

Dedicated to Professor Victor-Emanuel Sahini  
on the occasion of his 80th anniversary

# DIVERGENCE-FREE DESCRIPTION FOR MOLECULAR ROTATION IN CARTESIAN COORDINATES: THE AXIS-ROTATION FORMULA AND SOME OF ITS APPLICATIONS TO COMPUTATIONAL CHEMISTRY

Viorel CHIHAIA,<sup>a\*</sup> Godehard SUTMANN,<sup>b</sup> Chang-Sup LEE<sup>c</sup> and Soong-Hyuck SUH<sup>d</sup>

<sup>a</sup> Institute of Physical Chemistry “I.G. Murgulescu”, Roumanian Academy, Spl. Independentei 202, 060021, Bucharest, Roumania

<sup>b</sup> Central Institute for Applied Mathematics (ZAM) and John von Neumann Institute for Computing (NIC), Research Center Jülich, D-52425 Jülich, Germany, e-mail: g.sutmann@fz-juelich.de

<sup>c</sup> Department of Chemistry, Keimyung University, Daegu, 704-701, Korea, e-mail: surfkm@kmu.ac.kr

<sup>d</sup> Department of Chemical Engineering, Keimyung University, Daegu, 704-701, Korea, e-mail: shsuh@kmu.ac.kr

Received January 15, 2007

In this work, based on simple algebraic manipulations, the divergence-free description of molecular rotations is revisited using the axis-rotation formula for the rigid-body system. The so-called axis-rotation formula is useful in various fields of computational chemistry, including molecular simulations, graphical rendering and group theory, allowing more convenient ways to construct and to manipulate the atomic or fragment structures of rotations. It is shown that the analytical expression of the axis-rotation operator facilitates obtaining the symmetry operator in analytical form, which is useful in the determination of group symmetries of molecules and the adaptation to the symmetry of atomic and molecular orbitals.

## INTRODUCTION

In order to describe the orientation and the rotation of a rigid body in space, two mathematical approaches are frequently used, based on Euler angles<sup>1</sup> or quaternions.<sup>2-4</sup> The method of Euler angles consists of three successive rotations around the x-, y- and z-axes with three angles  $\phi$ ,  $\theta$  and  $\psi$ , known as the Euler angles. Different sets of Euler rotations exist depending on the succession of the three rotations. Quaternions, which obey the Pawley algebra<sup>5</sup>, are mathematical objects that generalize the complex numbers. Formally, the quaternion  $\mathbf{q}=(q_0, \mathbf{q}_v)^T$  is a vector with four components, *i.e.* one scalar component,  $q_0$ , and a vector  $\mathbf{q}_v=(q_1, q_2, q_3)^T$ . The quaternion of the form  $\mathbf{q}=(\cos(\varphi/2), \hat{\mathbf{n}}\sin(\varphi/2))^T$  is normalized

$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ , in which it is associated to the rotation of the angle  $\varphi$  around an arbitrary axis  $\hat{\mathbf{n}}=(n_1, n_2, n_3)^T$ .<sup>3</sup> The components of the quaternions are expressed as the products of sine and cosine functions of the Euler angles. Therefore Euler angles and quaternions can mutually be transformed into each other.<sup>4</sup>

Deriving the equations of motion it turns out that the time derivatives of the Euler angles are expressed through the angular frequencies around the principal axis of the molecule via a matrix, which is singular for the orientations  $\varphi=0$  and  $\varphi=\pi$ . This problem may be overcome either by successive rotations of smaller angles or by changing the coordinate system.<sup>6</sup> Both solutions require additional statements in the algorithm and introduce a computational overhead. The advantage of the quaternion description is that

\* Corresponding author: vchihaia@icf.ro

there are no special cases for the orientation of a molecule, i.e. it is divergence free. Frequently a computational problem is associated with quaternion algorithms, i.e. the norm of the quaternions is not conserved along the evolution in time.<sup>7-9</sup> The accumulation of the numerical errors is usually prevented by a renormalization of the quaternions at each time step. Only recently, Martyna and co-workers<sup>10</sup> proposed a symplectic method which conserves the norm of the quaternions in time. The method of Euler angles provides a physical description of the rigid body rotation. On the other hand, the quaternion formalism completely describes the rotation, but the quaternions are complex mathematical objects which do not directly reflect a physical picture of a rotation.

A third possible way to describe the rotation of a rigid body around an axis is the so-called axis-rotation or Rodrigue's formula.<sup>1,11-15</sup> For unknown reason, it is not popular in the field of molecular simulations, and to our best knowledge it is mentioned only in two works.<sup>14,15</sup> The axis-rotation formula can be derived from the quaternion formalism, and it can be directly obtained through the decomposition of the vector position in the perpendicular and the parallel components to the rotation axis. The axis-rotation formula provides a clear picture of the rotation in space. In the present article, based on simple algebraic manipulations, we will revisit the derivation of the axis-rotation formula and give some applications in molecular simulations, graphical rendering, and group theory. It is shown to allow a simple geometrical manipulation of molecules or molecular fragments, which is a possible facility for the preparation of the coordinates of molecular systems in areas like quantum chemistry calculations, molecular dynamics or Monte Carlo simulations. The rotational dynamics of a rigid body can be recovered and analysed from the trajectories of particles or fragments. Furthermore, the axis-rotation formula is useful to construct the symmetry operator in group theory.

### ROTATION OF A POINT AROUND AN AXIS

Consider the changes of the coordinates of a point after the rotation with the angle  $\varphi$  around an arbitrary axis, denoted by the unit vector  $\hat{\mathbf{n}}$ , that is going through the point O represented by the

vector  $\mathbf{s}_0$  with respect to the origin of a given reference frame. Two points, a bond or a direction (for example, the electric dipole, the principal inertia axis or the angular velocity) may define the rotation axis. The direction of rotation is given by the right-hand convention.

The vectors  $\mathbf{s} = \mathbf{s}_0 + \mathbf{r}$  and  $\mathbf{s}' = \mathbf{s}_0 + \mathbf{r}'$  give the position of a point P before and after the rotation, respectively (see Fig. 1). A simple formula that relates  $\mathbf{r}$  to  $\mathbf{r}'$  (the relative position vectors to  $\mathbf{s}_0$  before and after rotations, respectively) is given by

$$\mathbf{r}' = \mathbf{r} \cos \varphi + \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{r}) (1 - \cos \varphi) + \hat{\mathbf{n}} \times \mathbf{r} \sin \varphi. \quad (1)$$

(A derivation of the formula is shown in Appendix A1). The transformation of the position vector relative to O is given by  $\mathbf{r}' = \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{r}$ , where

$$\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} = \cos \varphi \mathbf{I} + (1 - \cos \varphi) \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot \mathbf{I}) + \sin \varphi \hat{\mathbf{n}} \times \mathbf{I}, \quad (2)$$

which is the associated axis-rotation operator and  $\mathbf{I}$  represents the identity operator. Applying the axis-rotation operator, the absolute position of the point P after rotation is expressed by

$$\mathbf{s}' = \mathbf{s}_0 + \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} (\mathbf{s} - \mathbf{s}_0). \quad (3)$$

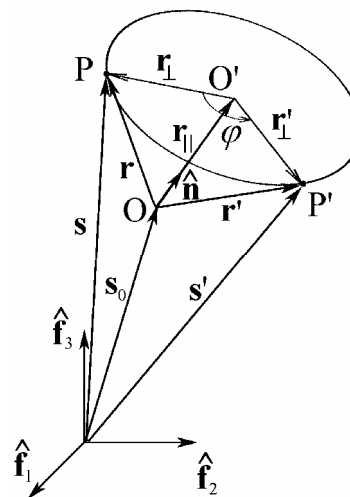


Fig. 1 – The definition of the absolute ( $\mathbf{s}$ ) and of the relative ( $\mathbf{r}$ ) position vectors, before and after the rotation of the point P to P' by the angle  $\varphi$  around the unit axis  $\hat{\mathbf{n}}$  which is going through the reference point O described by  $\mathbf{s}_0$ . The vectors  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  are the unit vectors of the absolute reference frame.

Furthermore, the velocity of a point P is given by

$$\dot{\mathbf{s}} = \dot{\mathbf{s}}_0 + \dot{\mathbf{r}}, \quad (4)$$

where

$$\dot{\mathbf{r}} = \omega \hat{\mathbf{n}} \times \mathbf{r} = \boldsymbol{\omega} \times \mathbf{r} \quad (5)$$

The above equation describes the rotation around the axis  $\hat{\mathbf{n}} = \boldsymbol{\omega} / \omega$  with the angular velocity  $\boldsymbol{\omega}$  (see Appendix A2). Since the description of positions, Eq. (3), and velocities, Eq. (4), can be performed relative to an arbitrary origin, it is convenient to introduce a local reference frame for the arbitrary point O. In the case of a set of atoms, the centre of mass system is a natural choice for the reference point O. The axes of the local reference frame can be chosen parallel to the axes of the absolute reference frame (i.e. not necessarily the principal inertia axes). With this choice of reference frame the translational motion is subtracted and consequently only the rotational part is considered here.

The rotation operator  $\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi}$  obeys the following relations associated with the rotation axes  $\hat{\mathbf{n}}$ :

$$\text{i) } \mathbf{R}_{\hat{\mathbf{n}}}^{\alpha} \mathbf{R}_{\hat{\mathbf{n}}}^{\beta} = \mathbf{R}_{\hat{\mathbf{n}}}^{\alpha+\beta} = \mathbf{R}_{\hat{\mathbf{n}}}^{\beta} \mathbf{R}_{\hat{\mathbf{n}}}^{\alpha}$$

$$\mathbf{R} = \begin{bmatrix} n_1 n_1 (1 - \cos \varphi) + \cos \varphi & n_1 n_2 (1 - \cos \varphi) - n_3 \sin \varphi & n_1 n_3 (1 - \cos \varphi) + n_2 \sin \varphi \\ n_1 n_2 (1 - \cos \varphi) + n_3 \sin \varphi & n_2 n_2 (1 - \cos \varphi) + \cos \varphi & n_2 n_3 (1 - \cos \varphi) - n_1 \sin \varphi \\ n_1 n_3 (1 - \cos \varphi) - n_2 \sin \varphi & n_2 n_3 (1 - \cos \varphi) + n_1 \sin \varphi & n_3 n_3 (1 - \cos \varphi) + \cos \varphi \end{bmatrix}, \quad (6)$$

and, using this matrix, only  $(15N+22)$  numerical operations are required for the computation of rotation (see Appendix A4).

## APPLICATIONS TO ORIENTATION AND SYMMETRY

### 1. Transformation between Cartesian and internal coordinates

The atomic and molecular positions in space can be described either by the Cartesian coordinates (CC) or by the internal coordinates (IC) representations of relative positions. Depending on the size and the topology of the system, the localization of the stationary points on the potential energy surface are largely affected by the coordinate system itself. The IC can be an efficient way for such cases. Changes of atomic positions in the CC are highly correlated in structural topologies, and, as a consequence, a larger number of iterations is required to find local minima. The harmonic and anharmonic couplings

$$\text{ii) } \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{R}_{\hat{\mathbf{n}}}^{-\varphi} = \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} = \mathbf{R}_{\hat{\mathbf{n}}}^0 = \mathbf{I} \text{ and}$$

$$(\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi})^{-1} = \mathbf{R}_{\hat{\mathbf{n}}}^{-\varphi} = \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi}$$

$$\text{iii) } \mathbf{R}_{\hat{\mathbf{n}}}^{\pi} = -\mathbf{I} + 2\hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{I})$$

$$\text{iv) } (\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{a}) \cdot (\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} \text{ and}$$

$$(\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{a}) \times (\mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \mathbf{b}) = \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} (\mathbf{a} \times \mathbf{b})$$

$$\text{v) } \hat{\mathbf{n}} \text{ can be the eigenvector of the rotational operator } \mathbf{R}_{\hat{\mathbf{n}}}^{\varphi} \hat{\mathbf{n}} = \hat{\mathbf{n}} \text{ for the eigenvalue 1.}$$

The algorithm of the rotation of an atom using the axis-rotation formula may be optimized so that it requires only 30 numerical operations (additions and multiplications), while the quaternion formalism requires at least 41 numerical operations. In the case of molecules or fragments of systems consisting of  $N$  atoms, the axis-rotation formula requires  $(26N+7)$  numerical operations, compared to  $41N$  numerical operations required by the quaternion formalism. The most economical way is to construct the rotation matrix (see, Eq. A3.2 in Appendix A3),

between ICs are important in cases of large systems with strong connectivity, increasing the number of optimization iterations, imposing difficulties for the convergence. The coordinates can be decoupled into linear combinations of distances, and bond and dihedral angles as in the natural,<sup>16</sup> the redundant<sup>17</sup> or delocalized<sup>18</sup> internal coordinates. Providing a good initial Hessian, especially when the initial geometry of the system is far from equilibrium, the optimization in CC is enforced, making the convergence in the CC comparable to the convergence in IC. The CC system allows the freezing of atoms in space or along some Cartesian directions eliminating the respective coordinates from the set of variables. Similarly, the IC representation facilitates directly the imposing of constraints for bonds and angles.

Because in the empirical force field methods the calculation of the energy and the gradients are not expensive, the number of cycles to converge to equilibrium is not an important criterion in choosing the optimization method. Usually, the optimization is done in the CC. In contrast, in quantum methods, the energy and its derivatives

are expensive to calculate and the optimization method has to be carefully chosen to reduce the number of optimization cycles. In most cases, the energy and its derivatives are analytically expressed in CCs. On the other hand, displacements toward a local minimum are performed in ICs and consequently mutual coordinate transformations are required between the two descriptions.

The conversion of CCs to ICs can be computed straightforwardly. Suppose e.g. a set of four connected atoms  $I$ ,  $J$ ,  $K$  and  $L$ . The internal coordinates corresponding to the atom  $L$  are the distance between the atoms  $K$  and  $L$ , the bond angle  $\alpha_{JKL}$  and the dihedral angle  $\beta_{IJKL}$ , respectively. Let us define two vectors

$$\hat{\mathbf{u}} = \frac{\hat{\mathbf{r}}_{JI} \times \hat{\mathbf{r}}_{JK}}{|\hat{\mathbf{r}}_{JI} \times \hat{\mathbf{r}}_{JK}|} \quad \text{and} \quad \hat{\mathbf{v}} = \frac{\hat{\mathbf{r}}_{KJ} \times \hat{\mathbf{r}}_{KL}}{|\hat{\mathbf{r}}_{KJ} \times \hat{\mathbf{r}}_{KL}|},$$

unit vectors defining the normal to the plane  $(I,J,K)$  and  $(J,K,L)$ , respectively, where the vector

$\hat{\mathbf{r}}_{XY} = \mathbf{r}_{XY} / |\mathbf{r}_{XY}| = (\mathbf{s}_X - \mathbf{s}_Y) / |\mathbf{s}_X - \mathbf{s}_Y|$  is the unit vector along the  $X$ - $Y$  bond. The intra-atomic distance is calculated directly as

$d_{KL} = |\mathbf{r}_{KL}| = |\mathbf{r}_L - \mathbf{r}_K|$ , where  $\mathbf{r}_{KL}$  is the relative vector from atom  $K$  to atom  $L$ . The oriented bond angle  $\alpha_{JKL}$  is measured from  $\hat{\mathbf{r}}_{KJ}$  to  $\hat{\mathbf{r}}_{KL}$ , around

$\hat{\mathbf{v}}$ .<sup>\*</sup> The oriented dihedral angle  $\beta_{IJKL}$  is measured from  $\hat{\mathbf{u}}$  to  $\hat{\mathbf{v}}$ , around  $\mathbf{r}_{JK}$ . A special case arises when the atoms  $I$ ,  $J$  and  $K$  (or  $J$ ,  $K$  and  $L$ ) are collinear: the vector  $\hat{\mathbf{u}}$  (or  $\hat{\mathbf{v}}$ ) can be any perpendicular unit vector to  $\hat{\mathbf{r}}_{KL}$ . For consistency it is chosen as in the previous iteration of optimisation. If this situation appears in the first optimization cycle, then the other set of atoms and angles is selected.

The inverse transformation from the IC to the CC system requires additional effort. For a set of

four atoms  $I$ ,  $J$ ,  $K$  and  $L$ , where the CCs of the atoms  $I$ - $J$  are already settled, the position of the atom  $L$  is established based on the length  $d_{LK}$  and the angles  $\alpha_{JKL}$  and  $\beta_{IJKL}$ . The usual procedure for this conversion employs three rotations around the three axes of the coordinates associated with the set. Here we present a procedure that requires only two rotations, and, moreover, an analytical formula, which relates the CCs to the ICs, is determined. The relation is useful for the determination of the inverse of the Wilson B matrix, which relates the displacements in ICs and CCs. For many atoms and a large number of conversions from internal to Cartesian coordinates, this procedure reduces the computational effort significantly (and therefore also round off errors are reduced). We will use the vectors  $\hat{\mathbf{r}}_{KJ}$  and  $\hat{\mathbf{u}}$ , defined above. Let us consider the point  $X$  along the axis  $\hat{\mathbf{r}}_{KJ}$  from  $K$  to  $J$ , at the distance  $d_{KL}$  from the atom  $K$ :  $\mathbf{r}_{KX} = \hat{\mathbf{r}}_{KJ} \cdot d_{KL}$ . Now the point  $X$  is rotated to the point  $Y$  with the angle  $\alpha_{JKL}$  around the axis  $\hat{\mathbf{u}}$ :  $\mathbf{r}_{KY} = \mathbf{R}_{\hat{\mathbf{u}}}^{\alpha_{JKL}} \mathbf{r}_{KX}$ . The point  $Y$ , contained in the plane  $(I,J,K)$ , is rotated around the axis  $\hat{\mathbf{r}}_{JK} = -\hat{\mathbf{r}}_{KJ}$  with the angle  $\beta_{IJKL}$  (or around the axis  $\hat{\mathbf{r}}_{KJ}$  with the angle  $-\beta_{IJKL}$ ) to the final position of the atom  $L$ , situated in the plane  $(J,K,L)$ :  $\mathbf{r}_{KL} = \mathbf{R}_{\hat{\mathbf{r}}_{JK}}^{\beta_{IJKL}} \mathbf{r}_{KY}$ . The absolute position of the atom  $L$  is  $\mathbf{s}_L = \mathbf{s}_K + \mathbf{r}_{KL}$ , where

$$\mathbf{r}_{KL} = d_{KL} [\hat{\mathbf{r}}_{KJ} \cos \alpha_{JKL} + \sin \alpha_{JKL} (\hat{\mathbf{u}} \times \hat{\mathbf{r}}_{KJ} \cos \beta_{IJKL} - \hat{\mathbf{u}} \sin \beta_{IJKL})] \quad (7)$$

is obtained by combining the axis-rotation formula for the two rotations and using properties iv) and v) of the axis-rotation operator.

The unit vector  $\hat{\mathbf{u}}$  is not determined when the atoms  $I$ ,  $J$  and  $K$  are collinear; it is settled from the previous iteration. Thus, the procedure is free of singularities and as the angles may have any values in  $\mathbb{R}$ , a standard procedure for optimization may be applied.

## 2. Orientation of a system

The orientation of a system is described by any set of three orthogonal unit vectors, that for a right-hand set rotate solidary with the system. The position vectors of three atoms of the system are

<sup>\*</sup> An angle  $\theta$  formed by two unit vectors  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$ , measured from  $\hat{\mathbf{a}}$  to  $\hat{\mathbf{b}}$  around the unit vector  $\hat{\mathbf{n}}$ , takes values between 0 and  $2\pi$ . Its measure is determined based on the signs of  $\cos \theta = \hat{\mathbf{a}} \cdot \hat{\mathbf{b}}$  and  $\sin \theta = \hat{\mathbf{n}} \cdot (\hat{\mathbf{a}} \times \hat{\mathbf{b}})$ . If the angle is measured from  $\hat{\mathbf{b}}$  to  $\hat{\mathbf{a}}$  or  $\hat{\mathbf{a}}$  to  $\hat{\mathbf{b}}$ , but round  $-\hat{\mathbf{n}}$  it will have the negative value  $-\theta$ , (therefore it takes values between  $-2\pi$  and 0) or  $2\pi - \theta$  because its periodicity (see the comments from Appendix A1).

enough to build such unit vectors: a first vector is chosen along the line that connects two atoms, a second one is considered perpendicular to this one lying in the plane of the three atoms and a third vector is perpendicular to the plane spanned by the first two vectors. However, the principal inertia axes  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  relative to the system's center-of-mass form the most convenient set. It characterizes the spatial arrangement of the constituent atoms in the local reference frame with the origin in O. Let us suppose that the eigenvalues and associated eigenvectors are ordered as  $I_1 \leq I_2 \leq I_3$ . In the case of a linear molecule,  $I_1 = 0$  and  $I_2 = I_3$  and consequently one principal axis  $\hat{\mathbf{f}}_1$  is oriented

along the molecule and the other two principal axes can be any orthogonal vectors that are perpendicular to the molecule. If  $I_1 \ll I_2 \leq I_3$ , then  $I_1$  is a measure of the deviation of the atoms from a linear molecule and the principal inertia axis  $\hat{\mathbf{f}}_1$  is close to the molecule axis. When the atoms are arranged in a plane, the principal moments of inertia satisfy the relation  $I_3 = I_1 + I_2$ . The deviation from this relation characterizes the deviation from the average plane of the atoms. In this case the principal axis  $\hat{\mathbf{f}}_3$  is identical to the normal vector of the plane spanned by the atoms while the other two principal axes lie within this plane.

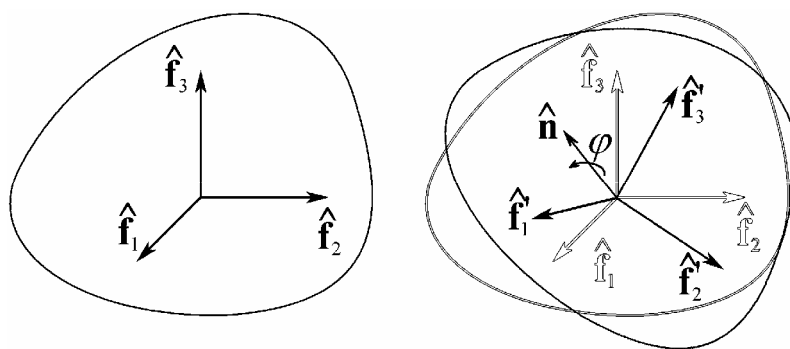


Fig. 2 – The body reference frames of a rigid-body before (left) and after his rotation (right). The vectors  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  and  $\hat{\mathbf{n}}$  are the principal inertia axes of the rigid-body before and after rotation.

When a molecule is rotated in space with angle  $\varphi$  around an axis  $\hat{\mathbf{n}}$ , which is going through the reference point O, the inertia matrix is changed as  $\mathbf{I}' = \mathbf{R}\mathbf{I}\mathbf{R}^T$ , where  $\mathbf{R}$  is the corresponding rotation matrix given in Eq. (6). The new principal axes  $\{\hat{\mathbf{f}}_i'\}_{i=1,3}$  can be determined calculating  $\mathbf{I}'$  and solving the eigenvalues and eigenvectors equation. It is easier to determine the new principal axes directly by the rotation of the initial principal axes  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$ :  $\hat{\mathbf{f}}_i' = \mathbf{R}_{\hat{\mathbf{n}}}^\varphi \hat{\mathbf{f}}_i$ . (see, Fig. 2) When the system is subject of some successive rotations and we do not need the coordinates of the atoms after each rotation, the new values of the principal axes after each rotation are calculated, and from the relative orientation of the last principal axis and of the initial principal axis the total rotation matrix is determined, as it is indicated in Appendix A3. Finally, using the total rotation matrix the coordinates of the atoms can be established.

### 3. Molecular motion from the trajectories of each atom

Suppose that we intend to evaluate the rotational motion of a molecule from the trajectories of each atom from a molecular system. According to the theorem of Chasles,<sup>19</sup> two configurations of a molecule may be described by a translational and a rotational transformation (note that this theorem was very recently proved with the help of quaternions<sup>20</sup>). Assume that the vectors of the basis set  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  are the principal inertia axes of a given molecule at the time step  $t_0$  and  $\{\hat{\mathbf{f}}_i'\}_{i=1,3}$  are those of the same molecule at the next time step  $t_0 + \Delta t$ . The translational motion is easily computed from the trajectories of the centre-of-mass position. The rotational motion can be

determined by the projection of the basis set  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  over  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$ . In the case of the flexible molecules, the vibrational motion contaminates the rotational motion determined in this way. Averaging the vibrational motion over a number of time steps,  $n$ , may eliminate this undesired effect. The rotational axes and the rotational angle, that correspond to the transformation of each molecule from one time step to the next step, are determined by the change of the principal axes between the two time steps as described in more detail in Appendix A3.

#### 4. Alignment of a system into a given direction

E.g. in graphical visualization, crystal surface construction, or the initial structural setup of a molecular system, fragments or even the complete system has to be oriented often under an angle  $\varphi$  formed by a fixed direction in space described by the unit vector  $\hat{\mathbf{u}}$  and a given direction from the system described by the unit vector  $\hat{\mathbf{v}}$ , measured around a unit vector  $\hat{\mathbf{m}}$  perpendicular to the plane

( $\hat{\mathbf{u}}, \hat{\mathbf{v}}$ ) (see, Fig. 3). Two atoms, a bond, a dipole moment, the normal to a plane or a special direction from space may determine each direction. For example, this procedure can be applied to display a system with a special direction  $\hat{\mathbf{v}}$  perpendicular to the screen or when an adsorbed molecule must be oriented with the dipole moment to make the angle  $\varphi$  with the normal of a surface. Moreover, some properties of the chemical system can be determined more efficiently by considering the system symmetry. Programmes that employ the symmetry of the molecule require the coordinates of the molecule to be given in the so-called standard symmetry orientation, which is defined based on the rules: (i) if the system has a symmetry  $C_s$ , then the z axis is chosen perpendicular to the symmetry plane, (ii) if the system has symmetry axes, then the z axis is oriented along the axis with the highest order, and (iii) the x axis is considered in one vertical plane if it exists or along an axis  $C_2$ , if it exists. The procedure presented below can be applied to orient a given system in the standard symmetry orientation.

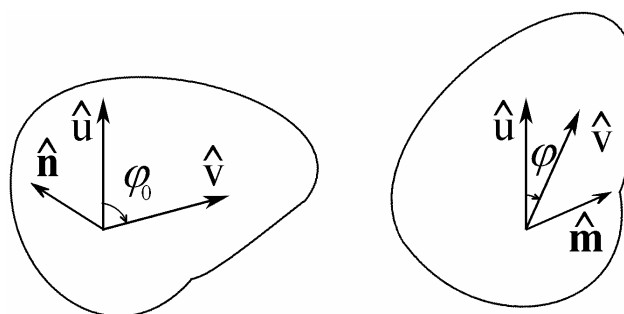


Fig. 3 – A rigid-body in initial orientation and after its reorientation to form the angle  $\varphi$  between the directions  $\hat{\mathbf{u}}$  (the given direction in space) and  $\hat{\mathbf{v}}$  (the fixed on the body).

The initial value of oriented-angle  $\varphi_0$  is measured from  $\hat{\mathbf{u}}$  to  $\hat{\mathbf{v}}$ , around the unit vector  $\hat{\mathbf{n}} = \frac{\hat{\mathbf{u}} \times \hat{\mathbf{v}}}{|\hat{\mathbf{u}} \times \hat{\mathbf{v}}|}$ . The system is translated with the centre of mass in the reference point O and is rotated by an angle  $-\varphi_0$  around the unit vector  $\hat{\mathbf{n}}$  to superpose the current direction  $\hat{\mathbf{v}}$  over  $\hat{\mathbf{u}}$ . The desired orientation of the system is obtained by rotation of angle  $\varphi$  around the unit vector  $\hat{\mathbf{m}}$ . Finally, the system is translated back, with its mass-centre position in the initial position.

#### 5. Merging and joining of two structures

Very often, two systems have to be arranged in a special relative orientation. An example is the adsorption of a molecule on an active site on a surface, which implies the geometrical matching between the parts of the molecule and the surface responsible for the frontier orbitals. A similar situation occurs in the case of drug interaction with receptor proteins. Such systems may be optimised as a supermolecule; however, due to the large

number of degrees of freedom, exploring all possible mutual orientations will result in expensive and long calculations. A significant reduction in computational operations can be achieved by investigating the alignment between the target molecule and the template, assuming in a first approximation the two systems as rigid-bodies. The relative positioning of the two structures is determined by 6 parameters given by the relative position of their centres-of-mass and their relative orientation. In the case of different geometries of the two structures or fragments of them, the number of iterations may be reduced by superimposing the principal inertia axis of the two (sub)structures, using the axis-rotation formula. (See Appendix A3.)

The design of new molecules or large systems requires the merging and/or joining of fragments or so called building blocks. This can be done using translations and rotations of the constituent parts. Let us suppose that the structures of two systems  $S1$  and  $S2$  are known and we intend to merge them into a new structure. The two structures are connected each other by at least a bond formed by two atoms,  $J$  from  $S1$  and  $K$  from  $S2$ . Two other atoms,  $I$  from  $S1$  and  $L$  from  $S2$ , connected to  $J$  and  $I$ , respectively, help us to define four internal coordinates: intra-atomic distance  $d_{JK}$ , the bond angles  $\alpha_{IJK}$  and  $\alpha_{JKL}$ , and the dihedral angle  $\beta_{IJKL}$ . In order to completely describe the relative arrangement of the two systems, two other internal coordinates are required. Such coordinates are two dihedral angles around the bonds  $I-J$  and  $K-L$  or two intra-atomic distances between other bonded atoms located in different systems.

When Cartesian coordinates describe the two systems  $S1$  and  $S2$  or when the final system has to be visualized, it is preferable to work directly in Cartesian coordinates. To establish the distance  $d_{JK}$  between the atoms  $J$  and  $K$ , all the atoms of the system  $S2$  are translated a distance  $d_{JK} - |\mathbf{r}_{JK}|$  along  $\hat{\mathbf{r}}_{KJ}$ . The new position of the atom  $K$  is  $\mathbf{s}_K = \mathbf{s}_J + d_{JK} \hat{\mathbf{r}}_{KJ}$ . The system  $S1$  is rotated an angle  $\alpha_{IJK} - \alpha_{IJK}^0$  around the axis  $\hat{\mathbf{u}} = \frac{\mathbf{r}_{JI} \times \mathbf{r}_{JK}}{|\mathbf{r}_{JI} \times \mathbf{r}_{JK}|}$  which goes through atom  $J$  and which is perpendicular to the plane spanned by the atoms  $(I, J, K)$ . The  $\alpha_{IJK}^0$  is the angle formed by the atoms  $I, J$  and  $K$  before the rotation, measured around  $\hat{\mathbf{u}}$ . In order to settle the angle  $\alpha_{JKL}$ , the

system  $S2$  is rotated an angle  $\alpha_{JKL} - \alpha_{JKL}^0$  in the new position around the axis  $\hat{\mathbf{v}} = \frac{\mathbf{r}_{KL} \times \mathbf{r}_{JK}}{|\mathbf{r}_{KL} \times \mathbf{r}_{JK}|}$ , which goes through  $K$  and which is perpendicular to the plane spanned by the atoms  $(J, K, L)$ . The  $\alpha_{JKL}^0$  is the value of the initial angle measured around  $\hat{\mathbf{v}}$ . To fix the dihedral angle, the current angle  $\beta_{IJKL}^0$  is determined and the system  $S2$  is rotated by the angle  $\beta_{IJKL} - \beta_{IJKL}^0$  around the axis  $\hat{\mathbf{r}}_{JK}$ . When the torsion angle  $\beta_{IJKL}$  is not predetermined, the stable conformers are established by determining the torsion angle that corresponds to the local energy minimums: one fragment is rotated around the axis  $J-K$  by an incremental angle. The dependency of the total energy function on the torsion angle is recorded and the minima are detected. For higher accuracy, the internal geometries of the two fragments are optimised for each torsion angle  $\beta_{IJKL}$ .

In the case of merging we have to identify the common domains of the two systems (skeleton, rings and bonds) and to superpose them excluding one of the two common domains. If they have no common domains, then we try with the enantiomer of one of them.

The procedure presented above is considered for the superposition of  $I$  over  $L$  and  $J$  over  $K$  imposing the constraints  $\beta_{IJKL} = \pi$ ,  $\alpha_{IJK} + \alpha_{JKL} = \pi$ , and  $d_{JK} = 0$ . With this constraints the systems have the freedom to rotate around the formed common bond  $I-J$ . If the systems have in common at least one more pair of common atoms  $X$  and  $Y$ , then the atoms are superposed including the internal coordinates  $d_{XY} = 0$ . If not, then the systems are rotated around the common axis  $I-J$  to identify the different possible isomers.

## 6. Application to symmetry

The symmetry of molecules, based on the concepts of group theory, allows us to select appropriate basis sets. Furthermore, the symmetry properties can be used to reduce the computational effort by block-diagonalization of the matrices that describe the interactions in a given system. The simple analytic formula of the axis-rotation operator makes it possible to build other symmetry operators for any orientations rather than the

standard symmetry orientation. The symmetry operators do not affect the molecular centre-of-mass, and, in the case that a given molecule presents the inversion symmetry, the inversion point coincides with the molecular centre-of-mass. Without loss of generality, the reference point O,

$$C_{\hat{n}}^n = \cos \varphi_n I + (1 - \cos \varphi_n) \hat{n}(\hat{n} \cdot I) + \sin \varphi_n \hat{n} \times I. \quad (8)$$

The reflection operator with respect to the symmetry plane is given by

$$\sigma_{\hat{n}} = I - 2\hat{n}(\hat{n} \cdot I), \quad (9)$$

which represents the rotational operator with  $\pi$  radians around an axis  $\hat{n}$  included in the symmetry plane.

$$S_{\hat{n}}^n = \cos \varphi_n I - (1 + \cos \varphi_n) \hat{n}(\hat{n} \cdot I) + \sin \varphi_n \hat{n} \times I. \quad (10)$$

The inversion operator, which transforms  $\mathbf{r}$  into  $-\mathbf{r}$ , with respect to a point O is

$$I_{inv} = -I. \quad (11)$$

Numerous algorithms have been developed to determine the symmetry of a system.<sup>21</sup> It is noted here that the operators in Eqs. (8) to (11) are helpful in detecting the symmetry elements effectively. Once the complete set of the symmetry elements for a given molecule is evaluated, then the sets of the equivalent atoms for each symmetry operator are identified. From these sets of equivalent atoms, a subset of atoms, the so-called unique atoms, can be determined. Using the coordinates of the unique atoms only, calculations are faster for the optimisation of molecular geometries. Different sets of functions (e.g. atomic orbitals, spin functions, molecular orbitals and vibration coordinates) can also be adapted to the symmetry of the system. In this situation, the matrix, which describes the electronic motion, the spin states, or the nuclear motion, will be block-diagonalized, and the calculations are done independently in each block, which may significantly reduce the size and the complexity of computations.

Applying the symmetry operators, say  $\mathbf{O}$  in Eqs. (8) to (11), to the functions of the basis set  $\{\mathbf{f}_{\lambda}\}_{\lambda=1,\omega}$ , the elements of the matrix for the

defining the rotation, may be chosen as the centre-of-mass position in the axis rotation formula.

The axis-rotation operator, corresponding to the symmetry axis  $\hat{n}$  and the quantized angle,  $\varphi_n = 2\pi/n$ , can be related with the rotation symmetry operator  $C_{\hat{n}}^n$  of degree  $n$  as

The rotation-reflection operator is determined by the composition of the rotational operator with respect to the axis  $\hat{n}$  and the reflection operator with respect to a plane containing the mass-centre and the unit vector  $\hat{n}$  as normal,

transformation of the function set  $\{\mathbf{f}_{\lambda}\}_{\lambda=1,\omega}$  to  $\{\mathbf{f}'_{\kappa}\}_{\kappa=1,\omega}$  becomes

$$\mathbf{f}'_{\kappa} = \sum_{\lambda=1}^{\omega} \mathbf{O} \mathbf{f}_{\lambda} = \sum_{\lambda=1}^{\omega} \mathbf{O}_{\kappa\lambda} \mathbf{f}_{\lambda}. \quad (12)$$

This determines a representation  $\Gamma$  for the group symmetry associated with the system. For the character of the symmetry operator in this representation, the equation of  $\chi_{\Gamma}(\mathbf{O}) = \sum_{\lambda=1}^{\omega} \mathbf{O}_{\lambda\lambda}$  is calculated as the trace of the corresponding matrix  $\mathbf{O}$ . The representation  $\Gamma$  is reducible, and can be written as a direct summation, i.e.  $\Gamma = \sum_i a_i \Gamma_i$ , of the irreducible representation  $\Gamma_i$  of the symmetry group. The multiplicity of each irreducible representation is given by

$$a_i = \frac{1}{h} \sum_{\mathbf{O}} \chi_{\Gamma}(\mathbf{O}) \chi_{\Gamma_i}^*(\mathbf{O}), \quad (13)$$

where the asterisk symbol denotes the complex conjugate of the character  $\chi$ , and  $h$  is the number of the symmetry operators  $\mathbf{O}$ , or the size of the point symmetry group.

For each irreducible representation  $\Gamma_i$ , an associated projection operator can be constructed as a combination of the symmetry operators  $\mathbf{O}$  as



$$\mathbf{P}^i = \frac{1}{h} \sum_{\mathbf{o}} \chi_i^*(\mathbf{o}) \mathbf{O}. \quad (14)$$

In this equation the coefficients are the complex conjugate characters of the operators in the irreducible representation. The symmetry-adapted functions corresponding to each irreducible representation can be built as a linear combination of the initial functions, over which the projection operator  $\mathbf{P}^i$  is applied.<sup>22</sup>

## CONCLUSION

In this work we have revisited the divergence-free description of molecular rotations using the axis-rotation formula for the rigid-body system. This formula, based on simple algebraic manipulations, can be applied in any coordinate system where rotations are involved without determining the Euler angles. The axis-rotation formula is also an alternative way to avoid mathematical complexities in the quaternion formalism. It can be useful in different fields of computational chemistry, allowing more convenient ways to construct and to manipulate the atoms or molecular fragment structures by axis-

rotations. We present some applications as: the transformation between Cartesian and internal coordinates, the orientation of a system, the molecular motion from the trajectories of each atom, the alignment of a system into a given direction, the merging and joining of two structures. Based on the axis-rotation formula the analytical expression of the axis-rotation operator is obtained. It facilitates the determination of the symmetry operator in analytical form, which is useful in the determination of the group symmetry for molecules and the adaptation to the symmetry of the atomic and molecular orbitals. We are currently investigating the possibilities of extending this formalism to various systems (including the buckyball and nanotube structures) with the bond-length or the bond-angle constraints. Another possible application will be the calculation of the electronic interaction integrals in quantum chemistry, where the transformations of the local reference frame of atoms to those of the standard reference frame can be avoided.

*Acknowledgements:* This work is partly supported by the Korea Research Foundation and the Korea Science and Engineering Foundation. VC gratefully acknowledges the financial support of the Romanian National University Council (the grant No. CNCSIS 27642/2005).

## APPENDIXES

## A1. The rotation formula

We present here the derivation of the axis-rotation formula in the different way from its original derivation in Ref. 1. The projections of the position vector  $\mathbf{r}$  and  $\mathbf{r}'$  (with the same original point O – see Fig. 1) to the rotational unit vector  $\hat{\mathbf{n}}=(n_1, n_2, n_3)^T$  should be the same scale each other

$$\hat{\mathbf{n}} \cdot \mathbf{r} = \hat{\mathbf{n}} \cdot \mathbf{r}', \quad (\text{A1.1})$$

and, in our case, the magnitudes of these vectors are both equal

$$|\mathbf{r}'| = |\mathbf{r}|. \quad (\text{A1.2})$$

The vectors  $\mathbf{r}$  and  $\mathbf{r}'$  can be decomposed into the parallel and the perpendicular components along the rotational unit vector  $\hat{\mathbf{n}}$ , e.g.

$$\mathbf{r}_\perp \cdot \mathbf{r}'_\perp = |\mathbf{r}_\perp| |\mathbf{r}'_\perp| \cos \varphi = [\mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2] \cos \varphi, \quad (\text{A1.6})$$

or,

$$\mathbf{r}_\perp \cdot \mathbf{r}'_\perp = [\mathbf{r} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})][\mathbf{r}' - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}')] = \mathbf{r} \mathbf{r}' - (\hat{\mathbf{n}} \cdot \mathbf{r})^2. \quad (\text{A1.7})$$

Eqs. (A1.6) and (A1.7) will give

$$\mathbf{r} \cdot \mathbf{r}' = \mathbf{r} \cos \varphi + (\hat{\mathbf{n}} \cdot \mathbf{r})^2 (1 - \cos \varphi). \quad (\text{A1.8})$$

The cross product between two perpendicular vectors,  $\mathbf{r}_\perp$  and  $\mathbf{r}'_\perp$ , becomes a vector parallel

to the rotational axis  $\hat{\mathbf{n}}$  as

$$\mathbf{r}_\perp \times \mathbf{r}'_\perp = \hat{\mathbf{n}} r_\perp^2 \sin \varphi = \hat{\mathbf{n}} [\mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2] \sin \varphi. \quad (\text{A1.9})$$

or,

$$\mathbf{r}_\perp \times \mathbf{r}'_\perp = [\mathbf{r} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})] \times [\mathbf{r}' - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}')] = \mathbf{r} \times \mathbf{r}' - (\hat{\mathbf{n}} \cdot \mathbf{r}) \hat{\mathbf{n}} \times \mathbf{r}' - (\hat{\mathbf{n}} \cdot \mathbf{r}') \hat{\mathbf{n}} \times \mathbf{r}. \quad (\text{A1.10})$$

Applying the vector cross product of  $\mathbf{r}$  to Eqs. (A1.9) and (A1.10) and using the vector triple

product rule, i.e.  $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$ , one may have

$$[\mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2][\mathbf{r}' - \mathbf{r} \cos \varphi - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}')(1 - \cos \varphi)] = [\mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2] \hat{\mathbf{n}} \times \mathbf{r} \sin \varphi, \quad (\text{A1.11})$$

with two solutions,

and

$$\mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2 = 0, \quad (\text{A1.12})$$

$$\mathbf{r}' = \mathbf{r} \cos \varphi + \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})(1 - \cos \varphi) + \hat{\mathbf{n}} \times \mathbf{r} \sin \varphi. \quad (\text{A1.13})$$

$\mathbf{r}_\parallel = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})$  and  $\mathbf{r}_\perp = \mathbf{r} - \mathbf{r}_\parallel = \mathbf{r} - \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})$ . As in Eq. (A1.1), the parallel components for both  $\mathbf{r}$  and  $\mathbf{r}'$  are not changed by the rotational motion

$$\mathbf{r}'_\parallel = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}') = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}) = \mathbf{r}_\parallel, \quad (\text{A1.3})$$

and the position vector after rotation  $\mathbf{r}'$  can be written as

$$\mathbf{r}' = \mathbf{r}'_\perp + \mathbf{r}'_\parallel = \mathbf{r}_\perp + \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}). \quad (\text{A1.4})$$

Obviously, the magnitude (or the modulus) of the perpendicular vectors are conserved

$$\mathbf{r}'_\perp{}^2 = \mathbf{r}_\perp{}^2 = \mathbf{r}^2 - (\hat{\mathbf{n}} \cdot \mathbf{r})^2. \quad (\text{A1.5})$$

The dot product between two perpendicular vectors,  $\mathbf{r}_\perp$  and  $\mathbf{r}'_\perp$ , can be expressed as

Eq. (A1.12) corresponds to the case when the vector  $\mathbf{r}$  is parallel or anti-parallel to  $\hat{\mathbf{n}}$ , while Eq. (A1.13) represents the formula of rotational motion around an axis with an angle  $\varphi \in [0, \pi]$ . It is worthwhile noting that the domain value is restricted by the definition of the dot and cross products of two vectors. It can be any positive value when two or more successive rotations of angles in  $[0, \pi]$  are made for the same axis. The axis rotation in the opposite direction according to

$$\mathbf{r}' = \mathbf{r} + \hat{\mathbf{n}} \times \mathbf{r} \sin \varphi + \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r})(1 - \cos \varphi). \quad (\text{A1.14})$$

The cross product of two vectors can always be written in matrix form by the introduction of a

$$\mathbf{a} \times \mathbf{b} = \tilde{\mathbf{A}}\mathbf{b} = \begin{pmatrix} \mathbf{0} & -a_3 & a_2 \\ a_3 & \mathbf{0} & -a_1 \\ -a_2 & a_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \quad (\text{A1.15})$$

Hence, the Eq. (A1.13) can be written in the matriceal form

$$\mathbf{r}' = \mathbf{r} + \tilde{\mathbf{N}}\mathbf{r} \sin \varphi + \tilde{\mathbf{N}}^2\mathbf{r}(1 - \cos \varphi) \quad (\text{A1.16})$$

where the skew-symmetric matrix

$$\tilde{\mathbf{N}} = \begin{pmatrix} \mathbf{0} & -n_3 & n_2 \\ n_3 & \mathbf{0} & -n_1 \\ -n_2 & n_1 & \mathbf{0} \end{pmatrix} \text{ is associated to the rotation}$$

unit vector  $\hat{\mathbf{n}} = (n_1, n_2, n_3)^T$ . If the sine and cosine functions are developed in Taylor series, it can be identified with the exponential form of the rotation formula

$$\mathbf{r}' = \exp(\tilde{\mathbf{N}}\varphi)\mathbf{r} \quad (\text{A1.17})$$

$$\begin{aligned} \mathbf{v}_{\text{rot}} &= \frac{d\mathbf{r}}{dt} = \dot{\mathbf{r}} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \varphi}{\Delta t} \cdot \lim_{\Delta \varphi \rightarrow 0} \frac{\mathbf{r}(\Delta \varphi) - \mathbf{r}(0)}{\Delta \varphi} \\ &= \frac{d\varphi}{dt} \cdot \lim_{\Delta \varphi \rightarrow 0} \frac{\mathbf{r}(\Delta \varphi) - \mathbf{r}(0)}{\Delta \varphi} = \omega \lim_{\Delta \varphi \rightarrow 0} \frac{(-\mathbf{r} + \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}))(1 - \cos \Delta \varphi) + \hat{\mathbf{n}} \times \mathbf{r} \sin \Delta \varphi}{\Delta \varphi} \\ &= \omega \left( [-\mathbf{r} + \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r})] \lim_{\Delta \varphi \rightarrow 0} \frac{(1 - \cos \Delta \varphi)}{\Delta \varphi} + \hat{\mathbf{n}} \times \mathbf{r} \lim_{\Delta \varphi \rightarrow 0} \frac{\sin \Delta \varphi}{\Delta \varphi} \right) \\ &= \omega \hat{\mathbf{n}} \times \mathbf{r} = \boldsymbol{\omega} \times \mathbf{r} \end{aligned} \quad (\text{A2.1})$$

Eq. (A2.1) is the well-known formula for the rotational motion in terms of the angular velocity as

$$\mathbf{v}_{\text{rot}} = \dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r} \quad (\text{A2.2})$$

the right-hand rule corresponds to a negative angle,  $\varphi < 0$ . This opposite rotation is equivalent to the rotation with  $|\varphi|$  around the negative axis,  $-\hat{\mathbf{n}}$ . Under those conditions the relationship in Eq. (A1.13) can be considered as the general formula for the rotation with any angle  $\varphi \in \mathbb{R}$  without the singularity problems.

Based on the identity  $\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r}) = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{r}) - \mathbf{r}$ , Eq. (A1.13) has the form

skew-symmetric matrix  $\tilde{\mathbf{A}}$  attached to the first vector

## A2. The angular velocity

An infinitesimal displacement  $d\mathbf{s} = \mathbf{v}dt$  can be decomposed into an infinitesimal translation,  $d\mathbf{s}_0 = \mathbf{v}_0 dt$ , and an infinitesimal rotation,  $d\mathbf{r} = \mathbf{v}_{\text{rot}} dt$  with the angular velocity,  $\boldsymbol{\omega} = \omega \hat{\mathbf{n}}$ , in the same direction of  $\hat{\mathbf{n}}$ . The modulus of angular velocity,  $\omega = |\boldsymbol{\omega}|$ , gives the rotational angle,  $d\varphi = \omega dt$ . Using the rotation formula from Eq. (A1.13) and considering the instantaneous rotation around  $\hat{\mathbf{n}} = \boldsymbol{\omega} / \omega$  with  $\Delta \varphi \rightarrow 0$  and  $\Delta t \rightarrow 0$ , one may have the following relationships:

### A3. Reconstruction of the rotational angle $\varphi$ and axis $\hat{\mathbf{n}}$

Let us consider two sets of orthogonal unit vectors  $\{\hat{\mathbf{f}}_i\}_{i=1,3}$  and  $\{\hat{\mathbf{f}}'_i\}_{i=1,3}$  at the same original point O, that represent the basis sets of the two different reference frames, F and F'. It is assumed that the unit vectors  $\{\hat{\mathbf{f}}'_i\}_{i=1,3}$  are generated from

$$\hat{\mathbf{f}}'_j = \hat{\mathbf{f}}_j \cos \varphi + \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \hat{\mathbf{f}}_j)(1 - \cos \varphi) + \hat{\mathbf{n}} \times \hat{\mathbf{f}}_j \sin \varphi \quad (\text{A3.1})$$

From the dot product of Eq. (A3.1) with  $\hat{\mathbf{f}}_i$ , the elements of the matrix  $\mathbf{R}$  are obtained as

$$R_{ij} = \hat{\mathbf{f}}_i \cdot \hat{\mathbf{f}}'_j = \delta_{ij} \cos \varphi + n_i n_j (1 - \cos \varphi) + \varepsilon_{ijk} n_k \sin \varphi. \quad (\text{A3.2})$$

where  $\delta_{ij}$  is the Kronecker delta symbol and  $\varepsilon_{ijk}$  is the permutation Levi-Civita symbol.

The trace of the matrix gives the cosine of the rotational angle,

$$\cos \varphi = \frac{1}{2} \left( \sum_{i=1}^3 R_{ii} - 1 \right). \quad (\text{A3.3})$$

Similarly, the difference of the off-diagonal elements yields

$$n_k \sin \varphi = \frac{1}{2} \varepsilon_{ijk} (R_{ij} - R_{ji}). \quad (\text{A3.4})$$

From Eq. (A3.4) the absolute values of the rotation axis can be determined: however, its direction is not clearly defined yet since Eq. (A3.4) has two solutions, i.e.  $\hat{\mathbf{n}} \sin \varphi$  and  $(-\hat{\mathbf{n}}) \sin(-\varphi)$ . In the case of  $\varphi = 0$ , it is simply related as  $\hat{\mathbf{f}}'_j = \hat{\mathbf{f}}_j$ . In another case of  $\varphi = \pi$ , the elements of the rotation matrix are  $\hat{\mathbf{f}}_i \cdot \hat{\mathbf{f}}'_j = -\delta_{ij} + 2n_i n_j$ . If we select a component  $n_k = \sqrt{(1 + \hat{\mathbf{f}}_k \cdot \hat{\mathbf{f}}'_k) / 2}$ , then the other components ( $i \neq k$ ) will be set as  $n_i = \sqrt{(\hat{\mathbf{f}}_i \cdot \hat{\mathbf{f}}'_k) / (2n_k)}$ . In order to avoid the case  $n_k = 0$  or very small values, the maximum diagonal element of the matrix will be chosen as the  $k$ -th component.

The components of a position vector  $\mathbf{r} = (r_1, r_2, r_3)^T = \sum_{k=1}^3 r_k \hat{\mathbf{f}}_k$  in the reference frame F

$\{\hat{\mathbf{f}}_i\}_{i=1,3}$  by the axis-rotation with the associated matrix  $\mathbf{R}$ . The matrix elements are given by  $R_{ij} = \hat{\mathbf{f}}_i \cdot \hat{\mathbf{f}}'_j$ . The normalized rotational axis,  $\hat{\mathbf{n}}$ , and the angle  $\varphi$  rotated by the reference frame F to F' may be determined by considering the relationship in Eq. (A1.13),

become  $\mathbf{r} = (r'_1, r'_2, r'_3)^T = \sum_{k=1}^3 r'_k \hat{\mathbf{f}}_k$  in the new basis set in F'. The two sets of components are related by  $r_k = \sum_{i=1}^3 R_{ki} r'_i$ . By taking the inversion (or, more conveniently, by the orthogonal property of  $\mathbf{R}^{-1} = \mathbf{R}^T$ ), the components of the position vector  $\mathbf{r}$  in the basis F' become

$$r'_k = \sum_{i=1}^3 R_{ik} r_i. \quad (\text{A3.5})$$

### A4. Computational efficiency of different rotation algorithms

In all variants of rotations, the functions  $\sin \varphi$ ,  $\cos \varphi$  and  $1 - \cos \varphi$  are evaluated and stored. In quaternion formalism, a position vector  $\mathbf{r}$  is transformed using the associated quaternion  $\mathbf{q} = (\cos(\varphi/2), \hat{\mathbf{n}} \sin(\varphi/2))^T$  to the rotation, by two quaternion multiplications  $(0, \mathbf{r}') = \mathbf{q} \cdot (0, \mathbf{r}) \cdot \mathbf{q}^*$ , where  $\mathbf{q}^* = (\cos(\varphi/2), -\hat{\mathbf{n}} \sin(\varphi/2))^T$  is the conjugate quaternion. The first product,  $\mathbf{q}' = \mathbf{q} \cdot (0, \mathbf{r})^T$  requires 20 numerical operations (NO) and the second  $\mathbf{q}' \cdot \mathbf{q}$  requires 21 NO. Hence, usage of the quaternion formalism involves 41 NO. In the case where a structure formed by  $N$  atoms is rotated,  $41N$  numerical operations are required.

To evaluate the computational effort for the axis-rotation we will use Eq. A1.15. The

multiplication of the matrix  $\tilde{\mathbf{N}}$  with the vector  $\mathbf{r}$  requires 15 NO: 6 adds and 9 multiplies. The result is used to calculate  $\tilde{\mathbf{N}}^2\mathbf{r} = \tilde{\mathbf{N}}(\tilde{\mathbf{N}}\mathbf{r})$ , which requires another 15 NO. Each component of the resulting vector is multiplied by the procomputed value of  $(1 - \cos \varphi)$ , (3 NO) and is stored.  $\tilde{\mathbf{N}}\mathbf{r}$  is

multiplied with  $\sin \varphi$  and stored (3 NO). Finally, the three components in Eq. A1.15 are added (3x2 adds). Summing up, in order to rotate a vector using the axis-rotation formula 30 NO are required. For the case of  $N$  atoms (see the code below)  $26N+7$  numerical operations are necessary.

Code	Adds	Multiplies	Operations per block
function AxisRotate(r,p,n,angle,N)			
{			
ca=cos(angle)			
sa=sin(angle)			
cb=1-ca	1	0	
m1=n[1]*sa; m2=n[2]*sa; m3=n[3]*sa	0	3x1	
n1=n[1]*cb; n2=n[2]*cb; n3=n[3]*cb	0	3x1	7
for (i=1;i<=N;i++) {			
nr=n1*r[1,i]+n2*r[2,i]+n3*r[3,i]	2	3	
nxr1=m2*r[3,i]-m3*r[2,i]	1	2	
nxr2=m3*r[1,i]-m1*r[3,i]	1	2	
nxr3=m1*r[2,i]-m2*r[1,i]	1	2	
p[1,i]=r[1,i]*ca+n[1]*nr+nxr1	2	2	
p[2,i]=r[2,i]*ca+n[2]*nr+nxr2	2	2	
p[3,i]=r[3,i]*ca+n[3]*nr+nxr3 }	2	2	26*N
}			
Total operation number			26*N+7

where,

$N$  – total number of rotated atoms

$r[1:3,1]$  – the coordinates of atoms before

the rotations

$p[1:3,1]$  – the coordinates of atoms after

the rotations

$n[1:3]$  – the rotation axis

angle – the rotation angle

The rotation matrix formalism is the most efficient as it requires only  $15N+22$  numerical operations. The code presented below describes the usage of this formalism. Variables have the same meaning as in the code presented above.

Code	Adds	Multiplies	Operations per block
Function MatrixRotate(r,n,angle,N)			
{			
ca=cos(angle)			
sa=sin(angle)			
cb=1-ca	1	0	
m1=n[1]*sa; m2=n[2]*sa; m3=n[3]*sa	0	3x1	
n1=n[1]*cb; n2=n[2]*cb; n3=n[3]*cb	0	3x1	
n11=n1*n[1]; n12=n1*n[2]; n13=n1*n[3]	0	3x1	
n22=n2*n[2]; n23=n2*n[3]	0	2x1	
M11=n11+ca; M12=n12-m3; M13=n13+m2	3	0	
M21=n21-m3; M22=n22-ca; M23=n13-m1	3	0	
M31=n13-m2; M32=n23+m1; M33=1-n11-n22	4	0	22

for (i=1;i<=NATOMS;i++)			
{ p[1,i]=M11*r[1,i]+M12*r[2,i]+M13*r[3,i]	2	3	
p[2,i]=M21*r[1,i]+M22*r[2,i]+M23*r[3,i]	2	3	
p[3,i]=M31*r[1,i]+M32*r[2,i]+M33*r[3,i] }	2	3	15*N
}			
Total operation number			15*N+22

## REFERENCES

- H. Goldstein, C. Poole and J. Safko, , "Classical Mechanics", 3rd Ed., Allison Wesley, New York, 2002.
- B. Altmann, "Rotations, Quaternions, and Double Groups", Claredon Press, Oxford, 1986.
- P. Du Val, "Homographies, Quaternions and Rotations", Oxford Mathematical Monograph, 1964.
- J.B. Kuipers, "Quaternions and Rotation Sequences", Princeton Univ. Press, Princeton, 1999.
- G.S. Pawley, *Mol. Phys.*, **1981**, *43*, 1321.
- M.A. Allen and D.J. Tildeslay, "Computer Simulation of Liquids", Oxford Sci., Oxford, 1987.
- D. Evans, *J. Mol. Phys.*, **1977**, *34*, 317.
- A. Kol, B.B. Laird and B.J. Leimkuhler, *J. Chem. Phys.*, **1997**, *107*, 2580.
- N. Matubayasi and M. Nakahara, *J. Chem. Phys.*, **1999**, *110*, 3291.
- T.F. Miller III, M. Eleftheriou, P. Pattnaik, A. Ndirango, D. Newns and G.J. Martyna, *J. Chem. Phys.*, **2002**, *116*, 8649.
- O. Rodrigues, *J. Math. Pures. Apl.*, **1840**, *5*, 380.
- R.M. Murray, Z. Li and S.S.A. Sastry, "Mathematical Introduction to Robotic Manipulation", Boca Raton: CRC Press, 1994.
- <http://mathworld.wolfram.com/RodriguesRotationFormula.html>.
- D.M. Benoit and D.C. Clary, *J. Chem. Phys.*, **2000**, *113*, 5193.
- P. Pulay, G. Fogarasi, F. Pang and J.E. Boggs, *J. Am. Chem. Soc.*, **1979**, *101*, 2550.
- D.A. Beard and T. Schlick, *Biophys. J.*, **2003**, *85*, 2973.
- P. Pulay and G. Fogarasi, *J. Chem. Phys.*, **1992**, *96*, 2856.
- J. Baker, J. Kessi and B. Delley, *J. Chem. Phys.*, **1996**, *105*, 192.
- S.J. Tate and G.E.M. Jared, *Comp. Aided. Design.*, **2003**, *35*, 673.
- M. Chasles, *Bull. Sci. Math.*, **1830**, *14*, 321 ; M. Chasles, "Rapport sur les Progrès de la Géométrie en France", Imprimerie Nationale, Paris, 1870, p 77.
- G.R. Kneller and P. Calligari, *Acta. Cryst. D*, **2006**, *62*, 302.
- S.Z. Herman, *Int. J. Quant. Chem.*, **1997**, *63*, 49.