VIRT&L-COMM.25.2023.6

Molecular networks by invariant shape coordinates and deformation indexes

Andrea Lombardi[1]

[1] Dipartimento di Chimica, Biologia e Biotecnologie, Unibersità di Perugia, Perugia, Italy
Corresponding author: andrea.lombardi@unipg.it

The classification of large molecules according to structural similarities is a relevant issue in biochemistry [1], even more in the upcoming big data era. The proper choice of parameters, containing invariant structural information, is a convenient way to induce mapping and grouping of structures, depending on predominant structural motifs, individual amino acid geometry or connectivity properties. Few-body hyperspherical coordinates provide the appropriate framework in which shape parameters and deformation indices [2-4] may be defined and applied to characterize complex structures.

Shape coordinates and invariant deformation indexes

The hyperspherical coordinates can be thought of as a generalization of the ordinary spherical coordinates, in the three-dimensional space, to spaces of higher dimensions. The radius is replaced by a "hyperradius", and the two *azimuthal* and *polar* spherical angles are replaced by a set of "hyperangles", by which the higher dimension Cartesian space can be parametrized.
The hyperspherical coordinates of a given system of $N \geq 2$ particles with masses $m_1, \cdots, m_N$ and Cartesian position vectors in the center of mass reference frame, $\mathbf{r}_1, \cdots, \mathbf{r}_N$, can be derived by introducing a new set of mass *scaled* vectors, $\mathbf{q}_\alpha = (m_\alpha/M)^{1/2}\mathbf{r}_\alpha$ $(1 \leq \alpha \leq N)$, where M is the total mass of the system, and building up a $3 \times N$ position matrix denoted by Z arranging column-wise the components of the $\mathbf{q}$ vectors, as follows:

$$Z = \begin{pmatrix} \mathbf{q}_{1,1} & \cdots & \mathbf{q}_{1,N} \\ \mathbf{q}_{2,1} & \cdots & \mathbf{q}_{2,N} \\ \mathbf{q}_{3,1} & \cdots & \mathbf{q}_{3,N} \end{pmatrix}$$

The next step is allowing to act rotations (left-multiplication) on the position matrix by an orthogonal matrix R^t, transpose of a matrix $R \in O(3)$. Then, alternative rotations can be performed in the so called *kinematic space* [5], by right-multiplication of Z by a $N \times N$ matrix $K \in O(N)$, $Z' = ZK$. In special cases, such as when the center of mass is the reference, the K matrices have a zero last column, and one only has to deal with a $3 \times (N-1)$ reduced Z matrix. This lower dimensional matrix Z is called *reduced position matrix*. The *singular value decomposition* theorem [5] applied to the $3 \times n$ position matrix Z (where $n = N$ or $n = N-1$) leads to a product of three matrices:

$$Z = R\Xi K^t$$

where $R \in O(3)$ and $K \in O(n)$ are 3×3 and $3 \times n$ orthogonal matrices, respectively, as previously outlined. The new $3 \times n$ matrix Ξ contains all zero entries, with the possible exception of the diagonal elements, $\Xi_{11} = \xi_1$, $\Xi_{22} = \xi_2$, $\Xi_{33} = \xi_3$ for which the inequality

$$0 \leq \xi_1 \leq \xi_2 \leq \xi_3$$

holds.
The values ξ_i , $(i = 1, 2, 3)$ are known as the *singular values* of the matrix Z and are uniquely determined, while the factors R and K in Eq. 3 are not, therefore they are invariant under both ordinary and kinematic rotations [5]. They are directly connected to the moments of the

ISSN: 2279-8773

inertia of the system and define the global shape of the system [5]. If $N \leq 3$ and Z is the full $3 \times N$ position matrix, then the smallest singular value ξ_1 is necessarily zero. The hyperradius ρ is defined as follows:

$$\rho^2 = \xi_1^2 + \xi_2^2 + \xi_3^2$$

This property can lead to applications to molecular dynamics and to the systematic study of the minimum energy structures of N-particle aggregates and large molecules such as proteins. Besides the hyperradius, two *deformation indices* can be defined by combining the ξ's, measuring deviations from the spherical top shape [5]:

$$\xi_- = \frac{\xi_1^2 - \xi_2^2}{\rho^2}$$

Which is zero for prolate top configurations, and

$$\xi_+ = \frac{\xi_2^2 - \xi_3^2}{\rho^2}$$

which is zero for oblate top configurations. By definition, when both indexes are zero, one has a spherical rotor. An additional remarkable property of the ξ's is that in the case $N = 4$, is that ξ_3 can take negative values and the chirality sign (mirror images) of the system depends on its sign.

Acknowledgments: Fondazione Cassa di Risparmio di Perugia, n.# 220.0513 to C.E. Fondo Ricerca di Base 2021 (RICBASE2021LOMBARDI).

References

[1] S. Konno, T. Namiki and. K. Ishimori, Quantitative description and classification of protein structures by a novel robust amino acid network: interaction selective network (ISN), Sci. Rep. 16654 (2019)
[2] A. Lombardi, AIP Conference Proceedings, Simmetry and deformations of clusters and biomolecules by invariant shape coordinates, 2343 (1), 020004 (2021)
[3] A. Lombardi, N. Faginas-Lago, N. and L. Pacifici, Protein Networks by Invariant Shape Coordinates and Deformation Indexes. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds) Computational Science and Its Applications – ICCSA 2022 Workshops. ICCSA 2022. Lecture Notes in Computer Science, vol 13382, 348-359 (2022)
[4] V. Aquilanti, A. Lombardi, E. Yurtsever, Global view of classical clusters: the hyperspherical approach to structure and dynamics, Phys.Chem. Chem. Phys. 4 (20), 5040 (2000)
[5] M. B. Sevryuk, A. Lombardi, and V. Aquilanti, Hyperangular momenta and energy partitions in multidimensional many-particle classical mechanics: The invariance approach to cluster dynamics, Physical Review A 72, p. 033201 (2005)