= (\mathbf{o}^\dagger + (\mathbf{b} - \mathbf{o})^\dagger \mathbf{u} \bar{\mathbf{T}}_t^\dagger \mathbf{u}^\dagger) \vec{e}, \quad (19)$$

where we make use of the matrix representation of the symmetry operator in local coordinates,

$$\hat{T}_t \vec{a}_\alpha = (\bar{\mathbf{T}}_t)_{\beta\alpha} \vec{a}_\beta \quad (20)$$

$$\hat{T}_t (\vec{b} = \bar{\mathbf{b}}^\dagger \vec{a}) = (\bar{\mathbf{T}}_t \bar{\mathbf{b}})^\dagger \vec{a} \quad (21)$$

$\bar{\mathbf{T}}_t$  is fixed by the nature of the symmetry, but  $\hat{T}_t \vec{b}$  depends on it through the parameters  $\mathbf{p}$ .

We can now express the image errors in terms of the local symmetry matrices,

$$d_{At}^2 = |\hat{T}_t \vec{R}_A - \vec{R}_{B_{At}}|^2 = \mathbf{d}_{At}^\dagger \mathbf{d}_{At} \quad (22)$$

$$\mathbf{d}_{At} = \mathbf{o} + (\mathbf{R}_A - \mathbf{o})^\dagger \mathbf{u} \bar{\mathbf{T}}_t^\dagger \mathbf{u}^\dagger - \mathbf{R}_{B_{At}}, \quad (23)$$

which differentiate as

$$\frac{\partial}{\partial o_\alpha} d_{At\beta} = \delta_{\alpha\beta} - (\mathbf{u} \bar{\mathbf{T}}_t^\dagger \mathbf{u}^\dagger)_{\alpha\beta} \quad (24)$$

$$\frac{\partial}{\partial q_\alpha} d_{At\beta} = (\mathbf{R}_A - \mathbf{o})^\dagger \left( \frac{\partial \mathbf{u}}{\partial q_\alpha} \bar{\mathbf{T}}_t^\dagger \mathbf{u}^\dagger + \mathbf{u} \bar{\mathbf{T}}_t^\dagger \frac{\partial \mathbf{u}^\dagger}{\partial q_\alpha} \right)_\beta. \quad (25)$$

We can then proceed with

$$\nabla d_{At} = d_{At}^{-1} \mathbf{d}_{At}^\dagger \nabla \mathbf{d}_{At} \quad (26)$$

$$\nabla F = \sum_A \sum_t Z_A f'(Z_A d_{At}) \nabla d_{At} \quad (27)$$

and then varying  $\mathbf{p}$  to minimize  $F$  using the Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm [27]. In practice, in order to increase the convexity of the objective function, we minimize  $F_1$  which uses  $f_1(x) = x^2$  instead of  $f_0(x)$  specified in (11); if the optimum  $F$  is small, the minima of  $F_0$  and  $F_1$  will be close.

Computational efficiency can be improved without loss of generality by summing, in (11), over a reduced set of operators that are a generator set, i.e. a minimal set of operators that when combined sufficient times generate the complete point group. This leads to a different symmetry measure, but one which will be zero for exact symmetry, but not otherwise. For any given group, there is typically more than one valid generator set, and each will give rise to a numerically different symmetry measure. For this reason, we use  $F_1$  with a generator set for optimisation, but for comparison and testing of symmetry measures,  $F_0$  with the full group is used.

### 2.3 Choice of coordinate frame: further detail

For some point groups, there are multiple feasible coordinate frame orientations for which a molecule with exact symmetry will give  $F = 0$ . In these cases it is desirable to introduce further criteria that lead to an unambiguous standard orientation. They include

**Asymmetric tops—axis permutations:** We first construct and diagonalise the second moment tensor

$$I_{\alpha\beta} = \sum_A Z_A R_{\alpha A} R_{\beta A} \quad (28)$$

Asymmetric top molecules give three distinct eigenvalues, and restrict the possible point groups to  $D_{2h}$  and its subgroups. The coordinate axes are defined by the eigenvectors for an exact structure, which form good starting guesses for minimising  $F$  otherwise. There are 6 possible coordinate axis permutations, of which 4 are infeasible for those groups with a unique  $C_2$  axis or unique mirror plane. For  $D_{2h}$  and planar  $C_{2v}$  molecules, we follow Recommendations 5b, 5a respectively of Ref. [28], but otherwise any remaining freedom is satisfied by assigning the  $x$  axis to the minimum eigensolution.

**Symmetric tops—choice of perpendicular axes:** Symmetric tops are characterised by double degeneracy in the second moment eigensolutions. The unique eigensolution maps to the  $z$  axis, and we follow Recommendations 5c and 5d of Ref. [28] where possible.

**Spherical tops:** In spherical tops, all three eigenvalues of the second moment tensor are equal, and the eigenvectors offer no help in finding the orientation that

matches the molecule to the point group. Furthermore, the optimisation of  $F$  with respect to coordinate axes can be very ill-conditioned; for example, in the icosahedral groups, the lowest rank non-zero multipole moment has angular momentum 6, meaning that the system is very close to spherical. For these systems, we proceed by first determining the maximum-order rotational axis of the point group, and then looking for an approximate regular polygon of that order amongst the atoms. The normal vector of this polygon defines an axis which is then aligned to one of the point group's axes. Before entering BFGS optimisation, a discrete scan of rotations about that axis is performed, in order to find a starting guess with the lowest  $F$ . This procedure is relatively costly, but does succeed in aligning even icosahedral molecules such as  $C_{60}$ .

## 2.4 Purification

Often, the application of this methodology will be to discover the point group, or to determine the extent of deviation of the structure from a given point group. But a further use is to identify a point group, and then refine the geometry so that it conforms as exactly as possible to the group. We will then seek the least-motion distortion of the coordinates that will result in zero  $F$ . We can proceed by minimising  $F$  with respect to the atomic coordinates using BFGS. The atomic-coordinate gradient of  $F$  is evaluated straightforwardly as

$$\begin{aligned}\sigma_{C\alpha} &= \frac{\partial}{\partial R_{C\alpha}} F = \sum_A \sum_t \frac{\partial}{\partial R_{C\alpha}} f(Z_A d_{At}) \\ &= \sum_t Z_C f'(Z_C d_{Ct}) d_{Ct}^{-1} (\mathbf{u} \bar{\mathbf{T}}_t^\dagger \mathbf{u}^\dagger \mathbf{d}_{Ct})_\alpha \\ &\quad - \sum_t Z_{\bar{B}_{Ct}} f'(Z_{\bar{B}_{Ct}} d_{\bar{B}_{Ct},t}) d_{\bar{B}_{Ct},t}^{-1} d_{\bar{B}_{Ct},t\alpha}\end{aligned}\quad (29)$$

where we define the inverse image map  $\bar{B}_{B_{At},t} = A$ . However, this is an ill-posed problem, since when the symmetry conditions are obeyed exactly, any arbitrary step in the direction of any combination of coordinates that is a basis for a totally symmetric irreducible representation of the group will retain  $F = 0$ .

In principle, the arbitrariness can be overcome by always moving orthogonal to the null space, thereby satisfying, at least conceptually, an objective that might be expressed as finding the symmetrised structure that is as close as possible to the original. However, until the minimum is reached, the null space is not pure and exact, because of the slightly broken symmetry, and the steps in the numerical optimisation may introduce unnecessary additional motion in the symmetric directions.

Instead, a simple way to remove the arbitrariness is via a tie-breaking penalty function. We define a measure of displacement of the structure  $\{\vec{R}_A\}$  from its unrefined starting point  $\{\vec{R}_A^0\}$  as



$$P = \frac{1}{N(N-1)/2} \sum_A \sum_{B < A} \left( f_0(|\vec{R}_A - \vec{R}_B|) - f_0(|\vec{R}_A^0 - \vec{R}_B^0|) \right)^2. \quad (30)$$

$P \in [0, 1)$ , and measures the changes in all interatomic distances;  $f_0$  is used to map the semi-infinite range to  $[0, 1)$  and has the effect of de-emphasising the weight of changes of the distances between very distant atoms. We then vary the coordinates to minimise  $F + \pi P$ , where  $\pi$  is a small chosen parameter.

## 2.5 Software implementation

All of the methodology described above is incorporated in a freely available software library [29]. The library is written in C++ with additional bindings for C (including Fortran-callable functions) and command line. Its principal functions are optimisation of coordinate frame for a given molecule and point group, calculation of the symmetry measure, point group discovery, and structure refinement.

## 3 Performance

Table 1 illustrates the effect of fully optimising the coordinate frame to minimise the symmetry measure. Exact  $D_{2h}$  atomic coordinates for ethene are displaced randomly using a uniform distribution of specified width. The table shows, for a number of values of the noise parameter, the mean  $D_{2h}$  symmetry measures obtained by (a) not adjusting the coordinate frame; (b) adopting centre of charge and inertial axes; (c) full frame optimisation. A sufficiently large sample was taken to ensure the convergence of the means to the three significant figures quoted. It is seen that the use of inertial axes reduces the deviation by a factor of approximately 4, and that a further smaller, but significant, improvement results from full optimisation.

**Table 1** Mean  $D_{2h}$  symmetry measures for ethene contaminated with random noise in atomic coordinates. Each atomic coordinate is displaced by a random value drawn from a uniform distribution between plus and minus ‘Noise’. ‘Unoptimised’ is the large-sample mean symmetry measure without any readjustment of coordinate frame. ‘Inertial’ and ‘Optimised’ give the mean symmetry measures after adjustment to centre of mass and inertial axes, and after full frame optimisation, respectively

Noise	Unoptimised	Inertial	Optimised
$1.0 \times 10^{-6}$	$1.69 \times 10^{-10}$	$4.24 \times 10^{-11}$	$3.06 \times 10^{-11}$
$1.0 \times 10^{-5}$	$1.69 \times 10^{-8}$	$4.25 \times 10^{-9}$	$3.06 \times 10^{-9}$
$1.0 \times 10^{-4}$	$1.69 \times 10^{-6}$	$4.24 \times 10^{-7}$	$3.05 \times 10^{-7}$
$1.0 \times 10^{-3}$	$1.69 \times 10^{-4}$	$4.25 \times 10^{-5}$	$2.62 \times 10^{-5}$
$1.0 \times 10^{-2}$	$1.68 \times 10^{-2}$	$4.24 \times 10^{-3}$	$2.49 \times 10^{-3}$
$1.0 \times 10^{-1}$	1.44	$4.10 \times 10^{-1}$	$2.45 \times 10^{-1}$
$2.0 \times 10^{-1}$	4.35	1.53	$9.34 \times 10^{-1}$

## 4 Conclusion

A new approach for fuzzy assignment of a point group to a molecule has been described. It produces the best match possible by choosing a coordinate frame that minimises a measure computed as a simple function of the molecular coordinates and point group specification. Except for improving algorithm speed and robustness in high symmetry cases, no inspection of the detail of the structure is needed to identify individual symmetry operations, with the consequence that the resources needed scale only linearly with the number of atoms and with the size of the group.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. H. Zabrodsky, S. Peleg, D. Avnir, Continuous symmetry measures. *J. Am. Chem. Soc.* **114**, 7843–7851 (1992). <https://doi.org/10.1021/ja00046a033>
2. H. Zabrodsky, S. Peleg, D. Avnir, Continuous symmetry measures. 2. Symmetry groups and the tetrahedron. *J. Am. Chem. Soc.* **115**, 8278–8289 (1993). <https://doi.org/10.1021/ja00071a042>
3. H. Zabrodsky, D. Avnir, Continuous symmetry measures. 4. Chirality. *J. Am. Chem. Soc.* (1995). <https://doi.org/10.1021/ja00106a053>
4. Y. Salomon, D. Avnir, Continuous symmetry measures: A note in proof of the folding/unfolding method. *J. Math. Chem.* (1999). <https://doi.org/10.1023/a:1019144702913>
5. Y. Salomon, D. Avnir, Continuous symmetry measures: Finding the closest C2-symmetric object or closest reflection-symmetric object using unit quaternions. *J. Comput. Chem.* (1999). [https://doi.org/10.1002/\(SICI\)1096-987X\(199906\)20:8<772::AID-JCC3>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1096-987X(199906)20:8<772::AID-JCC3>3.0.CO;2-U)
6. M. Pinsky, K.B. Lipkowitz, D. Avnir, Continuous symmetry measures. VI. The relations between polyhedral point-group/subgroup symmetries. *J. Math. Chem.* (2001). <https://doi.org/10.1023/A:1013133602531>
7. M. Pinsky, C. Dryzun, D. Casanova, P. Alemany, D. Avnir, Analytical methods for calculating continuous symmetry measures and the chirality measure. *J. Comput. Chem.* (2008). <https://doi.org/10.1002/jcc.20990>
8. C. Dryzun, A. Zait, D. Avnir, Quantitative symmetry and chirality—A fast computational algorithm for large structures: Proteins, macromolecules, nanotubes, and unit cells. *J. Comput. Chem.* (2011). <https://doi.org/10.1002/jcc.21828>
9. M. Pinsky, A. Zait, M. Bonjack, D. Avnir, Continuous symmetry analyses:  $C_m$  and  $D_n$  measures of molecules, complexes, and proteins. *J. Comput. Chem.* (2013). <https://doi.org/10.1002/jcc.23092>
10. P. Alemany, D. Casanova, S. Alvarez, C. Dryzun, D. Avnir, Continuous symmetry measures: a new tool in quantum chemistry, in *Reviews in Computational Chemistry*, vol. 30, chapter 7 (Wiley, London, 2017), pp. 289–352
11. R.J. Largent, W.F. Polik, J.R. Schmidt, Symmetrizer: Algorithmic determination of point groups in nearly symmetric molecules. *J. Comput. Chem.* **33**, 1637–1642 (2012). <https://doi.org/10.1002/jcc.22995>

12. P.M.W. Gill, B.G. Johnson, J.A. Pople, A standard grid for density functional calculations. *Chem. Phys. Lett.* **209**, 506–512 (1993). [https://doi.org/10.1016/0009-2614\(93\)80125-9](https://doi.org/10.1016/0009-2614(93)80125-9)
13. <https://www.pqs-chem.com>. Accessed 8 July 2021
14. H.-J. Werner, P.J. Knowles, F.R. Manby, J.A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D.A. Kreplin, Q. Ma, T.F. Müller, A. Mitrushchenkov, K.A. Peterson, I. Polyak, G. Rauhut, M. Sibaev, The Molpro quantum chemistry package. *J. Chem. Phys.* **152**, 144107 (2020). <https://doi.org/10.1063/5.0005081>
15. J. Maruani, P. Mezey, From symmetry to syntopy: A fuzzy-set approach to quasi-symmetric systems. *Journal de Chimie Physique* **87**, 1025–1047 (1990). <https://doi.org/10.1051/jcp/19908701025>
16. P.W. Fowler, Vocabulary for fuzzy symmetry. *Nature* **360**, 626 (1992). <https://doi.org/10.1038/360626a0>
17. P.G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (Wiley-VCH, 1996)
18. P.G. Mezey, A proof of the metric properties of the symmetric scaling-nesting dissimilarity measure and related symmetry deficiency measures. *Int. J. Quantum Chem.* **63**, 105–109 (1997). [https://doi.org/10.1002/\(SICI\)1097-461X\(1997\)63:1<105::AID-QUA14>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1097-461X(1997)63:1<105::AID-QUA14>3.0.CO;2-B)
19. P.G. Mezey, Generalized chirality and symmetry deficiency. *J. Math. Chem.* **23**, 65–84 (1998). <https://doi.org/10.1023/a:1019121208423>
20. D. Casanova, P. Alemany, S. Alvarez, Symmetry measures of the electron density. *J. Comput. Chem.* **31**, 2389–2404 (2010). <https://doi.org/10.1002/jcc.21532>
21. P.G. Mezey, J. Maruani, The concept of 'syntopy'. *Mol. Phys.* **69**, 97–113 (1990). <https://doi.org/10.1080/00268979000100071>
22. M.T. Oakley, R.L. Johnston, D.J. Wales, Symmetrisation schemes for global optimisation of atomic clusters. *Phys. Chem. Chem. Phys.* **15**, 3965–3976 (2013). <https://doi.org/10.1039/c3cp44332a>
23. S. Grimme, Continuous symmetry measures for electronic wavefunctions. *Chem. Phys. Lett.* **297**, 15–22 (1998). [https://doi.org/10.1016/S0009-2614\(98\)01101-4](https://doi.org/10.1016/S0009-2614(98)01101-4)
24. D. Casanova, P. Alemany, Revisiting the foundations of symmetry operation measures for electronic wavefunctions. *Chem. Phys. Lett.* **511**, 486–490 (2011). <https://doi.org/10.1016/j.cplett.2011.06.080>
25. D.M. Silver, K. Ruedenberg, Overlap integrals over Slater-type atomic orbitals. *J. Chem. Phys.* **49**, 4301–4305 (1968). <https://doi.org/10.1063/1.1669874>
26. D.J. Willock, *Molecular Symmetry* (John Wiley & Sons Ltd, Chichester, 2009)
27. J. Nocedal, S. Wright, *Numerical Optimization*, 2nd edn. (Springer, Berlin, 2006)
28. R.S. Mulliken, Report on notation for the spectra of polyatomic molecules. *J. Chem. Phys.* **23**, 1997–2011 (1955). <https://doi.org/10.1063/1.1740655>
29. [https://gitlab.com/molpro/point\\_charge\\_symmetry](https://gitlab.com/molpro/point_charge_symmetry). Accessed 8 July 2021

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.