Rigid Bodies

November 5, 2018

Contents

A rigid body is defined as one in which the distance between any two points in the body remains constant. If we think of the body as made up of individual particles, the distances between any two of these particles is fixed. We may equally well consider a continuum approximation. In either case the complete orientation of the body requires six parameters. Beginning from an arbitrary origin and coordinates, we may locate any one point, P_1, in the body by three coordinates. Picking any other point, P_2, the distance d_{12} between them is fixed, so P_2 lies on a sphere of radius d_{12} centered on P_1, and we may specify the position of P_2 on this sphere by giving two angles. Finally, with the positions of P_1 and P_2 fixed, any third point P_3 lies in a plane with P_1 and P_2. Since this plane contains the line connecting P_1 and P_2, the only freedom in specifying P_3 is a single angle, specifying the orientation of this plane. The positions of all further points are fixed by their distances from these three. We therefore require $3 + 2 + 1 = 6$ parameters to fully specify the position and orientation of a rigid body.

Fix an orthonormal coordinate system, with basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$, fixed in space, and a second set, $\mathbf{i}', \mathbf{j}', \mathbf{k}'$, fixed in the rigid body. Think of the origin of the second set as P_1 so that the origins of these two systems separated by the position vector $\mathbf{R}_1$ of the point P_1. The only remaining difference between the basis vectors will be determined by the three angular variables required to specify P_2, P_3.

1 Change of basis

We now seek the relationship between two orthonormal bases with a common origin.

The first key fact is that the transformation is *linear*, and this is immdediate by the definition of a vector basis. Given a set of basis vectors $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, every vector can be expanded as a linear combination

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

Since any other basis is comprised of vectors, the vectors of the new basis $(\mathbf{i}', \mathbf{j}', \mathbf{k}')$ may be expanded in the old:

$$\begin{aligned} \mathbf{i}' &= a_{11}\mathbf{i} + a_{12}\mathbf{j} + a_{13}\mathbf{k} \\ \mathbf{j}' &= a_{21}\mathbf{i} + a_{22}\mathbf{j} + a_{23}\mathbf{k} \\ \mathbf{k}' &= a_{31}\mathbf{i} + a_{32}\mathbf{j} + a_{33}\mathbf{k} \end{aligned} \tag{1}$$

and conversely,

$$\begin{aligned} \mathbf{i} &= b_{11}\mathbf{i}' + b_{12}\mathbf{j}' + b_{13}\mathbf{k}' \\ \mathbf{j} &= b_{21}\mathbf{i}' + b_{22}\mathbf{j}' + b_{23}\mathbf{k}' \\ \mathbf{k} &= b_{31}\mathbf{i}' + b_{32}\mathbf{j}' + b_{33}\mathbf{k}' \end{aligned} \tag{2}$$

We consider two types of transformation: passive and active. A passive transformation is one in which any given vector remains fixed while we transform the basis. An active transformation is one in which we leave the basis fixed, but transform all vectors. Thus, for a *passive* transformation and an arbitrary vector $\mathbf{v}$, we expand in each basis:

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ &= v_1'\mathbf{i}' + v_2'\mathbf{j}' + v_3'\mathbf{k}' \end{aligned}$$

where the new basis vectors expanded in the old as in Eq.(1) as above. For an active transformation, we consider two vectors

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ \mathbf{v}' &= v_1'\mathbf{i} + v_2'\mathbf{j} + v_3'\mathbf{k} \end{aligned}$$

where, since the transformation must still be linear, the components are related by

$$\begin{aligned} v_1' &= A_{11}v_1 + A_{12}v_2 + A_{13}v_3 \\ v_2' &= A_{21}v_1 + A_{22}v_2 + A_{23}v_3 \\ v_3' &= A_{31}v_1 + A_{32}v_2 + A_{33}v_3 \end{aligned} \tag{3}$$

Before looking at the relationship between active and passive transformations, we develop some simpler notation.

1.1 Einstein summation convention

All of this is much easier in index notation. Let the three basis vectors be denoted by $\hat{\mathbf{e}}_i$, $i = 1, 2, 3$, so that

$$\begin{aligned} \hat{\mathbf{e}}_1 &= \hat{\mathbf{i}} \\ \hat{\mathbf{e}}_2 &= \hat{\mathbf{j}} \\ \hat{\mathbf{e}}_3 &= \hat{\mathbf{k}} \end{aligned}$$

and similarly for $\hat{\mathbf{e}}_i^{'}$. Then the basis transformations, Eqs.(1) and (2), may be written as

$$\begin{aligned} \hat{\mathbf{e}}_i^{'} &= \sum_{j=1}^{3} a_{ij}\hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_i &= \sum_{j=1}^{3} b_{ij}\hat{\mathbf{e}}_j^{'} \end{aligned}$$

where a_{ij} is the matrix

$$a_{ij} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

This lets us write three long equations as a single short one. Note that the first index tells us which *row*, and the second which *column*.

The expansion of the vector $\mathbf{v}$ is

$$\mathbf{v} = \sum_{j=1}^{3} v_j \hat{\mathbf{e}}_j$$

and an active transformation as

$$\sum_{k=1}^{3} v_k^{'}\hat{\mathbf{e}}_k = \sum_{j=1}^{3}\sum_{k=1}^{3} A_{kj} v_j \hat{\mathbf{e}}_k$$

It is easy to see that this will lead us to write $\sum_{j=1}^{3}$ millions of times. The Einstein convention avoids this by noting that when there is a sum there is also a repeated index – j, in the cases above. If we are careful never to repeat an index that we do not sum, we may drop the summation sign. Thus,

$$\begin{aligned} \sum_{j=1}^{3} a_{ij}\hat{\mathbf{e}}_j &\Longrightarrow a_{ij}\hat{\mathbf{e}}_j \\ \sum_{j=1}^{3} b_{ij}\hat{\mathbf{e}}_j^{'} &\Longrightarrow b_{ij}\hat{\mathbf{e}}_j^{'} \\ \sum_{j=1}^{3} v_j\hat{\mathbf{e}}_j &\Longrightarrow v_j\hat{\mathbf{e}}_j \end{aligned}$$

The repeated index is called a dummy index, and it does not matter what letter we choose for it,

$$v_j \hat{\mathbf{e}}_j = v_k \hat{\mathbf{e}}_k$$

as long as we do not use an index that we have used elsewhere in the same expression. Thus, in the basis change examples above, we cannot use i as the dummy index because it is used to distinguish three independent equations:

$$\begin{aligned} \hat{\mathbf{e}}_1^{'} &= a_{1j} \hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_2^{'} &= a_{2j} \hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_3^{'} &= a_{3j} \hat{\mathbf{e}}_j \end{aligned}$$

Such an index is called a free index. Free indices must match in every term of an expression. Thus, dummy indices tell us that there is a sum, while free indices tell us which equation we are looking at.

Since the basis is orthonormal, we know that the dot product is given by

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

where δ_{ij} is the Kronecker delta, equal to 1 if $i = j$ and to 0 if $i \neq j$. Notice that the expression above represents nine separate equations. If we repeat the index, we have a single equation

$$\begin{aligned} \hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_k &= \delta_{kk} \\ &= 3 \end{aligned}$$

Be sure you understand why the result is 3.

We can find the relationship between the matrices a_{ij} and b_{ij}, since, substituting one basis change into the other,

$$\begin{aligned} \hat{\mathbf{e}}_i^{'} &= a_{ij} \hat{\mathbf{e}}_j \\ &= a_{ij} \left(b_{jk} \hat{\mathbf{e}}_k^{'} \right) \\ &= a_{ij} b_{jk} \hat{\mathbf{e}}_k^{'} \end{aligned}$$

Taking the dot product with $\hat{\mathbf{e}}_m^{'}$ (notice that we cannot use any of the indices i, j or k), we have

$$\begin{aligned} \hat{\mathbf{e}}_i^{'} &= a_{ij} b_{jk} \hat{\mathbf{e}}_k^{'} \\ \hat{\mathbf{e}}_m^{'} \cdot \hat{\mathbf{e}}_i^{'} &= \hat{\mathbf{e}}_m^{'} \cdot \left(a_{ij} b_{jk} \hat{\mathbf{e}}_k^{'} \right) \\ \delta_{mi} &= a_{ij} b_{jk} \hat{\mathbf{e}}_m^{'} \cdot \hat{\mathbf{e}}_k^{'} \\ &= a_{ij} b_{jk} \delta_{mk} \\ &= a_{ij} b_{jm} \end{aligned}$$

and since $\delta_{mi} = \delta_{im}$ is the identity matrix, 1, this shows that

$$AB = 1$$

so that the matrix B with components b_{ij} is inverse to A,

$$B = A^{-1}$$

This introduces matrix notation, where A and B are entire matrices with components

$$\begin{aligned} [A]_{ij} &= a_{ij} \\ [B]_{ij} &= b_{ij} \end{aligned}$$

When we write a matrix product as AB we understand this to be normal matrix multiplication, in which successive columns of b_{jm} are summed with successive rows of a_{ij}.

1.2 Passive transformation

Consider a passive transformation from $\hat{\mathbf{e}}_j$ to $\hat{\mathbf{e}}_i^{'}$. Substituting for the relationship between the basis vectors, we have

$$\begin{aligned} v_i'\hat{\mathbf{e}}_i^{'} &= v_i\hat{\mathbf{e}}_i \\ &= v_i\left(b_{ij}\hat{\mathbf{e}}_j^{'}\right) \\ &= (v_i b_{ij})\,\hat{\mathbf{e}}_j^{'} \end{aligned}$$

so that the components of the vector $\mathbf{v}$ in the new basis are given by

$$v_k^{'} = v_i b_{ik}$$

To prove this formally, we take the dot product of both sides of our equation with each of the three basis vectors, $\hat{\mathbf{e}}_k^{'}$:

$$\begin{aligned} v_i^{'}\hat{\mathbf{e}}_i^{'}\cdot\hat{\mathbf{e}}_k^{'} &= (v_i b_{ij})\,\hat{\mathbf{e}}_j^{'}\cdot\hat{\mathbf{e}}_k^{'} \\ v_i^{'}\delta_{ik} &= v_i b_{ij}\delta_{jk} \\ v_k^{'} &= v_i b_{ik} \end{aligned}$$

Notice how the identity matrix simply replaces one index with another.

1.3 Active transformation

What active transformation gives the same components for v' as a given passive one? That is, suppose we want an active transformation that takes $\mathbf{v} = v_i\hat{\mathbf{e}}^i$ to $\mathbf{v}' = v_i'\hat{\mathbf{e}}^i$ where $v_k^{'} = v_i b_{ik}$. This requires,

$$\begin{aligned} \sum_{k=1}^{3} v_k^{'}\hat{\mathbf{e}}_k &= \sum_{j=1}^{3}\sum_{k=1}^{3} A_{kj}v_j\hat{\mathbf{e}}_k \\ v_k^{'} &= A_{kj}v_j \\ v_j b_{jk} &= A_{kj}v_j \end{aligned}$$

The relationship here is the transpose.

Let A have components a_{ij}. Then the *transpose* of A, called A^t, has components a_{ji}:

$$\begin{aligned} [A]_{ij} &= a_{ij} \\ \left[A^t\right]_{ij} &= a_{ji} \end{aligned}$$

This allows us to write the relation

$$v_j b_{jk} = A_{kj}v_j$$

as

$$v_j\left[B^t\right]_{kj} = A_{kj}v_j$$

Since, for each value of j and k these are just ordinary numbers, we may write them in any order, so that

$$\left[B^t\right]_{kj} v_j = A_{kj}v_j$$

and therefore

$$B^t = A$$

1.4 Transformations

We will normally write active transformations,

$$v_k^{'} = [O]_{ki} v_i$$

in components, or

$$\mathbf{v}' = O\mathbf{v}$$

in matrix notation.

Now consider an active rotation of a vector $\mathbf{v}$ to a new vector $\mathbf{v}'$. Consider a counterclockwise rotation about the z axis through an angle θ. If $\mathbf{v}$ begins as a vector $(v, 0, 0)$ along the x-axis then it will end up as $\mathbf{v}' = (v\cos\theta, v\sin\theta, 0)$. Therefore, the first column of O must be:

$$[O]_{ij} = \begin{pmatrix} \cos\theta & \cdot & \cdot \\ \sin\theta & \cdot & \cdot \\ 0 & \cdot & \cdot \end{pmatrix}$$

If $\mathbf{v}$ were initially in the y-direction, $(0, v, 0)$, then it would rotate to a negative x-value, $\mathbf{v}' = (-v\sin\theta, v\cos\theta, 0)$. This becomes the second column of O, and since anything in the z direction is unaffected, the full transformation must be

$$[O]_{ij} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4}$$

To accomplish the same result with a passive rotation, $\hat{\mathbf{e}}_i^{'} = \sum_{j=1}^{3} a_{ij}\hat{\mathbf{e}}_j$, we would have rotated the basis clockwise by the same angle. The matrix a_{ij} takes the form

$$a_{ij} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that the matrix a_{ij} is the inverse of O.

1.5 The Levi-Civita tensor

In 3-dimensions, we define the Levi-Civita tensor, ε_{ijk}, to be totally antisymmetric, so we get a minus sign under interchange of any pair of indices. We work throughout in Cartesian coordinate. This means that most of the 27 components are zero, since, for example,

$$\varepsilon_{212} = -\varepsilon_{212}$$

if we imagine interchanging the two 2s. This means that the only nonzero components are the ones for which i, j and k all take different values. There are only six of these, and all of their values are determined once we choose any one of them. Define

$$\varepsilon_{123} \equiv 1$$

Then by antisymmetry it follows that

$$\begin{aligned} \varepsilon_{123} &= \varepsilon_{231} = \varepsilon_{312} = +1 \\ \varepsilon_{132} &= \varepsilon_{213} = \varepsilon_{321} = -1 \end{aligned}$$

All other components are zero.

Using ε_{ijk} we can write index expressions for the cross product and the curl. The i^{th} component of the cross product is given by

$$[\mathbf{u} \times \mathbf{v}]_i = \varepsilon_{ijk} u_j v_k \tag{5}$$

(note the two implicit sums on j and k) as we check by simply writing out the sums for each value of i,

$$\begin{aligned}
[\mathbf{u}\times\mathbf{v}]_1 &= \varepsilon_{1jk}u_jv_k \\
&= \varepsilon_{123}u_2v_3+\varepsilon_{132}u_3v_2+(\textit{all other terms are zero}) \\
&= u_2v_3-u_3v_2 \\
[\mathbf{u}\times\mathbf{v}]_2 &= \varepsilon_{2jk}u_jv_k \\
&= \varepsilon_{231}u_3v_1+\varepsilon_{213}u_1v_3 \\
&= u_3v_1-u_1v_3 \\
[\mathbf{u}\times\mathbf{v}]_3 &= \varepsilon_{3jk}u_jv_k \\
&= u_1v_2-u_2v_1
\end{aligned}$$

We get the curl simply by replacing u_j by $\nabla_j=\frac{\partial}{\partial x_j}$,

$$[\nabla\times\mathbf{v}]_i=\varepsilon_{ijk}\nabla_jv_k \tag{6}$$

If we sum Eqs.(5) and (6) with basis vectors $\mathbf{e}_i$, where $\mathbf{e}_1=\mathbf{i},\mathbf{e}_2=\mathbf{j},\mathbf{e}_3=\mathbf{k}$, we may write the curl and cross product as vectors:

$$\begin{aligned}
\mathbf{u}\times\mathbf{v} &= [\mathbf{u}\times\mathbf{v}]_i\,\mathbf{e}_i \\
&= \varepsilon_{ijk}u_jv_k\mathbf{e}_i \\
\nabla\times\mathbf{v} &= \varepsilon_{ijk}\mathbf{e}_i\nabla_jv_k
\end{aligned}$$

There are useful identities involving pairs of Levi-Civita tensors. The most general is

$$\varepsilon_{ijk}\varepsilon_{lmn}=\delta_{il}\delta_{jm}\delta_{kn}+\delta_{im}\delta_{jn}\delta_{kl}+\delta_{in}\delta_{jl}\delta_{km}-\delta_{il}\delta_{jn}\delta_{km}-\delta_{in}\delta_{jm}\delta_{kl}-\delta_{im}\delta_{jl}\delta_{kn} \tag{7}$$

This happens because on each ε_{ijk}, the i,j and k must be $1,2$ and 3 in some order, so any individual term may be written in terms of deltas. For example,

$$\varepsilon_{132}\varepsilon_{213} = +\delta_{11}\delta_{22}\delta_{33}$$

To check the whole identity explicitly, note that both sides are totally antisymmetric on ijk and on lmn. But in three dimensions there can be only one totally antisymmetric object with three indices, up to an overall multiple. Therefore, since the right side is antisymmetric on ijk and on lmn, it must be proportional to both ε_{ijk} and ε_{lmn},

$$\varepsilon_{ijk}\varepsilon_{lmn}=\lambda\left(\delta_{il}\delta_{jm}\delta_{kn}+\delta_{im}\delta_{jn}\delta_{kl}+\delta_{in}\delta_{jl}\delta_{km}-\delta_{il}\delta_{jn}\delta_{km}-\delta_{in}\delta_{jm}\delta_{kl}-\delta_{im}\delta_{jl}\delta_{kn}\right)$$

for some λ. To fix λ we only need to check one component, say $\varepsilon_{123}\varepsilon_{123}$:

$$\begin{aligned}
\varepsilon_{123}\varepsilon_{123} &= \lambda\left(\delta_{11}\delta_{22}\delta_{33}+\delta_{12}\delta_{23}\delta_{31}+\delta_{13}\delta_{22}\delta_{33}-\delta_{11}\delta_{23}\delta_{32}-\delta_{13}\delta_{22}\delta_{31}-\delta_{12}\delta_{21}\delta_{33}\right) \\
&= \lambda\delta_{11}\delta_{22}\delta_{33}
\end{aligned}$$

so we must have $\lambda=1$ and the full identity holds.

We get a second identity by setting $n=k$ and summing,

$$\begin{aligned}
\varepsilon_{ijk}\varepsilon_{lmk} &= \delta_{il}\delta_{jm}\delta_{kk}+\delta_{im}\delta_{jk}\delta_{kl}+\delta_{ik}\delta_{jl}\delta_{km}-\delta_{il}\delta_{jk}\delta_{km}-\delta_{ik}\delta_{jm}\delta_{kl}-\delta_{im}\delta_{jl}\delta_{kk} \\
&= 3\delta_{il}\delta_{jm}+\delta_{im}\delta_{jl}+\delta_{im}\delta_{jl}-\delta_{il}\delta_{jm}-\delta_{il}\delta_{jm}-3\delta_{im}\delta_{jl} \\
&= (3-1-1)\,\delta_{il}\delta_{jm}-(3-1-1)\,\delta_{im}\delta_{jl} \\
&= \delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}
\end{aligned}$$

so we have a much simpler, and very useful, relation

$$\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \tag{8}$$

A second sum gives another identity. Setting $m = j$ and summing again,

$$\begin{aligned} \varepsilon_{ijk}\varepsilon_{ljk} &= \delta_{il}\delta_{mm} - \delta_{im}\delta_{ml} \\ &= 3\delta_{il} - \delta_{il} \\ &= 2\delta_{il} \end{aligned} \tag{9}$$

Setting the last two indices equal and summing provides a check on our normalization,

$$\varepsilon_{ijk}\varepsilon_{ijk} = 2\delta_{ii} = 6$$

This is correct, since there are only six nonzero components and we are summing their squares.

Now we use these properties to prove some vector identities. First, consider the triple product (corresponding to the volume of the parallelipiped defined by the three vectors),

$$\begin{aligned} \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) &= u_i\left[\mathbf{v}\times\mathbf{w}\right]_i \\ &= u_i\varepsilon_{ijk}v_jw_k \\ &= \varepsilon_{ijk}u_iv_jw_k \end{aligned}$$

Because $\varepsilon_{ijk} = \varepsilon_{kij} = \varepsilon_{jki}$, it is unchanged by cyclically permuting the vectors

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) = \mathbf{v}\cdot(\mathbf{w}\times\mathbf{u})$$

Each of these gives the same form when written in components, e.g., $\mathbf{v}\cdot(\mathbf{w}\times\mathbf{u}) = v_i\varepsilon_{ijk}w_ju_k = \varepsilon_{ijk}u_kv_iw_j = \varepsilon_{kij}u_kv_iw_j$. Renaming the dummy indices gives the original expression.

Next, consider a double cross product:

$$\begin{aligned} \left[\mathbf{u}\times(\mathbf{v}\times\mathbf{w})\right]_i &= \varepsilon_{ijk}u_j\left[\mathbf{v}\times\mathbf{w}\right]_k \\ &= \varepsilon_{ijk}u_j\varepsilon_{klm}v_lw_m \\ &= \varepsilon_{ijk}\varepsilon_{klm}u_jv_lw_m \end{aligned}$$

Now using our second identity, Eq.(8), to replace $\varepsilon_{ijk}\varepsilon_{klm} = \varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$, and using the Kronecker deltas to change indices,

$$\begin{aligned} \left[\mathbf{u}\times(\mathbf{v}\times\mathbf{w})\right]_i &= \left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)u_jv_lw_m \\ &= \delta_{il}\delta_{jm}u_jv_lw_m - \delta_{im}\delta_{jl}u_jv_lw_m \\ &= u_jv_iw_j - u_jv_jw_i \\ &= v_i\left(\mathbf{u}\cdot\mathbf{w}\right) - w_i\left(\mathbf{u}\cdot\mathbf{v}\right) \end{aligned}$$

Returning to full vector notation, this is the $BAC - CAB$ rule,

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\,\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\,\mathbf{w}$$

As one final example of the usefulness of the Levi-Civita symbol, look at the curl of a cross product,

$$\begin{aligned} \left[\nabla\times(\mathbf{v}\times\mathbf{w})\right]_i &= \varepsilon_{ijk}\nabla_j\left[\mathbf{v}\times\mathbf{w}\right]_k \\ &= \varepsilon_{ijk}\nabla_j\left(\varepsilon_{klm}v_lw_m\right) \\ &= \varepsilon_{ijk}\varepsilon_{klm}\left(\left(\nabla_jv_l\right)w_m + v_l\nabla_jw_m\right) \\ &= \left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)\left(\left(\nabla_jv_l\right)w_m + v_l\nabla_jw_m\right) \\ &= \delta_{il}\delta_{jm}\left(\left(\nabla_jv_l\right)w_m + v_l\nabla_jw_m\right) - \delta_{im}\delta_{jl}\left(\left(\nabla_jv_l\right)w_m + v_l\nabla_jw_m\right) \\ &= \left(\nabla_mv_i\right)w_m + v_i\nabla_mw_m - \left(\nabla_jv_j\right)w_i - v_j\nabla_jw_i \end{aligned}$$

Restoring the vector notation, this becomes

$$\nabla\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{w}\cdot\nabla)\,\mathbf{v} + (\nabla\cdot\mathbf{w})\,\mathbf{v} - (\nabla\cdot\mathbf{v})\,\mathbf{w} - (\mathbf{v}\cdot\nabla)\,\mathbf{w}$$

If you doubt the advantages here, try to prove these identities by explicitly writing out all of the components!

2 Orthogonal transformations

2.1 Defining property

The squared length of a vector is given by taking the dot product of a vector with itself,

$$v^2 = \mathbf{v} \cdot \mathbf{v}$$

An *orthogonal transformation* is a linear transformation of a vector space that preserves lengths of vectors. This defining property may therefore be written as a linear transformation,

$$\mathbf{v}' = O\mathbf{v}$$

such that

$$\mathbf{v}' \cdot \mathbf{v}' = \mathbf{v} \cdot \mathbf{v}$$

Write this definition in terms of components using index notation. Setting the components of the transformation to $[O]_{ij} = O_{ij}$

$$v_i^{'} = O_{ij} v_j$$

we have

$$\begin{aligned} \mathbf{v}' \cdot \mathbf{v}' &= \mathbf{v} \cdot \mathbf{v} \\ v_i^{'} v_i^{'} &= v_i v_i \\ (O_{ij} v_j)(O_{ik} v_k) &= v_i v_i \end{aligned}$$

Notice that we change the name of the dummy indices so that we never have more than two indices repeated. Each term in the expression $(O_{ij} v_j)(O_{ik} v_k)$, for each value of i, j, k is just a real number, so we may rearrange the terms in any order we like. We may also write the dot product as a double sum, $v_i v_i = \delta_{jk} v_j v_k$. Then

$$\begin{aligned} v_i v_i &= O_{ij} v_j O_{ik} v_k \\ \delta_{jk} v_j v_k &= O_{ij} O_{ik} v_j v_k \end{aligned}$$

or

$$0 = (O_{ij} O_{ik} - \delta_{jk}) v_j v_k$$

If, instead of two copies of v_i, we had two arbitrary, independent vectors

$$0 = (O_{ij} O_{ik} - \delta_{jk}) u_j v_k$$

we could conclude that $O_{ij} O_{ik} - \delta_{jk}$ must vanish. However, since $v_i v_j = v_j v_i$ is symmetric, we need to derive a further general result. Suppose $A_{ij} = -A_{ji}$ are the components of any antisymmetric matrix, and $S_{ij} = S_{ji}$ are the components of an arbitrary symmetric matrix. Consider the trace of the product,

$$tr\,(AS) = tr\,(A_{ij} S_{jk}) = A_{ij} S_{ji}$$

We can evaluate this, just from the symmetries,

$$\begin{aligned} A_{ij} S_{ji} &= A_{ij} S_{ij} \\ &= -A_{ji} S_{ij} \end{aligned}$$

Since dummy indices can be named arbitrarily, we may rename i as j and j as i to write $A_{ji} S_{ij} = A_{ij} S_{ji}$. Then

$$\begin{aligned} A_{ij} S_{ji} &= -A_{ji} S_{ji} \\ 2A_{ij} S_{ji} &= 0 \\ A_{ij} S_{ji} &= 0 \end{aligned}$$

Therefore, the full contraction of a symmetric matrix with an antisymmetric matrix always vanishes.

Returning to our defining property of orthogonal transformations,

$$0 = (O_{ij}O_{ik} - \delta_{jk})\, v_j v_k$$

and recognizing that the matrix

$$M_{ij} = v_i v_j$$

is symmetric but otherwise arbitrary, we see that we can make no claim about the antisymmetric part of $(O_{ij}O_{ik} - \delta_{jk})$, since such a contraction vanishes identically in any case. What we can conclude is that the contraction of $(O_{ij}O_{ik} - \delta_{jk})$ with the arbitrary symmetric matrix $v_j v_k$ requires the vanishing of the *symmetric* part,

$$(O_{ij}O_{ik} - \delta_{jk}) + (O_{ik}O_{ij} - \delta_{kj}) = 0$$

This is enough, because

$$\begin{aligned} O_{ij}O_{ik} &= O_{ik}O_{ij} \\ \delta_{jk} &= \delta_{kj} \end{aligned}$$

and we just have two copies of the same thing,

$$\begin{aligned} O_{ij}O_{ik} - \delta_{jk} &= 0 \\ O_{ij}O_{ik} &= \delta_{jk} \end{aligned}$$

and since the components of O^t are just $[O^t]_{ij} = O_{ji}$, we may write this in matrix notation as

$$O^t O = 1$$

This is the defining property of an orthogonal transformation:

$$O^t = O^{-1} \tag{10}$$

2.2 All real, 3-dimensional representations of orthogonal transformations

While we have found one particular rotation matrix, Eq.(4), knowing the defining property Eq.(10) allows us to find an expression for any orthogonal transformation.

Notice that the transpose of the identity is the inverse of the identity, $\mathbf{1}^t = \mathbf{1} = \mathbf{1}^{-1}$, so we may consider linear transformations near the identity which also satisfy Eq.(10). Let

$$O = 1 + \varepsilon$$

where the components $[\varepsilon]_{ij} = \varepsilon_{ij}$ are all small, $|\varepsilon_{ij}| \ll 1$ for all i, j. Keeping only terms to first order in ε_{ij}, the transpose is just

$$O^t = 1 + \varepsilon^t$$

The inverse of a transformation near the identity must also be near the identity, so to find the right form for O^{-1} to first order, set

$$\begin{aligned} OO^{-1} &= (1+\varepsilon)(1+\delta) \\ &= 1 + \varepsilon + \delta - \varepsilon\delta \\ &\approx 1 + \varepsilon + \delta \end{aligned}$$

To have the inverse requires only $\delta = -\varepsilon$, correct to first order in ε. Now we impose Eq.(10),

$$\begin{aligned} O^t &= O^{-1} \\ 1+\varepsilon^t &= 1-\varepsilon \\ \varepsilon^t &= -\varepsilon \end{aligned}$$

so that the matrix ε must be antisymmetric. *Every infinitesimal orthogonal transformation differs from the identity by an antisymmetric matrix.*

Next, we write the most general antisymmetric 3×3 matrix as a linear combination of a convenient basis,

$$\begin{aligned} \varepsilon &= w_i J_i \\ &= w_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + w_2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + w_3 \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & w_2 & -w_3 \\ -w_2 & 0 & w_1 \\ w_3 & -w_1 & 0 \end{pmatrix} \end{aligned}$$

Notice that the components of the three matrices J_i are neatly summarized by

$$[J_i]_{jk} = \varepsilon_{ijk}$$

where ε_{ijk} is the totally antisymmetric Levi-Civita symbol. The matrices J_i are called the *generators* of the transformations.

Knowing the generators is enough to recover an arbitrary rotation. Starting with

$$\begin{aligned} O &= 1+\varepsilon \\ &= 1+\epsilon n_i J_i \end{aligned}$$

where ϵ is a small parameter and $\mathbf{n}$ any unit vector, we may apply O repeatedly, taking the limit

$$\begin{aligned} O(\theta) &= \lim_{n\longrightarrow\infty} O^n \\ &= \lim_{n\longrightarrow\infty} (1+\epsilon n_i J_i)^n \end{aligned}$$

where the limit is taken in such a way that we hold the product $n\epsilon$ constant,

$$n\epsilon = \theta$$

where θ is finite.

Quite generally, this limit gives rise to an exponential.

Theorem: For any matrix M,

$$\lim_{n\longrightarrow\infty} (1+\epsilon M)^n = \exp(\theta M)$$

where the limit is taken so that $n\epsilon = \theta$ remains finite.

Proof Noting that the calculation does not depend on the size or details of the matrix M, we use the binomial expansion, $(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^{n-k} b^k$ we have

$$\begin{aligned} O(\theta,\mathbf{n}) &= \lim_{n\longrightarrow\infty} O^n \\ &= \lim_{n\longrightarrow\infty} (1+\epsilon M)^n \\ &= \lim_{n\longrightarrow\infty} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (1)^{n-k} (\epsilon M)^k \\ &= \lim_{n\longrightarrow\infty} \sum_{k=0}^{n} \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \epsilon^k M^k \end{aligned}$$

Multiplying and dividing by n^k,

$$\begin{aligned} O(\theta,\mathbf{n}) &= \lim_{n\longrightarrow\infty}\sum_{k=0}^{n}\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\frac{(n\epsilon)^k}{n^k}M^k \\ & \lim_{n\longrightarrow\infty}\sum_{k=0}^{n}\frac{\frac{n}{n}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!}(\theta M)^k \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}(\theta M)^k \\ &\equiv \exp(\theta M) \end{aligned}$$

We define the exponential of a matrix by the power series for the exponential, which we can compute because we know how to find powers of the matrix.

To find the form of a general rotation, we apply the theorem with $M = n_i J_i$. We therefore need to find powers of $n_i J_i$. This turns out to be straightforward because of our choice for the three matrices J_i. Remembering that $n_i n_j \delta_{ij} = \mathbf{n}\cdot\mathbf{n} = \mathbf{1}$,

$$\begin{aligned} [n_i J_i]_{jk} &= n_i \varepsilon_{ijk} \\ \left[(n_i J_i)^2\right]_{mn} &= (n_i\varepsilon_{imk})(n_j\varepsilon_{jkn}) \\ &= n_i n_j \varepsilon_{imk}\varepsilon_{jkn} \\ &= -n_i n_j(\delta_{ij}\delta_{mn}-\delta_{in}\delta_{jm}) \\ &= -(n_i n_j\delta_{ij})\delta_{mn} + n_m n_n \\ &= n_m n_n - \delta_{mn} \\ \left[(n_i J_i)^3\right]_{mn} &= (n_m n_k - \delta_{mk}) n_i \varepsilon_{ikn} \\ &= n_m n_k n_i \varepsilon_{ikn} - \delta_{mk} n_i \varepsilon_{ikn} \\ &= -n_i\varepsilon_{imn} \\ &= -[n_i J_i]_{mn} \end{aligned}$$

The powers come back to $n_i J_i$ with only a sign change, so we can divide the series into even and odd powers,

$$\begin{aligned} \left[(n_i J_i)^{2l}\right]_{mn} &= (-1)^l (n_m n_n - \delta_{mn}) \quad \textit{for all } l > 0 \\ \left[(n_i J_i)^{2l+1}\right]_{mn} &= (-1)^l [n_i J_i]_{mn} \quad \textit{for all } l \geq 0 \end{aligned}$$

We can now compute the exponential explicitly:

$$\begin{aligned} [O(\theta,\hat{\mathbf{n}})]_{mn} &= [\exp(\theta n_i J_i)]_{mn} \\ &= \left[\sum_{k=0}^{\infty}\frac{1}{k!}(\theta n_i J_i)^k\right]_{mn} \\ &= \left[(n_i J_i)^0\right]_{mn} + \left[\sum_{l=1}^{\infty}\frac{\theta^{2l}}{(2l)!}(n_i J_i)^{2l}\right]_{mn} + \left[\sum_{l=0}^{\infty}\frac{\theta^{2l+1}}{(2l+1)!}(n_i J_i)^{2l+1}\right]_{mn} \\ &= \delta_{mn} + \sum_{l=1}^{\infty}\frac{(-1)^l}{(2l)!}(\delta_{mn} - n_m n_n) + \sum_{l=0}^{\infty}\frac{(-1)^l}{(2l+1)!}[n_i J_i]_{mn} \end{aligned}$$

This first sum is $(\cos\theta - 1)$ because the $l = 0$ term is missing from the first sum, while the second sum is $\sin\theta$, so a general rotation matrix takes the form

$$\begin{aligned} [O(\theta,\hat{\mathbf{n}})]_{mn} &= [\exp(\theta n_i J_i)]_{mn} \\ &= \delta_{mn} + (\cos\theta - 1)(\delta_{mn} - n_m n_n) + \sin\theta\, [n_i J_i]_{mn} \qquad (11) \end{aligned}$$

This is the matrix for a rotation through an angle θ around an axis in the direction of $\mathbf{n}$. To see this, let O act on an arbitrary vector $\mathbf{v}$, and write the result in normal vector notation,

$$\begin{aligned}
[O(\theta,\hat{\mathbf{n}})]_{mn} v_n &= (\delta_{mn} + (\cos\theta - 1)(\delta_{mn} - n_m n_n) + \sin\theta\,[n_i J_i]_{mn}) v_n \\
&= \delta_{mn} v_n + (\cos\theta - 1)(\delta_{mn} v_n - n_m n_n v_n) + \sin\theta\,[n_i J_i]_{mn} v_n \\
&= v_m + (\cos\theta - 1)(v_m - n_m (n_n v_n)) + \sin\theta \varepsilon_{imn} n_i v_n
\end{aligned}$$

Define the components of $\mathbf{v}$ parallel and perpendicular to the unit vector $\mathbf{n}$:

$$\begin{aligned}
\mathbf{v}_{\|} &= (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n} \\
\mathbf{v}_{\perp} &= \mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n}
\end{aligned}$$

In terms of these,

$$\begin{aligned}
O(\theta,\hat{\mathbf{n}})\,\mathbf{v} &= \mathbf{v} + (\mathbf{v} - (\mathbf{v}\cdot\mathbf{n})\,\mathbf{n})(\cos\theta - 1) - \sin\theta\,(\mathbf{n}\times\mathbf{v}) \\
&= \mathbf{v}_{\|} + \mathbf{v}_{\perp}\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v}) \\
&= \mathbf{v}_{\|} + \mathbf{v}_{\perp}\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{v}_{\perp})
\end{aligned}$$

since $\mathbf{n}\times\mathbf{v} = \mathbf{n}\times\left(\mathbf{v}_{\perp} + \mathbf{v}_{\|}\right) = \mathbf{n}\times\mathbf{v}_{\perp}$. This expresses the rotated vector in terms of three mutually perpendicular vectors, $\mathbf{v}_{\|}, \mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The direction $\mathbf{n}$ is the axis of the rotation. The part of $\mathbf{v}$ parallel to $\mathbf{n}$ is therefore unchanged. The rotation takes place in the plane perpendicular to $\mathbf{n}$, and this plane is spanned by $\mathbf{v}_{\perp}, (\mathbf{n}\times\mathbf{v})$. The rotation in this plane takes $\mathbf{v}_{\perp}$ into the linear combination $\mathbf{v}_{\perp}\cos\theta - (\mathbf{n}\times\mathbf{v})\sin\theta$, which is exactly what we expect for a rotation of $\mathbf{v}_{\perp}$ through an angle θ according to the form of Eq.(4). The rotation $O(\theta,\hat{\mathbf{n}})$ is therefore a rotation by θ around the axis $\hat{\mathbf{n}}$.

3 Unitary representations (optional)

3.1 The special unitary group in 2 dimensions

Every orthogonal group ($SO(n)$, (rotations in n real dimensions) may be written as special cases of rotations in a related complex space. For $SO(3)$, it turns out that unitary transformations of a complex, 2-dimensional space can be used to rotate real 3-vectors. We show that we can write a real, 3-dimensional vector as a complex, traceless hermitian matrix.

Consider a linear combinations of traceless, Hermitian 2×2 matrices,

$$\begin{aligned}
\mathbf{v} &= v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 \\
&= v_i\sigma_i
\end{aligned}$$

where σ_i are the three Pauli matrices,

$$\begin{aligned}
\sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\sigma_2 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned}$$

Prove, as an exercise, that these are the *only* traceless, Hermitian 2×2 matrics. This means that there is a one to one correspondence between 2-dimensional traceless Hermitian matrics and 3-dimensional real vectors,

$$\begin{aligned}
\mathbf{v} &= v_i\sigma_i \\
&= \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{pmatrix}
\end{aligned}$$

Notice that linear combinations of these matrices remain traceless and Hermitian. The squared length of the vector $\mathbf{v}$ may be written as the negative of the determinant,

$$\begin{aligned} \det \mathbf{v} &= \det \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{pmatrix} \\ &= \left[-v_3^2 - (v_1 + iv_2)(v_1 - iv_2)\right] \\ &= -\left(v_1^2 + v_2^2 + v_3^2\right) \\ &= -\mathbf{v}^2 \end{aligned}$$

These facts give us a second way to describe orthogonal transformations. For the usual real representation of vectors, we have already seen that a real, linear transformation satisfying $O^t = O^{-1}$ is an orthogonal transformation. For the matrix representation, we require a similarity transformation,

$$\mathbf{v}' = U\mathbf{v}U^{-1}$$

which preserves three properties:

1. Vanishing trace:
$$tr\left(\mathbf{v}'\right) = tr\left(U\mathbf{v}U^{-1}\right) = 0$$
We show below that $\mathbf{v}'$ will be traceless if $\mathbf{v}$ is provided the determinant of U is one, $\det(U) = 1$.

2. Hermiticity:
$$\begin{aligned} \mathbf{v}'^\dagger &= \left(U\mathbf{v}U^{-1}\right)^\dagger \\ &= U^{-1\dagger}\mathbf{v}^\dagger U^\dagger \\ &= U^{-1\dagger}\mathbf{v}U^\dagger \end{aligned}$$
For this to reproduce $\mathbf{v}'$, we need $U^{-1\dagger} = U$, or equivalently, $U^{-1} = U^\dagger$, the condition that U be unitary.

3. Determinant:
$$\begin{aligned} \det \mathbf{v}' &= \det\left(U\mathbf{v}U^{-1}\right) \\ &= \det(U) \cdot \det \mathbf{v} \cdot \det\left(U^{-1}\right) \\ &= \det \mathbf{v} \end{aligned}$$

These properties guarantee that $\mathbf{v}'$ has the form $v_i'\sigma_i$ for three real numbers v_i', if and only if $\mathbf{v}$ does. Furthermore, since the determinant is preserved, the vectors $\mathbf{v}'$ and $\mathbf{v}$ have the same length. Therefore, U produces an orthogonal transformation if and only if U is unitary, with determinant 1. The group of matrices with these two properties is called $SU(2)$, the group of special (i.e, $\det U = 1$) unitary $\left(U^\dagger = U^{-1}\right)$ transformations in 2-dimensions.

3.2 Rotations as SU(2)

3.2.1 Infinitesimal $SU(2)$ matrices

We now find the general form of an $SU(2)$ transformation. Starting from an infinitesimal rotation, $U = 1+\varepsilon$, we require

$$\begin{aligned} UU^\dagger &= 1 \\ (1+\varepsilon)(1+\varepsilon)^\dagger &= 1 \\ 1 + \varepsilon + \varepsilon^\dagger + O\left(\varepsilon^2\right) &= 1 \\ \varepsilon + \varepsilon^\dagger &= 0 \\ \varepsilon^\dagger &= -\varepsilon \end{aligned}$$

which means the generator must be anti-hermitian. Let $\varepsilon = ih$, where h is Hermitian, $h = h^{\dagger} = \begin{pmatrix} a+b & c-id \\ c+id & a-b \end{pmatrix}$. We also need the determinant of U to be 1. To first order, with

$$U = 1 + i \begin{pmatrix} a+b & c-id \\ c+id & a-b \end{pmatrix} = \begin{pmatrix} 1+ia+ib & ic+d \\ ic-d & 1+ia-ib \end{pmatrix}$$

and a, b, c, d all small, this gives

$$\begin{aligned} 1 &= \det U \\ &= \det \begin{pmatrix} 1+ia+ib & ic+d \\ ic-d & 1+ia-ib \end{pmatrix} \\ &= (1+ia+ib)(1+ia-ib) - (ic+d)(ic-d) \\ &= \left(1+ia-ib+\left(ia-a^2+ab\right)+\left(ib-ba+b^2\right)\right)+\left(c^2+d^2\right) \\ &= 1+ia-ib+ia+ib+\textit{second order} \end{aligned}$$

Dropping the second order terms and canceling the 1 on each side leaves

2ia=0

and therefore $a = 0$. This means that h is traceless, and U may be written as a linear combination of the Pauli matrices,

$$U = 1 + i\varepsilon \mathbf{n} \cdot \boldsymbol{\sigma}$$

with $\varepsilon \ll 1$ and $\mathbf{n}$ a unit vector.

3.2.2 Finite transformations

Finite $SU(2)$ transformations may be found by taking the limit of many infinitesimal transformations. Using our general result [2.2]

$$\begin{aligned} U(\varphi, \mathbf{n}) &= \lim_{k \to \infty} (1 + i\varepsilon \mathbf{n} \cdot \boldsymbol{\sigma})^k \\ &= \exp\left(\frac{i\varphi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}\right) \end{aligned}$$

where

$$\lim_{n \to \infty} (n\varepsilon) = \frac{\varphi}{2}$$

The exponential is defined as the power series,

$$\begin{aligned} U &= \exp\left(\frac{i\varphi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{i\varphi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}\right)^k \end{aligned}$$

so we need to compute powers of the Pauli matrices. For this it is helpful to have the product of any two Pauli matrices,

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{1} + i\varepsilon_{ijk} \sigma_k$$

which you are invited to prove. Then

$$(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = (\mathbf{n} \cdot \boldsymbol{\sigma})^2$$

$$
\begin{aligned}
&= (n_i\sigma_i)(n_j\sigma_j) \\
&= n_i n_j \sigma_i \sigma_j \\
&= n_i n_j (\delta_{ij}\mathbf{1} + i\varepsilon_{ijk}\sigma_k) \\
&= n_i n_j \delta_{ij}\mathbf{1} + i\varepsilon_{ijk} n_i n_j \sigma_k \\
&= (\mathbf{n}\cdot\mathbf{n})\,\mathbf{1} + i(\mathbf{n}\times\mathbf{n})\cdot\boldsymbol{\sigma} \\
&= \mathbf{1}
\end{aligned}
$$

Higher powers follow immediately. For all $k = 0, 1, 2, \ldots$

$$
\begin{aligned}
\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)^{2m+1} &= (-1)^k\, i\left(\frac{\varphi}{2}\right)^{2k+1}\mathbf{n}\cdot\boldsymbol{\sigma} \\
\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)^{2k} &= (-1)^k\left(\frac{\varphi}{2}\right)^{2k}\mathbf{1}
\end{aligned}
$$

The exponential becomes

$$
\begin{aligned}
U &= \sum_{k=0}^{\infty}\frac{1}{k!}(i\mathbf{a}\cdot\boldsymbol{\sigma})^k \\
&= \mathbf{1}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\left(\frac{\varphi}{2}\right)^{2k} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}\left(\frac{\varphi}{2}\right)^{2k+1} \\
&= \mathbf{1}\cos\frac{\varphi}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}
\end{aligned}
$$

There are important things to be gained from the $SU(2)$ representation of rotations. First, it is much easier to work with the Pauli matrices than it is with 3×3 matrices. Although the generators in the 2- and 3-dimensional cases are simple, the exponentials are not. The exponential of the J_i matrices is rather complicated, while the exponential of the Pauli matrices may again be expressed in terms of the Pauli matrices,

$$
e^{\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}} = \mathbf{1}\cos\frac{\varphi}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}
$$

and this substantially simplifies calculations.

3.2.3 Rotation of a real 3-vector using $SU(2)$

Now apply this to a 3-vector, written as

$$
X = \mathbf{x}\cdot\boldsymbol{\sigma}
$$

We have

$$
\begin{aligned}
X' &= \mathbf{x}'\cdot\boldsymbol{\sigma} \\
&= U(\mathbf{x}\cdot\boldsymbol{\sigma})U^\dagger \\
&= \left(\mathbf{1}\cos\frac{\varphi}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}\right)(\mathbf{x}\cdot\boldsymbol{\sigma})\left(\mathbf{1}\cos\frac{\varphi}{2} - i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}\right) \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} + i(\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} - i(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} \\
&\quad + (\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2} \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} + i n_i x_j[\sigma_i,\sigma_j]\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} + (\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2}
\end{aligned}
$$

Evaluating the products of Pauli matrices,

$$\begin{aligned}
in_i x_j [\sigma_i, \sigma_j] &= in_i x_j (2i\varepsilon_{ijk}\sigma_k) \\
&= -2(\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma} \\
(\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma}) &= (\mathbf{n}\cdot\boldsymbol{\sigma})\, x_i n_j (\delta_{ij}1 + i\varepsilon_{ijk}\sigma_k) \\
&= (\mathbf{n}\cdot\boldsymbol{\sigma})((\mathbf{x}\cdot\mathbf{n})1 + i(\mathbf{x}\times\mathbf{n})\cdot\boldsymbol{\sigma}) \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + i(\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\times\mathbf{n})\cdot\boldsymbol{\sigma} \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + in_i(\mathbf{x}\times\mathbf{n})_j\,\sigma_i\sigma_j \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + in_i(\mathbf{x}\times\mathbf{n})_j\,(\delta_{ij}1 + i\varepsilon_{ijk}\sigma_k) \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + i\mathbf{n}\cdot(\mathbf{x}\times\mathbf{n})\,1 - (\mathbf{n}\times(\mathbf{x}\times\mathbf{n}))\cdot\boldsymbol{\sigma} \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - (\mathbf{x}(\mathbf{n}\cdot\mathbf{n}) - \mathbf{n}(\mathbf{x}\cdot\mathbf{n}))\cdot\boldsymbol{\sigma} \\
&= 2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - \mathbf{x}\cdot\boldsymbol{\sigma}
\end{aligned}$$

Substituting,

$$\begin{aligned}
\mathbf{x}'\cdot\boldsymbol{\sigma} &= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} - 2(\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} + (2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - \mathbf{x}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2} \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})\left(\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} - \sin\frac{\varphi}{2}\sin\frac{\varphi}{2}\right) - (\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}2\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} + 2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2} \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\varphi - (\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}\sin\varphi + (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma})(1-\cos\varphi) \\
&= [\mathbf{x}\cos\varphi - \mathbf{n}\times\mathbf{x}\sin\varphi + (\mathbf{x}\cdot\mathbf{n})\mathbf{n}(1-\cos\varphi)]\cdot\boldsymbol{\sigma} \\
&= [(\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\mathbf{n})\cos\varphi - \mathbf{n}\times\mathbf{x}\sin\varphi + (\mathbf{x}\cdot\mathbf{n})\mathbf{n}]\cdot\boldsymbol{\sigma}
\end{aligned}$$

and equating coefficients,

$$\begin{aligned}
\mathbf{x}' &= (\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\mathbf{n})\cos\varphi - \mathbf{n}\times\mathbf{x}\sin\varphi + (\mathbf{x}\cdot\mathbf{n})\mathbf{n} \\
&= \mathbf{x}_\parallel + \mathbf{x}_\perp\cos\varphi - (\mathbf{n}\times\mathbf{x}_\perp)\sin\varphi
\end{aligned}$$

which is the same transformation as we derived from $SO(3)$.

3.2.4 Spinors

Because the transformation $\mathbf{x}' = U\mathbf{x}U^\dagger$ is quadratic in U, we find double angle formulas and the real rotation is through an angle φ, not $\frac{\varphi}{2}$, This means that as φ runs from 0 to 2π, the 3-dim angle only runs from 0 to π, and a complete cycle requires φ to climb to 4π. Notice that U and $-U$ give the same rotation of $\mathbf{x}$.

More importantly, there is a crucial physical insight. The transformations U act on our hermitian matrices by a similarity transformation, but they also act on some 2-dimensional complex vector space. Denote a vector in this space as $\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, so that the transformation of ψ is given by

$$\psi' = e^{\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}}\psi$$

This transformation preserves the hermitian norm of ψ, since

$$\begin{aligned}
\psi'^\dagger\psi' &= \left(\psi^\dagger U^\dagger\right)(U\psi) \\
&= \psi^\dagger\left(U^\dagger U\right)\psi \\
&= \psi^\dagger\psi
\end{aligned}$$

The complex vector ψ is called a *spinor*, with its first and second components being called "spin up" and "spin down". While spinors were not discovered physically until quantum mechanics, their existence is predictable classically from the properties of rotations. Also notice that as φ runs from 0 to 2π, ψ changes by only

$$\begin{aligned}\psi' &= e^{i\pi\mathbf{n}\cdot\boldsymbol{\sigma}}\psi \\ &= \left(\mathbf{1}\cos\pi + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\pi\right)\psi \\ &= -\psi\end{aligned}$$

so while a vector rotates by 2π, a spinor changes sign. A complete cycle of an $SU(2)$ transformation therefore requires φ to run through 4π.

3.3 Lorentz transformations

A simple generlization of $SU(2)$ allows us to describe the Lorentz group. With a general 4-vector given by

$$v^\alpha = \left(v^0, v^1, v^2, v^3\right) = \left(v^0, \mathbf{v}\right)$$

we define a Lorentz transformation as any transformation preserving the proper lengths of 4-vectors,

$$s^2 = -\left(v^0\right)^2 + \mathbf{v}^2$$

Consider the space of 2-dim hermitian matrices. A general complex matrix may we written as

$$A = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

If A is hermitian then we require

$$A = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = A^\dagger = \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ \bar{\gamma} & \bar{\delta} \end{pmatrix}$$

so that $\alpha = \bar{\alpha} = a$, $\delta = \bar{\delta} = b$ and $\beta = \bar{\gamma}$, where a, b are real and an overbar denotes complex conjugation. We therefore may write any 2-dimensional Hermitian matrix as

$$A = \begin{pmatrix} a & \bar{\beta} \\ \beta & b \end{pmatrix}$$

Introducing four real numbers, x, y, z, t, and setting $a = t + z, b = t - z$ and $\beta = x + iy$, A takes the form

$$\begin{aligned}A &= \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \\ &= t\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= t\mathbf{1} + \mathbf{x}\cdot\boldsymbol{\sigma}\end{aligned}$$

Hermitian matrices therefore form a real, 4-dimensional space. We choose the identity and the Pauli matrices as a basis for this 4-dim vector space.

Now consider the determinant of A,

$$\begin{aligned}\det A &= \det\begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \\ &= (t+z)(t-z) - (x+iy)(x-iy) \\ &= t^2 - z^2 - x^2 - y^2 \\ &= -s^2\end{aligned}$$

This is the proper length of a 4-vector in spacetime, which means that any transformation which preserves the hermiticity and determinant of A is a Lorentz transformation.

It is now easy to write the Lorentz transformations. The most general, linear transformation of a matrix is by similarity transformation, so we consider any transformation of the form

$$A' = LAL^{\dagger}$$

where we use $L^{\dagger}$ on the right so that the new matrix is hermitian whenever A is,

$$\begin{aligned}
(A')^{\dagger} &= \left(LAL^{\dagger}\right)^{\dagger} \\
&= L^{\dagger\dagger}A^{\dagger}L^{\dagger} \\
&= LAL^{\dagger} \\
&= A'
\end{aligned}$$

L also preserves the determinant provided

$$\begin{aligned}
1 &= \det A' \\
&= \det\left(LAL^{\dagger}\right) \\
&= \det L \det A \det L^{\dagger} \\
&= \det L \det L^{\dagger} \\
&= |\det L|^{2}
\end{aligned}$$

so that $\det L = \pm 1$. The positive determinant transformations preserve the direction of time and are called orthochronos, forming the special linear group in 2 complex dimensions, $SL(2, C)$. Since these transformations preserve $\tau^2 = t^2 - z^2 - x^2 - y^2$, they are Lorentz transformations.

4 Rigid Body Dynamics

So far the problems we have addressed make use of inertial frames of reference, i.e., those vector bases in which Newton's second law holds (even though we have used the principle of least action, which allows general coordinates). However, it is sometimes useful to use non-inertial frames, and particularly when a system is rotating. When we affix an orthonormal frame to the surface of Earth, for example, that frame rotates with Earth's motion and is therefore non-inertial. The effect of this is to add terms to the acceleration due to the acceleration of the reference frame. Typically, these terms can be brought to the force side of the equation, giving rise to the idea of fictitious forces – centrifugal force and the Coriolis force are examples.

Here we concern ourselves with rotating frames of reference.

4.1 Rotating frames of reference

Using our expression for a general rotation, it is fairly easy to include the effect of a rotating vector basis. Consider the change, $d\mathbf{b}$, of some physical quantity describing a rotating body. We write this in two different reference frames, one inertial and one rotating with the body. The difference between these will be the change due to the rotation,

$$(d\mathbf{b})_{inertial} = (d\mathbf{b})_{body} + (d\mathbf{b})_{rot}$$

Now consider an infinitesimal rotation. We found that the transformation matrix for an infinitesimal rotation must have the form

$$O(d\theta, \hat{\mathbf{n}}) = 1 - d\theta \hat{\mathbf{n}} \cdot \mathbf{J}$$

where

$$[J_i]_{jk} = \varepsilon_{ijk}$$

and we write $-d\theta$ so that the rotation is in the counterclockwise (positive) direction. To check the direction of rotation, we may choose $\hat{\mathbf{n}} = \hat{\mathbf{k}}$ and act on a unit vector $\hat{\mathbf{i}}$ in the x direction,

$$\begin{aligned}
\left[O\left(d\theta, \hat{\mathbf{n}}\right)\right]_{jk} i_k &= \left(\delta_{jk} - d\theta n_i \varepsilon_{ijk}\right) i_k \\
&= \left(i_j - d\theta \varepsilon_{3j1}\right) \\
&= (1,0,0) + d\theta\,(0,1,0)
\end{aligned}$$

since ε_{3j1} must have $j = 2$ to be nonzero, and $\varepsilon_{321} = -1$. The vector acquires a positive y-component, and has therefore rotated counterclockwise, as desired.

Suppose a vector at time t, $\mathbf{b}(t)$ is fixed in a body which rotates with angular velocity $\boldsymbol{\omega} = \frac{d\theta}{dt}\mathbf{n}$. Then after a time dt it will have rotated through an angle $d\theta = \omega dt$, so that at time $t + dt$ the vector is

$$\mathbf{b}(t + dt) = O\left(\omega t, \hat{\mathbf{n}}\right) \mathbf{b}(t)$$

In components,

$$\begin{aligned}
b_j(t + dt) &= \left(\delta_{jk} - \omega t n_i \varepsilon_{ijk}\right) b_k(t) \\
&= \delta_{jk} b_k(t) - \omega t \left(n_i \varepsilon_{ijk} b_k(t)\right) \\
&= b_j(t) - \omega dt \left(\varepsilon_{kij} b_k(t)\, n_i\right)
\end{aligned}$$

Therefore, returning to vector notation,

$$\mathbf{b}(t + dt) - \mathbf{b}(t) = -\omega dt \mathbf{b}(t) \times \mathbf{n}$$

Dividing by dt we get the rate of change due to the rotation,

$$\frac{d\mathbf{b}(t)}{dt} = \boldsymbol{\omega} \times \mathbf{b}(t)$$

If, instead of remaining fixed in the rotating system, $\mathbf{b}(t)$ moves relative to the rotating body, its rate of change is the sum of this change and the rate of change due to rotation,

$$\left(\frac{d\mathbf{b}}{dt}\right)_{inertial} = \left(\frac{d\mathbf{b}}{dt}\right)_{body} + \boldsymbol{\omega} \times \mathbf{b}(t)$$

and since $\mathbf{b}(t)$ is arbitrary, we can make the operator identification expressing the rate of change of any quantity relative to an inertial frame of reference in terms of its rate of change relative to the rotating body,

$$\left(\frac{d}{dt}\right)_{inertial} = \left(\frac{d}{dt}\right)_{body} + \boldsymbol{\omega} \times$$

4.2 Dynamics in a rotating frame of reference

Consider two frames of reference with a common origin – an inertial frame, and a rotating frame. Let $\mathbf{r}(t)$ be the position vector of a particle in the rotating frame of reference. Then the velocity of the particle in an inertial frame, $\mathbf{v}_{inertial}$, and the velocity in the rotating frame, $\mathbf{v}_{body}$, are related by

$$\begin{aligned}
\left(\frac{d\mathbf{r}}{dt}\right)_{inertial} &= \left(\frac{d\mathbf{r}}{dt}\right)_{body} + \boldsymbol{\omega} \times \mathbf{r} \\
\mathbf{v}_{inertial} &= \mathbf{v}_{body} + \boldsymbol{\omega} \times \mathbf{r}
\end{aligned}$$

To find the acceleration, we apply the operator again,

$$\mathbf{a}_{inertial} = \frac{d\mathbf{v}_{inertial}}{dt}$$

$$
\begin{aligned}
&= \left(\frac{d}{dt} + \boldsymbol{\omega}\times\right)(\mathbf{v}_{body} + \boldsymbol{\omega}\times\mathbf{r}) \\
&= \frac{d(\mathbf{v}_{body} + \boldsymbol{\omega}\times\mathbf{r})}{dt} + \boldsymbol{\omega}\times(\mathbf{v}_{body} + \boldsymbol{\omega}\times\mathbf{r}) \\
&= \left(\frac{d\mathbf{v}}{dt}\right)_{body} + \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} + \boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt} + \boldsymbol{\omega}\times\mathbf{v}_{body} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) \\
&= \left(\frac{d\mathbf{v}}{dt}\right)_{body} + \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} + 2\boldsymbol{\omega}\times\mathbf{v}_{body} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})
\end{aligned}
$$

The accelerations are therefore related by

$$
\mathbf{a}_{inertial} = \mathbf{a}_{body} + \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} + 2\boldsymbol{\omega}\times\mathbf{v}_{body} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})
$$

Since Newton's second law holds in the inertial frame, we have

$$
\mathbf{F} = m\mathbf{a}_{inertial}
$$

where $\mathbf{F}$ refers to any applied forces. Therefore, bringing the extra terms to the left, the acceleration in the rotating frame is given by

$$
\mathbf{F} - m\frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} - 2m\boldsymbol{\omega}\times\mathbf{v}_{body} - m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) = m\mathbf{a}_{body} \tag{12}
$$

This is the *Coriolis theorem*. We consider each term.

The first

$$
-m\frac{d\boldsymbol{\omega}}{dt}
$$

applies only if the rate of rotation is changing. The direction makes sense, because if the angular velocity is increasing, then $\frac{d\boldsymbol{\omega}}{dt}$ is in the direction of the rotation and the inertia of the particle will resist this change. The effective force is therefore in the opposite direction.

The second term

$$
-2m\boldsymbol{\omega}\times\mathbf{v}_{body}
$$

is called the *Coriolis force*. Notice that it is greatest if the velocity is perpendicular to the axis of rotation. This corresponds to motion which, for positive $\mathbf{v}_{body}$, moves the particle further from the axis of rotation. Since the velocity required to stay above a point on a rotating body increases with increasing distance from the axis, the particle will be moving too slow to keep up. It therefore appears that a force is acting in the direction opposite to the direction of rotation. For example, consider a particle at Earth's equator which is gaining altitude. Since Earth rotates from west to east, the rising particle will fall behind and therefore seem to accelerate from toward the west.

The final term

$$
-m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})
$$

is the familiar centrifugal force (that is, the centripetal acceleration treated as a force). For Earth's rotation, $\boldsymbol{\omega}\times\mathbf{r}$ is the direction of the velocity of a body rotating with Earth, and direction of the centrifugal force is therefore directly away from the axis of rotation. The effect is due to the tendency of the body to move in a straight line in the inertial frame, hence away from the axis. For a particle at the equator, the centrifugal force is directed radially outward, opposing the force of gravity. The net acceleration due to gravity and the centrifugal acceleration is therefore,

$$
\begin{aligned}
g_{eff} &= g - \omega^2 r \\
&= 9.8 - \left(7.29\times10^{-5}\right)^2\times 6.38\times10^6 \\
&= 9.8 - .0339 \\
&= g\,(1 - .035)
\end{aligned}
$$

so that the gravitational attraction is reduced by about 3.5%. Since the effect is absent near the poles, Earth is not a perfect sphere, but has an equatorial bulge.

5 Examples

5.1 Lagrange points

Consider the three body problem for gravitation, in which to much more massive bodies move about one another in a circular orbit, with the third, lighter body moving in their combined potential. Assume the motions of all three bodies lie in a single plane. The action for the full system is

$$S = \int_0^t \frac{1}{2}M_1\dot{\mathbf{X}}_1^2 + \frac{1}{2}M_2\dot{\mathbf{X}}_2^2 + \frac{1}{2}m\dot{\mathbf{x}}^2 + \frac{GM_1M_2}{|\mathbf{X}_1 - \mathbf{X}_2|} + \frac{GM_1m}{|\mathbf{X}_1 - \mathbf{x}|} + \frac{GM_2m}{|\mathbf{X}_2 - \mathbf{x}|}$$

We treat this as a perturbative problem, with $m \ll M_1, M_2$, so that terms of order $\frac{m}{M}$ are completely negligible – imagine a small rock or satellite in the neighborhood of the earth and the moon. Moving to center of mass coordinates, we set

$$\begin{aligned}
\mathbf{R} &= \frac{M_1\mathbf{X}_1 + M_2\mathbf{X}_2 + m\mathbf{x}}{M_1 + M_2 + m} \\
&\approx \frac{M_1\mathbf{X}_1 + M_2\mathbf{X}_2}{M} \\
\mathbf{r} &= \mathbf{X}_2 - \mathbf{X}_1
\end{aligned}$$

where $M = M_1 + M_2$. Inverting as we did for the two body problem,

$$\begin{aligned}
\mathbf{X}_2 &= \mathbf{R} + \frac{M_1}{M}\mathbf{r} \\
\mathbf{X}_1 &= \mathbf{R} - \frac{M_2}{M}\mathbf{r}
\end{aligned}$$

and the dominant terms of the action become

$$\frac{1}{2}M_1\dot{\mathbf{X}}_1^2 + \frac{1}{2}M_2\dot{\mathbf{X}}_2^2 + \frac{GM_1M_2}{|\mathbf{X}_1 - \mathbf{X}_2|} = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 + \frac{GM_1M_2}{|\mathbf{r}|}$$

Let our coordinate system be at rest at the center of mass, so the first term drops out. The action is now

$$\begin{aligned}
S &= \int_0^t \left(\frac{1}{2}\mu\dot{\mathbf{r}}^2 + \frac{1}{2}m\dot{\mathbf{x}}^2 + \frac{GM_1M_2}{|\mathbf{r}|} + \frac{GM_1m}{|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}|} + \frac{GM_2m}{|\frac{M_1}{M}\mathbf{r} - \mathbf{x}|}\right) \\
&= \int_0^t \left(\frac{1}{2}\mu\left(\dot{r} + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2\right) + \frac{1}{2}m\dot{\mathbf{x}}^2 + \frac{GM_1M_2}{|\mathbf{r}|} + \frac{GM_1m}{|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}|} + \frac{GM_2m}{|\frac{M_1}{M}\mathbf{r} - \mathbf{x}|}\right)
\end{aligned}$$

Let the common plane of the motion be specified by $\theta = \frac{\pi}{2}$. Then for a circular orbit at $r = r_0$,

$$S = \int_0^t \left(\frac{1}{2}\mu r^2\dot{\varphi}^2 + \frac{1}{2}m\left(\dot{x}^2 + x^2\dot{\varphi}_x^2\right) + \frac{GM_1M_2}{|\mathbf{r}|} + \frac{GM_1m}{|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}|} + \frac{GM_2m}{|\frac{M_1}{M}\mathbf{r} - \mathbf{x}|}\right)$$

The conserved energy is

$$E = \frac{1}{2}\mu r^2\dot{\varphi}^2 + \frac{1}{2}m\left(\dot{x}^2 + x^2\dot{\varphi}_x^2\right) - \frac{GM_1M_2}{|\mathbf{r}|} - \frac{GM_1m}{|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}|} - \frac{GM_2m}{|\frac{M_1}{M}\mathbf{r} - \mathbf{x}|}$$

For a circular orbit, we expect

$$\begin{aligned}\frac{GM_1M_2}{r_0^2} &= \mu r_0\dot{\varphi}^2 \\ \omega_0^2 = \dot{\varphi}^2 &= \frac{GM}{r_0^3}\end{aligned}$$

Move to a frame rotating with this angular velocity. Since we are looking for stationary points, $\dot{\varphi}_x^2$ will take this same value, ω_0^2, so the energy reduces to

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}\frac{GMm}{r_0^3}x^2 - \frac{GM_1m}{\left|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}\right|} - \frac{GM_2m}{\left|\frac{M_1}{M}\mathbf{r} - \mathbf{x}\right|}$$

and effective potential for the satellite is,

$$V_{eff} = \frac{GMm}{2r^3}x^2 - \frac{GM_1m}{\left|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}\right|} - \frac{GM_2m}{\left|\frac{M_1}{M}\mathbf{r} - \mathbf{x}\right|}$$

We need to expand the denominators. Define

$$\begin{aligned}L_1 &\equiv \left|-\frac{M_2}{M}\mathbf{r} - \mathbf{x}\right| \\ &= \sqrt{\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}rx\cos\varphi} \\ L_2 &= \sqrt{\frac{M_1^2}{M^2}r^2 + x^2 - \frac{2M_1}{M}rx\cos\varphi}\end{aligned}$$

remembering that $\mathbf{x}$ is the vector to m from the center of mass, while $\mathbf{r}$ is the separation of M_1 and M_2, with φ the angle between them. The effective potential is

$$V_{eff} = \frac{GMm}{2r^3}x^2 - \frac{GM_1m}{L_1} - \frac{GM_2m}{L_2}$$

We seek the extrema,

$$\begin{aligned}0 &= \boldsymbol{\nabla}V_{eff} \\ &= \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{\varphi}}\frac{1}{x}\frac{\partial}{\partial\varphi}\right)\left(\frac{GMm}{r^3}x^2 - \frac{GM_1m}{L_1} - \frac{GM_2m}{L_2}\right)\end{aligned}$$

so we will need

$$\begin{aligned}\boldsymbol{\nabla}\frac{1}{L_1} &= -\frac{1}{2L_1^3}\left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{\varphi}}\frac{1}{x}\frac{\partial}{\partial\varphi}\right)\left(\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}rx\cos\varphi\right) \\ &= -\frac{1}{2L_1^3}\left(\hat{\mathbf{x}}\left(2x + \frac{2M_2}{M}r\cos\varphi\right) + \hat{\boldsymbol{\varphi}}\left(-\frac{2M_2}{M}r\sin\varphi\right)\right) \\ &= -\frac{1}{L_1^3}\left(\hat{\mathbf{x}}\left(x + \frac{M_2}{M}r\cos\varphi\right) - \hat{\boldsymbol{\varphi}}\left(\frac{M_2}{M}r\sin\varphi\right)\right) \\ \boldsymbol{\nabla}\frac{1}{L_2} &= -\frac{1}{L_2^3}\left(\hat{\mathbf{x}}\left(x - \frac{M_1}{M}r\cos\varphi\right) - \hat{\boldsymbol{\varphi}}\left(\frac{M_1}{M}r\sin\varphi\right)\right)\end{aligned}$$

Therefore,

$$\begin{aligned}0 &= \boldsymbol{\nabla}V_{eff} \\ &= \hat{\mathbf{x}}\left(\frac{GMm}{r^3}x + \frac{GM_1m}{L_1^3}\left(x + \frac{M_2}{M}r\cos\varphi\right) + \frac{GM_2m}{L_2^3}\left(x - \frac{M_1}{M}r\cos\varphi\right)\right) \\ &\quad + \hat{\boldsymbol{\varphi}}\left(\frac{1}{L_2^3} - \frac{1}{L_1^3}\right)G\mu mr\sin\varphi\end{aligned}$$

There are two cases to consider, depending on whether $\sin\varphi = 0$ or not.

5.1.1 Case 1: $\varphi = 0$

If $\varphi = 0$ then only the $\hat{\mathbf{x}}$ term remains, and $L_{1,2}$ simplify,

$$\begin{aligned} L_1 &= \sqrt{\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}rx} \\ &= x + \frac{M_2}{M}r \\ L_2 &\equiv x - \frac{M_1}{M}r \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &= \frac{GMm}{r^3}x + \frac{GM_1m}{L_1^3}\left(x + \frac{M_2}{M}r\right) + \frac{GM_2m}{L_2^3}\left(x - \frac{M_1}{M}r\right) \\ &= Gm\left(\frac{M}{r^3}x + \frac{M_1}{L_1^2} - \frac{M_2}{L_2^2}\right) \\ 0 &= \frac{M}{r^3}xL_1^2L_2^2 + M_1L_2^2 - M_2L_1^2 \\ &= \frac{M}{r^3}x\left(\frac{M_1}{M}r - x\right)^2\left(\frac{M_2}{M}r + x\right)^2 + M_1\left(\frac{M_1}{M}r - x\right)^2 - M_2\left(\frac{M_2}{M}r + x\right)^2 \end{aligned}$$

This is a fifth order equation. Expanding the squares, the first term becomes

$$\begin{aligned} A &\equiv \frac{M}{r^3}x\left(\frac{M_1}{M}r - x\right)^2\left(\frac{M_2}{M}r + x\right)^2 \\ &= \frac{M}{r^3}x\left(\frac{M_1^2}{M^2}r^2 - 2\frac{M_1}{M}rx + x^2\right)\left(\frac{M_2^2}{M^2}r^2 + 2\frac{M_2}{M}rx + x^2\right) \\ &= \frac{M}{r^3}x\left(\frac{\mu^2}{M^2}r^4 + 2\frac{\mu M_1}{M^2}r^3x + \frac{M_1^2}{M^2}r^2x^2 - \frac{2\mu M_2}{M^2}r^3x - \frac{4\mu}{M}r^2x^2 - \frac{2M_1}{M}rx^3 + \frac{M_2^2}{M^2}r^2x^2 + 2\frac{M_2}{M}rx^3 + x^4\right) \\ &= \frac{M}{r^3}x\left(x^4 + \frac{2\left(M_2 - M_1\right)}{M}rx^3 + \frac{1}{M^2}\left(\left(M_1 - M_2\right)^2 - 2M_1M_2\right)r^2x^2 + \frac{2\mu\left(M_1 - M_2\right)}{M^2}r^3x + \frac{\mu^2}{M^2}r^4\right) \end{aligned}$$

The final pair of terms together become

$$\begin{aligned} B &\equiv M_1\left(\frac{M_1}{M}r - x\right)^2 - M_2\left(\frac{M_2}{M}r + x\right)^2 \\ &= M_1\left(\frac{M_1^2}{M^2}r^2 - 2\frac{M_1}{M}rx + x^2\right) - M_2\left(\frac{M_2^2}{M^2}r^2 + 2\frac{M_2}{M}rx + x^2\right) \\ &= \frac{M_1^3}{M^2}r^2 - 2\frac{M_1^2}{M}rx + M_1x^2 - \frac{M_2^3}{M^2}r^2 - 2\frac{M_2^2}{M}rx - M_2x^2 \\ &= \left(M_1 - M_2\right)x^2 + -2\frac{\left(M_2^2 + M_1^2\right)}{M}rx + \frac{M_1^3 - M_2^3}{M^2}r^2 \end{aligned}$$

Putting these together,

$$\begin{aligned} 0 &= A + B \\ &= \frac{M}{r^3}xx^5 + 2\left(M_2 - M_1\right)\frac{1}{r^2}x^4 + \frac{1}{M}\left(\left(M_1 - M_2\right)^2 - 2M_1M_2\right)\frac{1}{r}x^3 + \frac{2\mu\left(M_1 - M_2\right)}{M}x^2 + \frac{\mu^2}{M}rx \\ &\quad + \left(M_1 - M_2\right)x^2 + -2\frac{\left(M_2^2 + M_1^2\right)}{M}rx + \frac{M_1^3 - M_2^3}{M^2}r^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{M}{r^3}x^5 + 2\left(M_2 - M_1\right)\frac{1}{r^2}x^4 + \frac{1}{M}\left(\left(M_1 - M_2\right)^2 - 2M_1M_2\right)\frac{1}{r}x^3 + \frac{2\mu\left(M_1 - M_2\right)}{M}x^2 \\
&\quad + \left(M_1 - M_2\right)x^2 + \frac{\mu^2}{M}rx - 2\frac{\left(M_2^2 + M_1^2\right)}{M}rx + \frac{M_1^3 - M_2^3}{M^2}r^2 \\
0 &= M^2\left(\frac{x}{r}\right)^5 + 2M\left(M_2 - M_1\right)\left(\frac{x}{r}\right)^4 + \left(\left(M_1 - M_2\right)^2 - 2M_1M_2\right)\left(\frac{x}{r}\right)^3 \\
&\quad + \left(3\mu\left(M_1 - M_2\right) + \frac{1}{M}\left(M_1^3 - M_2^3\right)\right)\left(\frac{x}{r}\right)^2 + \left(\mu^2 - 2\left(M_2^2 + M_1^2\right)\right)\left(\frac{x}{r}\right) + \frac{M_1^3 - M_2^3}{M}
\end{aligned}
$$

We consider the much simpler case when $M_1 = M_2$, so that $\mu = \frac{M_1M_2}{M} = \frac{M}{4}$. Then the condition reduces to

$$
\begin{aligned}
0 &= M^2\left(\frac{x}{r}\right)^5 - \frac{1}{2}M^2\left(\frac{x}{r}\right)^3 - \frac{15}{16}M^2\left(\frac{x}{r}\right) \\
0 &= \left(\left(\frac{x}{r}\right)^4 - \frac{1}{2}\left(\frac{x}{r}\right)^2 - \frac{15}{16}\right)\left(\frac{x}{r}\right)
\end{aligned}
$$

and we can solve exactly. The first term vanishes if

$$
\begin{aligned}
\left(\frac{x}{r}\right)^2 &= \frac{1}{2}\left(\frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{15}{4}}\right) \\
&= \frac{1}{4}\left(1 \pm 4\right)
\end{aligned}
$$

We must have the positive root, so the full equation factors as

$$
0 = \left(\frac{x}{r}\right)\left(\frac{x}{r} + \frac{5}{4}\right)\left(\frac{x}{r} - \frac{5}{4}\right)
$$

5.1.2 Case 2: $\sin\varphi \neq 0$

In this case, we first solve the $\hat{\boldsymbol{\varphi}}$ equation,

$$
0 = \hat{\boldsymbol{\varphi}}\left(\frac{1}{L_2^3} - \frac{1}{L_1^3}\right)G\mu mr\sin\varphi
$$

If the sine does not vanish then we must have

$$
\begin{aligned}
L_1 &= L_2 \\
\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}rx\cos\varphi &= \frac{M_1^2}{M^2}r^2 + x^2 - \frac{2M_1}{M}rx\cos\varphi \\
\frac{2M_1}{M}rx\cos\varphi + \frac{2M_2}{M}rx\cos\varphi &= \frac{M_1^2}{M^2}r^2 - \frac{M_2^2}{M^2}r^2 \\
\frac{2\left(M_1 + M_2\right)}{M}x\cos\varphi &= \frac{M_1^2 - M_2^2}{M^2}r \\
x\cos\varphi &= \frac{M_1 - M_2}{2M}r
\end{aligned}
$$

Replacing $x\cos\varphi$ with this expression, $L = L_1 = L_2$, with

$$
\begin{aligned}
L &= \sqrt{\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}rx\cos\varphi} \\
&= \sqrt{\frac{M_2^2}{M^2}r^2 + x^2 + \frac{2M_2}{M}\frac{M_1 - M_2}{2M}r^2} \\
&= \sqrt{\frac{\mu}{M}r^2 + x^2}
\end{aligned}
$$

Multiply the $\hat{\mathbf{x}}$ equation by x and substitute for L and $x\cos\varphi$,

$$\begin{aligned}
0 &= \frac{M}{r^3}x^2 + \frac{M_1}{L_1^3}\left(x^2 + \frac{M_2}{M}rx\cos\varphi\right) + \frac{M_2}{L_2^3}\left(x^2 - \frac{M_1}{M}rx\cos\varphi\right) \\
&= \frac{M}{r^3}x^2 + \frac{1}{L^3}\left(M_1x^2 + M_2x^2 + \frac{M_1 - M_2}{2M}\mu r^2 - \frac{M_1 - M_2}{2M}\mu r^2\right) \\
&= \frac{M}{r^3}x^2 + \frac{1}{L^3}Mx^2
\end{aligned}$$

so that

$$\begin{aligned}
L^3x^2 &= -r^3x^2 \\
L^6 &= r^6 \\
\left(\frac{\mu}{M}r^2 + x^2\right)^3 &= r^6 \\
\left(\frac{\mu}{M} + \left(\frac{x}{r}\right)^2\right)^3 &= 1 \\
\frac{\mu}{M} + \left(\frac{x}{r}\right)^2 &= 1 \\
\frac{x}{r} &= \pm\sqrt{1 - \frac{\mu}{M}}
\end{aligned}$$

and we have two Lagrange points with nonzero φ.

5.2 Foucault Pendulum

We apply the Coriolis theorem, Eq.(12), to the motion of a pendulum at latitude α.

$$\mathbf{F} - m\frac{d\boldsymbol{\omega}}{dt} \times \mathbf{r} - 2m\boldsymbol{\omega} \times \mathbf{v}_{body} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = m\mathbf{a}_{body}$$

The rotating frame here is Earth's daily rotation.

It is important to have an idea of the relative magnitudes of the various terms. The external force $\mathbf{F}$ is due to gravity and tension, both of which are proportional to mg. The second term vanishes because the angular velocity of Earth is extremely close to constant. For the Coriolis and centripetal terms, we need the magnitude of the angular velocity of Earth, which is

$$\omega = \frac{2\pi}{1\,day} \times \frac{1\,day}{24 \times 3600\,sec} = 7.27 \times 10^{-5}Hz$$

The radius of Earth is about $6.37 \times 10^6\,m$ while the velocity of a $10m$ pendulum is about $v = \sqrt{gL} = 9.90\,\frac{m}{s}$. Therefore, the acceleration due to the Coriolis term is of order

$$v\omega \sim 9.90\,\frac{m}{s} \times 7.27 \times 10^{-5}Hz = 7.20 \times 10^{-4}\frac{m}{s^2}$$

and the centripetal acceleration is of order

$$\omega^2 R = .0337\,\frac{m}{s^2}$$

Therefore,

$$g \gg \omega^2 R \gg \omega v$$

5.2.1 Modification to gravity

Because the centripetal acceleration does not depend on the motion of the pendulum, we simply regard it as part of the acceleration of gravity. In fact, it is this combined force that determines what we regard as "downward", rather than the direction toward the center of mass of Earth. Therefore, if $\mathbf{g} = -\frac{GM}{R^2}\hat{\boldsymbol{r}}$ is the true *gravitational* acceleration, the acceleration we measure at the surface is

$$\mathbf{g}' = \mathbf{g} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$$

Choose both Cartesian $\left(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\right)$ and spherical $\left(\hat{\boldsymbol{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}\right)$ coordinates at rest in the Earth, with $r = R$ the radius of Earth, $\theta = \frac{\pi}{2} - \alpha$ the angle from the north pole, where α is the latitude. These are related by

$$\begin{aligned}
\hat{\mathbf{r}} &= \hat{\mathbf{i}} \sin\theta\cos\varphi + \hat{\mathbf{j}}\sin\theta\sin\varphi + \hat{\mathbf{k}}\cos\theta \\
\hat{\boldsymbol{\theta}} &= \hat{\mathbf{i}}\cos\theta\cos\varphi + \hat{\mathbf{j}}\cos\theta\sin\varphi - \hat{\mathbf{k}}\sin\theta \\
\hat{\boldsymbol{\varphi}} &= -\hat{\mathbf{i}}\sin\theta\sin\varphi + \hat{\mathbf{j}}\sin\theta\cos\varphi
\end{aligned}$$

The problem is symmetric in φ, so we rotate the coordinates so that the pendulum is at $\varphi = 0$. Then the unit normal to Earth's surface in the Cartesian basis is

$$\begin{aligned}
\hat{\mathbf{r}} &= \hat{\mathbf{i}}\sin\theta + \hat{\mathbf{k}}\cos\theta \\
&= \hat{\mathbf{i}}\cos\alpha + \hat{\mathbf{k}}\sin\alpha \\
\hat{\boldsymbol{\theta}} &= \hat{\mathbf{i}}\cos\theta - \hat{\mathbf{k}}\sin\theta \\
&= \hat{\mathbf{i}}\sin\alpha - \hat{\mathbf{k}}\cos\alpha
\end{aligned}$$

With the angular momentum given by $\boldsymbol{\omega} = \omega\hat{\mathbf{k}}$, the centripetal acceleration is

$$\begin{aligned}
\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) &= \omega\hat{\mathbf{k}} \times \left(\omega\hat{\mathbf{k}} \times R\hat{\mathbf{r}}\right) \\
&= \omega\hat{\mathbf{k}} \times \left(\omega\hat{\mathbf{k}} \times R\left(\hat{\mathbf{i}}\cos\alpha + \hat{\mathbf{k}}\sin\alpha\right)\right) \\
&= \omega\hat{\mathbf{k}} \times \left(\omega R\cos\alpha\hat{\mathbf{j}}\right) \\
&= -\omega^2 R\cos\alpha\hat{\mathbf{i}}
\end{aligned}$$

or equivalently,

$$\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = -\omega^2 R\cos\alpha\left(\hat{\mathbf{r}}\cos\alpha + \hat{\boldsymbol{\theta}}\sin\alpha\right)$$

Therefore, the actual downward direction is slightly away from the direction to the center of mass. This makes a fractional correction of about $\frac{\omega^2 R}{g} = \frac{.0337}{9.8} = .00344$ to the acceleration of gravity. The pendulum, when stationary, will hang in the direction of $\mathbf{g}'$ and have magnitude

$$\begin{aligned}
g' &= \sqrt{\left(g + \omega^2 R\cos^2\alpha\right)^2 + \omega^4 R^2\cos^2\alpha\sin^2\alpha} \\
&= \sqrt{g^2 + 2g\omega^2 R\cos^2\alpha + \omega^4 R^2\cos^2\alpha} \\
&\approx \sqrt{g^2 + 2g\omega^2 R\cos^2\alpha}
\end{aligned}$$

since $\omega^4 R^2$ is only of order .0009.

5.2.2 The motion of the pendulum

We now change to Cartesian coordinates centered on the pendulum. Let $\hat{\mathbf{t}}$ be a unit vector opposite the local "down", that is, $\hat{\mathbf{t}} \equiv \frac{\mathbf{g}'}{g'}$, and let $\hat{\mathbf{u}}, \hat{\mathbf{w}}$ complete an orthonormal basis, pointing to the East and North, respectively. The equation of motion is now

$$\mathbf{T} - m\mathbf{g}' - 2m\boldsymbol{\omega} \times \mathbf{v} = m\ddot{\mathbf{x}}$$

where $\mathbf{T}$ is the tension and $\ddot{\mathbf{x}}$ lies in the uw plane.

Let the angle β of oscillation of the pendulum be very small and the supporting cable of very long length L so that the arc of the pendulum is sufficiently close to horizontal motion. To be precise about the position, when the pendulum is displaced through an arc length βL its gain in height is $L - L\cos\beta \approx \frac{1}{2}L\beta^2$, so the position of the pendulum bob is

$$\begin{aligned}\mathbf{x} &= \frac{1}{2}L\beta^2\hat{\mathbf{t}} + L\sin\beta\left(\hat{\mathbf{u}}\cos\varphi + \hat{\mathbf{w}}\sin\varphi\right) \\ &\approx L\beta\left(\hat{\mathbf{u}}\cos\varphi + \hat{\mathbf{w}}\sin\varphi\right) \qquad (13)\end{aligned}$$

where φ is the angle in the uw plane measured from the u-axis. The tension is

$$\mathbf{T} = T\cos\beta\hat{\mathbf{t}} + T\sin\beta\hat{\mathbf{x}} = T\hat{\mathbf{T}}$$

where $\hat{\mathbf{x}}$ is the direction of the displacement from equilibrium given by Eq.(13).

Decompose the equation of motion in directions parallel and perpendicular to $\hat{\mathbf{T}}$. There is no acceleration in the along $\mathbf{T}$ so,

$$\begin{aligned}\hat{\mathbf{T}}\cdot\left(\mathbf{T} - m\mathbf{g}' - 2m\boldsymbol{\omega}\times\mathbf{v}\right) &= m\hat{\mathbf{T}}\cdot\ddot{\mathbf{x}} \\ T - mg'\cos\beta - 2m\hat{\mathbf{T}}\cdot\left(\boldsymbol{\omega}\times\mathbf{v}\right) &= 0\end{aligned}$$

and since ωv is of order 10^{-4} we may drop it and write

$$T = mg'\cos\beta$$

Noticing that to first order in β,

$$\begin{aligned}\hat{\mathbf{x}}\cdot\hat{\mathbf{T}} &= \left(\cos\beta\hat{\mathbf{t}} + \sin\beta\hat{\mathbf{x}}\right)\cdot\left(\frac{1}{2}L\beta^2\hat{\mathbf{t}} + L\beta\hat{\mathbf{x}}\right) \\ &= \frac{1}{2}L\beta^2 + L\beta^2 \\ &\approx 0\end{aligned}$$

so the remaining equation is in the horizontal plane.

Writing out the remaining equation in the uw plane:

$$-mg'\sin\beta\hat{\mathbf{x}} - 2m\boldsymbol{\omega}\times\mathbf{v} = m\ddot{\mathbf{x}}$$

we may write $g'\sin\beta\hat{\mathbf{x}} \approx g'\beta\hat{\mathbf{x}} = \frac{g'}{L}L\beta\hat{\mathbf{x}}$. Then defining $\omega_0 \equiv \sqrt{\frac{g'}{L}}$ and recognizing βL as the magnitude of the displacement, the first term becomes

$$-mg'\sin\beta\hat{\mathbf{x}} = -m\omega_0^2\mathbf{x}$$

To find the Coriolis force, we write the angular velocity of Earth in the new basis,

$$\begin{aligned}\boldsymbol{\omega} &= \hat{\mathbf{t}}\omega\cos\theta + \hat{\mathbf{w}}\omega\sin\theta \\ &= \hat{\mathbf{t}}\omega\sin\alpha + \hat{\mathbf{w}}\omega\cos\alpha\end{aligned}$$

Then

$$\begin{aligned}\boldsymbol{\omega}\times v\hat{\mathbf{x}} &= \left(\hat{\mathbf{t}}\omega\sin\alpha + \hat{\mathbf{w}}\omega\cos\alpha\right)\times\left(\dot{u}\hat{\mathbf{u}} + \dot{v}\hat{\mathbf{w}}\right) \\ &= \left(\dot{u}\hat{\mathbf{w}} - \dot{v}\hat{\mathbf{u}}\right)\omega\sin\alpha - \hat{\mathbf{t}}\dot{u}\omega\cos\alpha\end{aligned}$$

The $\hat{\mathbf{t}}$ component is a tiny contribution to the tension in the cable, and the tension always adjusts to balance any force applied in its own direction, so we may drop this. Expandint the resulting equations in the $(\hat{\mathbf{w}}, \hat{\mathbf{u}})$ basis and cancelling m, we have

$$-\omega_0^2 (u\hat{\mathbf{u}} + w\hat{\mathbf{w}}) - 2(\dot{u}\hat{\mathbf{w}} - \dot{w}\hat{\mathbf{u}})\,\omega\sin\alpha = (\ddot{u}\hat{\mathbf{u}} + \ddot{w}\hat{\mathbf{w}})$$

so we have two coupled equations of motion,

$$\begin{aligned} \ddot{u} - 2\dot{w}\omega\sin\alpha + \omega_0^2 u &= 0 \\ \ddot{w} + 2\dot{u}\omega\sin\alpha + \omega_0^2 w &= 0 \end{aligned}$$

We may write these as a single complex equation. Let $z \equiv u + iw$. Then multiplying the second by i and adding,

$$\ddot{z} + 2i\dot{z}\omega\sin\alpha + \omega_0^2 z = 0$$

To decouple the two directions, define the precession velocity, $\omega_p \equiv \omega\sin\alpha$, and let

$$z = se^{-i\omega_p t}$$

This means that z is just s together with a slow rotation in the plane. Finding derivatives,

$$\begin{aligned} z &= se^{-i\omega_p t} \\ \dot{z} &= (\dot{s} - i\omega_p s)\,e^{-i\omega_p t} \\ \ddot{z} &= (\ddot{s} - i\omega_p \dot{s})\,e^{-i\omega_p t} - i\omega_p (\dot{s} - i\omega_p s)\,e^{-i\omega_p t} \\ &= \left(\ddot{s} - 2i\omega_p \dot{s} - \omega_p^2 s\right) e^{-i\omega_p t} \end{aligned}$$

and substituting,

$$\begin{aligned} 0 &= \ddot{z} + 2i\dot{z}\omega\sin\alpha + \omega_0^2 z \\ &= \left(\left(\ddot{s} - 2i\omega_p\dot{s} - \omega_p^2 s\right) + 2i(\dot{s} - i\omega_p s)\,\omega_p + \omega_0^2 s\right) e^{i\omega_p t} \\ 0 &= \ddot{s} - 2i\omega_p\dot{s} - \omega_p^2 s + 2i\omega_p\dot{s} + 2\omega_p^2 s + \omega_0^2 s \\ &= \ddot{s} + \left(\omega_0^2 + \omega_p^2\right) s \end{aligned}$$

This is simple harmonic with

$$\begin{aligned} s &= s_1 e^{i\omega t} + s_2 e^{-i\omega t} \\ z &= \left(s_1 e^{i\omega t} + s_2 e^{-i\omega t}\right) e^{-i\omega_p t} \end{aligned}$$

where $\omega = \sqrt{\omega_0^2 + \omega_p^2} = \sqrt{\frac{g'}{L} + \omega^2\sin^2\alpha}$. For $L = 10\,m$ and $g' \approx 9.8\frac{m}{s^2}$, $\omega_0 \approx 1$ while $\omega \approx 7\times 10^{-5}$, so ω_P^2 is neglibible compared to ω_0^2.

Let $s(0) = s_0$ and $\dot{s}(0) = 0$, both real, at $t = 0$. Then

$$\begin{aligned} s &= r_1 e^{-i\varphi_1} e^{i\omega t} + r_2 e^{-i\varphi_2} e^{-i\omega t} \\ &= r_1 e^{i(\omega t - \varphi_1)} + r_2 e^{-i(\omega t + \varphi_2)} \\ s_0 &= r_1 e^{-i\varphi_1} + r_2 e^{-i\varphi_2} \\ &= r_1(\cos\varphi_1 - i\sin\varphi_1) + r_2(\cos\varphi_2 + i\sin\varphi_2) \\ s_0 &= r_1\cos\varphi_1 + r_2\cos\varphi_2 \\ 0 &= i(r_1\sin\varphi_1 - r_2\sin\varphi_2) \\ \dot{s}_0 &= i\omega r_1 e^{-i\varphi_1} - i\omega r_2 e^{-i\varphi_2} \\ &= i\omega r_1(\cos\varphi_1 - i\sin\varphi_1) - i\omega r_2(\cos\varphi_2 + i\sin\varphi_2) \\ \dot{s}_0 &= \omega r_1\sin\varphi_1 + \omega r_2\sin\varphi_2 \\ 0 &= i\omega(r_1\cos\varphi_1 - r_2\cos\varphi_2) \end{aligned}$$

Therefore,

$$\begin{aligned}
s_0 &= r_1 \cos\varphi_1 + r_2 \cos\varphi_2 \\
0 &= r_1 \sin\varphi_1 - r_2 \sin\varphi_2 \\
\dot{s}_0 &= r_1 \sin\varphi_1 + r_2 \sin\varphi_2 \\
0 &= r_1 \cos\varphi_1 - r_2 \cos\varphi_2
\end{aligned}$$

Solving

$$\begin{aligned}
r_1 \cos\varphi_1 &= r_2 \cos\varphi_2 \\
s_0 &= 2r_1 \cos\varphi_1 \\
r_1 \sin\varphi_1 &= r_2 \sin\varphi_2 \\
\dot{s}_0 &= 2r_1 \sin\varphi_1
\end{aligned}$$

If we take the initial velocity to be zero then $\varphi_1 = 0$ and we have

$$\begin{aligned}
r_1 &= r_2 \cos\varphi_2 \\
s_0 &= 2r_1 \\
0 &= r_2 \sin\varphi_2 \\
\dot{s}_0 &= 0
\end{aligned}$$

Then $\varphi_2 = 0$ and

$$\begin{aligned}
r_1 &= r_2 \\
s_0 &= 2r_1 \\
\dot{s}_0 &= 0
\end{aligned}$$

Finally

$$\begin{aligned}
s &= \frac{s_0}{2}\left(e^{i\omega t} + e^{-i\omega t}\right) \\
&= s_0 \cos\omega t
\end{aligned}$$

and the pendulum simply oscillates along the original displacement. To this, $e^{i\omega_p t}$ adds a very slow precession,

$$\begin{aligned}
z &= u + iv \\
&= (\cos\omega_p t - i\sin\omega_p t)\, s_0 \cos\omega t
\end{aligned}$$

so that

$$\begin{aligned}
u &= s_0 \cos\omega t \cos\omega_p t \\
v &= -s_0 \cos\omega t \sin\omega_p t
\end{aligned}$$

The basic $\sqrt{\frac{g}{L}}$ oscillation is relatively fast, so seeing it, we separate the motion into two parts – oscillation and precession – instead of a single looping motion.

6 Moment of Inertia

6.1 Total torque and total angular momentum

Fix an arbitrary inertial frame of reference, and consider a rigid body. Consider the total torque on the body. The torque on the i^{th} particle due to internal forces will be

$$\boldsymbol{\tau}_i = \sum_{j=1}^{N} \mathbf{r}_i \times \mathbf{F}_{ji}$$

where $\mathbf{F}_{ji}$ is the force exerted by the j^{th} particle on the i^{th} particle, with $\mathbf{F}_{ii} = 0$. The total torque on the body is therefore the double sum,

$$\begin{aligned}
\boldsymbol{\tau}_{internal} &= \sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{r}_i \times \mathbf{F}_{ji} \\
&= \frac{1}{2}\sum_{i<j}^{N}\sum_{j=1}^{N}\left(\mathbf{r}_i \times \mathbf{F}_{ji} + \mathbf{r}_j \times \mathbf{F}_{ij}\right) \\
&= \frac{1}{2}\sum_{i<1}^{N}\sum_{j=1}^{N}\left(\mathbf{r}_i - \mathbf{r}_j\right) \times \mathbf{F}_{ji}
\end{aligned}$$

where we use Newton's third law in the last step. However, we assume that the forces between particles within the rigid body are along the line joining the two particles, so we have

$$\mathbf{F}_{ji} = F_{ji}\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|}$$

and all the cross products vanish. The total internal torque is thus zero:

$$\boldsymbol{\tau}_{internal} = 0$$

Therefore, we consider only external forces acting on the body when we compute the torque.

For a macroscopic rigid body it is easier to work in the continuum limit. Let the density at each point of the body be $\rho(\mathbf{r})$ (for a discrete collection of masses, we may let ρ be a sum of Dirac delta functions and recover the discrete picture). The contribution to the total torque of an external force $d\mathbf{F}(\mathbf{r})$ acting at position $\mathbf{r}$ of the body is

$$d\boldsymbol{\tau} = \mathbf{r} \times d\mathbf{F}(\mathbf{r})$$

and the total follows by integrating this. For a small increment of mass, $dm = \rho(\mathbf{r})\, d^3x$, the force using Newton's second law is $d\mathbf{F}(\mathbf{r}) = dm\,\frac{d\mathbf{v}}{dt} = \frac{d\mathbf{v}}{dt}\rho(\mathbf{r})\, d^3x$. Integrating the resulting torque, we have

$$\begin{aligned}
\boldsymbol{\tau} &= \int \rho(\mathbf{r})\left(\mathbf{r} \times \frac{d\mathbf{v}}{dt}\right) d^3x \\
&= \int \rho(\mathbf{r})\left[\frac{d}{dt}(\mathbf{r} \times \mathbf{v}) - \left(\frac{d\mathbf{r}}{dt} \times \mathbf{v}\right)\right] d^3x
\end{aligned}$$

Since $\frac{d\mathbf{r}}{dt} \times \mathbf{v} = \mathbf{v} \times \mathbf{v} = 0$, and the density is independent of time,

$$\boldsymbol{\tau} = \frac{d}{dt}\int \rho(\mathbf{r})(\mathbf{r} \times \mathbf{v})\, d^3x$$

Compare this to the total angular momentum. The angular momentum $d\mathbf{L}$ for a small mass element, $dm = \rho(\mathbf{r})\, d^3x$, is $d\mathbf{L} = \rho(\mathbf{r})(\mathbf{r} \times \mathbf{v})\, d^3x$. Integrating over the whole body gives the total angular momentum,

$$\mathbf{L} = \int \rho(\mathbf{r})(\mathbf{r} \times \mathbf{v})\, d^3x$$

so the total external torque equals the rate of change of total angular momentum,

$$\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} \tag{14}$$

6.2 Moment of intertia

As it stands, the angular momentum of a rigid body depends on both the velocity of its parts and their arrangement. To describe the motion of a given rigid body, it is useful to separate properties of the body itself from the dynamical properties dependent only on its state of motion. We accomplish this as follows.

Suppose the rigid body rotates with angular velocity $\boldsymbol{\omega}$. Then the velocity of any point in the body is $\boldsymbol{\omega} \times \mathbf{r}$, so the angular momentum is

$$\begin{aligned} \mathbf{L} &= \int \rho(\mathbf{r})\left(\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})\right) d^3x \\ &= \int \rho(\mathbf{r})\left(\boldsymbol{\omega} r^2 - \mathbf{r}(\mathbf{r} \cdot \boldsymbol{\omega})\right) d^3x \end{aligned}$$

We would like to extract $\boldsymbol{\omega}$ from the integral above, but this requires index notation. Writing the equation in components,

$$\begin{aligned} L_i &= \int \rho(\mathbf{r})\left(\omega_i r^2 - r_i r_j \omega_j\right) d^3x \\ &= \int \rho(\mathbf{r})\left(\omega_j \delta_{ij} r^2 - r_i r_j \omega_j\right) d^3x \\ &= \omega_j \int \rho(\mathbf{r})\left(\delta_{ij} r^2 - r_i r_j\right) d^3x \end{aligned}$$

where we may pull ω_j out of the integral since it is independent of position. Notice how the use of dummy indices and the Kronecker delta allows us to get the same index on ω_j in both terms so that we can bring it outside, keeping the sum on j. Now define the *moment of inertia tensor*,

$$I_{ij} \equiv \int \rho(\mathbf{r})\left(\delta_{ij} r^2 - r_i r_j\right) d^3x \tag{15}$$

which depends only on the shape of the particular rigid body and not on its motion. This moment of inertia tensor is symmetric,

$$I_{ij} = I_{ji}$$

The torque equation may now be written as

$$\tau_i = \frac{d}{dt}\left(I_{ij}\omega_j\right)$$

We have therefore shown that the angluar momentum is

$$L_i = I_{ij}\omega_j \tag{16}$$

where equation of motion in an inertial frame is still $\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt}$. In general, I_{ij} is not proportional to the identity, so that the angular momentum and the angular velocity are not parallel.

6.3 Example: Moment of inertia of a thin rectangular sheet

For a uniform 2 dimensional rectangle of sides a and b along the x and y axes, respectively, we seek

$$I_{ij} = \int \rho(x, y)\left(\mathbf{x}^2 \delta_{ij} - x_i x_j\right) d^2x$$

about the center of mass. Since the body is uniform, the density (mass per unit area for a thin sheet) is constant,

$$\rho = \frac{M}{ab}$$

Also note that the center of mass is at $R_i = \left(\frac{a}{2}, \frac{b}{2}\right)$, so we compute relative to that point,

$$I_{ij} = \frac{M}{ab} \int \left((\mathbf{x} - \mathbf{R})^2 \delta_{ij} - (x_i - R_i)(x_j - R_j) \right) d^2x$$

We compute the integrals one at a time

$$\begin{aligned}
I_{11} &= \frac{M}{ab} \int \left((\mathbf{x} - \mathbf{R})^2 \delta_{11} - (x_1 - R_1)(x_1 - R_1) \right) d^2x \\
&= \frac{M}{ab} \int_0^b dy \int_0^a dx \left(\left(x - \frac{a}{2}\right)^2 + \left(y - \frac{b}{2}\right)^2 - \left(x - \frac{a}{2}\right)^2 \right) \\
&= \frac{M}{ab} \int_0^b dy \int_0^a dx \left(y - \frac{b}{2}\right)^2 \\
&= \frac{M}{ab} \int_0^b dy\, a \left(y - \frac{b}{2}\right)^2 \\
&= \frac{M}{b} \frac{1}{3} \left(y - \frac{b}{2}\right)^3 \Bigg|_0^b \\
&= \frac{M}{b} \left(\frac{1}{3} \left(\frac{b}{2}\right)^3 - \frac{1}{3} \left(-\frac{b}{2}\right)^3 \right) \\
&= \frac{M}{b} \frac{2}{3} \left(\frac{b^3}{8}\right) \\
&= \frac{Mb^2}{12}
\end{aligned}$$

For I_{22} the calculation is similar, yielding $I_{22} = \frac{Ma^2}{12}$. The off-diagonal component is

$$\begin{aligned}
I_{12} = I_{21} &= \frac{M}{ab} \int \left((\mathbf{x} - \mathbf{R})^2 \delta_{12} - \left(x - \frac{a}{2}\right)\left(y - \frac{b}{2}\right) \right) d^2x \\
&= -\frac{M}{ab} \int_0^b dy \left(y - \frac{b}{2}\right) \int_0^a dx \left(x - \frac{a}{2}\right) \\
&= -\frac{M}{ab} \left[\frac{1}{2} \left(y - \frac{b}{2}\right)^2 \right]_0^b \left[\frac{1}{2} \left(x - \frac{a}{2}\right)^2 \right]_0^a \\
&= -\frac{M}{4ab} \left[\left(\frac{b}{2}\right)^2 - \left(-\frac{b}{2}\right)^2 \right]_0^b \left[\left(\frac{a}{2}\right)^2 - \left(-\frac{a}{2}\right)^2 \right]_0^a \\
&= 0
\end{aligned}$$

6.4 Rotating reference frame and the Euler equation

Next, look at the equation of motion in a rotating frame of reference. We must replace the time derivative,

$$\left(\frac{d}{dt}\right)_{inertial} = \left(\frac{d}{dt}\right)_{body} + \boldsymbol{\omega} \times$$

and the equation of motion becomes

$$\boldsymbol{\tau} = \left(\frac{d\mathbf{L}}{dt}\right)_{body} + \boldsymbol{\omega} \times \mathbf{L}$$

This is the Euler equation.

In order to use the Euler equation, it is helpful to choose the right frame of reference. Given our rotating frame, any constant orthogonal transformation of the basis takes us to another equivalent rotating frame, but with a different orientation of the basis vectors. Furthermore, we know that any symmetric matrix may be diagonalized by an orthogonal transformation. Therefore, it is possible to rotate our basis to one in which I_{ij} is diagonal. In this basis, we have

$$[I]_{ij} = \begin{pmatrix} I_{11} & 0 & 0 \\ 0 & I_{22} & 0 \\ 0 & 0 & I_{33} \end{pmatrix} \tag{17}$$

The three eigenvalues, I_{11}, I_{22} and I_{33} are called the *principal moments of inertia.*

If we now write out the Euler equation in components using the principal moments, we have

$$\tau_i = \frac{d}{dt} I_{ij}\omega_j + \varepsilon_{ijk}\omega_j I_{km}\omega_m$$

so writing each component, $i = 1, 2, 3$, separately,

$$\begin{aligned} \tau_1 &= \frac{d}{dt} I_{1j}\omega_j + \varepsilon_{1jk}\omega_j I_{km}\omega_m \\ &= \frac{d}{dt} I_{11}\omega_1 + \varepsilon_{123}\omega_2 I_{3m}\omega_m + \varepsilon_{132}\omega_3 I_{2m}\omega_m \\ &= I_{11}\frac{d\omega_1}{dt} + \varepsilon_{123}\omega_2 I_{33}\omega_3 + \varepsilon_{132}\omega_3 I_{22}\omega_2 \\ &= I_{11}\frac{d\omega_1}{dt} + \omega_2\omega_3 \left(I_{33} - I_{22}\right) \end{aligned}$$

and similarly,

$$\begin{aligned} \tau_2 &= \frac{d}{dt} I_{22}\omega_2 + \varepsilon_{231}\omega_3 I_{11}\omega_1 + \varepsilon_{213}\omega_1 I_{33}\omega_3 \\ &= I_{22}\frac{d\omega_2}{dt} + \omega_3\omega_1 \left(I_{11} - I_{33}\right) \end{aligned}$$

and

$$\begin{aligned} \tau_3 &= \frac{d}{dt} I_{33}\omega_3 + \varepsilon_{312}\omega_1 I_{22}\omega_2 + \varepsilon_{321}\omega_2 I_{11}\omega_1 \\ &= I_{33}\frac{d\omega_3}{dt} + \omega_1\omega_2 \left(I_{22} - I_{11}\right) \end{aligned}$$

Introducing the briefer (but potentially misleading) notation

$$\begin{aligned} I_1 &= I_{11} \\ I_2 &= I_{22} \\ I_3 &= I_{33} \end{aligned}$$

we have the Euler equations in the form

$$\begin{aligned} \tau_1 &= I_1\dot{\omega}_1 - \omega_2\omega_3 \left(I_2 - I_3\right) \\ \tau_2 &= I_2\dot{\omega}_2 - \omega_3\omega_1 \left(I_3 - I_1\right) \\ \tau_3 &= I_3\dot{\omega}_3 - \omega_1\omega_2 \left(I_1 - I_2\right) \end{aligned} \tag{18}$$

6.5 Torque-free motion

When the torque vanishes, both the kinetic energy and the angular momentum are conserved. To find the kinetic energy, we write the action. There is no potential; in the inertial frame, the kinetic energy is the integral over the rigid body,

$$\begin{aligned}
T &= \frac{1}{2}\int \rho(\mathbf{r})(\mathbf{v}(\mathbf{r}))^2 d^3x \\
&= \frac{1}{2}\int \rho(\mathbf{r})(\boldsymbol{\omega}\times\mathbf{r})\cdot(\boldsymbol{\omega}\times\mathbf{r}) d^3x \\
&= \frac{1}{2}\int \rho(\mathbf{r})(\varepsilon_{imn}\omega_m r_n)(\varepsilon_{ijk}\omega_j r_k) d^3x \\
&= \frac{1}{2}\int \rho(\mathbf{r})(\varepsilon_{mni}\varepsilon_{jki})\omega_m r_n \omega_j r_k d^3x \\
&= \frac{1}{2}\int \rho(\mathbf{r})(\delta_{mj}\delta_{nk}-\delta_{mk}\delta_{nj})\omega_m r_n \omega_j r_k d^3x \\
&= \frac{1}{2}\omega_m\omega_j\int \rho(\mathbf{r})(\delta_{mj} r_n r_n - r_j r_m) d^3x \\
&= \frac{1}{2}\omega_m\omega_j I_{mj}
\end{aligned}$$

so we have

$$T = \frac{1}{2} I_{ij}\omega_i\omega_j \tag{19}$$

The force-free action is therefore

$$S = \int \frac{1}{2} I_{ij}\omega_i\omega_j dt$$

where $\omega_i = \dot{\varphi} n_i$. Since there is no explicit time dependence, the energy

$$\begin{aligned}
E &= \frac{\partial L}{\partial\dot{\varphi}}\dot{\varphi} - L \\
&= (I_{ij} n_i n_j \dot{\varphi})\dot{\varphi} - \frac{1}{2} I_{ij}\omega_i\omega_j \\
&= \frac{1}{2} I_{ij}\omega_i\omega_j
\end{aligned}$$

is conserved. Since there is no torque (or because φ is cyclic), we know that the total angular momentum is also conserved,

$$L_i = I_{ij}\omega_j$$

Suppose, for concreteness, that $I_1 > I_2 > I_3$. The case when two of the principal moments are equal is simpler and will be examined separately. Then for torque-free motion, the Euler equations, Eqs.(18), become

$$\begin{aligned}
I_1\dot{\omega}_1 &= \omega_2\omega_3(I_2 - I_3) \\
I_2\dot{\omega}_2 &= -\omega_3\omega_1(I_1 - I_3) \\
I_3\dot{\omega}_3 &= \omega_1\omega_2(I_1 - I_2)
\end{aligned}$$

where the differences between principal moments on the right are all non-negative. Add multiples of the first pair:

$$\begin{aligned}
I_1(I_1 - I_3)\omega_1\dot{\omega}_1 &= \omega_1\omega_2\omega_3(I_1 - I_3)(I_2 - I_3) \\
I_2(I_2 - I_3)\omega_2\dot{\omega}_2 &= -\omega_2\omega_3\omega_1(I_1 - I_3)(I_2 - I_3)
\end{aligned}$$

so that the right sides cancel, leaving the vanishing sum,

$$\begin{aligned} 0 &= I_1 (I_1 - I_3) \omega_1 \dot{\omega}_1 + I_2 (I_2 - I_3) \omega_2 \dot{\omega}_2 \\ &= \frac{1}{2} \frac{d}{dt} \left(I_1 (I_1 - I_3) \omega_1^2 + I_2 (I_2 - I_3) \omega_2^2 \right) \end{aligned}$$

so we find a constant, A, given by

$$A = I_1 (I_1 - I_3) \omega_1^2 + I_2 (I_2 - I_3) \omega_2^2$$

Similarly, we find a relation between ω_3^2 and ω_2^2,

$$\begin{aligned} I_2 (I_1 - I_2) \omega_2 \dot{\omega}_2 &= -\omega_3 \omega_1 \omega_2 (I_1 - I_3) (I_1 - I_2) \\ I_3 (I_1 - I_3) \omega_3 \dot{\omega}_3 &= \omega_1 \omega_2 \omega_3 (I_1 - I_2) (I_1 - I_3) \end{aligned}$$

so

$$0 = \frac{1}{2} \frac{d}{dt} \left(I_2 (I_1 - I_2) \omega_2^2 + I_3 (I_1 - I_3) \omega_3^2 \right)$$

Calling the second constant B, we solve for two of the components,

$$I_1 \omega_1^2 = \frac{1}{I_1 - I_3} \left[A - I_2 (I_2 - I_3) \omega_2^2 \right] \tag{20}$$

$$I_3 \omega_3^2 = \frac{1}{I_1 - I_3} \left[B - I_2 (I_1 - I_2) \omega_2^2 \right] \tag{21}$$

Substituting Eqs.(20) and (21) into the energy,

$$\begin{aligned} 2E &= I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2 \\ &= \frac{1}{I_1 - I_3} \left[A - I_2 (I_2 - I_3) \omega_2^2 \right] + I_2 \omega_2^2 + \frac{1}{I_1 - I_3} \left[B - I_2 (I_1 - I_2) \omega_2^2 \right] \\ &= \frac{1}{I_1 - I_3} \left(A + B - I_2 (I_2 - I_3) \omega_2^2 + I_2 (I_1 - I_3) \omega_2^2 - I_2 (I_1 - I_2) \omega_2^2 \right) \\ &= \frac{1}{I_1 - I_3} \left(A + B + (-I_2 I_2 + I_2 I_3 + I_1 I_2 - I_2 I_3 - I_1 I_2 + I_2 I_2) \omega_2^2 \right) \\ &= \frac{1}{I_1 - I_3} (A + B) \end{aligned}$$

so the sum of the constants is proportional to the energy,

$$A + B = 2 (I_1 - I_3) E$$

To find the remaining component, we solve for ω_1 and ω_3,

$$\begin{aligned} \omega_1 &= \sqrt{\frac{1}{I_1 (I_1 - I_3)} \left(A - I_2 (I_2 - I_3) \omega_2^2 \right)} \\ \omega_3 &= \sqrt{\frac{1}{I_3 (I_1 - I_3)} \left(B - I_2 (I_1 - I_2) \omega_2^2 \right)} \end{aligned}$$

and substitute into the Euler equation for ω_2,

$$I_2 \dot{\omega}_2 = -\omega_3 \omega_1 (I_1 - I_3)$$

Integrating gives an expression which can be written in terms of elliptic integrals,

$$-\sqrt{\frac{I_1 I_2 I_3}{AB}}t = \int \frac{d\omega_2}{\sqrt{\left(1 - \frac{I_2(I_2-I_3)}{A}\omega_2^2\right)\left(1 - \frac{I_2(I_1-I_2)}{B}\omega_2^2\right)}}$$

Rescale ω_2, letting

$$\begin{aligned} \chi &\equiv \sqrt{\frac{I_2 (I_2 - I_3)}{A}}\omega_2 \\ k^2 &\equiv \frac{A (I_1 - I_2)}{B (I_2 - I_3)} \end{aligned}$$

Then

$$-\sqrt{\frac{I_2 (I_2 - I_3)}{A}}\sqrt{\frac{I_1 I_2 I_3}{AB}}t = \int \frac{d\chi}{\sqrt{(1-\chi^2)(1-k^2\chi^2)}}$$

The right side is Jacobi's form of the elliptic integral of the first kind,

$$F(x,k) = \int_0^x \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

so we have

$$-\sqrt{\frac{I_2 (I_2 - I_3)}{A}}\sqrt{\frac{I_1 I_2 I_3}{AB}}t = F\left(\sqrt{\frac{I_2 (I_2 - I_3)}{A}}\omega_2, \frac{A (I_1 - I_2)}{B (I_2 - I_3)}\right)$$

which, in principle, gives the solution for $\omega_2(t)$.

We show below that for a symmetric body, the torque-free solution is much easier to understand.

6.6 Torque-free motion of a symmetric rigid body

Now consider the case when two of the moments of inertia are equal. This happens when the rigid body is rotationally symmetric around one axis. Let the z-axis be the axis of symmetry. Then $I_1 = I_2$, and the torque-free Euler equations become

$$\begin{aligned} 0 &= I_1\dot{\omega}_1 - \omega_2\omega_3 (I_1 - I_3) \\ 0 &= I_1\dot{\omega}_2 + \omega_1\omega_3 (I_1 - I_3) \\ 0 &= I_3\dot{\omega}_3 \end{aligned}$$

The final equation shows that ω_3 is constant. Defining the constant frequency $\Omega \equiv \omega_3\left(\frac{I_1-I_3}{I_1}\right)$ the remaining two equations are

$$\begin{aligned} \dot{\omega}_1 &= \Omega\omega_2 \\ \dot{\omega}_2 &= -\Omega\omega_1 \end{aligned}$$

We decouple these by differentiating the first and substituting the second,

$$\begin{aligned} \ddot{\omega}_1 &= \Omega\dot{\omega}_2 \\ &= -\Omega^2\omega_1 \end{aligned}$$

and similarly, by differentiating the second and substituting the first. This results in the pair of simple harmonic equations,

$$\begin{aligned} \ddot{\omega}_1 + \Omega^2\omega_1 &= 0 \\ \ddot{\omega}_2 + \Omega^2\omega_2 &= 0 \end{aligned}$$

with the immediate solution

$$\omega_1 = A\cos\Omega t + B\sin\Omega t$$

for ω_1 and, returning to the original equation $\dot{\omega}_1 = \Omega\omega_2$,

$$\omega_2 = -A\sin\Omega t + B\cos\Omega t$$

Notice that

$$\omega_1^2 + \omega_2^2 = A^2 + B^2$$

so the x and y components of the angular velocity together form a constant length vector that precesses around the z axis. If the angular velocity is dominated by ω_3, the remaining components give the object a "wobble" – it spins slightly off its symmetry axis, precessing. On the other hand, if ω_3 is small, the motion is a "tumble" – end over end rotation *of* its symmetry axis.

Remember that this analysis takes place in a frame of reference rotating with angular velocity $\boldsymbol{\omega}$. If all of the motion were about the z-axis, the object would be at rest in the rotating frame. The fact that we get time dependence of our solution for $\boldsymbol{\omega}$ means that even in a frame rotating with the body, the body precesses. If we transform back to the inertial frame, it is both spinning and precessing.

7 Lagrangian approach to rigid bodies

7.1 Angular velocity vector

To study symmetric rigid bodies with one point fixed and gravity acting – tops – we begin afresh and write an action for the problem. In order to do this, we require some set of coordinates. We take these to be the direction of a unit vector $\hat{\mathbf{n}} = \hat{\mathbf{n}}(\theta, \varphi)$ where θ and φ are the usual polar and azimuthal angles, respectively, and a third angle of rotation, ψ, about $\hat{\mathbf{n}}$. By specifying a triple of angles, (θ, φ, ψ), we may therefore specify the orientation of any rigid body.

Pick a Cartesian inertial frame of reference, $\left(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\right)$ with position vector given by $\mathbf{x}' = (x', y', z')$. To rotate to our body axes, we use $O(\hat{\mathbf{n}}(\theta,\varphi), \psi)$,

$$\begin{aligned}
\mathbf{x} &= O(\hat{\mathbf{n}}, \chi)\,\mathbf{x}' \\
&= \mathbf{x}'_{\parallel} + \mathbf{x}'_{\perp}\cos\theta - \sin\theta\,(\mathbf{n}\times\mathbf{x}'_{\perp})
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{x}'_{\parallel} &= \hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\mathbf{x}'\right) \\
\mathbf{x}'_{\perp} &= \mathbf{x}' - \hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\mathbf{x}'\right)
\end{aligned}$$

It is convenient to choose three orthonormal vectors for the body frame, $(\hat{\mathbf{n}}, \hat{\mathbf{m}}, \hat{\mathbf{s}})$ given by

$$\begin{aligned}
\hat{\mathbf{n}} &= \hat{\mathbf{n}} \\
\hat{\mathbf{m}} &= \frac{\mathbf{x}'_{\perp}}{x'_{\perp}} \\
\hat{\mathbf{s}} &= \hat{\mathbf{n}}\times\hat{\mathbf{m}}
\end{aligned}$$

In terms of these, the rotation is given by

$$O(\theta, \hat{\mathbf{n}})\,\mathbf{x}' = x'_{\parallel}\hat{\mathbf{n}} + x'_{\perp}\cos\theta\,\hat{\mathbf{m}} - x'_{\perp}\sin\theta\,\hat{\mathbf{s}}$$

We also will need unit vectors in the directions of increasing coordinate angles, (θ, φ, ψ),

$$\begin{aligned}
\hat{\boldsymbol{\theta}} &= \hat{\mathbf{i}}\cos\theta\cos\varphi + \hat{\mathbf{j}}\cos\theta\sin\varphi - \hat{\mathbf{k}}\sin\theta \\
\hat{\boldsymbol{\varphi}} &= -\hat{\mathbf{i}}\sin\varphi + \hat{\mathbf{j}}\cos\varphi \\
\hat{\mathbf{n}} &= \hat{\mathbf{i}}\sin\theta\cos\varphi + \hat{\mathbf{j}}\sin\theta\sin\varphi + \hat{\mathbf{k}}\cos\theta
\end{aligned}$$

We may use these to specify an arbitrary angular velocity vector. A change in θ produces an angular velocity around the axis $\hat{\mathbf{n}} \times \hat{\boldsymbol{\theta}}$, while a change in φproduces a rotation around $\hat{\mathbf{k}}$. Changes in $\psi$are rotations about $\hat{\mathbf{n}}$. A general change of all three angles, $\left(\dot{\theta}, \dot{\varphi}, \dot{\psi}\right)$ give a linear combination of all three, $\boldsymbol{\omega} = \dot{\theta}\hat{\mathbf{n}} \times \hat{\boldsymbol{\theta}} + \dot{\varphi}\hat{\mathbf{k}} + \dot{\psi}\hat{\mathbf{n}}$. Expanding in the original basis,

$$\begin{aligned}
\boldsymbol{\omega} &= \dot{\theta}\hat{\mathbf{n}} \times \hat{\boldsymbol{\theta}} + \dot{\varphi}\hat{\mathbf{k}} + \dot{\psi}\hat{\mathbf{n}} \\
&= \dot{\theta}\left(\hat{\mathbf{i}}\sin\theta\cos\varphi + \hat{\mathbf{j}}\sin\theta\sin\varphi + \hat{\mathbf{k}}\cos\theta\right) \times \left(\hat{\mathbf{i}}\cos\theta\cos\varphi + \hat{\mathbf{j}}\cos\theta\sin\varphi - \hat{\mathbf{k}}\sin\theta\right) \\
&\quad + \dot{\varphi}\hat{\mathbf{k}} + \dot{\psi}\left(\hat{\mathbf{i}}\sin\theta\cos\varphi + \hat{\mathbf{j}}\sin\theta\sin\varphi + \hat{\mathbf{k}}\cos\theta\right) \\
&= \dot{\theta}\left(\hat{\mathbf{k}}\cos\theta\sin\theta\cos\varphi\sin\varphi + \hat{\mathbf{j}}\sin\theta\sin\theta\cos\varphi - \hat{\mathbf{k}}\cos\theta\sin\theta\sin\varphi\cos\varphi - \hat{\mathbf{i}}\sin\theta\sin\theta\sin\varphi + \hat{\mathbf{j}}\cos\theta\cos\theta\cos\varphi - \hat{\mathbf{i}}\cos^2\theta\sin\right. \\
&\quad + \dot{\varphi}\hat{\mathbf{k}} + \dot{\psi}\left(\hat{\mathbf{i}}\sin\theta\cos\varphi + \hat{\mathbf{j}}\sin\theta\sin\varphi + \hat{\mathbf{k}}\cos\theta\right) \\
&= \left(\dot{\psi}\sin\theta\cos\varphi - \dot{\theta}\sin\varphi\right)\hat{\mathbf{i}} + \left(\dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi\right)\hat{\mathbf{j}} + \left(\dot{\psi}\cos\theta + \dot{\varphi}\right)\hat{\mathbf{k}}
\end{aligned}$$

so the angular velocity is

$$\begin{aligned}
\omega &= \left(\dot{\psi}\sin\theta\cos\varphi - \dot{\theta}\sin\varphi\right)\hat{\mathbf{i}} + \left(\dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi\right)\hat{\mathbf{j}} + \left(\dot{\psi}\cos\theta + \dot{\varphi}\right)\hat{\mathbf{k}} \\
&= \omega_1\hat{\mathbf{i}} + \omega_2\hat{\mathbf{j}} + \omega_3\hat{\mathbf{k}}
\end{aligned}$$

7.2 Action functional for rigid body motion

We are now in a position to write the Lagrangian and action for a rigid body. In terms of Euler coordinates, the kinetic energy is

$$T = \frac{1}{2}I_{ij}\omega_i\omega_j$$

Substituting for the angular velocity in the principal axis frame this becomes

$$T = \frac{1}{2}I_1\left(\dot{\psi}\sin\theta\cos\varphi - \dot{\theta}\sin\varphi\right)^2 + \frac{1}{2}I_2\left(\dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi\right)^2 + \frac{1}{2}I_3\left(\dot{\psi}\cos\theta + \dot{\varphi}\right)^2$$

We may also have the kinetic energy of the center of mass.

For a slowly changing force field, we may write the potential as a function of the center of mass only, but if there is a gradient or the forces are applied at specific points of the rigid body, there may be torques as well. If the body is in a gravitational field, the potential is found by integrating

$$dV = -dm\mathbf{g}(\mathbf{r}) \cdot d\mathbf{x}$$

where $\mathbf{g}(\mathbf{x})$ is the local gravitational acceleration. For a uniform gravitational field, $\mathbf{g} = -g\mathbf{k}$ is constant so $\int \mathbf{g}(\mathbf{r}) \cdot d\mathbf{x} = \mathbf{g} \cdot \mathbf{x}$

$$\begin{aligned}
dV &= -\left(\rho d^3x\right)\mathbf{g} \cdot \mathbf{x} \\
V &= -\mathbf{g} \cdot \int \rho\mathbf{x}d^3x \\
&= -M\mathbf{g} \cdot \mathbf{R}
\end{aligned}$$

since the center of mass is defined as

$$\mathbf{R} = \frac{1}{M}\int \rho \mathbf{x} d^3x$$

We now apply these considerations to the case of a rigid body symmetric about one axis, with one point fixed: tops.

7.3 Symmetric body with torque: tops

Now suppose the rotationally symmetric body rests on one point of the symmetry axis, like a top spinning on a tabletop. We take this point as fixed. Then, unless the top is perfectly vertical, there is a torque acting, produced by gravity acting at the center of mass. If the center of mass is a distance l above the fixed tip, then the potential is

$$V = Mgl\cos\theta$$

Taking the z-axis of the body frame as the symmetry axis, we have $I_2 = I_1$, and the first two terms of the kinetic energy simplify considerably.

$$\begin{aligned}
T &= \frac{1}{2}I_1\left(\left(\dot{\psi}\sin\theta\cos\varphi - \dot{\theta}\sin\varphi\right)^2 + \left(\dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi\right)^2\right) + \frac{1}{2}I_3\left(\dot{\psi}\cos\theta + \dot{\varphi}\right)^2 \\
T &= \frac{1}{2}I_1\left(\dot{\psi}^2\sin^2\theta\cos^2\varphi - 2\dot{\psi}\dot{\theta}\sin\theta\cos\varphi\sin\varphi + \dot{\theta}^2\sin^2\varphi\right) \\
&\quad + \frac{1}{2}I_1\left(\dot{\theta}^2\cos^2\varphi + 2\dot{\theta}\dot{\psi}\cos\varphi\sin\theta\sin\varphi + \dot{\psi}^2\sin^2\theta\sin^2\varphi\right) + \frac{1}{2}I_3\left(\dot{\psi}\cos\theta + \dot{\varphi}\right)^2 \\
T &= \frac{1}{2}I_1\left(\dot{\psi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\psi}\cos\theta + \dot{\varphi}\right)^2
\end{aligned}$$

Then the action becomes

$$S = \int dt\left(\frac{1}{2}I_1\left(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\psi} + \dot{\varphi}\cos\theta\right)^2 - Mgl\cos\theta\right)$$

7.3.1 Conserved quantities

We first look for conserved quantities. Two angles, φ and ψ, are cyclic, so their conjugate momenta are conserved:

$$\begin{aligned}
p_\varphi &= \frac{\partial L}{\partial \dot{\varphi}} \\
&= I_1\dot{\varphi}\sin^2\theta + I_3\left(\dot{\psi} + \dot{\varphi}\cos\theta\right)\cos\theta \\
&= \dot{\varphi}\left(I_1\sin^2\theta + I_3\cos^2\theta\right) + I_3\dot{\psi}\cos\theta
\end{aligned}$$

and

$$\begin{aligned}
p_\psi &= \frac{\partial L}{\partial \dot{\psi}} \\
&= I_3\left(\dot{\psi} + \dot{\varphi}\cos\theta\right)
\end{aligned}$$

The energy provides a third constant of the motion since $\frac{\partial L}{\partial t} = 0$. Since the Lagrangian is quadratic in the velocities, we have $E = T + V$,

$$\begin{aligned}
E &= \frac{1}{2}I_1\left(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\psi} + \dot{\varphi}\cos\theta\right)^2 + Mgl\cos\theta \\
&= \frac{1}{2}I_1\left(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{p_\psi^2}{2I_3} + Mgl\cos\theta
\end{aligned}$$

7.3.2 Solving

To solve, we first eliminate $\dot{\psi}$,

$$\dot{\psi} = \frac{p_\psi}{I_3} - \dot{\varphi}\cos\theta$$

then substitute this into p_φ,

$$\begin{aligned} p_\varphi &= \dot{\varphi}\left(I_1\sin^2\theta + I_3\cos^2\theta\right) + I_3\left(\frac{p_\psi}{I_3} - \dot{\varphi}\cos\theta\right)\cos\theta \\ &= \dot{\varphi}\left(I_1\sin^2\theta + I_3\cos^2\theta - I_3\cos^2\theta\right) + p_\psi\cos\theta \\ &= \dot{\varphi}I_1\sin^2\theta + p_\psi\cos\theta \end{aligned}$$

so that we may also solve for $\dot{\varphi}$. This gives

$$\dot{\varphi} = \frac{p_\varphi - p_\psi\cos\theta}{I_1\sin^2\theta}$$

Finally, we use the energy to express θ as an integral,

$$\begin{aligned} E &= \frac{1}{2}I_1\left(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{p_\psi^2}{2I_3} + Mgl\cos\theta \\ &= \frac{1}{2}I_1\left(\left(\frac{p_\varphi - p_\psi\cos\theta}{I_1\sin^2\theta}\right)^2\sin^2\theta + \dot{\theta}^2\right) + \frac{p_\psi^2}{2I_3} + Mgl\cos\theta \\ &= \frac{1}{2}I_1\dot{\theta}^2 + \frac{\left(p_\varphi - p_\psi\cos\theta\right)^2}{2I_1\sin^2\theta} + \frac{p_\psi^2}{2I_3} + Mgl\cos\theta \end{aligned}$$

We may drop the constant term, $\frac{p_\psi^2}{2I_3}$. Then, solving for $\dot{\theta}$ to integrate,

$$\begin{aligned} t &= \int\frac{d\theta}{\sqrt{\frac{2E}{I_1} - \frac{(p_\varphi - p_\psi\cos\theta)^2}{I_1^2\sin^2\theta} - \frac{2Mgl}{I_1}\cos\theta}} \\ &= \int\frac{I_1\sin\theta d\theta}{\sqrt{2I_1E\sin^2\theta - \left(p_\varphi - p_\psi\cos\theta\right)^2 - 2I_1Mgl\sin^2\theta\cos\theta}} \end{aligned}$$

or, setting $x = \cos\theta$,

$$t = -\int\frac{dx}{\sqrt{\frac{2E}{I_1}\left(1 - x^2\right) - \frac{1}{I_1^2}\left(p_\varphi - p_\psi x\right)^2 - \frac{2Mgl}{I_1}\left(x - x^3\right)}}$$

The cubic in under the root makes this difficult, but it can be expressed in terms of elliptic integrals or numerically integrated. The resulting $\theta(t)$ then allows us to integrate to find $\varphi(t)$ and $\psi(t)$.

7.3.3 Potential energy

A simpler way to approach the qualitative behavior is to view the energy as that of a 1-dimensional problem with kinetic energy $\frac{1}{2}I_1\dot{\theta}^2$ and an effective potential

$$V_{eff} = \frac{\left(p_\varphi - p_\psi\cos\theta\right)^2}{2I_1\sin^2\theta} + Mgl\cos\theta$$

where we again drop the irrelevant constant, $\frac{p_\psi^2}{2I_3}$.

To explore the motion in this potential, again set $x = \cos\theta$. Then

$$V_{eff} = \frac{(p_\varphi - p_\psi x)^2}{2I_1(1-x^2)} + Mglx$$

This has extrema when

$$\begin{aligned}
0 &= \frac{dV_{eff}}{dx} \\
&= \frac{-2p_\psi(p_\varphi - p_\psi x)}{2I_1(1-x^2)} - \frac{(p_\varphi - p_\psi x)^2}{2I_1(1-x^2)^2}(-2x) + Mgl \\
&= \frac{1}{2I_1(1-x^2)^2}\left(-p_\psi 2(p_\varphi - p_\psi x)\left(1-x^2\right) + 2x(p_\varphi - p_\psi x)^2 + 2I_1 Mgl\left(1-x^2\right)^2\right) \\
0 &= -2p_\psi p_\varphi\left(1-x^2\right) + 2p_\psi^2 x\left(1-x^2\right) + 2p_\varphi^2 x - 4p_\varphi p_\psi x^2 + 2p_\psi^2 x^3 + 2I_1 Mgl\left(1-2x^2+x^4\right) \\
0 &= (2I_1 Mgl - 2p_\psi p_\varphi) + \left(2p_\varphi^2 + 2p_\psi^2\right)x + (2p_\psi p_\varphi - 4p_\varphi p_\psi - 4I_1 Mgl)x^2 + \left(2p_\psi^2 - 2p_\psi^2\right)x^3 + 2I_1 Mglx^4 \\
0 &= (I_1 Mgl - p_\psi p_\varphi) + \left(p_\varphi^2 + p_\psi^2\right)x - (p_\varphi p_\psi + 2I_1 Mgl)x^2 + (I_1 Mgl)x^4 \\
0 &= I_1 Mgl\left(1-x^2\right)^2 - p_\psi p_\varphi\left(1+x^2\right) + \left(p_\varphi^2 + p_\psi^2\right)x
\end{aligned}$$

For x near 1, so that the top is nearly vertical, the first term may be neglected and we have approximately

$$\begin{aligned}
0 &= p_\psi p_\varphi - \left(p_\varphi^2 + p_\psi^2\right)x + p_\psi p_\varphi x^2 \\
x &= \frac{1}{2p_\psi p_\varphi}\left(\left(p_\varphi^2 + p_\psi^2\right) \pm \sqrt{\left(p_\varphi^2 + p_\psi^2\right)^2 - 4p_\psi^2 p_\varphi^2}\right) \\
&= \frac{1}{2p_\psi p_\varphi}\left(p_\varphi^2 + p_\psi^2 \pm \left(p_\varphi^2 - p_\psi^2\right)\right)
\end{aligned}$$

yielding two solutions, $x = \frac{p_\varphi}{p_\psi}, \frac{p_\psi}{p_\varphi}$. Only one of these values is less than one, so there is a unique angle $\cos\theta_0 = \min\left(\frac{p_\varphi}{p_\psi}, \frac{p_\psi}{p_\varphi}\right)$ which remains approximately constant. With $x_0 = \cos\theta_0$, the remaining angles satisfy

$$\begin{aligned}
\dot\psi &= \frac{p_\psi}{I_3} - \dot\varphi x_0 \\
\dot\varphi &= \frac{p_\varphi - p_\psi x_0}{I_1\left(1-x_0^2\right)} \\
&= \frac{p_\varphi - p_\psi \min\left(\frac{p_\varphi}{p_\psi}, \frac{p_\psi}{p_\varphi}\right)}{I_1\left(1-x_0^2\right)}
\end{aligned}$$

If $p_\varphi < p_\psi$ then $\dot\varphi = 0$ and the top spins in a fixed (or very slowly changing) position, while for $p_\psi < p_\varphi$, the top precesses in a nearly vertical position at a speed $\frac{p_\varphi - \frac{p_\psi^2}{p_\varphi}}{I_1\left(1-\frac{p_\psi^2}{p_\varphi^2}\right)} = \frac{p_\varphi}{I_1}$. In this case,

$$\dot\psi = \frac{p_\psi}{I_3} - \frac{p_\varphi}{I_1}\frac{p_\psi}{p_\varphi} = p_\psi\left(\frac{1}{I_3} - \frac{1}{I_1}\right)$$

Now consider small x, so we neglect the x^4 term. Then

$$0 = (p_\psi p_\varphi + 2I_1 Mgl)x^2 - \left(p_\varphi^2 + p_\psi^2\right)x - (I_1 Mgl - p_\psi p_\varphi)$$

$$x = \frac{1}{2\left(p_\psi p_\varphi + 2I_1 Mgl\right)}\left(\left(p_\varphi^2 + p_\psi^2\right) \pm \sqrt{\left(p_\psi^2 + p_\varphi^2\right)^2 + 4\left(p_\psi p_\varphi + 2I_1 Mgl\right)\left(I_1 Mgl - p_\psi p_\varphi\right)}\right)$$

$$x = \frac{1}{2\left(p_\psi p_\varphi + 2I_1 Mgl\right)}\left(\left(p_\varphi^2 + p_\psi^2\right) \pm \sqrt{\left(p_\psi^2 - p_\varphi^2\right)^2 - 4I_1 Mgl p_\psi p_\varphi + 8I_1^2 M^2 g^2 l^2}\right)$$

This has real solutions as long as

$$\left(p_\psi^2 - p_\varphi^2\right)^2 + 8I_1^2 M^2 g^2 l^2 \geq 4I_1 Mgl p_\psi p_\varphi$$

When the spin is fast, we have $p_\psi \gg p_\varphi$ and may approximate

$$p_\psi^4 + 8I_1^2 M^2 g^2 l^2 \geq 4I_1 Mgl p_\psi^2 \frac{p_\varphi}{p_\psi}$$

But then

$$\begin{aligned} 0 &\leq \left(p_\psi^2 - 2I_1 Mgl\right)^2 \\ &= p_\psi^4 - 4I_1 Mgl p_\psi^2 + 4I_1^2 M^2 g^2 l^2 \\ p_\psi^4 + 4I_1^2 M^2 g^2 l^2 &\geq 4I_1 Mgl p_\psi \end{aligned}$$

and since $\frac{p_\varphi}{p_\psi} < 1$, there is always a solution. When the energy equals the potential at such a minimum, the top will precess in a circle at a fixed angle θ. We may the consider perturbations around this solution. Various solutions are depicted in Figure 5.9 of Goldstein, depending on the relative frequencies in the θ and φ oscillations.

7.3.4 Slowly precessing top

Consider the case when we have $\dot\psi \gg \dot\theta \gg \dot\varphi$. Then we may make the following approximations:

$$\begin{aligned} \dot\psi &= \frac{p_\psi}{I_3} - \dot\varphi\cos\theta \\ &\approx \frac{p_\psi}{I_3} \end{aligned}$$

and

$$\dot\varphi = \frac{p_\varphi - p_\psi\cos\theta}{I_1\sin^2\theta} \ll \dot\psi$$

while the energy is approximately

$$\begin{aligned} E &= \frac{1}{2}I_1\left(\dot\varphi^2\sin^2\theta + \dot\theta^2\right) + Mgl\cos\theta \\ &\approx \frac{1}{2}I_1\dot\theta^2 + Mgl\cos\theta \end{aligned}$$

Then, solving for $\dot\theta$ to integrate, we find an elliptic integral:

$$\begin{aligned} t &= \int \frac{d\theta}{\sqrt{\frac{2E'}{I_1} - \frac{2Mgl}{I_1}\cos\theta}} \\ &= \sqrt{\frac{2I_1}{E' - Mgl}}F\left(\frac{\theta}{2}; -\frac{4Mgl}{2E' - 2Mgl}\right) \end{aligned}$$

7.4 Gyroscopes

Gyroscopes are typically mounted on freely turning frames so that there is no external torque. In this case, the potential vanishes and we have the simpler system

$$\begin{aligned}
\dot{\psi} &= \frac{p_\psi}{I_3} - \dot{\varphi}\cos\theta \\
\dot{\varphi} &= \frac{p_\varphi - p_\psi \cos\theta}{I_1 \sin^2\theta} \\
E &= \frac{1}{2} I_1 \left(\dot{\varphi}^2 \sin^2\theta + \dot{\theta}^2\right) \\
&= \frac{1}{2} I_1 \left(\left(\frac{p_\varphi - p_\psi \cos\theta}{I_1 \sin^2\theta}\right)^2 \sin^2\theta + \dot{\theta}^2\right) \\
&= \frac{1}{2} I_1 \dot{\theta}^2 + \frac{\left(p_\varphi - p_\psi \cos\theta\right)^2}{2 I_1 \sin^2\theta}
\end{aligned}$$

where we have dropped the extraneous constant $\frac{p_\psi^2}{2I_3}$ from the energy. The effective potential is now

$$V_{eff} = \frac{\left(p_\varphi - p_\psi \cos\theta\right)^2}{2 I_1 \sin^2\theta}$$

The extrema at

$$x = \cos\theta = \min\left(\frac{p_\varphi}{p_\psi}, \frac{p_\psi}{p_\varphi}\right)$$

where the minimum selects for the value which gives $x \leq 1$. There is therefore exactly one solution

$$\begin{aligned}
t &= \int \frac{d\theta}{\sqrt{\frac{2E}{I_1} - \frac{\left(p_\varphi - p_\psi \cos\theta\right)^2}{I_1^2 \sin^2\theta}}} \\
&= \int \frac{\sin\theta d\theta}{\sqrt{\frac{2E}{I_1}\sin^2\theta - \frac{1}{I_1^2}\left(p_\varphi - p_\psi \cos\theta\right)^2}} \\
&= \int \frac{I_1 dx}{\sqrt{\left(2I_1 E - p_\varphi^2\right) + 2 p_\varphi p_\psi x - \left(2 I_1 E + p_\psi^2\right) x^2}}
\end{aligned}$$

This time the root is quadratic, and we may complete the square,

$$\left(2I_1 E - p_\varphi^2\right) + 2 p_\varphi p_\psi x - \left(2 I_1 E + p_\psi^2\right) x^2 = -\left(\sqrt{2I_1 E + p_\psi^2}\, x - \frac{p_\varphi p_\psi}{\sqrt{2 I_1 E + p_\psi^2}}\right)^2 + \frac{p_\varphi^2 p_\psi^2}{2 I_1 E + p_\psi^2} + 2 I_1 E - p_\varphi^2$$

Setting

$$\begin{aligned}
\xi &= \sqrt{2I_1 E + p_\psi^2}\, x - \frac{p_\varphi p_\psi}{\sqrt{2 I_1 E + p_\psi^2}} \\
dx &= \frac{d\xi}{\sqrt{2 I_1 E + p_\psi^2}} \\
A^2 &= \frac{p_\varphi^2 p_\psi^2}{2 I_1 E + p_\psi^2} + 2 I_1 E - p_\varphi^2 \\
\Omega &= \frac{1}{I_1}\sqrt{2 I_1 E + p_\psi^2}
\end{aligned}$$

we have

$$t = \frac{I_1}{\sqrt{2I_1E + p_\psi^2}} \int \frac{d\xi}{\sqrt{A^2 - \xi^2}}$$

so we set

$$\xi = A \sin\alpha$$

and the integral is

$$\begin{aligned}
t &= \frac{I_1}{\sqrt{2I_1E + p_\psi^2}} \sin^{-1}\frac{\xi}{A} \\
\sqrt{2I_1E + p_\psi^2}\cos\theta - \frac{p_\varphi p_\psi}{\sqrt{2I_1E + p_\psi^2}} &= A \sin\Omega t \\
\cos\theta &= \frac{p_\varphi p_\psi}{2I_1E + p_\psi^2} + \frac{A}{\sqrt{2I_1E + p_\psi^2}} \sin\Omega t \\
&= \cos\theta_0 + b\sin\Omega t
\end{aligned}$$

This displays nutation clearly: the tip angle of the gyroscope oscillates up and down around the angle θ_0 with period Ω. From $\Omega = \frac{1}{I_1}\sqrt{2I_1E + p_\psi^2}$ we see that Ω may have any magnitude, depending on the size of I_1. Then, from

$$\dot{\varphi} = \frac{p_\varphi - p_\psi \cos\theta}{I_1 \sin^2\theta}$$

we see that the value of p_φ makes $\dot{\varphi}$ independent of Ω. This means that the rate of precession and the rate of nutation are independent of one another.