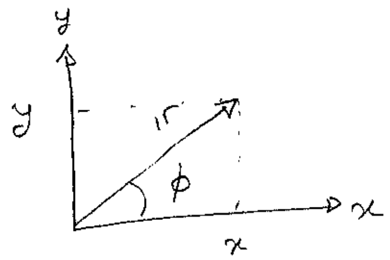
Rotations and angular momentum

We have learned that invariance under time translations imply the conservation of energy. Similarly, invariance under space translations imply the conservation of linear momentum. These are two manifestations of Nöether's theorem. We now reach the third big manifestation: invariance under rotations, as we will see, imply the conservation of angular momentum.

Rotations play a really big role in physics. And they are also a bit more complicated than translations. For this reason, we will first study rotations in detail.

Rotations in 2D

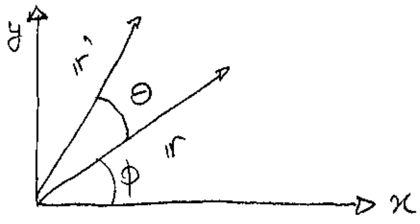
The figure below shows a 2D vector $\mathbf{r} = (x, y)$ in its natural habitat: the cartesian plane



The components of this vector, x and y, are related to $r = |\mathbf{r}|$ and ϕ, as

$$x = r\cos\phi \qquad (1)$$
$$y = r\sin\phi$$

Now I want to manually rotate the vector. This means I want to produce a new vector $r' = (x', y')$ with the same length but rotated by an angle θ with respect to r, in the counter-clockwise direction



The defining property of a rotation is that it leaves the length of the vector unchanged

$$(r')^2 = (x')^2 + (y')^2 = x^2 + y^2 = r^2 \tag{2}$$

Let us now relate x' and y' with x and y. The components x' and y' are similar to x and y in Eq (1). The length r is the same. The only difference is that the angle is now $\phi + \theta$. Thus

$$\begin{aligned} x' &= r\cos(\phi + \theta) \\ y' &= r\sin(\phi + \theta) \end{aligned} \tag{3}$$

Expanding the trigonometric functions we get

$$x' = r(\cos\phi\cos\theta - \sin\phi\sin\theta)$$

$$y' = r(\sin\phi\cos\theta + \cos\phi\sin\theta)$$

Looking at this equation, we recognize x and y as defined in Eq. (1). Thus

$$\boxed{\begin{aligned} x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \end{aligned}} \tag{4}$$

This is the formula we are looking for. It relates the new coordinates to the old coordinates when the vector is rotated by an angle θ.

What we just did is called an <u>active rotation</u> because <u>we</u> manually rotated the vector. It is a convention to always adopt $\theta > 0$ when the rotation is in the counter-clockwise direction. Soon we will study <u>passive rotations</u>, where the vector doesn't move, but the frame of reference does.

What is the inverse operation of Eq (4)? That is, how do we get (x, y) as a function of (x', y')? Well, it is a system of linear equations so we can certainly solve it. But if we think about it for a second, to go from $\mathbf{r}'$ to $\mathbf{r}$ we just need to rotate by an angle $-\theta$. Recalling that $\cos\theta$ is even and $\sin\theta$ is odd, we may then write the inverse of (4) as

$$\begin{aligned} x &= x'\cos\theta + y'\sin\theta \\ y &= -x'\sin\theta + y'\cos\theta \end{aligned} \tag{5}$$

Eqs (4) or (5) are <u>linear</u> transformations and hence may be more conveniently written using matrix notation. We treat $\mathbb{r}$ and $\mathbb{r}'$ as <u>column vectors</u>

$$\mathbb{r} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \mathbb{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{6}$$

then Eq (4) may be written as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{7}$$

The new vector $\mathbb{r}'$ is obtained from the old one by multiplying it with a matrix. this matrix,

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{8}$$

is called a <u>rotation matrix</u>. Then (7) may also be written as

$$\mathbb{r}' = R(\theta)\, \mathbb{r} \tag{9}$$

or, in terms of components,

$$x_i' = \sum_j R_{ij}(\theta) x_j \tag{10}$$

Eqs (4), (7), (9) and (10) are all equivalent ways of writting exactly the same thing.

The inverse transformation (5) may be written more compactly as

$$\mathbb{r} = R(-\theta)\, \mathbb{r}' \tag{11}$$

But using (9) we find that

$$\mathbb{r} = R(-\theta)\, R(\theta)\, \mathbb{r} \tag{12}$$

Hence we conclude that

$$R(-\theta)\, R(\theta) = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{13}$$

The matrix I is called the identity matrix. Eq (13) has a simpley physical interpretation: if you rotate by θ and then by $-\theta$, you get back where you started. This means that $R(-\theta)$ is the <u>matrix inverse</u> of $R(\theta)$:

$$R(-\theta) = R(\theta)^{-1}$$

But looking at the definition of $R(\theta)$ in Eq (8) we see that

$$R(-\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = R(\theta)^T \tag{14}$$

where T is the transpose (you invert rows and columns). Thus

$$R(-\theta) = R(\theta)^T = R(\theta)^{-1} \tag{15}$$

A matrix A which is such that $A^T = A^{-1}$ is called an <u>orthogonal matrix</u>. All rotations (even in 3D) are described by orthogonal matrices. And orthogonal matrices are very special (people who work with numerical linear algebra absolutely love them).

To see why they are special let me first note that the transpose of a column vector is a row vector

$$\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \Rightarrow \quad \mathbf{r}^T = \begin{pmatrix} x & y \end{pmatrix} \qquad (16)$$

Please note that $\mathbf{r}^T \neq (x, y)$. The notation (x, y) with a comma, is just an abbreviation of $\begin{pmatrix} x \\ y \end{pmatrix}$. It is used simply because writting columns in the text is not very convenient.

With (16) in mind we may define the dot product as

$$\mathbf{r} \cdot \mathbf{r} = \mathbf{r}^T \mathbf{r} \qquad (17)$$

See with your own eyes:

$$\mathbf{r}^T \mathbf{r} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 = \mathbf{r} \cdot \mathbf{r}. \qquad (18)$$

Recalling also that $(AB)^T = B^T A^T$, we finally may write

$$(\mathbf{r}')^T (\mathbf{r}') = (R\mathbf{r})^T (R\mathbf{r}) = \mathbf{r}^T R^T R \, \mathbf{r}$$

But $R^T R = I$ so

$$(\mathbf{r}')^T (\mathbf{r}') = \mathbf{r}^T \mathbf{r} \qquad (19)$$

This shows why orthogonal matrices are important: they preserve the length of vectors. This is why rotations are always associated with orthogonal matrices.

Another very important idea is that of composition of rotations: Rotate first by θ_1 then by θ_2. The net result is the same as a single rotation with angle $\theta_1 + \theta_2$. This must reflect in the rotation matrix

$$R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2) \tag{20}$$

Let us check that this is indeed true:

$$R(\theta_1)R(\theta_2) = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}\begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -\cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2 \\ \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2 & -\sin\theta_1\sin\theta_2 + \cos\theta_1\cos\theta_2 \end{pmatrix}$$

$$= \begin{pmatrix} \cos(\theta_1 + \theta_2) & -\sin(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) & \cos(\theta_1 + \theta_2) \end{pmatrix}$$

$$= R(\theta_1 + \theta_2)$$

Yay! Mathematics works! No witchcraft required.

Eq (20) also shows that

$$R(\theta_1)R(\theta_2) = R(\theta_2)R(\theta_1) \tag{21}$$

We say that $R(\theta_1)$ and $R(\theta_2)$ <u>commute</u>. This is a special property of rotations in 2D. In 3D they will not be true in general.

This idea of composing rotations opens up the interesting possibility of decomposing a finite rotation into many infinitesimal rotations. Infinitesimal operations are usually much simpler to work with.

Let us expand in a Taylor series, for a small angle $\delta\theta$,

$$\sin \delta\theta \simeq \delta\theta$$
$$\cos \delta\theta \simeq 1 - \frac{\delta\theta^2}{2}$$

then we see that

$$R(\delta\theta) \simeq \begin{pmatrix} 1 & -\delta\theta \\ \delta\theta & 1 \end{pmatrix} + \mathcal{O}(\delta\theta)^2 \tag{22}$$

the transformations (4) then become

$$\begin{aligned} x' &\simeq x - y\,\delta\theta \\ y' &\simeq y + x\,\delta\theta \end{aligned} \tag{23}$$

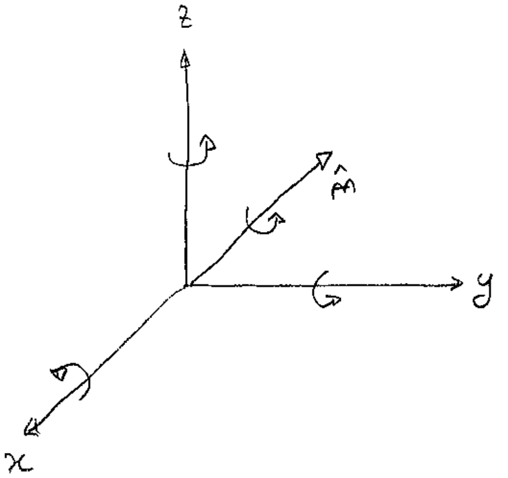
we therefore see that a rotation by $\delta\theta$ cause a change

$$\begin{aligned} \delta x &= -y\,\delta\theta \\ \delta y &= x\,\delta\theta \end{aligned} \tag{24}$$

This follows the usual rule that the response is proportional to the stimulus.

Rotations in 3D

Rotations in 3D are trickier because we have to specify around which axis we are rotating.

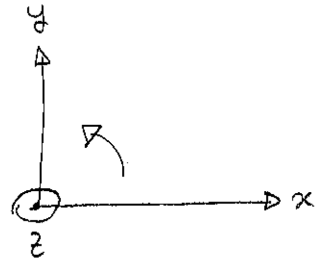


We can rotate around $\hat{x}$, $\hat{y}$, $\hat{z}$ or around any direction specified by a unit vector $\hat{n}$. So to completely specify a rotation in 3D we need 3 parameters: two specifying the direction $\hat{n}$, plus the angle of rotation θ.

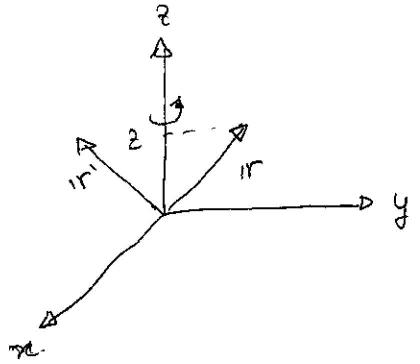
You can also think about rotations in terms of the possible planes at which the rotation occurs. Any rotation may be decomposed as a rotation in the xz plane, another in the xy plane and another in the yz plane. Hence we need 3 parameters.

In 4D we have four axes, $\hat{x}$, $\hat{y}$, $\hat{z}$ and $\hat{w}$. So we need to specify rotations in the planes xy, xz, xw, yz, yw, zw; ie, rotations in 4D require 6 parameters (there is nothing silly about 4D; relativity operates in 4D).

what we have been doing so far in 2D was a rotation in the xy plane, which is the same as a rotation around the z axis



The z component of a vector does not change in this rotation



the x and y components, on the other hand, transform exactly like in (4). Thus

$$\boxed{\begin{aligned} x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \\ z' &= z \end{aligned}} \qquad (25)$$

The infinitesimal version of this translation is readily obtained:

$$x' = x - y\,\delta\theta$$
$$y' = y + x\,\delta\theta \qquad (26)$$
$$z' = z$$

It would be very nice to be able to write this in vector notation. For, if we could do that, we would be able to figure out how to rotate around the other axes.

The way to write (26) in vector notation is to use the vector product:

$$\mathbf{r}' = \mathbf{r} + \delta\theta\,(\hat{z} \times \mathbf{r}) \qquad (27)$$

Let us check that this is true:

$$\hat{z} \times \mathbf{r} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 0 & 0 & 1 \\ x & y & z \end{vmatrix} = \hat{i}\,(-y) + \hat{j}\,(x)$$

Thus, Eq (27) is indeed the same as (26).

Eq (27) has a very clear physical interpretation: the right-hand side contains $\delta\theta$ (the amount you rotate) and $\hat{z}$, the direction around which you rotate.

So if we want to rotate around some other axis, you just replace $\hat{z}$ by the unit vector you desire. For instance, let us look at a rotation around x. We should have

$$\mathbf{r}' = \mathbf{r} + \delta\theta \, (\hat{x} \times \mathbf{r}) \tag{28}$$

But

$$\hat{x} \times \mathbf{r} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 0 & 0 \\ x & y & z \end{vmatrix} = \hat{j}\,(-z) + \hat{k}\,(y)$$

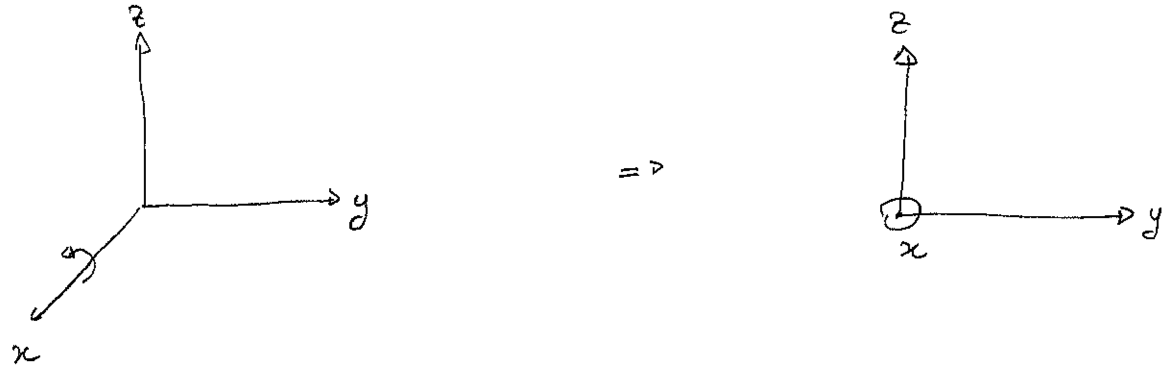
Hence, in components Eq (28) reads

$$\begin{aligned} x' &= x \\ y' &= y - z\,\delta\theta \\ z' &= z + y\,\delta\theta \end{aligned} \tag{29}$$

Looking at (29) we can even be so bold and try to reconstruct the finite rotation, based on the similarity with (25) and (26):

$$\begin{aligned} x' &= x \\ y' &= y\cos\theta - z\sin\theta \\ z' &= y\sin\theta + z\cos\theta \end{aligned} \tag{30}$$

As a consistency check, let us derive this result in another way:



We therefore see that a rotation around x is very similar to a rotation around z. We just need to relabel the coordinates as

$$
\begin{aligned}
z &\rightarrow x \\
x &\rightarrow y \qquad (31)\\
y &\rightarrow z
\end{aligned}
$$

This is what is known as a <u>cyclic</u> <u>permutation</u>. Due to the way we define our cartesian system, we can always use cyclic permutations to go from one axis to another. But the permutation must always be cyclic!

$$x \rightarrow y \rightarrow z$$

You can check that (30) is the same as (25) under a cyclic permutation.

To finish, let us find the formula for a rotation around the y axis. We will do it in two ways, to practice. First, using Eq (27):

$$\mathbf{r}' = \mathbf{r} + \delta\theta\,(\hat{y} \times \mathbf{r})$$

$$\hat{y} \times \mathbf{r} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 0 & 1 & 0 \\ x & y & z \end{vmatrix} = \hat{i}\,(z) + \hat{k}\,(-x)$$

Thus

$$\begin{aligned} x' &= x + \delta\theta\, z \\ y' &= y \\ z' &= z - \delta\theta\, x \end{aligned} \tag{32}$$

chocolate

The corresponding finite rotation will then be

$$\begin{aligned} x' &= x\cos\theta + z\sin\theta \\ y' &= y \\ z' &= -x\sin\theta + z\cos\theta \end{aligned} \tag{33}$$

Alternatively, we may use Eq (30) and do a cyclic permutation

$$\begin{aligned} x' &= x \\ y' &= y\cos\theta - z\sin\theta \\ z' &= y\sin\theta + z\cos\theta \end{aligned} \quad \Longrightarrow \quad \begin{aligned} y' &= y \\ z' &= z\cos\theta - x\sin\theta \\ x' &= z\sin\theta + x\cos\theta \end{aligned}$$

This is exactly the same as (33) (only the order is mixed up). Cyclic permutations is a very powerful tool. It saves a lot of work and can be used as a consistency check.

Now let us try to write our basic rotation formulas using matrices. Start with (25). The corresponding matrix will be

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{34}$$

This is the matrix which rotates by θ around z. Similarly, from (30) we get

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{35}$$

Note the 2×2 blocks which are just the 2D rotation matrix (8).

Finally, we write $R_y(\theta)$ using (33):

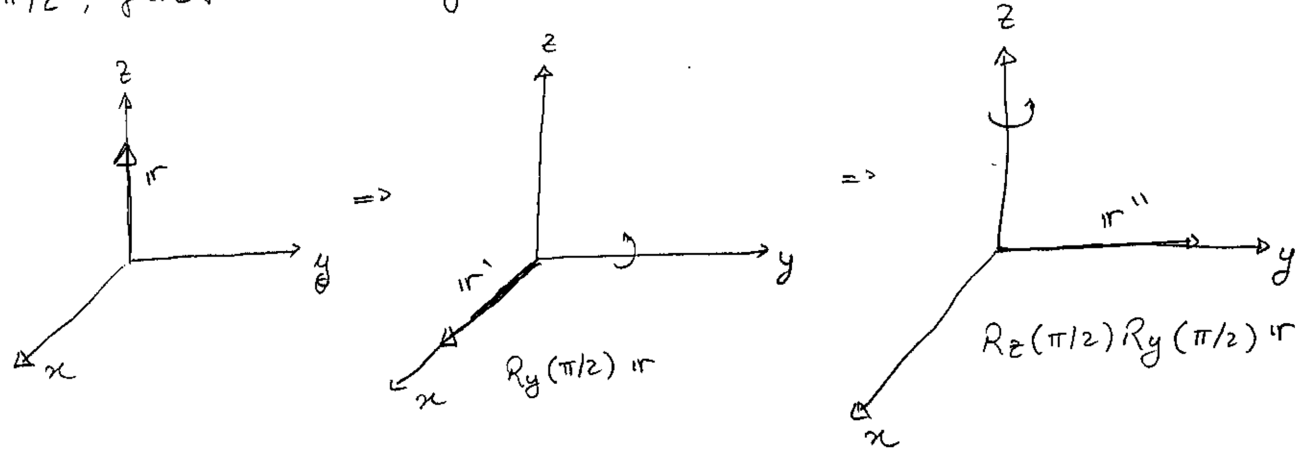
$$R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{36}$$

This is a little more weird (in particular because the minus sign is in a different position, which is not a typo). But that is life.

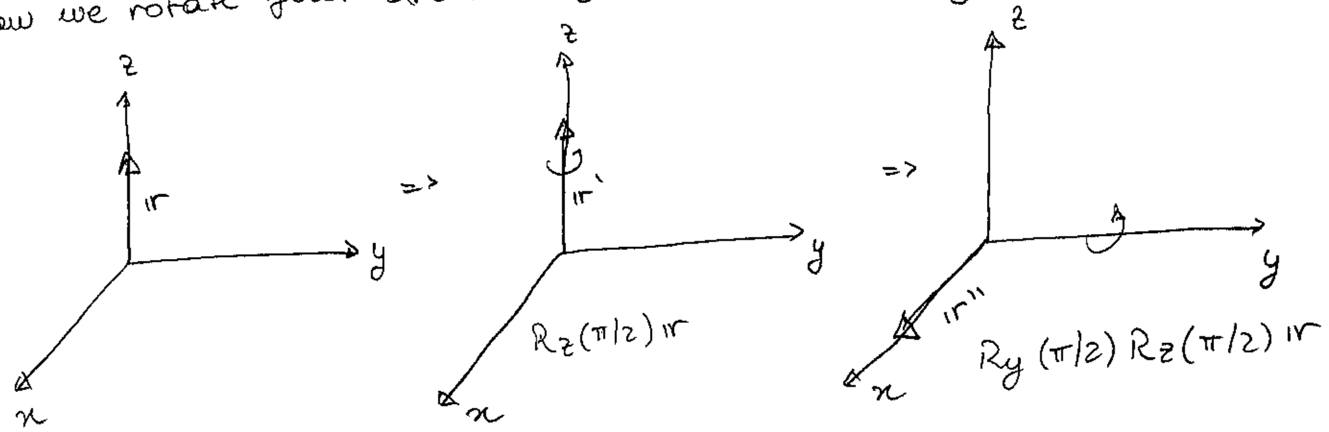
Now comes a really important fact: <u>rotations</u> <u>around</u> <u>different</u> <u>axes</u> <u>do</u> <u>not</u> <u>commute</u>. For instance

$$R_z(\theta_1)\,R_y(\theta_2) \neq R_y(\theta_2)\,R_z(\theta_1)$$

We can see why by considering an example. Suppose we rotate by $\pi/2$, first around y then around z



Now we rotate first around z then around y



We clearly got to different places.

Conclusion: when we perform rotations around different axis, great care must be taken with the <u>order</u> in which the rotations are performed.

Rotations around the same axis can be decomposed into smaller rotations. But rotations around different axes generally cannot.

Let us now go back to infinitesimal rotations. We have seen that a rotation around $\hat{z}$,

$$\mathbb{r}' = \mathbb{r} + \delta\theta\,(\hat{z} \times \mathbb{r})$$

and similarly for $\hat{x}$ and $\hat{y}$. There is nothing special about the 3 cartesian axes. So we may consider a rotation around some arbitrary direction described by a unit vector $\hat{n}$. We will then have

$$\mathbb{r}' = \mathbb{r} + \delta\theta\,(\hat{n} \times \mathbb{r}) \tag{37}$$

We can actually go even a step further and introduce a convenient notation. Let us define a vector

$$\vec{\delta\theta} = \delta\theta\,\hat{n} \tag{38}$$

In words

$$\vec{\delta\theta} : \begin{cases} \text{direction} = \text{direction around which you rotate} \\ \text{magnitude} = \text{angle you rotate} \end{cases}$$

Eq (37) then becomes

$$\boxed{\mathbb{r}' = \mathbb{r} + \vec{\delta\theta} \times \mathbb{r}} \tag{39}$$

Or

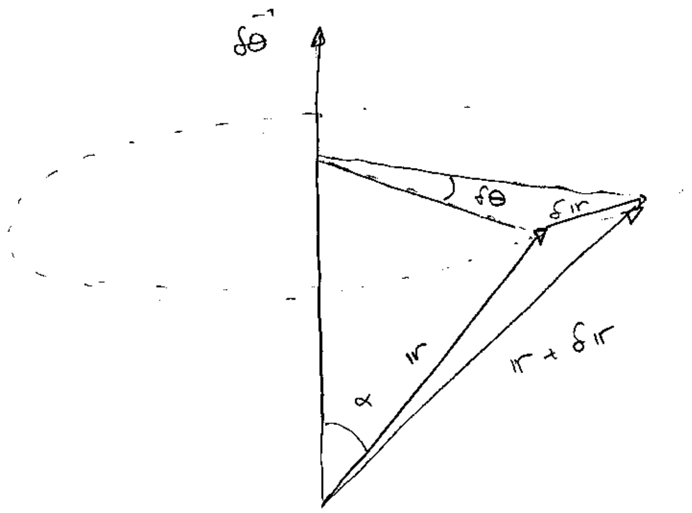
$$\mathbb{r}' = \mathbb{r} + \delta\mathbb{r}$$

and

$$\boxed{\delta\mathbb{r} = \vec{\delta\theta} \times \mathbb{r}} \tag{40}$$

So the change in the vector $\mathbf{r}$ due to an infinitesimal rotation $\vec{\delta\theta}$ is $\vec{\delta\theta} \times \mathbf{r}$.

If you look back, we proved Eq (40) based on Eq (4). So our proof is quite solid since every step is justified. But we may also give a simpler, and more direct proof of (40) based only on geometry.



We see from the (beautifully drawn) picture that $\delta\mathbf{r}$ is a vector which is perpendicular to both $\mathbf{r}$ and $\vec{\delta\theta}$. Moreover its magnitude must be

$$|\delta\mathbf{r}| = \delta\theta \, r \sin\alpha \tag{41}$$

where α is the angle between $\mathbf{r}$ and $\vec{\delta\theta}$ (set $\alpha = 0$ and you will see that $\delta\mathbf{r} \to 0$). Therefore, $\delta\mathbf{r}$ coincides both in magnitude and direction with $\vec{\delta\theta} \times \mathbf{r}$. This is the more direct proof of Eq (40).

We have seen that finite rotations do not commute: the order at which you apply different rotations matters a lot. It turns out, however, that infinitesimal rotations do commute.

Consider two rotations, one by $\delta\vec{\Theta}_1$ and the other by $\delta\vec{\Theta}_2$. After the first rotation we arrive at

$$\mathbb{r}' = \mathbb{r} + \delta\vec{\Theta}_1 \times \mathbb{r}$$

Then, after the second rotation

$$\begin{aligned}\mathbb{r}'' &= \mathbb{r}' + \delta\vec{\Theta}_2 \times \mathbb{r}' \\ &= (\mathbb{r} + \delta\vec{\Theta}_1 \times \mathbb{r}) + \delta\vec{\Theta}_2 \times [\mathbb{r} + \delta\vec{\Theta}_1 \times \mathbb{r}]\end{aligned}$$

The term $\delta\vec{\Theta}_2 \times (\delta\vec{\Theta}_1 \times \mathbb{r})$ is a second order differential and may thus be neglected (it makes no sense in keeping it since we started with a first order formula). Thus

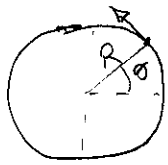
$$\mathbb{r}'' = \mathbb{r} + (\delta\vec{\Theta}_1 + \delta\vec{\Theta}_2) \times \mathbb{r} \tag{42}$$

We therefore see that for infinitesimal rotations, the order at which we apply each step is unimportant.

Angular velocity

this is a good opportunity to rediscuss with more powerful mathematics, a concept you certainly have seen before: angular velocity.

Consider a particle moving in a circle



The angle θ is changing with time. So we define the angular velocity as

$$\omega = \frac{d\theta}{dt} = \dot{\theta} \qquad (43)$$

It has units of rad/s. we may relate the angular velocity with the actual velocity by noting that if θ changes by an amount $d\theta$, the position of the particle changed by $R\,d\theta$. Thus

$$v = R\frac{d\theta}{dt} = R\omega \qquad (44)$$

the concept of angular velocity may also be employed in situations where the particle is not moving in a circle. we do this as follows. Consider two instants of time sepparated by an interval δt. Let $\mathbf{r}$ and $\mathbf{r} + \delta\mathbf{r}$ be the position of the particle in these two instants.

The two points, $\mathbf{r}$ and $\mathbf{r} + \delta\mathbf{r}$, define a plane in space. Thus, we may say that at any given instant of time, the particle is moving in a certain plane. The vector $\delta\mathbf{r}$ lies in this plane.

So, at any given time we may, based on the Fig. in page 18, define a vector $\vec{\delta\theta}$ such that

$$\delta\mathbf{r} = \vec{\delta\theta} \times \mathbf{r}$$

Dividing by δt:

$$\frac{\delta\mathbf{r}}{\delta t} = \frac{\vec{\delta\theta}}{\delta t} \times \mathbf{r}$$

We identify the velocity $\mathbf{v} = \frac{\delta\mathbf{r}}{\delta t}$. Thus, we define the <u>angular</u> <u>velocity</u> as

$$\boxed{\boldsymbol{\omega} = \frac{\vec{\delta\theta}}{\delta t}} \tag{45}$$

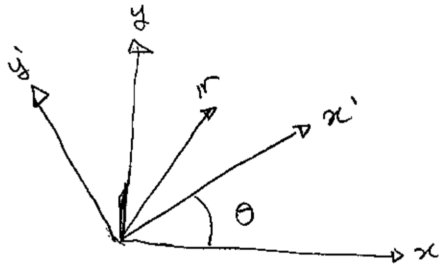
which then gives

$$\boxed{\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}} \tag{46}$$

In this way, we may define the angular velocity for any type of motion. Of course, if it is useful or not, that is a whole different story.

Passive rotations

When we rotate a vector, we call it an active rotation. This is what we have been doing so far. But there is also another possibility which is to rotate the reference frame, while leaving the vector untouched. This is called a <u>passive rotation</u>



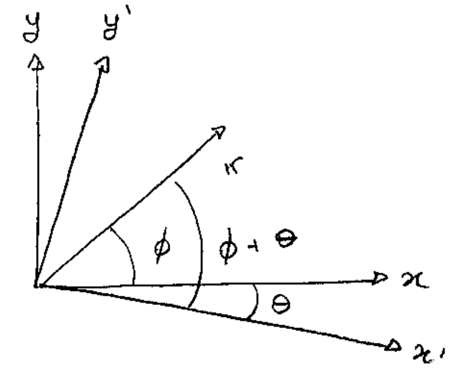
In frame K the vector $\mathbf{r}$ has coordinates (x, y, z). In another frame K' it has coordinates (x', y', z'). Our goal is then to relate the coordinates in the two frames. Note now, however, that we don't have two vectors, $\mathbf{r}$ and $\mathbf{r}'$. We have just one vector, $\mathbf{r}$. But this vector has different components in different frames.

It is <u>very</u> <u>important</u> that you learn to appreciate the difference between the vector $\mathbf{r}$, as a <u>geometrical</u> <u>object</u>, and its description in terms of a coordinate system (i.e, the components of the vector).

The vector $\mathbf{r}$ is a geometrical object pointing in some direction of space. It does not require a reference frame and is thus independent of our choice of frame. Its components, however, change from frame to frame.

In a passive rotation the vector $\mathbf{r}$, as a geometrical object, remains untouched. All we change is the reference frame we use to describe it.

Let us now go back to our main problem of finding (x', y', z') in terms of (x, y, z). Luckly, we will not require any extra work. To see why, consider a clockwise rotation:



From the figure we see that in the K' frame, it is as if we rotated r by an angle θ. (compare this with the figure in page 2).

So a clockwise rotation of the reference frame will have the same transformation law as a counter-clockwise rotation of the vector itself. This makes sense. Please think about it.

Therefore, a counter-clockwise passive rotation is the same as a counter-clockwise active rotation, but by an angle $-\theta$.

For instance, the rotation in the figure reads

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \\ z' &= z \end{aligned} \qquad (47)$$

[Compare with Eq (4)].

In conclusion, we may use all the results we developed for active rotations. We simply need to replace θ with $-\theta$. The inverse of (47) is

$$\begin{aligned} x &= x'\cos\theta - y'\sin\theta \\ y &= x'\sin\theta + y'\cos\theta \end{aligned} \qquad (47')$$

Now let me ask you a deep question: what is a vector? We learn that a set of 3 numbers (a, b, c) is a vector. And while that may be true from a mathematical standpoint, in physics we require a more rigorous definition. In physics a vector is not a set of numbers (a, b, c). These are the components of the vector. For us a vector is a geometrical object, independent of its components.

And this geometrical object must obey a transformation rule. Namely, it must transform under rotations like Eq (47).

Now, I understand that this may seem confusing if you think about a vector such as $(1, 7, -3)$. We should instead, think about abstract objects. Consider, for instance, the force

$$\mathbb{F} = -\nabla U \tag{48}$$

We always take for granted that $\mathbb{F}$ is a vector because we write ∇U in vector notation:

$$\nabla U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z} \right) \tag{49}$$

But for ∇U to actually be a vector, it must correctly transform under rotations. Not all combinations of 3 quantities form a vector. For instance

$$I = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial x}, \frac{\partial U}{\partial x} \right) \tag{50}$$

is not a vector as we will show in a moment, Let us start with ∇U. To show that it is a vector we must check that its components correctly transform like (47). So, using (47')

$$\frac{\partial U}{\partial x'} = \frac{\partial U}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial U}{\partial y}\frac{\partial y}{\partial x'}$$

$$= \frac{\partial U}{\partial x}\cos\theta + \frac{\partial U}{\partial y}\sin\theta$$

$$\frac{\partial U}{\partial y'} = \frac{\partial U}{\partial x}\frac{\partial x}{\partial y'} + \frac{\partial U}{\partial y}\frac{\partial y}{\partial y'}$$

$$= -\frac{\partial U}{\partial x}\sin\theta + \frac{\partial U}{\partial y}\cos\theta$$

Thus,

$$(\nabla U)_{x'} = (\nabla U)_x \cos\theta + (\nabla U)_y \sin\theta$$

$$(\nabla U)_{y'} = -(\nabla U)_x \sin\theta + (\nabla U)_y \cos\theta$$

$$(\nabla U)_{z'} = (\nabla U)_z$$

Hence, ∇U transforms exactly like Eq (47) and may therefore be considered a vector, as does $\mathbb{F}$.

If we look at I in (50), on the other hand, we have

$$I_x = \frac{\partial U}{\partial x} \qquad I_y = \frac{\partial U}{\partial x} \qquad I_z = \frac{\partial U}{\partial x}$$

$$I_{x'} = \frac{\partial U}{\partial x'} = \frac{\partial U}{\partial x}\cos\theta + \frac{\partial U}{\partial y}\sin\theta$$

But $\partial U/\partial y$ is not even a part of I, so I will <u>not</u> transform like Eq (47). Hence (50) is not a vector.

We therefore conclude that ∇U is a vector, because it correctly transforms under rotations. This is why we call nabla,

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \qquad (51)$$

as a vector operator.

When a quantity is invariant under rotations, we call it a scalar. The dot product between two vectors is a scalar

$$\mathbb{A} \cdot \mathbb{B} = \text{scalar} \qquad (52)$$

Let us see why: in the K frame

$$\mathbb{A} \cdot \mathbb{B} = A_x B_x + A_y B_y + A_z B_z$$

and in the K' frame

$$\mathbb{A} \cdot \mathbb{B} = A_x' B_x' + A_y' B_y' + A_z' B_z'$$

Since we are assuming $\mathbb{A}$ and $\mathbb{B}$ are vectors, they must correctly transform under rotations, ie, they must obey (47). Then

$$\begin{aligned} A_x' B_x' &= (A_x \cos\theta + A_y \sin\theta)(B_x \cos\theta + B_y \sin\theta) \\ &= A_x B_x \cos^2\theta + A_y B_y \sin^2\theta + \\ &\quad + \sin\theta \cos\theta \, (A_x B_y + A_y B_x) \end{aligned}$$

$$\begin{aligned} A_y' B_y' &= (-A_x \sin\theta + A_y \cos\theta)(-B_x \sin\theta + B_y \cos\theta) \\ &= A_x B_x \sin^2\theta + A_y B_y \cos^2\theta \\ &\quad - \sin\theta \cos\theta \, (A_x B_y + A_y B_x) \end{aligned}$$

$$A_z' B_z' = A_z B_z.$$

Thus

$$A_x' B_x' + A_y' B_y' + A_z' B_z' = A_x B_x + A_y B_y + A_z B_z$$

conclusion: $\mathbb{A} \cdot \mathbb{B}$ is invariant under rotations, and hence is a <u>scalar</u>.

Since ∇ is a vector, given another vector field $\mathbb{E}$,

$$\nabla \cdot \mathbb{E} = \text{scalar}. \tag{53}$$

Lastly, let us look at the cross product

$$\mathbb{C} = \mathbb{A} \times \mathbb{B} \tag{54}$$

[cf. (46)]. In components

$$\begin{aligned} C_x &= A_y B_z - A_z B_y \\ C_y &= A_z B_x - A_x B_z \\ C_z &= A_x B_y - A_y B_x \end{aligned} \tag{55}$$

Let us check that one of the components transforms as a vector. then by symmetry the same will be true for the others.

$$C_x' = A_y' B_z' - A_z' B_y'$$

$$= (-A_x \sin\theta + A_y \cos\theta) B_z - A_z (-B_x \sin\theta + B_y \cos\theta)$$

$$= \underbrace{(A_y B_z - A_z B_y)}_{C_x} \cos\theta + \underbrace{(A_z B_x - A_x B_z)}_{C_y} \sin\theta$$

$$C_x' = C_x \cos\theta + C_y \sin\theta$$

which is the correct transformation rule.

Thus, $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ correctly transforms like a vector.

However, $\mathbf{C}$ is not entirely like a vector, but has a small difference. To see it we must consider the operation of <u>inversion</u>.

Given a vector $\mathbf{A}$, inversion means

$$(A_x, A_y, A_z) \to (-A_x, -A_y, -A_z)$$

or

$$\mathbf{A} \to -\mathbf{A} \qquad (56)$$

First, it is important that inversion is not the same as a rotation by $180°$. You can find a rotation which takes $x \to -x$ and $y \to -y$. But never one which takes all 3 components simultaneously.

Now, in an inversion, $\mathbb{A} \to -\mathbb{A}$ and $\mathbb{B} \to -\mathbb{B}$. Hence

$$\mathbb{C} \to (-\mathbb{A}) \times (-\mathbb{B}) = \mathbb{C}$$

Vectors change sign under inversions. $\mathbb{A} \times \mathbb{B}$ does not. The cross product is, for this reason, called a pseudo-vector or axial vector.

The same is true for

$$\mathbb{G} \cdot (\mathbb{A} \times \mathbb{B})$$

We saw that $\mathbb{A} \cdot \mathbb{B}$ is a scalar. And, under inversion,

$$\mathbb{A} \cdot \mathbb{B} \to (-\mathbb{A}) \cdot (-\mathbb{B}) = \mathbb{A} \cdot \mathbb{B}$$

But

$$\mathbb{G} \cdot (\mathbb{A} \times \mathbb{B}) \to (-\mathbb{G}) \cdot \left[(-\mathbb{A}) \times (-\mathbb{B})\right] = -\mathbb{G} \cdot (\mathbb{A} \times \mathbb{B})$$

This quantity is therefore called a pseudo scalar

Pseudo scalars appear in the theory of the electroweak interaction.

Angular momentum

We will now consider the last big manifestation of Nöether's theorem: symmetry under rotations. Consider a Lagrangian of the form

$$\mathcal{L} = \sum_a \frac{1}{2} m_a v_a^2 - U(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{57}$$

representing N interacting particles. Let us derive the condition for $\mathcal{L}$ to be invariant under rotations. As we did with linear momentum, we may consider infinitesimal rotations

$$\begin{aligned} \mathbf{r}_a' &= \mathbf{r}_a + \delta \mathbf{r}_a \\ \delta \mathbf{r}_a &= \vec{\delta\theta} \times \mathbf{r}_a \end{aligned} \tag{58}$$

The new Lagrangian is

$$\begin{aligned} \mathcal{L}' &= \mathcal{L}(\mathbf{r}_a + \delta \mathbf{r}_a, \mathbf{v}_a + \delta \mathbf{v}_a) \\ &\simeq \mathcal{L} + \sum_a \left\{ \frac{\partial \mathcal{L}}{\partial \mathbf{r}_a} \cdot \delta \mathbf{r}_a + \frac{\partial \mathcal{L}}{\partial \mathbf{v}_a} \cdot \delta \mathbf{v}_a \right\} \end{aligned} \tag{59}$$

where $\delta \mathbf{v}_a = \delta \dot{\mathbf{r}}_a = \vec{\delta\theta} \times \dot{\mathbf{r}}_a$. Imposing that $\mathcal{L}$ and $\mathcal{L}'$ be equal, we get

$$\delta \mathcal{L} = \sum_a \left\{ \frac{\partial \mathcal{L}}{\partial \mathbf{r}_a} \cdot \delta \mathbf{r}_a + \frac{\partial \mathcal{L}}{\partial \mathbf{v}_a} \cdot \delta \mathbf{v}_a \right\} = 0 \tag{60}$$

We can now work on this formula a bit to make it more intuitive. First we use the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_a} = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \mathbf{v}_a}\right) = \frac{d \mathbf{p}_a}{dt}$$

We then get

$$\delta \mathcal{L} = \sum_a \left\{ \dot{\mathbf{p}}_a \cdot (\vec{\delta\theta} \times \mathbf{r}_a) + \mathbf{p}_a \cdot (\vec{\delta\theta} \times \dot{\mathbf{r}}_a) \right\}$$

We may now group the two terms under a single derivative

$$\delta \mathcal{L} = \frac{d}{dt} \sum_a \mathbf{p}_a \cdot (\vec{\delta\theta} \times \mathbf{r}_a) \tag{61}$$

(Please check that this is true). Finally, we use the cyclic property of the triple product

$$\mathbb{A} \cdot (\mathbb{B} \times \mathbb{C}) = \mathbb{C} \cdot (\mathbb{A} \times \mathbb{B}) = \mathbb{B} \cdot (\mathbb{C} \times \mathbb{A})$$

We then get

$$\delta \mathcal{L} = \frac{d}{dt} \sum_a \vec{\delta\theta} \cdot (\mathbf{r}_a \times \mathbf{p}_a) \tag{62}$$

Example : isolated system

$$\mathcal{L} = \sum_a \frac{1}{2} m_a v_a^2 - U(\mathbf{r}_1, \dots, \mathbf{r}_N) \tag{66}$$

$$U = \frac{1}{2} \sum_{i \neq j} \Phi(|\mathbf{r}_i - \mathbf{r}_j|)$$

the potential depends only on the relative distance between the particles, which are scalars. Hence $\mathcal{L}$ is invariant under rotations.

In this type of system we have seen that the energy E and the total linear momentum of the system $\mathbb{P}$ are conserved. these quantities, like $\mathbb{J}$, are additive in the sense that if you have two independent systems,

$$\begin{aligned} E &= E_1 + E_2 \\ \mathbb{P} &= \mathbb{P}_1 + \mathbb{P}_2 \\ \mathbb{J} &= \mathbb{J}_1 + \mathbb{J}_2 \end{aligned} \tag{67}$$

In fact

> " For an isolated physical system, E, $\mathbb{P}$ and $\mathbb{J}$ are the only additive conserved quantities."

Hence there are only 7 conserved quantities in a mechanical system which are additive.

We now define a quantity

$$\boxed{\mathbb{J} = \sum_a \mathbb{r}_a \times \mathbb{p}_a} \qquad (63)$$

called the angular momentum. Eq (62) then becomes

$$\delta \mathcal{L} = \vec{\delta\theta} \cdot \frac{d\mathbb{J}}{dt} = 0 \qquad (64)$$

Since this must be true for any $\vec{\delta\theta}$, we conclude that

$$\frac{d\mathbb{J}}{dt} = 0$$

Hence, we finally arrive at the law of conservation of angular momentum

If a system is invariant under rotations

$$\Downarrow$$

$$\mathbb{J} = \sum_a \mathbb{r}_a \times \mathbb{p}_a = \text{constant} \qquad (65)$$

Example: central force

$$\mathcal{L} = \frac{1}{2} m v^2 - U(r)$$

Both r and v are scalars and hence are invariant under rotations.

Let us now compare our result with what we know from Newtonian mechanics. For simplicity, we consider just a single particle. Newton's law gives

$$\frac{d\mathbb{p}}{dt} = \mathbb{F} \qquad (68)$$

Multiplying by "$\mathbb{r}\times$" on both sides we get

$$\mathbb{r}\times\frac{d\mathbb{p}}{dt} = \mathbb{r}\times\mathbb{F} \qquad (69)$$

For one particle, $\mathbb{J} = \mathbb{r}\times\mathbb{p}$ so

$$\frac{d\mathbb{J}}{dt} = \dot{\mathbb{r}}\times\mathbb{p} + \mathbb{r}\times\dot{\mathbb{p}} = \mathbb{r}\times\dot{\mathbb{p}}$$

since $\dot{\mathbb{r}} \parallel \mathbb{p}$. We also define the <u>Torque</u> as

$$\mathbb{T} = \mathbb{r}\times\mathbb{F} \qquad (70)$$

so that Eq (69) becomes

$$\frac{d\mathbb{J}}{dt} = \mathbb{T} \qquad (71)$$

We may now state our conservation law as saying that

"If the torque acting on a mechanical system is zero, the angular momentum is conserved"

It is at first not intuitive to make the connection between "Torque is zero" and "$\mathcal{L}$ is invariant under rotations". But we can see this explicitly as follows. Going back to (60), but thinking about a single particle, we have

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{r}} \cdot (\delta \vec{\theta} \times \mathbf{r}) + \frac{\partial \mathcal{L}}{\partial \mathbf{v}} \cdot (\delta \vec{\theta} \times \mathbf{v})$$

well, $\frac{\partial \mathcal{L}}{\partial \mathbf{v}} = \mathbf{p}$ and $\mathbf{p} \parallel \mathbf{v}$ so the last term is zero.

Moreover, $\frac{\partial \mathcal{L}}{\partial \mathbf{r}} = \mathbf{F}$ so we get

$$\delta \mathcal{L} = \mathbf{F} \cdot (\delta \vec{\theta} \times \mathbf{r}) = \delta \vec{\theta} \cdot (\mathbf{r} \times \mathbf{F})$$

Hence, we see that the variation in the Lagrangian due to an infinitesimal rotation is

$$\delta \mathcal{L} = \delta \vec{\theta} \cdot \mathbb{T} \tag{72}$$

So the torque is what quantifies how strongly will the Lagrangian change (the response) due to a rotation $\delta \vec{\theta}$ (the stimulus).

It is then clear that if $\mathbb{T} = 0$ then $\delta \mathcal{L} = 0$; ie, the Lagrangian will not be affected by a rotation.

Spin and orbital angular momentum

Let us now study how the angular momentum changes under a Galileo transformation. We have

$$\mathbb{J} = \sum_a \mathbb{r}_a \times \mathbb{p}_a = \sum_a m_a \mathbb{r}_a \times \mathbb{v}_a \qquad (73)$$

In a Galileo transformation, if a second frame K' is moving with a velocity $\mathbb{V}$ then

$$\mathbb{r}_a' = \mathbb{r}_a - \mathbb{V} t \qquad (74)$$
$$\mathbb{v}_a' = \mathbb{v}_a - \mathbb{V}$$

Substituting in (73) we get

$$\mathbb{J}' = \sum_a m_a \mathbb{r}_a' \times \mathbb{v}_a' = \sum_a m_a (\mathbb{r}_a - \mathbb{V} t) \times (\mathbb{v}_a - \mathbb{V})$$

$$= \sum_a m_a \left[\mathbb{r}_a \times \mathbb{v}_a - \mathbb{r}_a \times \mathbb{V} \quad \mathbb{V} t \times \mathbb{v}_a + 0 \right]$$

$$= \mathbb{J} + \left(\sum_a m_a \mathbb{r}_a \right) \times \mathbb{V} + t \, \mathbb{V} \times \left(\sum_a m_a \mathbb{v}_a \right)$$

Now recall that

$$\sum_a m_a \mathbb{r}_a = M \mathbb{R}, \qquad M = \sum_a m_a$$

$$\sum_a m_a \mathbb{v}_a = \mathbb{P}$$

are the position and momentum of the center of mass in the K frame

We then get

$$\mathbb{J}' = \mathbb{J} - M\mathbb{R} \times \mathbb{V} - t\, \mathbb{V} \times \mathbb{P} \qquad (75)$$

this holds for any arbitrary Galileo transformation. Now let us choose $\mathbb{V}$ in such a way that the system as a whole is at rest in the frame K'; ie such that $\mathbb{P}' = 0$. We have seen that to accomplish this we must set

$$\mathbb{V} = \frac{\mathbb{P}}{M}$$

then the last term in (75) vanishes and we are left with

$$\boxed{\mathbb{J} = \mathbb{J}' + \mathbb{R} \times \mathbb{P}} \qquad (76)$$

This has a very important physical interpretation. The first term $\mathbb{J}'$ is the <u>intrinsic</u> (or <u>spin</u>) <u>angular momentum</u>. It is the angular momentum the system has even when it is at rest. The second term $\mathbb{R} \times \mathbb{P}$ is the <u>orbital angular momentum</u> which describes the rotation of the system as a whole. The earth has a $\mathbb{J}'$ related to it spinning around its own axis, and a $\mathbb{R} \times \mathbb{P}$ related to the orbit around the sun.

It was one of the greatest discoveries of physics when, around 1930, it was found that elementary particles such as the electron also have a spin, even though they are <u>not</u> orbiting around themselves. The spin of the electron, is a fundamental property of the particle.