Rotation matrices and their infinitesimal generators

Markus Deserno
Department of Chemistry and Biochemistry, UCLA, USA
(Dated: February 5, 2004)

We present two alternative derivations for the rotation matrix corresponding to a specific rotation vector. The first is based on simple geometric considerations, the second uses the representation of a rotation matrix by its infinitesimal generator.

I. GEOMETRIC DERIVATION

Let the vector $\boldsymbol{\alpha}$ describe a rotation in the following way: The modulus $\alpha = |\boldsymbol{\alpha}|$ describes the rotation angle and the direction describes the rotation axis which is supposed to pass through the origin of the coordinate system. Let us denote the corresponding unit vector by $\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha}/\alpha$. What is the matrix $\mathsf{M}(\boldsymbol{\alpha})$ corresponding to this rotation?

Fig. 1 shows a sketch from which we can determine this matrix using geometrical considerations. Let some vector $\boldsymbol{r}$ be rotated into the vector $\boldsymbol{r}'$. We can write $\boldsymbol{r}'$ as the sum of three contributions: Start with $\boldsymbol{r}$, which leads one to the circular plane. Next, one moves towards the axis until one hits the base-point of the vertical which starts at $\boldsymbol{r}'$. Finally, move towards $\boldsymbol{r}'$. Let us first determine the directions of the second and third step. Define the vector $\boldsymbol{\rho}$ as the radius vector which points from the center of the circle towards the endpoint of $\boldsymbol{r}$. Obviously we have $\hat{\boldsymbol{\alpha}} r \cos\theta + \boldsymbol{\rho} = \boldsymbol{r}$. Since $\cos\theta = \hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r}/r$, we have $\boldsymbol{\rho} = \boldsymbol{r} - \hat{\boldsymbol{\alpha}}\ \hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r}$. Let us further call the direction of the second step $\hat{\boldsymbol{n}}$. Apparently this is in the direction of the cross-product of the vectors $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{r}$ (in this order!). Thus we have $\hat{\boldsymbol{n}} = \hat{\boldsymbol{\alpha}} \times \boldsymbol{r}/|\hat{\boldsymbol{\alpha}} \times \boldsymbol{r}| = \hat{\boldsymbol{\alpha}} \times \boldsymbol{r}/(r \sin\theta)$.

Next we have to know how far to move in steps 2 and 3. From Fig. 1 we see easily that in step 2 we have to move a distance $\rho(1 - \cos\alpha)$ and in step 3 a distance $\rho \sin\alpha$.

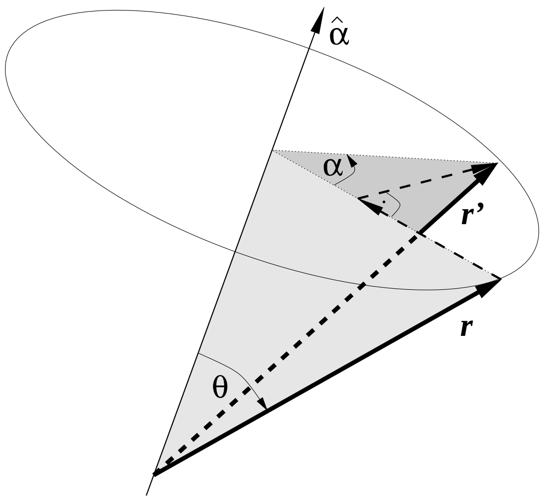

FIG. 1: Sketch of the geometrical relations that permit the determination of the rotated vector $\boldsymbol{r}'$.

Thus we get

$$\begin{aligned} \boldsymbol{r}' &= \boldsymbol{r} - (\boldsymbol{r} - \hat{\boldsymbol{\alpha}}\ \hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r})(1 - \cos\alpha) + \hat{\boldsymbol{\alpha}} \times \boldsymbol{r} \sin\alpha \\ &= \mathsf{M}(\boldsymbol{\alpha})\, \boldsymbol{r} \end{aligned} \tag{1}$$

with the rotation matrix [1]

$$\mathsf{M}(\boldsymbol{\alpha}) = \mathbb{I} - (\mathbb{I} - \hat{\boldsymbol{\alpha}} \otimes \hat{\boldsymbol{\alpha}})(1 - \cos\alpha) + \mathsf{C}(\hat{\boldsymbol{\alpha}}) \sin\alpha, \tag{2}$$

and the cross-product-matrix $\mathsf{C}(\hat{\boldsymbol{\alpha}})$ defined by [2]

$$\mathsf{C}(\hat{\boldsymbol{\alpha}}) = \begin{pmatrix} 0 & -\hat{\alpha}_3 & \hat{\alpha}_2 \\ \hat{\alpha}_3 & 0 & -\hat{\alpha}_1 \\ -\hat{\alpha}_2 & \hat{\alpha}_1 & 0 \end{pmatrix}. \tag{3}$$

II. ALGEBRAIC DERIVATION

We can also derive the general rotation matrix in an algebraic way. For this we first define the three generator matrices L_i by

$$\mathsf{L}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad , \quad \mathsf{L}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

$$\text{and} \quad \mathsf{L}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Using the totally antisymmetric epsilon-tensor we can concisely write this as

$$(\mathsf{L}_i)_{mn} = -\varepsilon_{imn}. \tag{4}$$

Note that the L_i satisfy an angular momentum algebra $[\mathsf{L}_i, \mathsf{L}_j] = \varepsilon_{ijk}\mathsf{L}_k$. From quantum mechanics we know that these matrices generate rotations. In fact, they can be used to very concisely write the rotation matrix:

$$\mathsf{M}(\boldsymbol{\alpha}) = \exp\{\boldsymbol{\alpha} \cdot \mathbf{L}\}, \tag{5}$$

where $\mathbf{L} = (\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_3)^\top$. We now want to evaluate the exponential function explicitly. For this we first need to find the even powers of $\boldsymbol{\alpha} \cdot \mathbf{L}$. The (m, n) component of the square is given by

$$\begin{aligned} (\boldsymbol{\alpha} \cdot \mathbf{L})^2_{mn} &= (\boldsymbol{\alpha} \cdot \mathbf{L})_{mk}(\boldsymbol{\alpha} \cdot \mathbf{L})_{kn} = (\alpha_i \mathsf{L}_i)_{mk}(\alpha_j \mathsf{L}_j)_{kn} \\ &= \alpha_i \alpha_j \varepsilon_{imk}\varepsilon_{jkn} = -\alpha_i\alpha_j(\delta_{ij}\delta_{mn} - \delta_{in}\delta_{jm}) \\ &= -\alpha^2(\mathbb{I} - \hat{\boldsymbol{\alpha}} \otimes \hat{\boldsymbol{\alpha}})_{mn}. \end{aligned} \tag{6}$$

The matrix in brackets is a projector, which means it reproduces itself upon squaring. Hence we have for all even powers $(\boldsymbol{\alpha}\cdot\mathsf{L})^{2n} = (-1)^n\alpha^{2n}(\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})$ *except* for $n=0$, where we of course get the identity matrix.

We are now in the position to evaluate the exponential function:

$$
\begin{aligned}
\exp\{\boldsymbol{\alpha}\cdot\mathsf{L}\} &= \sum_{n=0}^{\infty}\frac{(\boldsymbol{\alpha}\cdot\mathsf{L})^n}{n!} \\
&= \sum_{n=0}^{\infty}\frac{(\boldsymbol{\alpha}\cdot\mathsf{L})^{2n}}{(2n)!} + \boldsymbol{\alpha}\cdot\mathsf{L}\sum_{n=0}^{\infty}\frac{(\boldsymbol{\alpha}\cdot\mathsf{L})^{2n}}{(2n+1)!} \\
&= (\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})\sum_{n=0}^{\infty}\frac{(-1)^n\alpha^{2n}}{(2n)!} \;+\; \hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}} \\
&\quad + \hat{\boldsymbol{\alpha}}\cdot\mathsf{L}\left[(\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})\sum_{n=0}^{\infty}\frac{(-1)^n\alpha^{2n+1}}{(2n+1)!} \;+\; \hat{\boldsymbol{\alpha}}\otimes\boldsymbol{\alpha}\right] \\
&= (\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})\cos\alpha + \;\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}} \\
&\quad + \hat{\boldsymbol{\alpha}}\cdot\mathsf{L}\left[(\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})\sin\alpha \;+\; \hat{\boldsymbol{\alpha}}\otimes\boldsymbol{\alpha}\right]. \qquad (7)
\end{aligned}
$$

The second term simplifies further since

$$(\hat{\boldsymbol{\alpha}}\cdot\mathsf{L}\;\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})_{mn} = (\hat{\alpha}_i\mathsf{L}_i)_{mk}(\hat{\alpha}_k\hat{\alpha}_n) = \varepsilon_{imk}\hat{\alpha}_i\hat{\alpha}_n\hat{\alpha}_k \;=\; 0$$

by virtue of the usual symmetry-antisymmetry argument. We thus obtain for the rotation matrix

$$
\begin{aligned}
\mathsf{M}(\boldsymbol{\alpha}) &= \exp\{\boldsymbol{\alpha}\cdot\mathsf{L}\} \\
&= \mathbb{I} - (\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})(1-\cos\alpha) + \hat{\boldsymbol{\alpha}}\cdot\mathsf{L}\sin\alpha, \qquad (8)
\end{aligned}
$$

which is easily seen to coincide with Eqn. (2), since obviously $\mathsf{C}(\hat{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\alpha}}\cdot\mathsf{L}$.

There is one more check we can do: Calculating the trace of the rotation matrix:

$$
\begin{aligned}
\mathrm{Tr}(\mathsf{M}(\boldsymbol{\alpha})) &= \underbrace{\mathrm{Tr}(\mathbb{I})}_{=3} - \underbrace{\mathrm{Tr}(\mathbb{I}-\hat{\boldsymbol{\alpha}}\otimes\hat{\boldsymbol{\alpha}})}_{=2}(1-\cos\alpha) \\
&\quad + \underbrace{\mathrm{Tr}(\hat{\boldsymbol{\alpha}}\cdot\mathsf{L})}_{=0}\sin\alpha \\
&= 3 - 2(1-\cos\alpha) \;=\; 1 + 2\,\cos\alpha, \qquad (9)
\end{aligned}
$$

which is the expected result for a rotation matrix with rotation angle α. (Since the trace is invariant under a similarity transformation, we can just imagine a simple rotation about the z-axis with an angle α and we see that this is true.)

---

[1] It should be clear from the derivation that $\mathsf{M}(\boldsymbol{\alpha})$ corresponds to an *active* rotation. The corresponding *passive* rotation is given by the inverse matrix $\mathsf{M}^{-1}(\boldsymbol{\alpha}) = \mathsf{M}(-\boldsymbol{\alpha})$.

[2] The matrix $\mathsf{C}(\hat{\boldsymbol{\alpha}})$ has eigenvalues 0, i, and −i. The eigenvector corresponding to the eigenvalue 0 is $\hat{\boldsymbol{\alpha}}$. Since $\hat{\boldsymbol{\alpha}}\times\boldsymbol{r} = \mathsf{C}(\hat{\boldsymbol{\alpha}})\,\boldsymbol{r}$ and $\hat{\boldsymbol{\alpha}}\times\hat{\boldsymbol{\alpha}} = \boldsymbol{0}$ this is obvious. The two other eigenvalues indicate that any vector in the plane orthogonal to $\hat{\boldsymbol{\alpha}}$ is rotated about $\hat{\boldsymbol{\alpha}}$ by 90°. Hence, $\mathsf{C}(\hat{\boldsymbol{\alpha}})$ does something "similar" to a rotation matrix: it combines a rotation about 90° with a projection into the plane perpendicular to the rotation axis.