Physics 106a 12 November, 2013

Lecture 13: Rotations

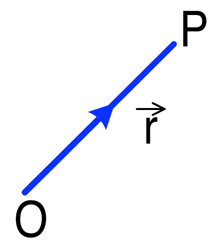
Rotational Symmetry

We use the rotational symmetry of space to write laws of motion in vector notation.

$$m\vec{a} = \vec{F} \; . \tag{1}$$

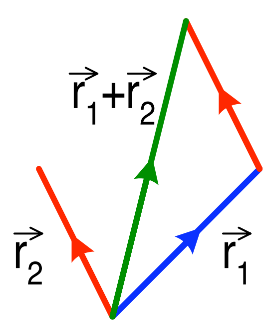
Whatever direction we apply the force, we get the same magnitude acceleration and in the direction of the force. This brings us to the question:

What is a vector?



Our template for a vector is the straight arrow OP joining point O to point P. This is the displacement vector $\vec{r}$. Often we will call the point O the origin. The vector has a magnitude (which we can measure with a calibrated ruler): for a displacement we call the magnitude the length r, and write $\vec{r}^2 = r^2$. The vector also has a direction: if we rotate the vector we get a different vector (pointing in a different direction). Thus the use of vectors is intimately connected with the physics of rotations. A vector is a quantity that behaves in the same way as $\vec{r}$ under rotations: the length is unchanged, and if the direction starts off along some $\vec{r}$, the rotated direction is along the rotated $\vec{r}$. A scalar on the other hand only has a magnitude and is unchanged by rotations. Other quantities (tensors) are changed by rotations in a more complicated way (see later).

We add two vectors $\vec{r}_3 = \vec{r}_1 + \vec{r}_2$ by laying them tail to head and forming the third leg of the triangle.



Multiplication by a scalar gives a vector in the same (positive scalar) or opposite (negative scalar) direction with length multiplied by the magnitude of the scalar.

We may define the dot product of two vectors by a construction such as

$$\vec{r}_1 \cdot \vec{r}_2 = \frac{1}{2}\left[(\vec{r}_1 + \vec{r}_2)^2 - \vec{r}_1^{\,2} - \vec{r}_2^{\,2}\right]. \tag{2}$$

Alternatively we can calculate the product of the lengths times the cosine of the angle between the the vectors. The length of a vector can be expressed as $r^2 = \vec{r} \cdot \vec{r}$.

By subtraction and multiplication by the scalar dt^{-1} we can form the velocity and acceleration vectors

$$\vec{v} = \lim_{\delta t \to 0} \left[\frac{\vec{r}(t+\delta t) - \vec{r}(t)}{\delta t} \right] = \frac{d\vec{r}}{dt}, \qquad \vec{a} = \frac{d\vec{v}}{dt}. \tag{3}$$

Newton tells us the law of motion

$$m\vec{a} = \vec{F} \tag{4}$$

where m is a scalar and $\vec{F}$ (which we must know how to evaluate). For example for two particles separated by the vector $\vec{r}$ we might have

$$\vec{F} = -\frac{GMm}{r^3}\vec{r} \qquad \text{or} \qquad \vec{F} = -K\vec{r} \tag{5}$$

for gravity or a spring. These are clearly vectors too (scalars $\times\ \vec{r}$). The interpretation of this law is given in the first section. The use of vectors, simplifies the formulation of the law, relying on the isotropy of space. Without the symmetry, vectors wouldn't be any help.

Vectors thus provide a coordinate-axis independent way of specifying Newton's law of motion, and other physics. We think of $\vec{r}$, $\vec{v}$, $\vec{F}$, $\vec{E}$, $\vec{B}\ldots$ as having a physical significance independent of any particular coordinate representation.

Components

Often it is convenient to use components of vectors with respect to a particular choice of Cartesian axes. Define $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$ as three unit length ($\hat{\imath}\cdot\hat{\imath} = 1$, etc.) orthogonal ($\hat{\imath}\cdot\hat{\jmath} = 0$, etc.) reference vectors—an *orthonormal basis*. Sometimes I'll use the notation $\hat{e}_i$ with $\hat{e}_1 = \hat{\imath}$, $\hat{e}_2 = \hat{\jmath}$, $\hat{e}_3 = \hat{k}$. Then for a general vector $\vec{a}$ (not meant to imply acceleration) we can write

$$\vec{a} = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}. \tag{6}$$

The a_i, $i = 1, 2, 3$ are called the components of the vector with respect to the basis. We might also use the notation a_x for a_1, $a_y = a_2$ etc. Since the axes are orthonormal the components are also given by projection

$$a_1 = \vec{a}\cdot\hat{\imath}, \text{ etc.} \tag{7}$$

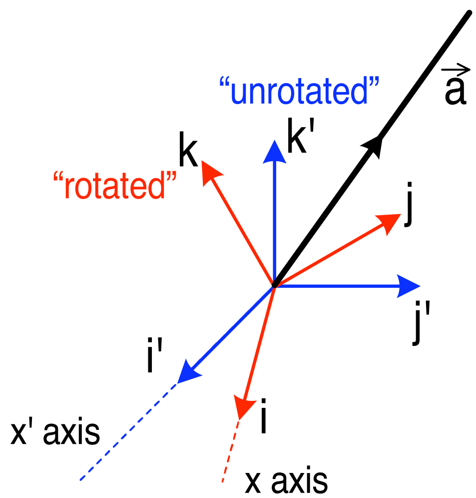
We often think of the three components (a_1, a_2, a_3) and $\vec{a}$ as equivalent.

Using the orthonormality of the basis vectors, the dot product of two vectors can be written in terms of the components

$$\vec{a}\cdot\vec{b} = \sum_{i=1}^{3} a_i b_i. \tag{8}$$

The isotropy of space means that no particular choice of basis is singled out. We are led to ask for a given vector $\vec{a}$ how the components with respect to different bases are related.

Rotation of coordinate axes



Consider a set of basis vectors $\hat{\imath}'$, $\hat{\jmath}'$, $\hat{k}'$ and a *rotated* set $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$. The components a_i' and a_i with respect to these axes are given by

$$\vec{a} = a_1'\hat{\imath}' + a_2'\hat{\jmath}' + a_3'\hat{k}' = a_1\hat{\imath} + a_2\hat{\jmath} + a_3\hat{k}\ . \tag{9}$$

Using the orthonormality we can calculate e.g.

$$a_1' = a_1\hat{\imath}'\cdot\hat{\imath} + a_2\hat{\imath}'\cdot\hat{\jmath} + a_3\hat{\imath}'\cdot\hat{k} \tag{10}$$

or in matrix notation

$$\begin{pmatrix} a_1' \\ a_2' \\ a_3' \end{pmatrix} = U \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{11}$$

with U the *rotation matrix*

$$U = \begin{pmatrix} \hat{\imath}'\cdot\hat{\imath} & \hat{\imath}'\cdot\hat{\jmath} & \hat{\imath}'\cdot\hat{k} \\ \hat{\jmath}'\cdot\hat{\imath} & \hat{\jmath}'\cdot\hat{\jmath} & \hat{\jmath}'\cdot\hat{k} \\ \hat{k}'\cdot\hat{\imath} & \hat{k}'\cdot\hat{\jmath} & \hat{k}'\cdot\hat{k} \end{pmatrix}. \tag{12}$$

This defines the rotation matrix in terms of the *direction cosines* between the basis vectors. The matrix notation means

$$a_i' = \sum_{j=1}^{3} U_{ij}a_j \equiv U_{ij}a_j \tag{13}$$

where the last expression uses the *Einstein repeated subscript convection* that we are to sum over any (pairwise) repeated subscript. *I will use this convention from now on*.

Note that we have kept the vector fixed, and rotated the basis vectors. This is called a *passive* rotation,

Rotation matrix

The rotation matrix U is an *orthogonal matrix*. Define the *transpose* $\tilde{U}$ by $\tilde{U}_{ij} = U_{ji}$. Then you can check

$$U\tilde{U} = \tilde{U}U = I \tag{14}$$

with I the unit matrix (ones down the diagonal, zeroes elsewhere)

$$I_{ij} = \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{15}$$

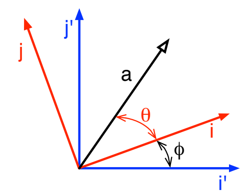
(the Kronecker delta). Equivalently the rows (and columns) are orthonormal

$$U_{ik}U_{jk} = \delta_{ij} \qquad \text{etc. (remember } k\text{-summed)} \tag{16}$$

The matrix U_{ij} contains 9 numbers and Eqs. (16) are six constraints (the similar statements for columns turn out not to be independent). Thus there are 3 numbers that specify U. One choice of parameters is the direction of the axis of rotation $\hat{n}$ and the angle of the rotation ϕ. Another choice is the *Euler angles* which we will consider in Lecture 16.

Example: Consider a rotation by ϕ about the z direction. The z component of a vector is unchanged, and the transformation of the x, y components is easily calculated by trigonometry. Suppose the vector projected onto the x, y plane makes an angle θ to the (rotated) $\hat{\imath}$ axis. Then it will make an angle $\theta + \phi$ to the (unrotated) $\hat{\imath}'$ axis.

$$a_1' = a\cos(\theta+\phi) = a(\cos\theta\cos\phi - \sin\theta\sin\phi) = a_1\cos\phi - a_2\sin\phi$$
$$a_2' = a\sin(\theta+\phi) = a(\cos\theta\sin\phi + \sin\theta\cos\phi) = a_1\sin\phi + a_2\cos\phi$$



This gives the rotation matrix

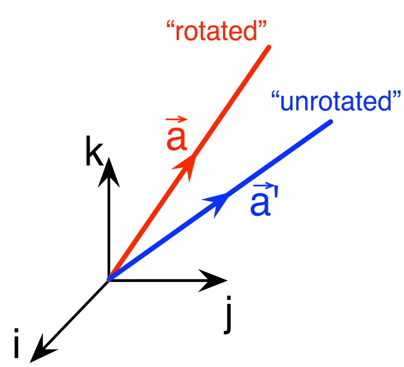
$$U(\phi, \hat{k}) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{17}$$

The general expression for the rotation matrix for an angle ϕ about an axis $\hat{n}$ is (see Assignment 6)

$$U_{ij} = (1-\cos\phi)\hat{n}_i\hat{n}_j + \cos\phi\,\delta_{ij} - \sin\phi\,\epsilon_{ijk}\hat{n}_k. \tag{18}$$

From this it follows that $\mathrm{Tr}\,U = 1 + 2\cos\phi$. Also $\hat{n}$ is the eigenvector of U with eigenvalue 1. (See the slides for a different argument leading to these results.)

Rotation of vector



Instead of rotating the axes/basis vectors we may discuss rotations of the vector, and then evaluate components with respect to one choice of basis. This is called an *active rotation*. Suppose the vector $\vec{a}'$ is rotated into a new vector $\vec{a}$ by the same physical rotation as we used for the basis vectors (e.g. the same rotation axis and angle of rotation). We now have the components $(\vec{a})_i$ of $\vec{a}$ and $(\vec{a}')_i$ of $\vec{a}'$ defined by

$$\vec{a} = (\vec{a})_1\hat{\imath} + (\vec{a})_2\hat{\jmath} + (\vec{a})_3\hat{k} \tag{19}$$
$$\vec{a}' = (\vec{a}')_1\hat{\imath} + (\vec{a}')_2\hat{\jmath} + (\vec{a}')_3\hat{k} \tag{20}$$

Using the fact that the components of the rotated vector with respect to rotated axes are the same as the components of the unrotated vector with respect to unrotated axes you can show

$$(\vec{a})_i = U_{ij}(\vec{a}')_j \tag{21}$$

with U the same rotation matrix. Notice however that the expression relates the rotated to the unrotated, wheres for the passive rotation the expression relates the unrotated to the rotated. As is clear intuitively, rotating the coordinate axes has the same effects on the components as the negative rotation acting on the vector.

Applied to both sides of $\vec{F} = m\vec{a}$ a passive rotation says that we get the same component expression e.g. $F_x = ma_x$ with respect to any choice of basis vectors, whereas an active rotation says that a rotated force gives a rotated acceleration.

Alternative definition of a vector

Some texts prefer to *define* a vector as a quantity defined by three numbers (a_1, a_2, a_3) that transform according to the rules Eq. (13). I prefer the geometric definition. It seems inelegant to define a quantity used to eliminate the need for components in terms of components! For an exposition of the geometric definition of *4-vectors* in special relativity see the first few pages of the notes to Ph136a (we will return to this next term).

Group properties

Rotations form a *group*. A group is defined as a collection of elements (here different rotations) with the properties:

multiplication: the product U_2U_1 is defined as first apply U_1 and then apply U_2. Clearly

$$U_2U_1 = U_3 \tag{22}$$

with U_3 some other rotation;

associative rule: it can be shown

$$(U_1U_2)U_3 = U_1(U_2U_3)\ ; \tag{23}$$

identity: the identity exists—do nothing!

inverse: the inverse rotation is a rotation about the same axis through the negative angle.

Note that multiplication of rotations is not commutative in general

$$U_1U_2 \neq U_2U_1 \tag{24}$$

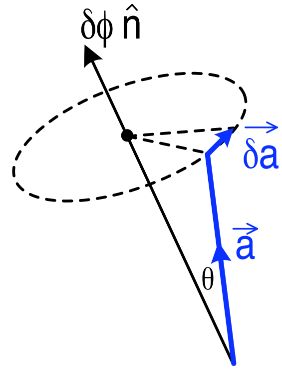
the order of rotations about different axes does matter. (Try $\pi/2$ rotations about two axes at $\pi/2$.) Multiplication is also rather complicated: it is not easy to see what U_3 is if U_1, U_2 are specified by axis/angles $\hat{n}_1, \theta_1$ and $\hat{n}_2, \theta_2$.

Rotations form a *continuous* or *Lie* group since the elements are specified by continuous parameters (direction of axis and angle).

The U here denote physical rotation operations. The matrices U_{ij} give a *representation* of the group—the effect of the operations on an orthonormal triad. We could also look at the effect of rotations on other quantities, such as a tensor (see later) or the $2p$ or $3d$ wave functions in hydrogen (see Ph125). These would give different representations.

Infinitesimal rotation

Infinitesimal rotations are a vector



Rotations by an infinitesimal angle $\delta\phi$ about some axis $\hat{n}$ are easier to deal with: they can in fact be described in terms of an infinitesimal *rotation vector* $\delta\vec{\phi} = \delta\phi\,\hat{n}$. Let's look at this in active form. Consider the change $\delta\vec{a}$ in a vector $\vec{a}$ due to an infinitesimal rotation $\delta\phi$ about an axis specified by the unit vector $\hat{n}$ (see figure). The vector $\vec{a}$ is at an angle θ to $\hat{n}$. The change $\delta\vec{a}$ is perpendicular to $\vec{a}$ and $\hat{n}$ and is of magnitude $a\sin\theta\delta\phi$, and so

$$\delta\vec{a} = \delta\phi\hat{n}\times\vec{a} = \delta\vec{\phi}\times\vec{a} \qquad \text{with } \delta\vec{\phi} = \delta\phi\hat{n}. \tag{25}$$

$\delta\vec{\phi}$ is a *pseudovector* — it behaves like $\hat{n}$ a vector under rotations, but under inversions $\vec{r}\to-\vec{r}$ it does *not* change sign (make some sketches). We also know this because the $\times$ operation requires a "right hand rule" which changes under inversion.

For successive infinitesimal rotations $\vec{a}''\to\vec{a}'\to\vec{a}$ we just add the $\delta\vec{\phi}$ vectors:

$$\vec{a} = \vec{a}' + \delta\vec{\phi}_2\times\vec{a}' = (\vec{a}'' + \delta\vec{\phi}_1\times\vec{a}'') + \delta\vec{\phi}_2\times(\vec{a}'' + \delta\vec{\phi}_1\times\vec{a}'') \tag{26}$$

$$= \vec{a}'' + (\delta\vec{\phi}_1 + \delta\vec{\phi}_2)\times\vec{a}'' + O(\delta\vec{\phi}^2). \tag{27}$$

This also implies that infinitesimal rotations commute—the order does not matter. These results are *not* true for finite rotations $\phi_1, \hat{n}_1$ and $\phi_2, \hat{n}_2$.

Angular velocity Rotation rate is defined as a small rotation divided by a small time increment and can be characterized in terms of the *angular velocity* $\vec{\omega} = d\vec{\phi}/dt$, which is also a vector (vector $\delta\vec{\phi}$ divided by scalar δt).

Rotation matrix for infinitesimal rotation: The rotated vector is $\vec{a}$ and the unrotated one $\vec{a}'$ in our notation convention, so that

$$\vec{a} = \vec{a}' + \delta\vec{\phi}\times\vec{a}' \tag{28}$$

or in component notation with respect to a fixed basis

$$(\vec{a})_i = (\delta_{ij} - \epsilon_{ijk}\delta\phi_k)(\vec{a}')_j \tag{29}$$

with ϵ_{ijk} the Levi-Civita symbol

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any repeated indices} \\ 1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123} \end{cases} \tag{30}$$

The corresponding rotation matrix is

$$\delta U = I + \delta\vec{\phi}\cdot\vec{M} \tag{31}$$

with $\vec{M}$ a vector of matrices with components

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{32}$$

or

$$(M_k)_{ij} = -\epsilon_{ijk}. \tag{33}$$

(You can check the general U given in terms of direction cosines Eq. (12) reduces to Eq. (31) for small rotations.)

We can build up a finite rotation from a succession of infinitesimal rotations about a fixed axis: for a rotation ϕ about $\hat{n}$ build up as N rotations through $\delta\phi = \phi/N$ for $N \to \infty$

$$U = \lim_{N\to\infty} \left(I + \frac{\phi}{N} \hat{n} \cdot \vec{M} \right)^N = e^{\phi \hat{n} \cdot \vec{M}}, \tag{34}$$

where the exponential of a matrix is just a compact way of writing the power series

$$e^A = 1 + A + \tfrac{1}{2!} A^2 + \cdots \tag{35}$$

You can check that finite rotations about the same axis *do* commute. Since we can form a general rotation from the matrices $\vec{M}$ giving an infinitesimal rotation the $\vec{M}$ are called *generators* of the Lie group (actually representations of the generators). You can show that

$$M_i M_j - M_j M_i = c_{ij}^k M_k \tag{36}$$

with the *structure constant* c_{ij}^k is this case given by $c_{ij}^k = \epsilon_{ijk}$. This type of expression for the *commutation rule* of the generators is called the *Lie algebra* and is a fundamental property of the group.

Noether's theorem for rotation revisited Define the rotation transformed paths in coordinate notation

$$Ri(\delta\vec{\phi}, t) = \delta U_{ij} r_j \tag{37}$$

with δU given by Eq. (31). For a system with rotational symmetry Noether's theorem gives the conserved quantities

$$I_k = p_i \left. \frac{\partial R_i}{\partial \delta\phi_k} \right|_{\delta\phi=0} = p_i (M_k)_{ij} r_j \qquad \text{(remember, } i, j \text{ summed)} \tag{38}$$

$$= -\epsilon_{ijk} p_i r_j = (\vec{r} \times \vec{p})_k \tag{39}$$

so that $\vec{I}$ is the angular momentum vector $\vec{r} \times \vec{p}$.

I've given the argument for a single particle: summing over many particles is a trivial extension.

Rigid body motion

For a rigid body, if we fix one reference point in the body (e.g. by going to an inertial frame instantaneously comoving with this point), the only possible motion is a rotation, and the time derivative of this rotation is the angular velocity $\vec{\omega}$. This is known as *Euler's theorem*. Now return to the original inertial frame, the velocity of the ith point or mass element is (*Chasles' theorem*)

$$\vec{v}_i = \vec{V} + \vec{\omega} \times \vec{r}_i, \tag{40}$$

where $\vec{V} = \dot{\vec{R}}$ is the translational velocity of the reference point K in the body at $\vec{R}$ relative to a fixed origin, $\vec{r}_i$ is the displacement of the ith point relative to K, and $\vec{\omega}$ is the angular velocity.

If we change the reference point to K' at $\vec{R}' = \vec{R} + \vec{a}$ so that the vector from the ith point to K' is $\vec{r}_i' = \vec{r}_i - \vec{a}$ then the same expression holds

$$\vec{v}_i = \vec{V}' + \vec{\omega} \times \vec{r}_i'. \tag{41}$$

with a modified translational velocity $\vec{V}' = \vec{V} + \vec{\omega} \times \vec{a}$, but the *same angular velocity*. The angular velocity is therefore *independent of the reference point in the body*, and we can talk about the angular velocity of a solid body without reference to a particular origin.

We may, if we wish, choose the reference point to eliminate the translational motion perpendicular to $\vec{\omega}$, so that the motion is rotation plus translation *along* the direction of $\vec{\omega}$ — starting off with K as the reference point, move it to K' with $\vec{a}$ chosen so that $\vec{\omega} \times \vec{a}$ eliminates the component perpendicular to $\vec{\omega}$ in $\vec{V}'$. Shifting the origin along the direction of $\vec{\omega}$ does not change $\vec{V}'$ or $\vec{\omega}$. This means the motion is instantaneously rotation about some line (not necessarily intersecting the body) we call the *axis of rotation*, plus translation along that axis

$$\vec{v}_i = V\hat{\omega} + \vec{\omega} \times \vec{r}_{\perp i}, \tag{42}$$

with $\hat{\omega}$ the unit vector in the direction of ω and $\vec{r}_{\perp i}$ the distance of the ith point from the axis.

Rolling motion In rolling problems, there is no sliding motion of the contact point or line, so for a stationary rolling surface the motion is pure rotation about some rotation axis that passes through the point of contact or about a line of contact. For a point contact we can often find the rotation axis by seeking another stationary point: the rotation axis must pass through this point and the contact point.

What is a tensor?

Consider a physical example of a tensor such as the electric conductivity. In a crystal, the electric current produced by an electric field is not necessarily along the field. The conductivity *tensor* outputs the current *vector* $\vec{j}$ when we input the electric field *vector* $\vec{E}$. Or the moment of inertia tensor we will discuss next week outputs the angular momentum vector given the angular velocity vector. These are *second rank* tensors: a second rank tensor takes a vector as an input and produces a vector (which is a linear function of the input) as the output.

Kip Thorne (see Ph136a notes) favors a "slot notation" for tensors that implements this idea. He would write the conductivity tensor as $\sigma(\_,\_)$ such that if you put the electric field in one slot you get the current vector

$$\vec{j} = \sigma(\_, \vec{E}) , \tag{43}$$

Alternatively, if you input vectors into both slots, you get a scalar, in fact

$$\frac{1}{2}\sigma(\vec{E}, \vec{E}) = \text{ rate of energy dissipation.} \tag{44}$$

A *second rank* tensor has two slots; an nth rank tensor has n slots, and produces a scalar from n input vectors. Since we know how vectors transform under rotations (and scalars do not) either of these relationships defines how the tensor changes under rotations. A tensor is linear in each of its inputs:

$$\sigma(\_, \alpha\vec{E}_1 + \beta\vec{E}_2) = \alpha\sigma(\_, \vec{E}_1) + \beta\sigma(\_, \vec{E}_2) \tag{45}$$

These properties form a coordinate independent way of defining tensors.

More conventionally tensors are defined in component form, e.g. if the current is linear in the electric field we must have

$$j_i = \sigma_{ij} E_j . \tag{46}$$

For this to be true also in the unrotated basis $j'_i = \sigma'_{ij} E'_j$ we have

$$U_{i\alpha} j_a = \sigma'_{ij} U_{j\beta} E_\beta . \tag{47}$$

The left hand side is $U_{i\alpha}\sigma_{\alpha\beta}E_\beta$ so that (equating coefficients of E_β and using $\tilde{U} = U^{-1}$)

$$\sigma'_{ij} = U_{i\alpha} U_{j\beta} \sigma_{\alpha\beta} . \tag{48}$$

Correspondingly, in terms of components, an nth rank (or rank-n) tensor $T_{i_1 i_2 \cdots i_n}$ is a quantity with 3^n components which transform as

$$T'_{i_1 i_2 \cdots i_n} = U_{i_1 j_1} U_{i_2 j_2} \ldots U_{i_n j_n} T_{j_1 j_2 \cdots j_n} . \tag{49}$$

A vector is a 1st-rank tensor. The outer product $a_i b_j$ of two vectors is a second rank tensor. δ_{ij} and ϵ_{ijk} are *isotropic* second and third rank tensors — the definitions apply in any coordinate basis. This can be seen directly by applying the rotation matrices.

Michael Cross November 13, 2013