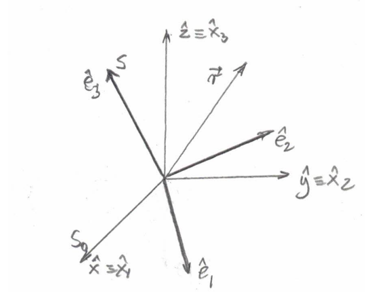
Classical Mechanics Lecture Notes: Rotation Axis Changing in Time

Taylor and other textbooks derive the fundamental formula for the velocity of a point at a position $\boldsymbol{r}$ rotating around an axis $\boldsymbol{\Omega}$ with the angular velocity $\Omega = |\boldsymbol{\Omega}|$

$$\left.\frac{d\boldsymbol{r}}{dt}\right|_{S_0} = \boldsymbol{\Omega} \times \boldsymbol{r}, \tag{1}$$

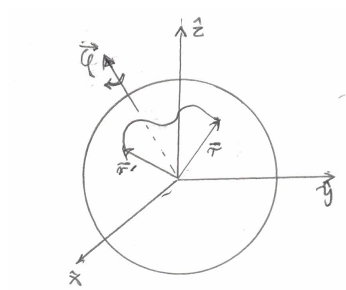
where S_0 is an inertial coordinate system (frame) and $\boldsymbol{\Omega}$ goes through its origin. All these textbooks tacitly assume that the axis $\boldsymbol{\Omega}$ is fixed in S_0 (does not depend on time). Some textbooks (for example, Symon) state that the derivation is valid also if $\boldsymbol{\Omega} = \boldsymbol{\Omega}(t)$, but do not provide a proof for such a case, at least the textbooks I checked. However, all these textbooks later use Eq. (1) in the derivation of Euler's rotational equations, where $\boldsymbol{\Omega}$ is evidently time dependent. One may doubt if the proof with a constant $\boldsymbol{\Omega}$ can apply to the time-dependent case. After all, as the vector $\boldsymbol{r}$ changes in time dt, the change of the axis $\boldsymbol{\Omega}$ should lead to effects of the same order.



Below a hopefully rigorous derivation of Eq. (1) is presented that explicitly accounts for $\boldsymbol{\Omega} = \boldsymbol{\Omega}(t)$. To set the stage, consider in addition to S_0 a noninertial coordinate system S with axes $\hat{\boldsymbol{e}}_i$ (these are, in general, not principal axes of rotation of any body), see the figure. The origins of the two systems overlap. System S rotates about the point $(0, 0, 0)$. We invoke here the most general definition of rotation as a transformation of a rigid body or of a coordinate system that leaves one point in place. Consider next a point at $\boldsymbol{r}$ that is fixed in S. This point has, of course, different coordinates in the two systems

$$\boldsymbol{r} = \sum_i x_i \hat{\boldsymbol{x}}_i = \sum_i x_i' \hat{\boldsymbol{e}}_i,$$

and x_i' are constant. One can also think about a rigid body fixed in S, with $\boldsymbol{r}$ representing a point of this body. Then $\hat{\boldsymbol{e}}_i$ can be chosen as its principal axes of inertia.



Since $\boldsymbol{r}$ is constant in S and S rotates, $\boldsymbol{r}$ moves on on the surface of the sphere of radius $r = |\boldsymbol{r}|$, as shown on the figure, with the path being an arbitrary curve. When $\boldsymbol{r}$ arrives at the position $\boldsymbol{r}'$, this rotation can be represented, as stated by Euler's theorem, by a simple rotation around an axis $\hat{\boldsymbol{\phi}}$ by an angle ϕ. Thus, there exists a rotation matrix $\mathcal{R}$ such that

$$[r'] = \mathcal{R}[r], \tag{2}$$

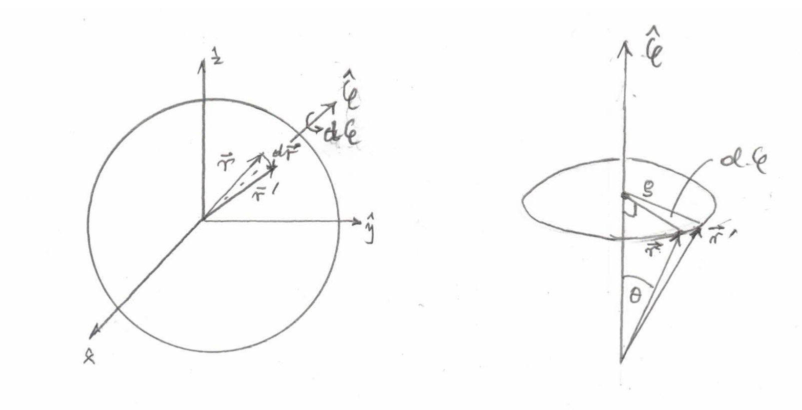
were $[r] = [x, y, z]^T$. As the position $\boldsymbol{r}'$ changes in time, $\boldsymbol{r}' = \boldsymbol{r}'(t)$, the direction and magnitude of the rotation axis also changes in time: $\boldsymbol{\phi} = \boldsymbol{\phi}(t)$.

Now consider the case when the length of the path on the surface of the sphere is infinitesimally small, i.e., the point moves from $\boldsymbol{r}$ to $\boldsymbol{r}'$ such that $d\boldsymbol{r} = \boldsymbol{r}' - \boldsymbol{r}$ is infinitesimally small, see the figure. According to the Euler theorem, there exists an axis $\hat{\boldsymbol{\phi}}$ such that the rotation around this axis by an angle $d\phi$ takes the point from $\boldsymbol{r}$ to $\boldsymbol{r}'$. In contrast to the standard derivation, we do not know a priori the direction of $\hat{\boldsymbol{\phi}}$ and the angle $d\phi$. However, we can find both using the elements of the proof of Euler's theorem. First, the knowledge of $\boldsymbol{r}$ and $\boldsymbol{r}'$ allows one to find the matrix $\mathcal{R}$ from Eq. (2) (3 equations) and from the condition that the rotation matrix is orthogonal: $\mathcal{R}^T\mathcal{R} = \mathbb{1}$ (6 equations). Then $\hat{\boldsymbol{\phi}}$ is obtained as an eigenvector of $\mathcal{R}$ with the eigenvalue 1

$$\mathcal{R}[\hat{\phi}] = [\hat{\phi}]. \tag{3}$$

Since the length of the eigenvector is arbitrary, we have chosen it to be a unit vector.

Now consider an infinitesimal change of $\boldsymbol{r}$, such that $\boldsymbol{r}' - \boldsymbol{r} = d\boldsymbol{r}$. Knowledge of both vectors determines $\mathcal{R}$ and therefore Eq. (3) determines $\hat{\boldsymbol{\phi}}$. Now, we can construct the figure as shown below on left by first drawing $\boldsymbol{r}$ and $\boldsymbol{r}'$ and then placing $\hat{\boldsymbol{\phi}}$ determined by these two vectors. Once $\hat{\boldsymbol{\phi}}$ is known, we know the angle it makes with $\boldsymbol{r}$ and can compute $d\phi$ from geometrical considerations, as shown in the figure on the left. This figure is analogous to the ones found in most textbooks, but was arrived at in a different way (since traditionally the figure is build by first drawing $\hat{\boldsymbol{\phi}}$ and $\boldsymbol{r}$, with $\boldsymbol{r}'$ determined by these two vectors). We



have $\cos\theta = \boldsymbol{r}\cdot\hat{\boldsymbol{\phi}}/r$, $\rho = r\sin\theta$, $d\phi = |d\boldsymbol{r}|/\rho$, and the directio of $d\boldsymbol{r}$ is $\hat{\boldsymbol{\phi}}\times\boldsymbol{r}/|\hat{\boldsymbol{\phi}}\times\boldsymbol{r}|$. Thus, all quantities are now known. We can now write

$$d\boldsymbol{r} = \frac{\hat{\boldsymbol{\phi}}\times\boldsymbol{r}}{|\hat{\boldsymbol{\phi}}\times\boldsymbol{r}|}dr = \frac{\hat{\boldsymbol{\phi}}\times\boldsymbol{r}}{r\sin\theta}\rho d\phi = \hat{\boldsymbol{\phi}}\times\boldsymbol{r}\,d\phi$$

and after division by dt

$$\left.\frac{d\boldsymbol{r}}{dt}\right|_{S_0} = \hat{\boldsymbol{\phi}}\times\boldsymbol{r}\,\frac{d\phi}{dt} = \hat{\boldsymbol{\phi}}\times\boldsymbol{r}\,\omega = \boldsymbol{\Omega}\times\boldsymbol{r}, \tag{4}$$

where $d\phi/dt = \Omega$ and $\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\Omega}}$. Thus, we did obtain Eq. (1) without the assumption that $\boldsymbol{\Omega}(t)$ is constant in time: obviously, at each infinitesimal step on a finite rotation path $\boldsymbol{\Omega}$ is, in general, different. The reason that the change of $\boldsymbol{\Omega}$ during the time dt does not affect the formula is that this change is not independent of the change $d\boldsymbol{r}$, but is defined by this change. One may say that it is the tail ($\boldsymbol{r}$) that wags the dog ($\boldsymbol{\Omega}$).

One may wonder why there is no $d\hat{\boldsymbol{\phi}}/dt$ present. The reason is that in $d\boldsymbol{\phi} = (d\boldsymbol{\phi}/dt)dt$, if chain rule is used to differentiate $\boldsymbol{\phi}$, the term with $d\hat{\boldsymbol{\phi}}/dt$ is of second order in dt (since $\phi \to d\phi = \omega dt$.