Chloe Elliott

# The Spinning Top

April 30th 2009

# Abstract

In this project I will look into the phenomenon of precession, in particular into the motion of the spinning top, and further onto how precession allows the Levitron top achieve stable levitation over a magnetic base.

# Contents

# Chapter 1

# Introduction

## 1.1 Context and Motivation

In this project, I am going to be considering the spinning motions of rigid bodies. I will begin by discussing the kinematics of the motion of rigid bodies, and then go on to look at their equations of motion. I am most interested in the phenomenon of precession.

The Spinning Top is a toy that can be spun on an axis, balancing on a point. It is one of the oldest recognisable toys found on archaeological sites, and it seems to have originated independently in cultures all over the world. The action of the spinning top is reliant on a gyroscopic effect for its motion. This effect can sometimes suggest counter-intuitive ideas, and it certainly amazed physicists Niels Bohr and Wolfgang Pauli!

**Figure 1.1: Pauli & Bohr marvel at the spinning top**[1]

All rotating objects can undergo precession. The gyroscope is a device which is used for measuring or maintaining orientation. The classic image of a gyroscope is a relatively large rotor suspended in light supporting rings, known as *gimbals*, which have frictionless bearings and ensure that the central rotor is

[1] **image obtained from www.damtp.cam.ac.uk/user/tong/dynamics/three.pdf**

isolated from external torques. The gyroscope is capable of extreme balanced stability when at high speeds, and can maintain the direction of the high speed rotation axis of the central rotor.

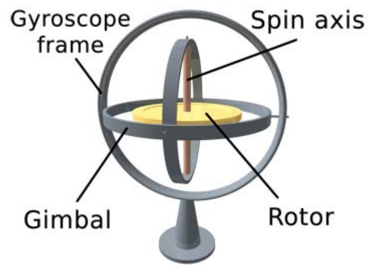


**Figure 1.2: A gyroscope**[2]

The angular momentum of the rotor maintains its magnitude in absence of external torques, and the orientation also remains almost fixed, regardless of the motion of the platform on which it is mounted.

The classical gyroscope has several applications, with the gyrocompass being arguably the most important. Gyrocompasses were first developed around 1906. They make use of a fast-spinning wheel in addition to friction forces in order to exploit the Earth's rotation and enable true north to be found. They are widely used as navigation devices on ships and aircrafts, and were particularly important in the early 1900's when they had significant military uses.

There are many more examples of gyroscopic motion: The wheels of bicycles, the spin of the Earth in space and even the behaviour of a boomerang all exhibit this type of motion.

The Levitron is an "amazing anti-gravity top" which is exceptional in that Earnshaw's theorem seems to indicate that its operation is impossible, and the majority of scientists thought that it would never be able to work. Precession is in fact responsible for the Levitron being able to apparently "float" in mid-air over a magnetic base. I will be looking closely into the Levitron and attempting to explain the principles behind its operation in this project.

[2]**image obtained from http://commons.wikimedia.org/wiki/File:3D_Gyroscope.png**

# Chapter 2

## Rigid Bodies and the Kinematics of their Motion

A rigid body can be described as a system of mass points subject to the constraint that the distance between any 2 points of the system must remain permanently fixed. In this chapter we will be considering the nature of rigid body motion.

### 2.1 Independent Coordinates of a Rigid Body

It is convenient to consider a coordinate system whose axes point along fixed directions in the body. We shall call this the "body-fixed" system. The origin is chosen to be some point held fixed in the body.

A rigid body of N particles has up to 3N degrees of freedom, however, there is the constraint that the distance between the particles must remain fixed:

$$| r_i - r_j | = r_{ij} = c_{ij} \tag{2.1}$$

where $r_{ij}$ is the distance between the $i$th and $j$th particles and $c_{ij}$ is a constant.

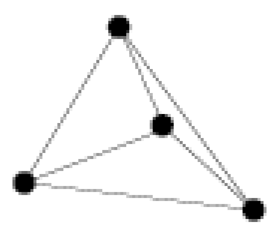

**Figure 2.1: A rigid body has fixed distances between its points**[3]

For N particles, $\frac{1}{2}N(N$ - 1) constraint equations exist, but these constraint equations are not all independent. If we fix a point in a rigid body, we only need to specify distances to any three non-collinear points. Once these first three points are fixed, any additional particle in the body gives 3 new coordinates, but also three new constraint equations, so the net increase in degrees of freedom is zero. There is therefore a maximum of 9 degrees of freedom for a rigid body.

The first three particles we chose have fixed distances between them:

$$r_{12} = c_{12} \qquad r_{23} = c_{23} \qquad r_{13} = c_{13}.$$

[3]**image obtained from http://www.damtp.cam.ac.uk/user/tong/dynamics/three.pdf**

So now the number of degrees of freedom is reduced to six.

We now know that six coordinates are required to completely specify the position of an unconstrained rigid body in 3-dimensional space: three of these are cartesian coordinates which we need to specify position of the centre of mass of the body axes and three are angles needed to specify the orientation of the body axes relative to a co-incident external set of axes.

Additional constraints on the body could reduce the number of independent coordinates further.

Many different ways exist of specifying the orientation of one cartesian system of axes with respect to another. A popular and useful method is to specify the direction cosines of a set of primed axes relative to unprimed axes. If we set the external axes to be labelled as $x$, $y$, $z$ axes, the $x'$ axis in the body frame can then be defined by its three direction cosines.

Convention is to label the three unit vectors along $x$, $y$, $z$ as **i**, **j** and **k**, and the unit vectors in the body system as **i**', **j**' and **k**'. We then define the direction cosines for $x'$ as:

$$\begin{aligned} cos\theta_{11} &= cos\theta_{i'i} = \mathbf{i}'\cdot\,\mathbf{i} \\ cos\theta_{12} &= cos\theta_{i'j} = \mathbf{i}'\cdot\,\mathbf{j} \\ cos\theta_{13} &= cos\theta_{i'k} = \mathbf{i}'\cdot\,\mathbf{k}. \end{aligned} \tag{2.2}$$

The relation between unit vectors in the two systems can be expressed as:

$$\begin{aligned} \mathbf{i}' &= cos\theta_{11}\mathbf{i} + cos\theta_{12}\mathbf{j} + cos\theta_{13}\mathbf{k} \\ \mathbf{j}' &= cos\theta_{21}\mathbf{i} + cos\theta_{22}\mathbf{j} + cos\theta_{23}\mathbf{k} \\ \mathbf{k}' &= cos\theta_{31}\mathbf{i} + cos\theta_{32}\mathbf{j} + cos\theta_{33}\mathbf{k}. \end{aligned} \tag{2.3}$$

An arbitrary vector, **v**, can be expressed in the two systems as:

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}' \tag{2.4}$$

and the relationship between primed and unprimed coordinates is:

$$\begin{aligned} x' &= cos\theta_{11}x + cos\theta_{12}y + cos\theta_{13}z \\ y' &= cos\theta_{21}x + cos\theta_{22}y + cos\theta_{23}z \\ z' &= cos\theta_{31}x + cos\theta_{32}y + cos\theta_{33}z \end{aligned} \tag{2.5}$$

with similar inverse relations.

Because the primed axis are fixed in the body, the nine direction cosines will change in time as the body rotates. The direction cosines can be thought of as generalised coordinates describing the orientation of the body, however, they cannot all be independent since we know that only three coordinates are required to specify an orientation. The relations between the direction cosines are due to the basis vectors being orthogonal to each other in each of the two coordinate systems, and all having unit magnitude:

$$\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{i}\cdot\mathbf{k} = 0$$

$$\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1 \tag{2.6}$$

with similar relations for $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$. Combining the set of equations (4.3) with the dot-product relations, we can then define six orthogonality relations between the direction cosines:

$$\sum_{q=1}^{3} cos\theta_{pq} cos\theta_{p'q} = 0 \qquad p,q = 1,\ 2,\ 3,\ p \neq p'$$

$$\sum_{q=1}^{3} cos^2\theta_{pq} = 1. \tag{2.7}$$

We can use the Kronecker delta function, $\delta_{pq}$ , defined by

$$\delta_{p'p} = 1 \quad p = p'$$

$$= 0 \quad p \neq p'$$

to reduce the orthogonality relations to:

$$\sum_{p=1}^{3} cos\theta_{pq} cos\theta_{p'q} = \delta_{p'p}. \tag{2.8}$$

Due to there being six orthogonality parameters between the nine direction cosines, the number of independent coordinates has now been reduced to just three.

The direction cosines are therefore convenient and useful to use as a method of specifying the relative orientation of a cartesian coordinate system to another coincident system, but they cannot be used as generalized coordinates to set up a Lagrangian. We must instead choose some set of three independent functions of the direction cosines. Although not unique, the Euler angles are a popular choice for this set, and we will look at these in more detail later.

### 2.2 Orthogonal transformations

A general linear transformation equation can be written as

$$\mathbf{x}' = \mathbf{A}\mathbf{x} \tag{2.9}$$

where $\mathbf{A}$ is the matrix of transformation which can be considered as an operator which transforms the unprimed system into the primed system. The vector remains unchanged; $\mathbf{A}$ acts on the coordinate system only, yielding the coordinates of the vector components in the new, primed frame. $\mathbf{A}$ has matrix elements $a_{ij.}$

Denoting the $x$, $y$, $z$ axes instead as the $x_{1,}$ $x_{2,}$ $x_3$ axes will prove to be advantageous here. In matrix form, the linear transformation can be written as:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{2.10}$$

to give three transformation equations. The set of equations (2.5) is a special case of a linear transformation, as there is not independence of all the direction cosines. We can re-derive the connection between the cosine coefficients in terms of our new $a_{ij}$ notation.

We are able to express each of the primed coordinates as a summation:

$$x'_i = \sum_{i=1}^{3} a_{ij} x_j \qquad i = 1,\ 2,\ 3. \tag{2.11}$$

The actual vector $\mathbf{x}$ remains unchanged in both the old and new coordinate systems, and so magnitude of the vector $\mathbf{x}$ (given in terms of the sum of squares of the components) therefore also remains unchanged in both systems:

$$\sum_{i=1}^{3} x_i'^2 = \sum_{i=1}^{3} x_i^2 \ . \tag{2.12}$$

Using the transformation equation (2.11), this now becomes

$$\sum_{i=1}^{3} \left( \sum_{i=1}^{3} a_{ij} x_j \right) \left( \sum_{i=1}^{3} a_{ik} x_k \right) = \sum_{i=1}^{3} \sum_{j,k=1}^{3} a_{ij} a_{ik} x_j x_k. \tag{2.13}$$

We can rearrange the summations and write this expression as

$$\sum_{j,k=1}^{3} \left( \sum_{i=1}^{3} a_{ij} a_{ik} \right) x_j x_k \tag{2.14}$$

which is equal to the right hand side of equation (2.12) if and only if

$$\sum_{i=1}^{3} a_{ij} a_{ik} = \delta_{jk} \qquad \text{for } j, k = 1,\ 2,\ 3 \tag{2.15}$$

where $\delta_{jk}$ is the Kronecker delta. If we express the $a_{ij}$ components in terms of the cosine coefficients, we now obtain a set of equations identical to equations (2.8).

Equation (2.15) is known as the orthogonality condition. The transition between coordinates fixed in space and coordinates fixed in the rigid body is achieved by means of an orthogonal transformation.

### 2.21 Transformation Matrix Properties

We will now think about what will happen when we perform two successive orthogonal transformations. Physically, this corresponds to the rigid body undergoing two successive displacements.

We can denote the first transformation as $\mathbf{B}$ :

$$\mathbf{x}' = \mathbf{B}\mathbf{x}. \tag{2.16}$$

In matrix form,

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{2.17}$$

We will denote the second transformation from $\mathbf{x}'$ to the next coordinate set $\mathbf{x}''$ by $\mathbf{A}$:

$$\mathbf{x}'' = \mathbf{A}\mathbf{x}' = \mathbf{A}\mathbf{B}\mathbf{x} \tag{2.18}$$

Or in matrix form:

$$\begin{aligned} \begin{pmatrix} x_1'' \\ x_2'' \\ x_3'' \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned} \tag{2.19}$$

We can also use component form to show that the application of $\mathbf{B}$ on system $\mathbf{x}$ is:

$$x_k = \sum_j b_{kj} x_j \tag{2.20}$$

and the application of $\mathbf{A}$ on $\mathbf{x}''$ is:

$$x_i'' = \sum_k a_{ik} x_k'$$

so combining these two equations above gives:

$$x_i'' = \sum_k a_{ik} \sum_j b_{kj} x_j \tag{2.21}$$

which is the same as

$$x_i'' = \sum_j \left( \sum_k a_{ik} b_{kj} \right) x_j. \tag{2.22}$$

If we set

$$c_{ij} = \sum_k a_{ik} b_{kj} \tag{2.23}$$

then we are able to rewrite equation (2.21) as:

$$x_i'' = \sum_j c_{ij} x_j. \tag{2.24}$$

We can therefore define an orthogonal transformation $\mathbf{C} = \mathbf{AB}$ where the elements of the square matrix $\mathbf{C}$ are given by the equation (2.23). This tells us that the successive transformation of orthogonal matrices $\mathbf{A}$ and $\mathbf{B}$ is completely equivalent to a third linear transformation $\mathbf{C}$.

It is worth noting that this matrix multiplication is generally not commutative:

$$\mathbf{AB} \neq \mathbf{BA}. \tag{2.25}$$

Matrix multiplication is, however, associative:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}). \tag{2.26}$$

In the addition of two transformation matrices:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{2.27}$$

and the elements of the matrix are $c_{ij} = a_{ij} + b_{ij}$.

The inverse transformation is one by which the new, primed system is transformed into the old, unprimed system, as shown by the following equation:

$$x_i = \sum_j a'_{ij} x'_j \tag{2.28}$$

where $a'_{ij}$ are the elements of the inverse transformation matrix $\mathbf{A}^{-1}$ .

This must be consistent with the set of equations:

$$x'_k = \sum_i a_{ki} x_i. \tag{2.29}$$

We can then substitute the $x_i$ from (2.28) into (2.29) to obtain:

$$\begin{aligned} x'_k &= \sum_i a_{ki} \sum_j a'_{ij} x'_j \\ &= \sum_j \left( \sum_i a_{ki} a'_{ij} \right) x_j \end{aligned} \tag{2.30}$$

which, since the $\mathbf{x}'$components are independent, can only possibly be correct if:

$$\sum_i a_{ki} a'_{ij} = \delta_{kj}. \tag{2.31}$$

If we let $\mathbf{A}$ and $\mathbf{A}^{-1}$ be the matrices with elements denoted by $a_{ki}$ and $a'_{ij}$ respectively, we see that equation (2.31) can instead be written as:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{1} \tag{2.32}$$

where $\mathbf{1}$ is known as the identity transformation:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.33}$$

Because $\mathbf{x} = \mathbf{1x}$, no change is produced in the coordinate system.

It can be shown that $\mathbf{A}$ and $\mathbf{A}^{-1}$ are commutative:

$$\mathbf{A} = \mathbf{1A} = (\mathbf{A}\mathbf{A}^{-1})\mathbf{A} = \mathbf{A}(\mathbf{A}^{-1}\mathbf{A}) = \mathbf{A1} = \mathbf{A} \tag{2.34}$$

and therefore

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{1}. \tag{2.35}$$

The transpose, $\mathbf{A}^T$ of matrix $\mathbf{A}$ can be obtained by interchanging the rows and columns:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}. \tag{2.36}$$

For an orthogonal transformation, it can be shown that $\mathbf{A}^{-1} = \mathbf{A}^T$ . We can prove this by first considering the sum:

$$\sum_{k,i} a_{kl} a_{ki} a'_{ij}.$$

If we first sum over $k$, we obtain the expression:

$$\sum_i \left( \sum_k a_{kl} a_{ki} \right) a'_{ij}.$$

If instead we sum over $i$ first, we get:

$$\sum_k \left( \sum_i a_{ki} a'_{ij} \right) a_{kl}.$$

We can then use the earlier obtained orthogonality condition to reduce the first sum to

$$\sum_i \delta_{il} a'_{ij} = a'_{lj}.$$

The second sum can also being reduced, using equation (2.31):

$$\sum_k \delta_{kj} a_{kl} = a_{jl}.$$

This indicates that the elements $a_{ij}$ of $\mathbf{A}$ and $a'_{ji}$ of the reciprocal $\mathbf{A}^{-1}$ are related by the equation

$$a'_{ij} = a_{ji} \tag{2.37}$$

and we have thus proved that $\mathbf{A}^{-1} = \mathbf{A}^T$ for orthogonal matrices. We can then substitute this result into the equation (2.35) to get the relation

$$\mathbf{A}^T\mathbf{A} = \mathbf{1} = \mathbf{A}\mathbf{A}^\mathbf{T}. \tag{2.38}$$

**2.22 Further Properties of Transformation Matrices**

We already know that if $\mathbf{A}$ is an orthogonal transformation matrix, then $\mathbf{x}' = \mathbf{A}\mathbf{x}$. If we then transform the coordinate system by another matrix, $\mathbf{B}$, the vector $\mathbf{x}'$ in this new coordinate system would now be:

$$\mathbf{B}\mathbf{x}' = \mathbf{B}\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}(\mathbf{B}^{-1}\mathbf{B})\mathbf{x} = (\mathbf{B}\mathbf{A}\mathbf{B}^{-1})\mathbf{B}\mathbf{x} \tag{2.39}$$

where $\mathbf{B}\mathbf{x}'$ is the vector $\mathbf{x}'$ in the new coordinate system, and $\mathbf{B}\mathbf{x}$ is the vector $\mathbf{r}$ in the new system. The matrix $\mathbf{A}' = (\mathbf{B}\mathbf{A}\mathbf{B}^{-1})$ is the transformation matrix $\mathbf{A}$ under the new coordinate system $\mathbf{B}$. Any transformation of a matrix having this form is known as a similarity transformation.

The determinant of a matrix is notated as:

$$\det(\mathbf{A}) = \mid \mathbf{A} \mid. \tag{2.40}$$

For two square matrices, $\mathbf{A}$ and $\mathbf{B}$, the following rule applies for their determinants:

$$\mid \mathbf{A}\mathbf{B} \mid = \mid \mathbf{A} \mid\mid \mathbf{B} \mid \tag{2.41}$$

If $\mathbf{A}$ is an orthogonal transformation matrix, then since the determinant of a unit matrix is 1 and the orthogonality condition indicates that $\mathbf{A}\mathbf{A}^\mathbf{T} = \mathbf{A}^\mathbf{T}\mathbf{A} = \mathbf{1}$, it becomes apparent that

$$\mid \mathbf{A} \mid\mid \mathbf{A}^\mathbf{T} \mid = \mid \mathbf{A}^\mathbf{T} \mid\mid \mathbf{A} \mid = 1, \tag{2.42}$$

implying that interchanging columns or rows of an orthogonal matrix does not have an effect on the determinant, i.e. $\mid \mathbf{A} \mid = \mid \mathbf{A}^\mathbf{T} \mid$. This then tells us that we can write:

$$\mid \mathbf{A} \mid^2 = 1 \tag{2.43}$$

which then allows us to conclude that $\mid \mathbf{A} \mid = \pm 1$ : there are only 2 possibilities for the determinant of an orthogonal matrix: $+1$ or -1.

Finally, we can show that the value of a determinant is invariant under a similarity transformation, by finding the determinant of both sides of a similarity transformation equation. For orthogonal matrices $\mathbf{A}$ and $\mathbf{B}$, where $\mathbf{A}' = (\mathbf{B}\mathbf{A}\mathbf{B}^{-1})$, this gives:

$$| \mathbf{A}' | = | \mathbf{B} || \mathbf{A} || \mathbf{B}^{-1} | = | \mathbf{A} | (| \mathbf{B} || \mathbf{B}^{-1} |) = | \mathbf{A} |. \tag{2.44}$$

### 2.3 The Euler Angles

As discussed earlier, we only require three independent parameters to describe the orientation of a rigid body. There are many choices of generalised coordinates that could be used to describe an arbitrary rotation of a coordinate system from one orientation to another, but Eularian angles are the most common and useful choice.

The transformation from an initial cartesian coordinate system to another can be achieved by means of three successive rotations performed in a specific sequence, with the three successive angles of rotation being defined as the Eularian angles.

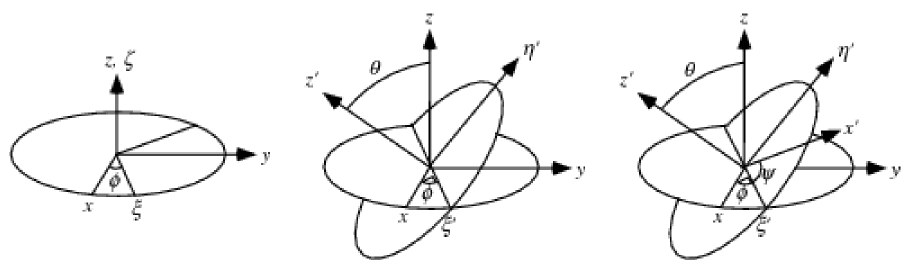


**Figure 4.2: The rotations which define the Eularian angles**[4]

The initial system of axes, $xyz$, is initially rotated anticlockwise by an angle $\phi$ about the z-axis, giving a transformation matrix $\mathbf{D}$:

$$\mathbf{D} = \begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.45}$$

This transformation results in a new system of axes, which will be labelled the $\xi\eta\zeta$ axes. The new set of axes are then rotated anticlockwise about the $\xi$ axis by an angle $\theta$. This transformation is given by the matrix $\mathbf{C}$:

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & sin\theta \\ 0 & -sin\theta & cos\theta \end{pmatrix} \tag{2.46}$$

This produces the intermediate set of axes, to be labelled the $\xi'\eta'\zeta'$axes. The intersection of the $xy$ and $\xi'\eta'$ planes is the $\xi'$ axis, and it is called the line of nodes.

The $\xi'\eta'\zeta'$ axes are then rotated anticlockwise by an angle $\psi$ about the $\zeta'$ axis, this transformation is given by the matrix $\mathbf{B}$:

$$\mathbf{B} = \begin{pmatrix} cos\psi & sin\psi & 0 \\ -sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.47}$$

[4]**image obtained from http://tabitha.phas.ubc.ca/wiki/index.php/Rigid_Bodies**

This gives the final set of axes, the $x'y'z'$system.

The Eularian angles $\phi$, $\theta$ and $\psi$ therefore specify the relative orientation of the $x'y'z'$ system to the $xyz$ system completely, and the the complete transformation matrix $\mathbf{A}$,where

$$\mathbf{x'} = \mathbf{Ax} \tag{2.48}$$

is simply the product of the three successive matrices:

$$\mathbf{A} = \mathbf{BCD} \tag{2.49}$$

where

$$\mathbf{A} = \begin{pmatrix} cos\psi cos\phi - cos\theta sin\phi sin\psi & cos\psi sin\phi + cos\theta cos\phi sin\psi & sin\psi sin\theta \\ -sin\psi cos\phi - cos\theta sin\phi cos\psi & -sin\psi sin\phi + cos\theta cos\phi cos\psi & cos\psi sin\theta \\ sin\theta sin\phi & -sin\theta cos\phi & cos\theta \end{pmatrix}. \tag{2.50}$$

### 2.4 Euler's Theorem on the Motion of a Rigid Body

At any instant, the orientation of a body can be specified by an orthogonal transformation matrix, $\mathbf{R}$. Orientation will change as time progresses, and the matrix is therefore a function of time, $\mathbf{R}(t)$.

We can choose the body-fixed axis to be coincident with the space axis at $t = 0$, so that $\mathbf{R}(0) = \mathbf{1}$ where $\mathbf{1}$ is the identity matrix as before. Physically, the motion must be continuous, so $\mathbf{R}(t)$ is a continuous function of time and it evolves continuously from the identity transformation.

Euler's theorem states that *any transformation in the 3-dimensional real space which has at least one fixed point can be described as a simple rotation about a single axis.* This axis, through the fixed point, means that the result of several consecutive rotations can be replaced by a single rotation. The axis of rotation is unaffected by the operation and so any vector lying along the axis of rotation must have identical components in both the initial and final axis.

### 2.41 Proof of Euler's Theorem

To prove Euler's theorem, we must show that there exists a vector $\mathbf{v}$ having the same components in both initial and final systems i.e. there must exist vector $\mathbf{v}$ unchanged by transformation $\mathbf{R}$.

$$\mathbf{v'} = \mathbf{Rv} = \mathbf{v}. \tag{2.51}$$

This is a special case of the more general equation,

$$\mathbf{v'} = \mathbf{Rv} = \lambda\mathbf{v} \tag{2.52}$$

where $\lambda$ is an unspecified constant, which is allowed to be complex. The values of $\lambda$ which allow equation (2.52) to be solved are called the eigenvectors of the matrix. The vector solutions are known as the eigenvectors of $\mathbf{R}$.

We are now able to state an alternative form of Euler's theorem:

*The real orthogonal matrix specifying the physical motion of a rigid body with one point fixed always has eigenvalue ± 1.*

We can therefore rewrite equation (2.52) as

$$(\mathbf{R} - \lambda\mathbf{1})\mathbf{v} = 0 \tag{2.53}$$

and restate it with $\mathbf{v} = (X_1\ X_2\ X_3)^{\mathrm{T}}$ (where we have chosen to denote the coordinates $x, y, z$ as $X_1, X_2, X_3$) to give 3 homogeneous simultaneous equations with 4 unknowns:

$$\begin{aligned}(r_{11} - \lambda)X_1 + r_{12}X_2 + r_{13}X_3 &= 0\\ r_{21}X_1 + (r_{22} - \lambda)X_2 + r_{23}X_3 &= 0\\ r_{31}X_1 + r_{32}X_2 + (r_{33} - \lambda)X_3 &= 0.\end{aligned} \tag{2.54}$$

From this set of equations (2.54), we are unable to exactly specify $X_1$, $X_2$ and $X_3$, but they do allow us to find ratios of the components. This corresponds to the situation where only the direction of the eigenvector can be fixed and the magnitude remains unknown, i.e. if $\mathbf{v}$ is an eigenvector then so is $t\mathbf{v}$ where $t$ is any scalar constant.

We only have non-trivial solutions to the set of equations (2.54) when the determinant of the coefficients is zero:

$$|\mathbf{R} - \lambda\mathbf{1}| = \begin{vmatrix} r_{11} - \lambda & r_{12} & r_{13} \\ r_{21} & r_{22} - \lambda & r_{23} \\ r_{31} & r_{32} & r_{33} - \lambda \end{vmatrix} = 0 \tag{2.55}$$

Equation (2.55) is known as the characteristic equation of the matrix. The values of $\lambda$ which allow the characteristic equation to be satisfied are the required eigenvalues.

We are now once again able to restate Euler's theorem: *the characteristic equation must have the root* $\lambda = \pm 1$. Generally, we will have 3 roots with 3 eigenvectors. The components of the eigenvectors may be labelled $X_{ik}$ , where the first subscript indicates which particular component is being considered, and the second subscript indicates which of the three eigenvectors is involved.

The set of equations (2.54) would then have typical members written as

$$\sum_j r_{ij} X_{jk} = \lambda_k X_{ik}. \tag{2.56}$$

We can then expand this for each $k$:

$$
\begin{aligned}
r_{11}X_{1k} + r_{12}X_{2k} + r_{13}X_{3k} &= \lambda_k X_{1k} \\
r_{21}X_{1k} + r_{22}X_{2k} + r_{23}X_{3k} &= \lambda_k X_{2k} \\
r_{31}X_{1k} + r_{32}X_{2k} + r_{33}X_{3k} &= \lambda_k X_{3k}.
\end{aligned} \tag{2.57}
$$

We are now able to put this into matrix form:

$$
\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \tag{2.58}
$$

and

$$
\boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{2.59}
$$

to give a new expression for the matrix equation:

$$
\mathbf{RX} = \mathbf{X}\boldsymbol{\lambda}. \tag{2.60}
$$

This allows us to write

$$
\mathbf{X^{-1}RX} = \boldsymbol{\lambda} \tag{2.61}
$$

by simply multiplying equation (2.60) by $\mathbf{X^{-1}}$ on the left hand side. Here, we are trying to diagonalise $\mathbf{R}$ by a similarity transformation equation. The elements of the diagonal matrix obtained by transforming $\mathbf{R}$ will then be the eigenvectors we are looking for. Euler's theorem for rigid body motion with one point fixed can thus be proved by using the orthogonality properties of $\mathbf{R}$.

We are trying to solve the characteristic equation, so we will begin by considering

$$
(\mathbf{R} - \mathbf{1})\mathbf{R}^T = \mathbf{1} - \mathbf{R}^T \tag{2.62}
$$

$$
\mathbf{R}^T\mathbf{R} = \mathbf{1} = \mathbf{R}\mathbf{R}^T \tag{2.63}
$$

and then take the determinant of both sides of equation (2.62):

$$
\mid \mathbf{R} - \mathbf{1} \mid\mid \mathbf{R}^T \mid = \mid \mathbf{1} - \mathbf{R}^T \mid. \tag{2.64}
$$

In order to describe rigid body motion, the matrix $\mathbf{R}$ must correspond to a proper rotation, i.e.

$$
\mid \mathbf{R} \mid = \mid \mathbf{R}^T \mid = +1 \tag{2.65}
$$

and we can therefore rewrite equation (2.64) as

$$
\mid \mathbf{R} \text{ - } \mathbf{1} \mid = \mid \mathbf{1} - \mathbf{R} \mid. \tag{2.66}
$$

We can set $\mathbf{D} = \mathbf{R}$ - $\mathbf{1}$, and then rewrite the above equation again as:

$$\mid \mathbf{D} \mid = \mid -\mathbf{D} \mid. \tag{2.67}$$

If $\mathbf{D}$ is an n x n matrix, then it has the well-known property that

$$\mid -\mathbf{D} \mid = (-1)^n \mid \mathbf{D} \mid \tag{2.68}$$

and so for our 3 x 3 matrix,

$$\mid -\mathbf{D} \mid = - \mid \mathbf{D} \mid \tag{2.69}$$

and therefore equation (2.67) can now be written as

$$\mid \mathbf{D} \mid = - \mid \mathbf{D} \mid. \tag{2.70}$$

Obviously this can only be true for $\mid \mathbf{D} \mid = 0$ and therefore, $|\mathbf{R}$ - $\mathbf{1}| = 0$ . We can then solve the characteristic equation (2.55) with $\lambda = +1$:

$$|\mathbf{R} - \lambda\mathbf{1}| = |\mathbf{R} - (+1)\mathbf{1}| = 0. \tag{2.71}$$

This implies that $\lambda = +1$ is an eigenvalue of $\mathbf{R}$. We shall now seek to find the remaining two eigenvalues. The determinant of the matrix $\boldsymbol{\lambda}$ is the product of the three eigenvalues $\lambda_1\lambda_2\lambda_3$. The determinant of a matrix is unaffected by a similarity transformation. $\mathbf{X^{-1}RX} = \boldsymbol{\lambda}$ is a similarity transformation and therefore the determinant of $\mathbf{R}$ is equal to the determinant of $\boldsymbol{\lambda}$:

$$\mid \mathbf{R} \mid = \lambda_1\lambda_2\lambda_3 \ . \tag{2.72}$$

We already know that $\mid \mathbf{R} \mid = +1$ and that one of the eigenvectors, $\lambda_3$ say, is equal to $+1$, and we are left with the equation:

$$\lambda_1\lambda_2 = +1. \tag{2.73}$$

We know that $\mathbf{R}$ is real, and so if $\lambda$ is a solution to the characteristic equation, then its complex conjugate $\lambda^*$ is also a solution.

It is a property of complex numbers that $\parallel c \parallel = \parallel c^* \parallel$for $c \in \mathbb{C}$. From this we know that all the eigenvalues have unit magnitude.

All three eigenvalues of the real orthogonal matrix with determinant $+1$ may be real. There is the trivial solution that $\lambda_1 = \lambda_2 = \lambda_3 = 1$. There is also the possibility for there to be one real root and two complex roots. The real root is $\lambda_3 = +1$ and the two complex roots must be complex conjugates and have product $+1$, to give the correct determinant. There is thus one and only one $+1$ eigenvalue in any non-trivial physical transformation. This is the statement of Euler's theorem.

We can obtain the direction cosines of the axes of rotation by finding the eigenvector associated with $\lambda = +1$ in equation (2.53) and scaling it to unit length. We can find the angle of rotation, $\phi$, without too much difficulty. We are able to use a similarity transformation to transform $\mathbf{R}$ to a coordinate system where the z-axis lies along the axis of rotation. The matrix $\mathbf{R'}$ represents the rotation through the angle $\phi$ about the z-axis. It has the form

$$\mathbf{R}' = \begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.74}$$

The trace of $\mathbf{R}'$ is then

$$1 + 2\cos\phi. \tag{2.75}$$

Under a similarity transformation, the trace is always invariant. This must therefore also be the value of the trace of $\mathbf{R}$ :

$$\sum_i r_{ii} = 1 + 2\cos\phi. \tag{2.76}$$

We can then solve for $\phi$.

Chalses' theorem immediately follows from Euler's theorem, and it states that *the most general displacement of a rigid body is a translation plus a rotation.*

### 2.5 Infinitesimal Rotations

A finite rotation about the z-axis has a rotation matrix of the form

$$\mathbf{A} = \begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.77}$$

Two rotations performed one after the other, i.e. the addition of two rotations, corresponds to the product of the two matrices, **AB**. Non-commutativity of matrix multiplication allows us to conclude that **A**,**B** are not commutative in addition and thus cannot be considered as vectors. This leads to the conclusion that the sum of finite rotations is dependent on the order in which the rotations are carried out. The same is not true if we consider infinitesimal rotations. Goldstein defines an infinitesimal rotation as: *"an orthogonal transformation of coordinate axes in which the components of a vector are almost the same in both sets of axes - the change is infinitesimal"* .

An infinitesimal transformation of a vector $\mathbf{x}'$ is given by

$$\mathbf{x}' = (\mathbf{1} + \boldsymbol{\varepsilon})\mathbf{x} \tag{2.78}$$

where the vector $\mathbf{1}$ is the identity matrix and $\boldsymbol{\varepsilon}$ is infinitesimal.

It can be shown that two infinitesimal transformations $\varepsilon_1$and $\varepsilon_2$ are commutative, by the equivalence of

$$(\mathbf{1} + \varepsilon_1)(\mathbf{1} + \varepsilon_2) = \mathbf{1}^2 + \varepsilon_1\mathbf{1} + \mathbf{1}\boldsymbol{\varepsilon_2} + \boldsymbol{\varepsilon_1}\boldsymbol{\varepsilon_2} \approx \mathbf{1} + \varepsilon_1 + \boldsymbol{\varepsilon_2}$$
$$(\mathbf{1} + \varepsilon_2)(\mathbf{1} + \varepsilon_1) = \mathbf{1}^2 + \varepsilon_2\mathbf{1} + \mathbf{1}\boldsymbol{\varepsilon_1} + \boldsymbol{\varepsilon_2}\boldsymbol{\varepsilon_1} \approx \mathbf{1} + \varepsilon_2 + \boldsymbol{\varepsilon_1} \tag{2.79}$$

We can define $\mathbf{A} \equiv \mathbf{1}+\boldsymbol{\varepsilon}$ as the transformation matrix, and the inverse of this is

$$\mathbf{A}^{-1} = \mathbf{1} - \boldsymbol{\varepsilon} \tag{2.80}$$

since

$$\mathbf{A}\mathbf{A}^{-1} = (\mathbf{1}+\boldsymbol{\varepsilon})(\mathbf{1}-\boldsymbol{\varepsilon}) = \mathbf{1}-\boldsymbol{\varepsilon}^2 \approx \mathbf{1}. \tag{2.81}$$

For small angles, $\sin\phi \approx d\phi$ and $\cos\phi \approx 1$. From this we can obtain the matrix of infinitesimal rotation:

$$\mathbf{1}+\boldsymbol{\varepsilon} = \begin{pmatrix} 1 & d\phi & 0 \\ -d\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.82}$$

And the infinitesimal matrix $\boldsymbol{\varepsilon}$ is therefore

$$\begin{aligned} \boldsymbol{\varepsilon} &= \begin{pmatrix} 0 & d\phi & 0 \\ -d\phi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= d\phi \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{2.83}$$

Since the infinitesimal transformation is defined to be a rotation, orthogonality of rotation matrices requires

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1} \tag{2.84}$$

but $\mathbf{A}^{-1} = \mathbf{1} - \boldsymbol{\varepsilon}$, so

$$(\mathbf{1}+\boldsymbol{\varepsilon})^{\mathrm{T}} = \mathbf{1}^{\mathrm{T}} + \boldsymbol{\varepsilon}^{\mathrm{T}} = \mathbf{1}+\boldsymbol{\varepsilon}^{\mathrm{T}} \tag{2.85}$$

and so

$$\boldsymbol{\varepsilon} = -\boldsymbol{\varepsilon}^{\mathrm{T}} \tag{2.86}$$

which is the definition of an antisymmetric matrix.

Given that $\boldsymbol{\varepsilon}$ is an antisymmetric matrix, the diagonal elements are therefore equal to zero and there can only possibly be three distinct elements in $\boldsymbol{\varepsilon}$.

The matrix $\boldsymbol{\varepsilon}$ must therefore be of the form

$$\boldsymbol{\varepsilon} = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \tag{2.87}$$

where $d\Omega_1$, $d\Omega_2$ and $d\Omega_3$ are quantities associated with the three independent parameters responsible for specifying the rotation of the body.

When the infinitesimal rotation matrix is applied to the vector $\mathbf{x}$, the differential change in $\mathbf{x}$ is given by

$$d\mathbf{x} = \mathbf{x}' - \mathbf{x} = (\mathbf{1}+\varepsilon)\mathbf{x} - \mathbf{x} = \varepsilon\mathbf{x}. \tag{2.88}$$

In matrix form, this is

$$\begin{aligned} d\mathbf{x} &= \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} yd\Omega_3 - zd\Omega_2 \\ zd\Omega_1 - xd\Omega_3 \\ xd\Omega_2 - yd\Omega_1 \end{pmatrix} \end{aligned}$$

and so

$$\begin{aligned} d\mathbf{x} &= (yd\Omega_3 - xd\Omega_2)\mathbf{i} + (zd\Omega_1 - xd\Omega_3)\mathbf{j} + (xd\Omega_2 - yd\Omega_1)\mathbf{k} \\ &= \mathbf{x} \times d\mathbf{\Omega} \end{aligned} \tag{2.89}$$

This implies that

$$\begin{aligned} \left(\tfrac{d\mathbf{x}}{dt}\right)_{\text{rotation}} &= \mathbf{x} \times \left(\tfrac{d\mathbf{\Omega}}{dt}\right) \\ &= \mathbf{x} \times \boldsymbol{\omega} \end{aligned} \tag{2.90}$$

The vector $\boldsymbol{\omega}$ here is the angular velocity of the body. It is defined as the instantaneous angular rate of rotation of the body,

$$\boldsymbol{\omega} = \tfrac{d\mathbf{\Omega}}{dt}. \tag{2.91}$$

## 2.6 Rate of Change of a Vector

If we set $\mathbf{R}$ as an arbitrary vector involved in describing motion of a rigid body with time, how the vector will vary in time as the body is in motion will be dependent on the coordinate system of observation. Only the effects of rotation of the body axes will result in the components of $\mathbf{R}$ with respect to the body axes differing from the components of $\mathbf{R}$ with respect to the space axes.

This allows us to write:

$$(d\mathbf{R})_{\text{body}} = (d\mathbf{R})_{\text{space}} + (d\mathbf{R})_{\text{rotation}}. \tag{2.92}$$

But as we saw before,

$$(d\mathbf{R})_{\text{rotation}} = \mathbf{R} \times d\mathbf{\Omega}. \tag{2.93}$$

The differential $d\mathbf{R}$ observed in the space set therefore is related to the same differential when observed in the body set by the following equation:

$$(d\mathbf{R})_{\text{space}} = (d\mathbf{R})_{body} + d\mathbf{\Omega} \times \mathbf{R}. \tag{2.94}$$

We can then obtain the rate of change of the vector $\mathbf{R}$ with respect to time:

$$\left(\tfrac{d\mathbf{R}}{dt}\right)_{\text{space}} = \left(\tfrac{d\mathbf{R}}{dt}\right)_{\text{body}} + \boldsymbol{\omega}\times \mathbf{R}. \tag{2.95}$$

This can be written as an operator equation due to the arbitrary nature of $\mathbf{R}$ :

$$\left(\tfrac{d}{dt}\right)_{\text{space}} = \left(\tfrac{d}{dt}\right)_{\text{body}} + \boldsymbol{\omega} \times \ . \tag{2.96}$$

Caution should be taken when taking the time derivative of a vector with respect to one coordinate system, as components can only be taken along a different set of coordinate axes *after* differentiation has been carried out.

It will prove useful to be able to use the Euler angles and their time derivatives to express the angular velocity vector, $\boldsymbol{\omega}$.We can consider the general infinitesimal rotation associated with the angular velocity vector to be made up of three sequential infinitesimal rotations with the angular velocities $\omega_\phi = \dot{\phi}$, $\omega_\theta = \dot{\theta}$ and $\omega_\psi = \dot{\psi}$. The angular velocity vectors do not have symmetrically placed directions: $\boldsymbol{\omega}_\phi$ is along the space z-axis, $\boldsymbol{\omega}_\theta$ points along the line of nodes and $\boldsymbol{\omega}_\psi$ is along the body $z'$-axis. In order to obtain $\boldsymbol{\omega}$ as the sum of the three angular velocities we need to use the orthogonal transformations described in section (2.3) to get the components of the vectors $\boldsymbol{\omega}_\phi$, $\boldsymbol{\omega}_\theta$ and $\boldsymbol{\omega}_\psi$ along any set of axes we desire.

We shall obtain the components of $\boldsymbol{\omega}$ for the set of body axes, as this will be the most useful coordinate system to consider.

Being parallel to the space z-axis, $\boldsymbol{\omega}_\phi$ requires the complete orthogonal transformation $\mathbf{A} = \mathbf{BCD}$ to be applied to obtain its components along the body axes:

$$(\boldsymbol{\omega}_\phi)_{x'} = \dot{\phi} sin\theta sin\psi \qquad (\boldsymbol{\omega}_\phi)_{y'} = \dot{\phi} sin\theta cos\psi \qquad (\boldsymbol{\omega}_\phi)_{z'} = \dot{\phi} cos\theta. \tag{2.97}$$

Since the direction of $\boldsymbol{\omega}_\theta$ lies along the line of nodes, which coincides with the $\xi'$-axis, we can find the components of $\omega_\theta$ along the body axes by applying the last orthogonal transformation in the sequence, $\mathbf{B}$ (2.47):

$$(\boldsymbol{\omega}_\theta)_{x'} = \dot{\theta} cos\psi \qquad (\boldsymbol{\omega}_\theta)_{y'} = -\dot{\theta} sin\psi \qquad (\boldsymbol{\omega}_\theta)_{z'} = 0. \tag{2.98}$$

The vector $\boldsymbol{\omega}_\psi$ lies along the $z'$-axis and so does not require an orthogonal transformation. We now are able to add the separate angular velocity components with respect to the body axes and we obtain:

$$\begin{aligned} \omega_{x'} &= \dot{\phi} sin\theta sin\psi + \dot{\theta} cos\psi \\ \omega_{y'} &= \dot{\phi} sin\theta cos\psi - \dot{\theta} sin\psi \\ \omega_{z'} &= \dot{\phi} cos\theta + \dot{\psi}. \end{aligned} \tag{2.99}$$

# Chapter 3

# Equations of Motion of Rigid Bodies

In this chapter we will look into the nature of rigid body motion, using the ideas and techniques explained in the previous chapter. We will be using the Euler angles as generalised coordinates and will apply other tools previously described to obtain the equations of motion for rigid bodies.

## 3.1 Angular Momentum of a Rigid Body About a Fixed Point

In general rigid body rotation, every particle lies a fixed distance from the origin and a fixed angle to the rotation axis. The velocity $\mathbf{v}_i$ of a rigid body about a fixed point is therefore

$$\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i \tag{3.1}$$

where $\mathbf{r}_i$ is the radius vector of the $i$th particle, relative to the fixed point. Thus the total angular momentum of the rigid body about the fixed point, given by

$$\mathbf{J} = \sum_i m_i (\mathbf{r}_i \times \mathbf{v}_i) \tag{3.2}$$

can now be written as

$$\mathbf{J} = \sum_i m_i (\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i )). \tag{3.3}$$

We can expand the triple cross product to obtain

$$\mathbf{J} = \sum_i m_i (\mathbf{r}_i{}^2 \boldsymbol{\omega} - (\mathbf{r}_i \cdot \boldsymbol{\omega})\, \mathbf{r}_i ). \tag{3.4}$$

Since $\mathbf{J}$ lies in the same plane as $\mathbf{r}$, it rotates as $\mathbf{r}$ rotates.

The components of $\mathbf{J}$ are given by

$$\begin{aligned}
J_x &= \omega_x \sum_i m_i (r_i^2 - x_i^2) - \omega_y \sum_i m_i x_i y_i - \omega_z \sum_i m_i x_i z_i \\
J_y &= \omega_y \sum_i m_i (r_i^2 - y_i^2) - \omega_x \sum_i m_i y_i x_i - \omega_z \sum_i m_i y_i z_i \\
J_z &= \omega_z \sum_i m_i (r_i^2 - z_i^2) - \omega_x \sum_i m_i z_i x_i - \omega_y \sum_i m_i z_i y\ .
\end{aligned} \tag{3.5}$$

Angular momentum of the body is therefore linearly related to its angular velocity.

The linear transformation can be written in matrix form:

$$\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \tag{3.6}$$

so that $J_x = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z$ and similarly for $J_y$ and $J_z$.

The elements $I_{xx}$, $I_{yy}$ and $I_{zz}$ are known as the *moment of inertia coefficients*. For a rigid body made up of discrete particles,

$$I_{xx} = \sum_i m_i(r_i^2 - x_i^2). \tag{3.7}$$

with similar relations for $I_{yy}$ and $I_{zz}$. For continuous bodies,

$$I_{xx} = \int_{dV} \rho(r)(r^2 - x^2)dV \tag{3.8}$$

where $\rho(r)$ is the mass density of the body.

The rest of the elements of the elements are known as *products of inertia* and they have the form given by

$$I_{xy} = -\sum_i m_i x_i y_i. \tag{3.9}$$

$\mathbf{J}$ and $\boldsymbol{\omega}$are vectors that are independent of our choice of axes. The relation between them can therefore be expressed as

$$\mathbf{J} = \underline{\mathbf{I}} \cdot \boldsymbol{\omega} \tag{3.10}$$

where $\underline{\mathbf{I}}$ is defined as the moment of inertia tensor.

### 3.2 Tensors

A tensor acts on a vector to produce a new vector which is linearly related to the old one but will, in general, have a different direction. The action of a tensor on a vector $\mathbf{a}$ can be denoted by

$$\mathbf{b} = \underline{\mathbf{T}} \cdot \mathbf{a} \tag{3.11}$$

which can be written as

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{3.12}$$

or

$$T_i = \sum_j T_{ij} a_j. \tag{3.13}$$

Familiar methods of multiplying two vectors are by means of a dot product or cross product, but there is a third type of multiplication, a dyadic product, which will be useful here.

A dyadic product, denoted $\otimes$, produces a tensor:

$$\mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix}. \tag{3.14}$$

A tensor of this type is known as a tensor of the 2nd rank. A tensor of zero rank has only one component, which is invariant under orthogonal transformation. A scalar is a tensor of zero rank. A tensor of first rank is completely equivalent to a vector.

The most important property of a vector is the way its components transform under a rotation of the coordinate axes; keeping its geometrical or physical meaning invariant.

### 3.3 Moment of Inertia Tensor

The kinetic energy of energy about a point is given by

$$T = \tfrac{1}{2} \sum_i m_i v_i^2 \tag{3.15}$$

where $\mathbf{v_i}$ is the velocity relative to a fixed point and $\mathbf{v_i} = \boldsymbol{\omega} \times \mathbf{r_i}$.

This allows us to write the kinetic energy as

$$T = \tfrac{1}{2} \sum_i m_i \mathbf{v_i} \cdot (\boldsymbol{\omega} \times \mathbf{r_i}) \tag{3.16}$$

which is equivalent to

$$T = \tfrac{\boldsymbol{\omega}}{2} \cdot \sum_i m_i \cdot (\mathbf{r_i} \cdot \mathbf{v_i}) \tag{3.17}$$

and so the rotation kinetic energy corresponding to equation (3.10) has the form:

$$T = \tfrac{1}{2}\, \boldsymbol{\omega} \cdot \underline{\mathbf{I}} \cdot \boldsymbol{\omega} \tag{3.18}$$

If we set $\mathbf{n}$ to be a unit vector in the direction of $\boldsymbol{\omega}$, we are then able to define $\boldsymbol{\omega} = \omega \mathbf{n}$ , and can rewrite the kinetic energy in the form

$$T = \tfrac{1}{2} \omega^2\ \mathbf{n} \cdot \underline{\mathbf{I}} \cdot \mathbf{n} = \tfrac{1}{2} I \omega^2 \tag{3.19}$$

where we define $I$, the moment of inertia about the axis of rotation, as

$$I = \mathbf{n} \cdot \underline{\mathbf{I}} \cdot \mathbf{n} = \sum_i m_i (r_i^2 \text{ - } (\mathbf{r_i} \cdot \mathbf{n})^2). \tag{3.20}$$

The direction of the axis of rotation has an effect on the value of the moment of inertia. The direction of $\boldsymbol{\omega}$ changes in time with respect to the body, and so the moment of inertia must also be a function of time. It is possible to keep the moment of inertia constant by constraining the body to allow it to rotate only about a fixed axis. If this constraint is applied then the form of kinetic energy in equation (3.19) is very close to the form we need to set up the Lagrangian for the system. The last step required is to express $\boldsymbol{\omega}$ as the time derivative of some angle. This is usually possible without too much difficulty.

### 3.4 The Principal Axis

We seek to show that any rigid body has three principal axes that can be chosen so that they are always mutually orthogonal. The equation (3.9) defining the products of inertia shows that the components of the inertia tensor are in fact symmetrical, that is

$$I_{ij} = I_{ji}. \tag{3.21}$$

This indicates that the only six coordinates of the tensor are independent.

We have already seen the relationship between a rigid body's angular momentum and angular velocity in equation (3.10), where $\underline{\mathbf{I}}$ is now known to be a tensor of the second rank, known as the moment of inertia tensor. If the angular momentum of a body is parallel to its angular velocity, then the rigid body is described as being *dynamically balanced*:

$$\mathbf{J} = \underline{\mathbf{I}} \cdot \boldsymbol{\omega} = I\boldsymbol{\omega} \tag{3.22}$$

where $I$ is some scalar number. The angular velocity $\boldsymbol{\omega}$ must point along a principal axis of the moment of inertia tensor for this to be true. The value of $I$ is then called the *principal moment of inertia*.

The eigenvectors of the tensor $\underline{\mathbf{I}}$ provide the principal axes, while the eigenvalues give the principal moments. We know that the eigenvectors satisfy the linear set of equations:

$$\begin{pmatrix} I_{xx} - \lambda & I_{xy} & I_{zx} \\ I_{xy} & I_{yy} - \lambda & I_{yz} \\ I_{zx} & I_{yz} & I_{zz} - \lambda \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = 0. \tag{3.23}$$

Here, the symmetry of $\underline{\mathbf{I}}$ has explicitly been displayed.

Non-trivial solutions for $\boldsymbol{\omega}$ will exist only for the case where the determinant vanishes:

$$\begin{vmatrix} I_{xx}-\lambda & I_{xy} & I_{zx} \\ I_{xy} & I_{yy}-\lambda & I_{yz} \\ I_{zx} & I_{yz} & I_{zz}-\lambda \end{vmatrix} = 0. \tag{3.24}$$

The characteristic equation is cubic in $\lambda$, and there will be three solutions: the desired principal moments $I_1$, $I_2$ and $I_3$. The eigenvectors can then be obtained by substituting each eigenvalue back into the set of linear equations in turn. Only two of the equations we get will be independent and we can then use these to work out the direction of the corresponding eigenvector. Unit eigenvectors, $\mathbf{e_i}$, $\mathbf{e_j}$ and $\mathbf{e_k}$ will be used to specify the directions of the principal axes, as the magnitudes of the eigenvectors are not determined.

Symmetry of the moment of inertia tensor means that its eigenvalues and eigenvectors have similar properties to those of a real symmetric matrix. The eigenvalues of any symmetric matrix are real. The principal moments are therefore also real.

If $\mathbf{e_i}$ and $\mathbf{e_j}$ are the eigenvectors corresponding to different principal moments, then these eigenvectors (principal axes) are orthogonal:

$$\mathbf{e_i} \cdot \mathbf{e_j} = 0. \tag{3.25}$$

If there is a repeated root of the characteristic equation then the rigid body has degenerate principal moments and any vector that is in the plane of $\mathbf{e_i}$ and $\mathbf{e_j}$ can be a principal axis and any pair of suitable vectors in this plane can be chosen as principal axes. If a body has an axis of symmetry then that axis is a principal axis and rotations around that axis will be dynamically balanced. We can thus use the symmetries of a rigid body in order to recognise the principal axes: If there is an axis of symmetry through the origin of the body then this axis is a principal axis for rotations about the origin. Another principal axis is the normal to a plane of reflection symmetry through the origin.

We can use these properties along with the property that any rigid body has three principal axes that can always be chosen so that they are mutually orthogonal, to determine the three principal axes for many rigid bodies. If the coordinate axes are chosen to lie along the directions of the principal axes, then the principal moments can be found more easily, they are just the moments of inertia about our three principal axes.

There is another method that can be used to approach the idea of principal axes. Earlier, we defined the moment of inertia about a given axis as $I = \mathbf{n} \cdot \underline{\mathbf{I}} \cdot \mathbf{n}$ . If we let the direction cosines of the axis be $\alpha$, $\beta$ and $\gamma$, and the unit vectors be denoted by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, then

$$\mathbf{n} = \alpha\mathbf{i} + \beta\mathbf{j} + \gamma\mathbf{k}. \tag{3.26}$$

We can then write $I$ as

$$I = I_{xx}\alpha^2 + I_{yy}\beta^2 + I_{zz}\gamma^2 + 2I_{xy}\alpha\beta + 2I_{yz}\beta\gamma + 2I_{zx}\gamma\alpha, \tag{3.27}$$

once again using the symmetry of $\underline{\mathbf{I}}$ explicitly.

It is useful to define a new vector $\boldsymbol{\rho}$ as

$$\boldsymbol{\rho} = \frac{\mathbf{n}}{\sqrt{I}} \tag{3.28}$$

and its magnitude is related to the moment of inertia about the axis whose direction is provided by $\mathbf{n}$. The equation of some surface in $\rho$ space is a function of the three variables $\rho_1$, $\rho_2$ and $\rho_3$:

$$1 = I_{xx}\rho_1^2 + I_{yy}\rho_2^2 + I_{zz}\rho_{33}^2 + 2I_{xy}\rho_1\rho_2 + 2I_{yz}\rho_2\rho_3 + 2I_{zx}\rho_3\rho_1. \tag{3.29}$$

This is the equation of the inertia ellipsoid. It is possible to transform to a set of cartesian axes so that the ellipsoid equation takes on the normal form with the principal axes of the ellipsoid along the new coordinate axes:

$$1 = I_1{\rho'_1}^2 + I_2{\rho'_2}^2 + I_3{\rho'_3}^2 \ . \tag{3.30}$$

Equation (3.30) is the same as equation (3.29) when the inertia tensor is diagonal. The coordinate transformation which turns the ellipsoid equation into normal form is therefore the same as the previously discussed principal axis transformation. The lengths of the inertia ellipsoid axes are determined by the principal moments of inertia. The inertia ellipsoid will be an ellipsoid of revolution if two of the roots of the characteristic equation are equal because the inertia ellipsoid will therefore have two equal axes.

In the case of all principal moments being equal, the inertia ellipsoid is in fact a sphere.

### 3.5 Euler's Equations of Motion

Six generalised coordinates are required to describe the motion of an unconstrained rigid body: 3 cartesian coordinates to describe the translational motion and three Euler angles to describe the rotational motion.

Kinetic energy can be written as the combination of the translational energy of the centre of mass and the rotational energy about the centre of mass:

$$T = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2. \tag{3.31}$$

If we work in the principal axes, rotational energy has the form

$$\tfrac{1}{2}I_1\omega_x^2 + \tfrac{1}{2}I_2\omega_y^2 + \tfrac{1}{2}I_3\omega_z^2 \tag{3.32}$$

and if one point in the body is fixed then the kinetic energy only contains rotational terms. If the forces are conservative, we can write the Lagrangian as

$$L = T - V = \tfrac{1}{2}(I_1\omega_x^2 + I_2\omega_y^2 + I_3\omega_z^2) - V(\theta,\phi,\psi) \tag{3.33}$$

with Euler angles being used to express the components of $\boldsymbol{\omega}$ .

The components of the tensor $\underline{\mathbf{I}}$ form a diagonal matrix:

$$\underline{\mathbf{I}} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{3.34}$$

and $I_{1,}I_2,I_3$ are the principal moments of inertia.

When we are considering the motion of a rigid body with one point fixed, it is useful to use Euler's equations of motion rather than the Lagrange equations of motion. We can derive the Euler equations of motion by considering the definition of *torque*. In an inertial frame, torque is equal to rate of change of angular momentum:

$$\frac{d\mathbf{J}}{dt} = \boldsymbol{\tau}. \tag{3.35}$$

In this case, the time derivative is referring to the space axes. In a body-fixed frame, this equation of motion becomes

$$\left(\frac{d\mathbf{J}}{dt}\right)_{body} + \boldsymbol{\omega} \times \mathbf{J} = \boldsymbol{\tau}. \tag{3.36}$$

The component of the equation of motion along the $x$ principal axis is then

$$\frac{d\mathbf{J_x}}{dt} + \omega_y J_z - \omega_z J_y = \tau_x \tag{3.37}$$

In this body-fixed frame, the tensor $\underline{\mathbf{I}}$ only depends on the positions of atoms in the body, and due to the rigidity condition on the body, these particles do not move and $\underline{\mathbf{I}}$ is therefore constant.

The equation $\mathbf{J} = \underline{\mathbf{I}} \cdot \boldsymbol{\omega}$ can then be used to re-write the equation of motion as

$$\underline{\mathbf{I}} \cdot \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (\underline{\mathbf{I}} \cdot \boldsymbol{\omega}) = \boldsymbol{\tau}. \tag{3.38}$$

This allows us to reduce the equation (3.37) to

$$I_1\dot{\omega}_x - \omega_y\omega_z(I_2 - I_3) = \tau_x.$$

### 3.6 Force Free Motion of a Rigid Body

If we consider a rigid body subject to no net forces or torques, the centre of mass of the body must be either at rest or moving uniformly. Without loss of generality, we are able to consider the frame where the centre of mass is stationary. In this case, rotation about the centre of mass is the sole motion from which the angular momentum arises.

Euler's equations are thus the equations of motion of the system, and without the presence of net torques they become:

$$
\begin{aligned}
I_1\dot{\omega}_x &= \omega_y\omega_z(I_2 - I_3) \\
I_2\dot{\omega}_y &= \omega_z\omega_x(I_3 - I_1) \\
I_3\dot{\omega}_z &= \omega_x\omega_y(I_1 - I_2).
\end{aligned}
\tag{3.39}
$$

These equations will also apply to the motion of a rigid body with no net torques and with one point fixed.

### 3.61 Poinsot's Construction

Poinsot's construction is a geometric description of the motion which we are able to derive without needing a complete solution to the problem above. It is a perfectly adequate method of completely describing force-free rigid body motion.

Earlier, we defined a vector $\boldsymbol{\rho}$, equation (3.28), whose components are measured by the principal axes of a body, with the coordinate system also oriented along the principal axes of the body. We can define a function in this space:

$$F(\rho) = \boldsymbol{\rho} \cdot \underline{\mathbf{I}} \cdot \boldsymbol{\rho} \tag{3.40}$$

and the surfaces of constant $F$ are ellipsoids, with the inertia ellipsoid being defined by the surface with $F = 1$. The vector $\boldsymbol{\rho}$ is parallel to the axis of rotation and it moves accordingly as the rotation axis changes in time. There is always a point on the inertia ellipsoid that is defined by the tip of $\boldsymbol{\rho}$, and if the gradient of $F$ is evaluated at this point then it gives the direction of the normal vector to the inertia ellipsoid.

Earlier we saw that $\underline{\mathbf{I}}$ has diagonal form in the principal axes. This makes it simple for us to evaluate the partial derivative of $F$ with respect to $\rho_1$:

$$\frac{\partial F}{\partial \rho_1} = 2I_1\rho_1. \tag{3.41}$$

We can also define $\boldsymbol{\rho}$ as

$$\boldsymbol{\rho} = \frac{\boldsymbol{\omega}}{\omega\sqrt{I}} \tag{3.42}$$

which then allows us to write

$$\frac{\partial F}{\partial \rho_1} = \frac{2}{\omega\sqrt{I}} I_1\omega_x = \frac{2}{\omega\sqrt{I}} J_x \tag{3.43}$$

and similar results follow for the other components of the gradient of $F$ :

$$
\begin{aligned}
\frac{\partial F}{\partial \rho_2} &= \frac{2}{\omega\sqrt{I}} J_y \\
\frac{\partial F}{\partial \rho_3} &= \frac{2}{\omega\sqrt{I}} J_z.
\end{aligned}
\tag{3.44}
$$

The angular velocity vector $\boldsymbol{\omega}$ always moves so that the normal corresponding to the inertia ellipsoid is always in the same direction as the angular momentum, $\mathbf{J}$. In this case, it is the inertia ellipsoid which moves in space, while the direction of $\mathbf{J}$ is fixed to preserve the relationship between the angular momentum and angular velocity.

The distance between the origin of the ellipsoid and its tangent plane at $\boldsymbol{\rho}$ must always be fixed. This fixed distance is given by

$$\frac{\boldsymbol{\rho}\cdot\mathbf{J}}{J} = \frac{\boldsymbol{\omega}\cdot\mathbf{J}}{\omega J\sqrt{I}} = \frac{2T}{J\sqrt{I\omega^2}} \tag{3.45}$$

and is the projection of $\boldsymbol{\rho}$ on $\mathbf{J}$.

This can also be written as

$$\frac{\boldsymbol{\rho}\cdot\mathbf{J}}{J} = \frac{\sqrt{2T}}{J} \tag{3.46}$$

where the angular momentum $J$ and the kinetic energy $T$ are just constants of motion. The normal to the plane has a fixed direction along the angular momentum vector $\mathbf{J}$, and so the tangent plane is known as the invariable plane. The motion of the force-free rigid body can be pictured as being such that the inertia ellipsoid rolls, on the invariable plane, without slipping, with the ellipsoid's centre always remaining a constant height above the plane. The point of contact of the ellipsoid with the plane is defined by the position of $\boldsymbol{\rho}$**.** Because $\boldsymbol{\rho}$ is along the instantaneous axis of rotation, the body is momentarily at rest in this direction.

The non-plane curve traced on the surface of the inertia ellipsoid with fixed centre by its point of contact with the fixed plane on which it rolls is known as the *polhode*. The curve traced out on the invariable plane by the point of contact between the plane and the inertia is known as the *herpolhode*.

The values of the kinetic energy $T$ and the angular momentum $\mathbf{J}$ determine the direction of the invariable plane and the height of the inertia ellipsoid above the plane. The direction of $\boldsymbol{\rho}$ furnishes the direction of the angular velocity in space and the orientation of the inertia ellipsoid, which is body-fixed, provides the instantaneous direction of the body.

### 3.62 Symmetric Rigid Bodies

If the body we are considering is symmetrical, the inertia ellipsoid is an ellipsoid of revolution and the polhode is a circle on the axis of symmetry. The vector $\boldsymbol{\omega}$ then moves on the surface of a cone correspondingly, and its direction is said to *precess* in time about the symmetry axis of the body.

If we allow the axis of symmetry of a rigid body to be the z principal axes of the moment of inertia tensor, then the other two principal axes can be chosen to be any two orthogonal vectors in the plane orthogonal to the z principal axis.

The principal moments about these two axis are equal:

$$I_1 = I_2 \tag{3.47}$$

and the equations of motion for a symmetric body can then be simplified to

$$\begin{aligned} I_1\dot{\omega}_x &= (I_1 - I_3)\omega_z\omega_y \\ I_2\dot{\omega}_y &= -(I_1 - I_3)\omega_z\omega_x \\ I_3\dot{\omega}_z &= 0. \end{aligned} \tag{3.48}$$

The third equation shows us that the rate of rotation around a symmetry axis is constant and we are therefore able to treat it as one of the initial known conditions. The two remaining equations are a pair of coupled linear differential equations for $\omega_x$ and $\omega_y$. We are able to define a constant

$$\Omega = \tfrac{I_1 - I_3}{I_1}\omega_z, \tag{3.49}$$

known as the angular frequency. It is then possible to rewrite the equations for $\omega_x$ and $\omega_y$ as

$$\begin{aligned} \dot{\omega}_x + \Omega\omega_y &= 0 \\ \dot{\omega}_y - \Omega\omega_x &= 0. \end{aligned} \tag{3.50}$$

There are various methods we could employ to solve these equations. One technique is to find the time derivative of the first equation:

$$\ddot{\omega}_x = \Omega\dot{\omega}_y \tag{3.51}$$

and to then use the second equation to substitute for $\dot{\omega}_y$:

$$\ddot{\omega}_x = \text{-}\ \Omega^2\omega_x. \tag{3.52}$$

A typical solution for $\omega_x$ can be written as

$$\omega_x = A\sin\,\Omega t \tag{3.53}$$

where $A$ is some constant and if we substitute this solution for $\omega_x$ into the first of the set of equations (3.48) we can solve for $\omega_y$ to obtain:

$$\omega_y = A\cos\,\Omega t. \tag{3.54}$$

The magnitude of the vector $\omega_x\mathbf{i} + \omega_y\mathbf{j}$ is constant and it rotates uniformly about the z-axis of the symmetrical body with angular frequency $\Omega$. The magnitude of the total angular velocity $\boldsymbol{\omega}$ of the body is also constant and in the body-fixed frame $\boldsymbol{\omega}$ *precesses* about the symmetry axis with the same frequency, $\Omega$. It is worth noting that the body axes are themselves rotating in space, at the greater frequency $\boldsymbol{\omega}$.

Equation (3.49) tells us that smaller the difference between $I_1$ and $I_3$ is, the slower the precession frequency $\Omega$ will be in comparison with the frequency of rotation, $\omega$. The kinetic energy and the magnitude of the angular momentum can be used to evaluate the constants $A$ and $\omega_z$:

$$T = \frac{1}{2} I_1 A^2 + \frac{1}{2} I_3 \omega_z^2$$
$$J^2 = I_1^2 A^2 + I_3^2 \omega_z^2. \tag{3.55}$$

Here, $A$ is actually the amplitude of precession. It is possible to solve these equations for both $A$ and $\omega_z$ in terms of $T$ and $L$.

### 3.7 The Heavy Symmetrical Top with One Point Fixed

A heavy symmetrical top is a symmetrical body, pivoted at a point on its axis of symmetry, moving in a gravitational field. In the previous section, we chose the z-axis of the body-fixed system to the the symmetry axis of the body. The symmetry axis is also one of the principal axes of the body.

The constraint that one point of the top must be fixed reduces the number of degrees of freedom of the system to 3, and the Euler angles are able to completely specify the motion of the body.

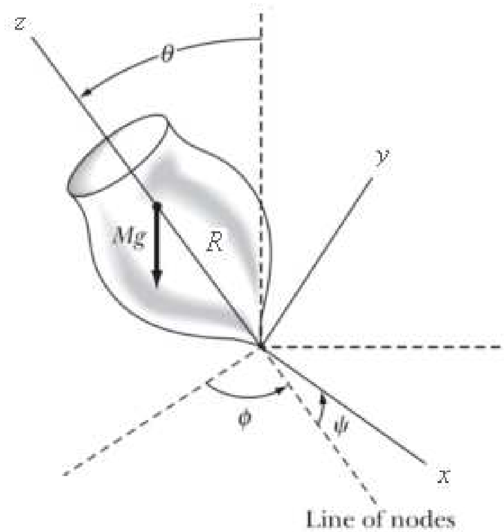


**Figure 3.1: The orientation of a spinning top is completely specified by the Euler Angles[5].**

### 5.71 The Lagrangian and the Generalised Momenta

The centre of gravity of the body lies along its symmetry axis, and if we denote the distance from here to the fixed point by $R$, we can then write the potential energy of the body as

$$V = MgRcos\theta. \tag{3.56}$$

[5]**image adapted from http://teacher.pas.rochester.edu/PHY235/LectureNotes/Chapter11.htm**

The kinetic energy of the symmetrical body can be written as

$$T = \tfrac{1}{2}I_1(\omega_x^2 + \omega_y^2) + \tfrac{1}{2}I_3\omega_z^2. \tag{3.57}$$

We can write this in a different form, using Euler angles, by using the set of equations (2.99), obtaining

$$T = \tfrac{1}{2}I_1(\dot{\theta}^2 + \dot{\phi}^2 sin^2\theta) + \tfrac{1}{2}I_3(\dot{\psi} + \dot{\phi}cos\theta)^2 \tag{3.58}$$

because the cross terms in $\omega_x^2$ and $\omega_y^2$ cancel each other out.

We can then write the Lagrangian for the system:

$$L = T - V = \tfrac{1}{2}I_1(\dot{\theta}^2 + \dot{\phi}^2 sin^2\theta) + \tfrac{1}{2}I_3(\dot{\psi} + \dot{\phi}cos\theta)^2 - MgRcos\theta. \tag{3.59}$$

The corresponding generalized momenta are then:

$$p_\phi = \tfrac{\partial L}{\partial \dot{\phi}} = (I_1 sin^2\theta + I_3 cos^2\theta)\dot{\phi} + I_3\dot{\psi}cos\theta \tag{3.60}$$

$$p_\theta = \tfrac{\partial L}{\partial \dot{\theta}} = I_1\dot{\theta} \tag{3.61}$$

$$p_\psi = \tfrac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = I_3\omega_z. \tag{3.62}$$

It can be seen that the Lagrangian does not involve the Euler angles $\phi$ and $\psi$ explicitly. This indicates that these coordinates are ignorable and the generalised momenta corresponding to these angles are constant in time and the system has only one degree of freedom, $\theta$.

**5.72 The Energy Equation**

As well as the generalised momenta for $\phi$ and $\psi$, the total energy $E$ of the conservative system will also be conserved:

$$E = T + V = \tfrac{I_1}{2}(\dot{\theta} + \dot{\phi}^2 sin^2\theta) + \tfrac{I_3}{2}\omega_z^2 + MgRcos\theta. \tag{3.63}$$

Since the momenta $p_\psi$ and $p_\phi$ are constants of motion, we can define two constants $a$ and $b$,

$$a = \tfrac{I_3\omega_3}{I_1}, \qquad b = \tfrac{p_\phi}{I_1} \tag{3.64}$$

which we can use to write

$$\dot{\phi} = \tfrac{b - acos\theta}{sin^2\theta}. \tag{3.65}$$

We can also obtain an equivalent expression for $\dot{\psi}$ :

$$\dot{\psi} = \tfrac{I_1 a}{I_3} - cos\theta\tfrac{b - acos\theta}{sin^2\theta}. \tag{3.66}$$

If we know $\theta$ as a function of time then it is possible to integrate the equations to give the dependence of $\phi$ and $\psi$ in time.

We can use equation (3.65) and equation (3.66) to eliminate $\dot{\phi}$ and $\dot{\psi}$ from the equation for the total energy, leaving a differential equation that is dependent only on $\theta$. First we need to be aware that $\omega_z$ is constant in time, as shown by equation (3.62), and equal to $\frac{I_1}{I_3}a$. We can then define a constant of motion, the *reduced energy,* $E'$ as:

$$E' \equiv E - \tfrac{1}{2}I_3\omega_z^2. \tag{3.67}$$

It is then possible for us the rewrite the total energy equation as

$$E' = \tfrac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2 sin^2\theta) + MgRcos\theta. \tag{3.68}$$

If we define two constants as

$$\alpha = \tfrac{2E'}{I_1}, \qquad \beta = \tfrac{2MgR}{I_1} \tag{3.69}$$

then we can substitute in equation (3.65) into equation (3.68) and rearrange the terms to give

$$sin^2\theta\ \dot{\theta}^2 = sin^2\theta(\alpha - \beta cos\theta) - (b - acos\theta)^2. \tag{3.70}$$

To simplify the analysis, we will set $u = cos\theta$ and then it is possible to rewrite the above equation as

$$\dot{u}^2 = (1-u^2)(\alpha - \beta u) - (a - bu)^2 = f(u). \tag{3.71}$$

We are also able to reduce the equations (3.65) and (3.66) to

$$\dot{\phi} = \tfrac{b-au}{1-u^2} \tag{3.72}$$

and

$$\dot{\psi} = \tfrac{I_1 a}{I_3} - \tfrac{u(b-au)}{1-u^2} \tag{3.73}$$

We can immediately take the square root of equation (3.71) and integrate it to get a quadrature:

$$t = \int\limits_{u(0)}^{u(t)} \frac{du}{\sqrt{(1-u^2)(\alpha-\beta u)-(b-au)^2}} \tag{3.74}$$

We could find similar integrals for $\phi$ and $\psi$, but we do not need to perform these integrations to get a general idea of the motion. The equation for $f(u)$ is cubic and the roots of the cubic polynomial provide the *turning angles* of $\theta$, at which the sign of $\dot{\theta}$ changes. .When $u$ is large, $\beta u^3$ becomes the dominant term of $f(u)$, and so $f(u) \to \infty$ as $u \to \infty$, and $f(u) \to -\infty$ as $u \to -\infty$.

When $u$ is equal to $\pm 1$, $f(u) = -(b \mp a)^2 \leq 0$, except in the instance when $u = \pm 1$ is a root of $f(u)$, which physically corresponds to a vertical top. Physically, the top can only be in motion when $\dot{u}^2$ is positive and $u$ must lie in the interval $u =$ -1 to $u = +1$, so, since $u = cos\theta$, we see that $\theta$ must lie between -$\pi$ and $+\pi$.

A plot of the function $f(u)$ looks like Figure (3.2):

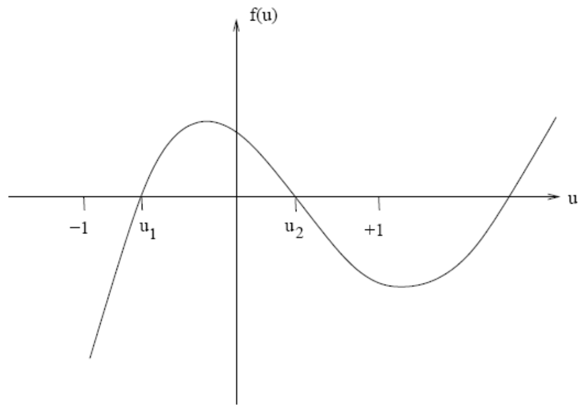


**Figure 3.2: A plot of the function $f(u)$**[6]

For any real top, $f(u)$ therefore has two roots $u_1$ and $u_2$, that lie between -1 and $+1$. The top must move so that $cos\theta$ always stays within this region.

### 5.73 Motion of the Spinning Top

The motion of the top can be depicted by tracing the curve of the intersection of the figure axis (the *locus* of the figure axis) onto a unit sphere about the fixed point. The locus lies between the co-latitude bounding circles $\theta_1 = arc\ cosu_1$ and $\theta_2 = arc\ cosu_2$. Any point on the locus has polar coordinates which are identical to the Euler angles for the body system. The value of the root $b - au$ determines the locus curve shape.

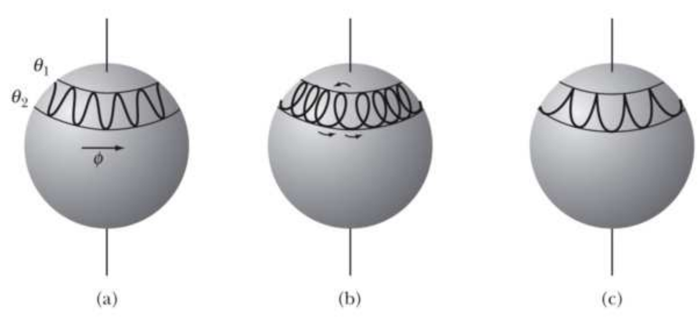


**Figure 3.3: The different types of motion depend on the direction of precession at the extrema** [7]

[6] image obtained from http://www.damtp.cam.ac.uk/user/tong/dynamics/three.pdf
[7] image obtained from http://teacher.pas.rochester.edu/PHY235/LectureNotes/Chapter11.htm

There are three possibilities for the motion, and they are dependent on the sign of $\dot{\phi}$, as determined by equation (3.72) at the roots $u = u_1$ and $u = u_2$. Motion as depicted in Figure 3.3(a) occurs if $\dot{\phi} > 0$ at both $u = u_1$ and $u = u_2$. If $\dot{\phi} > 0$ at $u = u_1$, but $\dot{\phi} < 0$ at $u = u_2$, then the type of motion as shown in Figure 3.3(b) will occur. Figure 3.3(c) shows the path of motion that will come about if $\dot{\phi} > 0$ at $u = u_1$, but $\dot{\phi} = 0$ at $u = u_2$.

Motion in $\phi$ is known as *precession*, whilst motion in $\theta$ is called *nutation*, and can be visualised as the figure axis nodding up and down between the bounding angles $\theta_1$ and $\theta_2$ as it goes around.

The motion as shown in Figure 3.3(c) is not as unlikely as it may seem. If we spin the top and let it go at some angle $\theta_0$, we have the initial conditions $t = 0$, $\theta = \theta_0$ and $\dot{\theta} = \dot{\phi} = 0$. Also, we should remember that the quantity

$$p_\phi = I_1 \dot{\phi} sin^2\theta + I_3 \omega_z cos\theta = I_3 \omega_z cos\theta_0 \tag{3.75}$$

is constant. We now have sufficient information to describe the qualitative motion of the top. It will initially begin to fall under gravity, and so $\theta$ increases. As the top is falling, $\dot{\phi}$ must turn and increase so that $p_\phi$ always remains constant. The direction of precession, $\dot{\phi}$ must be in the same direction as $\omega_z$, and we get the type of motion as shown by Figure 3.3(c).

When $\dot{\theta}$ and $\dot{\phi}$ are initially equal to zero, and if we assume that the initial kinetic energy of rotation is large compared with the maximum change in potential energy:

$$\tfrac{1}{2} I_3 \omega_z^2 >> 2MgR, \tag{3.76}$$

then we are able to quantitatively predict how the top will move. The precession and the nutation will only be small disturbances to the rotation of the top about its figure axis. In this case, we would call the top a "fast top". For a fast top, the angular momentum is along the axis of spin of the top.

**5.74 Uniform Precession**

If $\dot{\theta}$ is equal to zero, and if $\dot{\phi}$ is constant, then the top will precess without bobbing up and down. This situation requires $f(u)$ to have a single real root $u_0$ lying between -1 and +1. This root must satisfy the following equations:

$$f(u_0) = (1 - u_0^2)\ (\alpha - \beta u_0) - (b - a u_0)^2 = 0 \tag{3.77}$$

and

$$f'(u_0) = -2u_0(\alpha - \beta u_0) - \beta(1 - u_0^2) + 2a(b - a u_0) = 0. \tag{3.78}$$

Combining these two equations, we find that

$$\tfrac{\beta}{2} = a\dot{\phi} - \dot{\phi}^2 u_0. \tag{3.79}$$

Substituting in the relation $I_1 a = I_3\omega_z$ and the definition for $\beta$ as seen in equations (3.69), this gives us the equation

$$MgR = \dot{\phi}(I_3\omega_z - I_1\dot{\phi}cos\theta_0). \tag{3.80}$$

For a fixed value of $\omega_z$ and fixed $\theta_0$, giving the top exactly the right initial shove $\dot{\phi}$ will allow the top to spin without nutation. The quadratic nature of equation (3.80) in $\dot{\phi}$ indicates that there are two frequencies with which it is possible for the top to precess without nutation, known as "slow" precession and "fast" precession. We can also see that equation (3.80) can never be satisfied by $\dot{\phi} = 0$, so an initial shove must always be given to the top for it to be able to precess uniformly. The "fast" and "slow" precessions will only exist if equation (3.80) can actually be solved. This requires

$$\omega_z > \tfrac{2}{I_3}\sqrt{MgRI_1 cos\theta_0}. \tag{3.81}$$

For a given $\theta_0$, the top therefore has to be spinning fast enough to have uniform solutions, otherwise it will topple over.

### 5.75 The Sleeping Top

We can start a top spinning with its axis vertical, with

$$\theta = \dot{\theta} = 0. \tag{3.82}$$

When it is spinning quietly about the vertical, we call it a *sleeping top*. In this situation, $f(u)$ has a root at $\theta = 0$ (or $u = +1$), so that $f(1) = 0$. We can use equations (3.64) and (3.69) to check that $a = b$ and $\alpha = \beta$ here. The function has two roots of $u = +1$,

$$f(u) = (1-u)^2(\alpha(1+u) - a^2) \tag{3.83}$$

and a root of $u_2 = \frac{\alpha^2}{\alpha-1}$. If $u_2$ is greater than 1, then $\omega_z$ will be greater than a critical angular velocity $\omega'$, where

$$\omega' = \tfrac{4I_1 MgR}{I_3^2} \tag{3.84}$$

and the graph of $f(u)$ looks like Figure 3.4(a). This motion is stable and the only possible motion is for $u = 1$, so the top just continues to spin about the vertical.

If $u_2 < 1$ , then $\omega_z < \omega'$ and $f(u)$ looks like Figure 3.4(b) and the top is unstable.

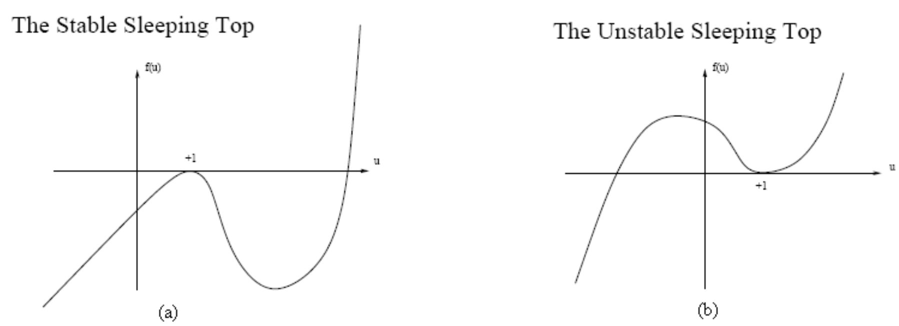


**Figure 3.5: A plot of $f(u)$ for the stable and unstable sleeping top**[8]

Practically, the top will spin about the vertical until friction gradually reduces the frequency of rotation to below the critical angular velocity and the top will then start to wobble as it slows down and will eventually fall.

[8]**image obtained from http://www.damtp.cam.ac.uk/user/tong/dynamics/three.pdf**

# Chapter 4

# The Levitron

## 4.1 Introduction to the Levitron

The Levitron is a popular mechanical toy invented by Mr. R Harrigan and developed by Mr. W Hones. It consists of a small, magnetic spinning top and a permanently magnetized ceramic base plate. The magnetic forces and torques couple uniquely with the gyroscopic motion of the spinning top, allowing stable levitation to occur. The top is seen to float above the base in mid-air, precessing and nutating about an equilibrium point until air resistance slows the top down sufficiently and the top then becomes unstable and falls.

**Figure 4.1: The Levitron**[9]

Knowledgeable physicists were adamant that Harrigan was wasting his time with the Levitron, because stable magnetic levitation for permanent static magnetic dipoles is forbidden by Earnshaw's theorem, but in this chapter I will discuss the mechanics of the device and attempt to explain the mechanical principles which allow it to work, providing a strict set of conditions are satisfied.

## 4.2 An Overview of How the Levitron Works

The heavy, symmetrical top is a rigid body, with mass $m$ and angular momentum $\mathbf{J}$. The mass of the top can be altered by adding small washers to it. The centre of mass of the top is located at $\mathbf{r} = (x, y, z)$. The top can be thought of as a magnetic dipole, with vector moment $\boldsymbol{\mu}$ which is located at the centre of mass and points along the symmetry axis.

A magnetic field $\mathbf{B}(\mathbf{r})$ is provided by the base plate, and its gradients provide a repulsive force which opposes the gravitational force, $mg$, on the top. This

[9]**image obtained from http://www.vnix.nl**

repulsive force is responsible allowing the top to achieve stability whilst spinning above the base; it acts on the vector moment in the presence of the spin $\mathbf{J}$.

It is not sufficient to just stabilize the top against flipping. If we assume that the magnetic dipole moment of the top, $\boldsymbol{\mu}$, is always oriented approximately vertically, in the downwards $-z$ direction and the repulsive field $\mathbf{B}(\mathbf{r})$ is also approximately vertical, but pointing in the $+z$ direction, then the magnetic energy $-\boldsymbol{\mu} \cdot \mathbf{B}(\mathbf{r})$ is approximately equal to $-\mu B_z$. The total potential energy $V$ is given by:

$$V = -\boldsymbol{\mu} \cdot \mathbf{B}(\mathbf{r}) + mgz \tag{4.1}$$

$$V = \mu B_z + mgz. \tag{4.2}$$

The first of two conditions that must be satisfied for stable levitation to occur is that the upward repulsive force, which is approximately equal to $\mu \partial_z B_z$, must balance the force of gravity on the top to allow it to float in equilibrium. The potential energy must have a critical point at the equilibrium point, and the second condition for stable levitation is that this critical point must be a minimum. This is actually impossible, by Earnshaw's theorem[10] since both gravity and $B_z$ are both harmonic potential fields. The energy minimum condition cannot be satisfied in both the vertical and horizontal directions simultaneously and instead there is a saddle point.

It is, however, still possible for the top to float stably in a time-independent potential field. If we take the potential to be the sum of the gravity and the magnetic energy of $\boldsymbol{\mu}$ averaged over its precession around $\mathbf{B}$ then the average energy makes use of the magnitude $B$ of $\mathbf{B}$ rather than the $z$-component. A minimum of the potential then exists for a narrow range of mass $m$ only. The range is narrow because it is the small deviations of $\mathbf{B}$ from vertical near the axis which are responsible for allowing the potential to posses a minimum. The form of $\mathbf{B}(\mathbf{r})$ varies with temperature, and is responsible for dictating the range of $m$. The mass required to allow the top to stably levitate may need to be frequently adjusted.

We need to make a correction to the adiabatic averaging underlying the static stability, as it is not exact. The first adjustment is to introduce an additional force called geometric magnetism which has the form $\mathbf{v} \times \mathbf{B}_{eff}(\mathbf{r})$ where $\mathbf{v} = \dot{\mathbf{r}}$ is the velocity at which the top is moving through the magnetic field, and $\mathbf{B}_{eff}(\mathbf{r})$ is the effective field constructed from component derivatives of $\mathbf{B}(\mathbf{r})$.

Three speeds must be very different to allow the top to float stably in the static potential field. The spin angular velocity of the top must be the fastest,

[10] Earnshaw's Theorem states that it is not possible to achieve stable suspension of an object against gravity, using any combination of electric charges and fixed magnets. The proof is fairly straight forward: The static force as a function of position $\mathbf{F}(\mathbf{x})$ which acts on a body due to the combination of gravitational, electrostatic and magnetic fields will always have divergence equal to zero, $\triangledown \cdot \mathbf{F}(\mathbf{x}) = 0$. The force at a point of equilibrium is zero. In the case of stable equilibrium, the force must point inwards towards the equilibrium point on some small sphere around the point. Gauss's theorem implies, however, $\int_S \mathbf{F}(\mathbf{x}) \cdot dS = \int_V \triangledown \cdot \mathbf{F}\, dV$, and since the divergence of the force over the volume inside is equal to zero, the radial component of the force over the surface must also be equal to zero.

the precession angular velocity of the top, and the rate at which **B** changes in the frame of the top must be the slowest.

If the top and base were both made from metal, the top would fall faster, because eddy currents would be induced, which would be an additional source of dissipation (energy is also dissipated by resistance from the air) and the top would fall more quickly.

### 4.3 Conditions for Stable Equilibrium

We are able to write the potential energy of top is given in equation (4.1). The magnetic torque on the top is equal to $\boldsymbol{\mu} \times \mathbf{B}$ and the spin of the top changes as the result of this torque,

$$\dot{\mathbf{J}}(t) = \boldsymbol{\mu}(t) \times \mathbf{B}(\mathbf{r}(t)). \tag{4.3}$$

We can make two assumptions for the top which will allow us to simplify the problem. Assuming that the top is small allows us to approximate its magnetism as a point dipole., located at coordinates $\mathbf{x}$. The second assumption is that the top is "fast", so that the angular momentum is along the axis of spin of the top, and this also coincides with the magnetic moment axis. The condition for a fast top is that the spin must be much faster than the precession and we shall see later that this condition is indeed satisfied.

We can write an equation for the magnetic field so that it is expressed in terms of its magnitude $B$ and direction $\mathbf{b}$ as seen from the spinning top's frame:

$$\mathbf{B}(\mathbf{r}(t)) \equiv B(t)\mathbf{b}(t). \tag{4.4}$$

We are now able to write a new equation for the spin as:

$$\dot{\mathbf{J}}(t) = \Omega(t)\mathbf{b}(t) \times \mathbf{J}(t) \tag{4.5}$$

where $\Omega$ is the angular precession frequency with which the top rotates about the instantaneous direction of the magnetic field as seen in the body-fixed frame:

$$\Omega = -\tfrac{\mu B}{J}. \tag{4.6}$$

As seen in the previous chapter, saying that the precession is fast is different to saying that the top is fast, and here the condition that the precession is fast is equivalent to the expression $\mid \Omega \mid \gg \mid \dot{\mathbf{b}} \mid$, and here the *adiabatic invariant* $\mathbf{J} \cdot \mathbf{b}$ must be approximately conserved, and this relation is known as an adiabatic slaving of $\mathbf{J}$ to $\mathbf{b}$. The component

$$\mu_B \equiv \boldsymbol{\mu}(t) \cdot \mathbf{b}(t) \tag{4.7}$$

must then also be an adiabatic invariant, and this allows us to rewrite the potential energy as:

$$V = V(r) = -\mu_B B(\mathbf{r}) + mgz. \tag{4.8}$$

Later, we will examine this approximation more closely.

With the assumptions made above, the top can how achieve stable levitation above the base if $V(\mathbf{r})$ has a possesses a minimum there. This requires the following three conditions to be satisfied:

$$
\begin{aligned}
&\text{(i) equilibrium:} && \nabla V(\mathbf{r}) = 0 \\
&\text{(ii) vertical stability:} && \partial_z^2 V(\mathbf{r}) > 0 \\
&\text{(iii) horizontal stability :} && \partial_x^2 V(\mathbf{r}) > 0 \text{ and } \partial_y^2 V(\mathbf{r}) > 0.
\end{aligned}
\tag{4.9}
$$

I am only going to be considering the situation where the minimum lies on the vertical symmetry axis of the base and neglecting the possibility of any off-axis minima.

Since the centre of mass of the top is not located inside the base, the base does not contribute any currents to $\mathbf{B}(\mathbf{r})$. $\mathbf{B}$ is curl and divergence free, so we are able to write the field as

$$\mathbf{B}(\mathbf{r}) = -\nabla\Phi(\mathbf{r}) \tag{4.10}$$

with

$$\nabla^2\Phi(\mathbf{r}) = 0\ . \tag{4.11}$$

The potential $\Phi$ is stationary at $x = y = 0$ in horizontal planes, and has circular symmetry to second order in $x$ and $y$, which is true for the Levitron with a square base.

If we define

$$\mathbf{R} \equiv (x, y) \tag{4.12}$$

and

$$\mathbf{R} = \mid \mathbf{R} \mid. \tag{4.13}$$

then we can write the potential close to the axis as

$$\Phi(\mathbf{r}) = \Phi(0,0,z) + \tfrac{1}{2}\partial_x^2\Phi(0,0,z)R^2 \ +...\ . \tag{4.14}$$

Conveniently, we can use the notation

$$\phi_n(z) = \partial_z^n\Phi(0,0,z). \tag{4.15}$$

Laplace's equation (4.11) must be satisfied by equation (4.14), and this gives

$$\Phi(\mathbf{r}) = \phi_0(z) - \tfrac{1}{4}\phi_2(z)R^2 + ...\ . \tag{4.16}$$

The adiabatic energy equation (4.8) makes use of the magnitude of the vector $\mathbf{B}$, and we can use equations (4.10) and (4.11) with (4.16) to write $B(\mathbf{r})$ to the second order in $R$:

$$B(\mathbf{r}) = \phi_1 \mathrm{sgn}\phi_1 \left[1 + \tfrac{R^2}{8}\left(\tfrac{\phi_2^2}{\phi_1^2} - 2\tfrac{\phi_3}{\phi_1}\right)\right] ..., \tag{4.17}$$

without explicitly writing the $z$ dependence of the $\phi_n$.

In the first of the set of equations (4.9), horizontal equilibrium on the axis is guaranteed by symmetry. For the system to also be in vertical equilibrium, gravity must be balanced by the repulsive magnetic force determined by the gradients of the magnitude of the field provided by the base plate, which is

$$mg = \mu_B \partial_z B = \mu_B \phi_2 \mathrm{sgn}\phi_1. \tag{4.18}$$

The gravitational force $mg$ is always positive, and so we can combine this fact with the second and third stability equations of set (4.9) to give three new conditions:

$$\begin{aligned} &\text{(i) equilibrium: } \mu_B \phi_2 \mathrm{sgn}\phi_1 > 0 \\ &\text{(ii) vertical stability: } \mu_B \phi_3 \mathrm{sgn}\phi_1 < 0 \\ &\text{(iii) horizontal stability: } \mu_B \mathrm{sgn}\phi_1 (2\phi_3 \tfrac{\phi_2^2}{\phi_1}) > 0. \end{aligned} \tag{4.19}$$

We cannot possibly satisfy these conditions when $\mu_B > 0$ because then (ii) will imply that $\phi_1$ and $\phi_3$ have opposite signs, and clearly equation (4.19(iii)) is then disobeyed. It is then a necessity that $\mu_B < 0$, which in practice means that the projection of $\boldsymbol{\mu}$ along $\mathbf{B}$ is anti-parallel to $\mathbf{B}$, and we get three new conditions for stable equilibrium:

$$\begin{aligned} &\text{(i) equilibrium: } \phi_1 \text{ and } \phi_2 \text{ have opposite signs} \\ &\text{(ii) vertical stability: } \phi_1 \text{ and } \phi_3 \text{ have the same signs} \\ &\text{(iii) horizontal stability: } \phi_2^2 - 2\phi_3\phi_1 > 0. \end{aligned} \tag{4.20}$$

The difference between $\mathbf{B}$ and $B$ gives rise to the term $\phi_2^2$ in (iii), and it allows both equations (ii) and (iii) to be simultaneously satisfied, confirming that the adiabatic potential in equation (4.8) does actually permit the top to achieve stable levitation over the base in spite of Earnshaw's seemingly contradictory theorem.

### 4.4 The Magnetic Field on the Axis

The Levitron has a square base which, with the exception of an unmagnetized central hole, is uniformly magnetized. The purpose of the hole is to allow a region which is almost field-free, so that the top can be spun by hand before it is lifted on a plastic plate to a position where it is able to float in equilibrium. We can consider the base to be a planar distribution of dipole sources which are vertically oriented, with density $\rho(\mathbf{R})$ and the following formula gives us the potential of a dipole:

$$\phi_0(z) = z \iint\limits_{base} d^2\mathbf{R} \frac{\rho(\mathbf{R})}{(R^2+z^2)^{\frac{3}{2}}} \tag{4.21}$$

where $\mathbf{R} = 0$ corresponds to the centre of the hole. We can assume that the dipoles in the base point up, then $\rho$ is positive. The dipoles in the spinning top must then point down, so $\mu_z$ is negative, and we can authenticate this assumption by seeing that the an upright top will repel the top of the base when held close; unlike dipoles will repel.

The analysis of stability for the top will almost be the same regardless of whether the base is a square with a hole in the centre or just a circular disc. It is simpler to carry out analysis for the case when the base is a uniformly magnetized disc, so this is what we will do. If the disc has a radius $a$ and $\rho$ is constant, then equation (4.21) gives:

$$\phi_0(z) = 2\pi\rho \left(1 - \frac{z}{\sqrt{a^2+z^2}}\right). \tag{4.22}$$

The condition given in equation (4.20(i)) requires $\phi_1$ and $\phi_2$ to have opposite signs. Equation (4.18) makes it seem as if magnetic repulsion can balance gravity at any height above the base if an appropriate mass is chosen. When $\phi_3 > 0$, $z < \frac{1}{2}a$ and when $\phi_3 < 0$, $z > \frac{1}{2}a$ and so vertical stability is guaranteed by by condition (4.20(ii)) for $z > \frac{1}{2}a$. The function $\phi_2^2 - 2\phi_3\phi_1$ from equation (4.20(iii)) is directly proportional to $2a^2 - 5z^2$ and so for the top to be horizontally stable, we need $z$ to be less than $a\sqrt{\frac{2}{5}}$ . Stable equilibrium can therefore occur in the range

$$\frac{1}{2} < \frac{z}{a} < \sqrt{\frac{2}{5}}. \tag{4.23}$$

It is also possible to find the range of masses which are allowed by the above condition (4.23). We can define three new coordinates,

$$\alpha = \frac{x}{a} \qquad \beta = \frac{y}{a} \qquad \gamma = \frac{z}{a} \tag{4.24}$$

which are all dimensionless. We can also define energy and mass by the following relations:

$$V \equiv \frac{Va}{2\pi|\mu_B|\rho}, \qquad M \equiv \frac{mga^2}{2\pi|\mu_B|\rho}. \tag{4.25}$$

We can combine equations (4.8), (4.17) and (4.22) to find that the potential energy close to the axis is equal to

$$V(\alpha, \beta, \gamma) = M\gamma + \frac{1}{(1+\gamma^2)^{\frac{3}{2}}} + \frac{3(\alpha^2+\beta^2)(2-5\gamma^2)}{8(1+\gamma^2)^{\frac{7}{2}}}. \tag{4.26}$$

The condition for equilibrium (4.18) determines the range of masses, $M$, which are allowed by condition (4.23), giving

$$M = \frac{3\gamma}{(1+\gamma^2)^{\frac{3}{2}}}. \tag{4.27}$$

$M$ has a maximum $M_+$ when $\gamma = \frac{1}{2}$, which is the lower stability limit for $\gamma$. If the mass is greater than this maximum, stable equilibrium is simply not possible, and the top will not be able to float. The maximum is equal to

$$M_+ = \frac{48}{5^{\frac{5}{2}}} = 0.85865010... \tag{4.28}$$

The upper stability limit for $\gamma$ is equal to $\sqrt{\frac{2}{5}}$ and the mass has a minimum $M_-$ here:

$$M_- = \frac{75\sqrt{2}}{7^{\frac{5}{2}}} = 0.818146658... \ . \tag{4.29}$$

If the mass is less than the minimum, vertical stability can actually be achieved, but the condition for horizontal equilibrium is not satisfied and so stable equilibrium can only occur when the mass is in the interval between $M_-$ and $M_+$. This interval is small - it is approximately only 5% of the mean stable mass.

It is possible to find a mass which yields the "most stable motion" of the top, and this is $M_S = 0.847837$.

If we no longer consider the base to be a magnetized disc but instead the more accurate model of a uniformly magnetized square slab with sides of length $2a$ and with an unmagnetized hole of radius $w$ in the centre, then we get a new equation:

$$\phi_0(z) = 2\pi\rho \frac{z}{\sqrt{z^2+w^2}} - 8\rho arcsin \left\{ \frac{z}{\sqrt{2(z^2+a^2)}} \right\}. \tag{4.30}$$

If we carry out stability analysis on this function for the Levitron, which has the relation that $\frac{w}{a} \approx \frac{1}{4}$, we obtain the result that the top is able to achieve stable levitation in the region

$$3.976 < \frac{z}{w} < 4.360 \tag{4.31}$$

which we can see is very narrow. We must make sure that we choose the mass of the top carefully so that equation (4.18) has a solution which lies in the stable range. A change $dm$ in the mass of the top results in a change in the height of equilibrium $d\gamma$ and we can use equation (4.18) to find that this gives

$$d\gamma = -\frac{dm}{m} \mid \frac{\phi_2}{\phi_3} \mid \tag{4.32}$$

where $\mid \frac{\phi_2}{\phi_3} \mid$ is known as an $amplification\ factor$ which decreases in the stable interval from infinity to a minimum value. When we use the base model of a square slab with a hole in as before, and $\frac{w}{a} \approx \frac{1}{4}$ , this minimum value is 7.05. Even the lightest of the washers which we can add to the top to increase its mass will result in the height of levitation changing by around one tenth of the interval of stability.

One feature of the Levitron which may seem puzzling is the need to constantly adjust the weight of the top over short periods of only a few minutes.

This has been shown experimentally to be due to handling of the top and ambient temperature changes resulting in temperature variation which has an effect on the mass of the top required to allow stable levitation to occur.

### 4.5 Adiabatic and Geometric Magnetism

In the Levitron, we know that in addition to gravity, magnetism also contributes to the motion of the top. From the equation for the potential energy of the top (4.1), we find the sum of these forces on the top to be

$$\mathbf{F} = -mg\mathbf{k} + \mathbf{F}_M \tag{4.33}$$

where $\mathbf{k}$ is the unit vector in the upward $z-$direction and $\mathbf{F}_M$ is the magnetic force, which we can rewrite to get:

$$\mathbf{F} = -mg\mathbf{k} + \nabla\boldsymbol{\mu}(t) \cdot \mathbf{B}(\mathbf{r}). \tag{4.34}$$

It is possible to split $\boldsymbol{\mu}$ into its components which are parallel to and perpendicular to the $\mathbf{B}$ field at any instant, to get

$$\boldsymbol{\mu}(t) = \mu_B \mathbf{b}(t) + \boldsymbol{\mu}_\perp(t) \tag{4.35}$$

where $\boldsymbol{\mu}_\perp$ is the perpendicular component. We can now write the magnetic force as:

$$\mathbf{F}_M \equiv \mathbf{F}_A + \mathbf{F}_G \tag{4.36}$$

with $\mathbf{F}_A$ being the *adiabatic* component of the magnetic force, and $\mathbf{F}_G$ being the *geometric* component. We can write this as

$$\mathbf{F}_M = \mu_B \nabla B(\mathbf{r}) + \boldsymbol{\mu}_\perp(t) \cdot \nabla\mathbf{B}(\mathbf{r}) \tag{4.37}$$

where $\boldsymbol{\mu}_\perp$ and $\mathbf{B}$ are connected by the dot product.

Since we have assumed that both the top and the precession are fast, the equation of motion (4.5) for $\mathbf{J}$ will also be satisfied by $\boldsymbol{\mu}$. Separating this into to parallel and perpendicular components of $\boldsymbol{\mu}$, we obtain

$$\begin{aligned} \dot{\boldsymbol{\mu}} &= \dot{\mu}_B \mathbf{b} + \mu_B \dot{\mathbf{b}} + \dot{\boldsymbol{\mu}}_\perp \\ &= \boldsymbol{\Omega}\mathbf{b} \times \boldsymbol{\mu}_\perp \end{aligned} \tag{4.38}$$

We need to make some adiabatic approximations. It is necessary to allow the precession of the top to be about a direction which is slightly different from the instantaneous field direction $\mathbf{b}(t)$. We also need to make the approximation that we can set the precession averaged velocity $\dot{\boldsymbol{\mu}}_\perp$ to zero, but not the component $\boldsymbol{\mu}_\perp$. We now find the parallel and perpendicular components of equation (4.38) give:

$$\dot{\mu}_B \approx 0 \tag{4.39}$$

and

$$\boldsymbol{\mu}_\perp \approx -\frac{\mu_B}{\boldsymbol{\Omega}}\mathbf{b} \times \dot{\mathbf{b}}. \tag{4.40}$$

Equation (4.39) is the conservation of the adiabatic invariant, and the component $\mathbf{F}_A$ in equation (4.36) becomes the adiabatic force of the lowest order that would be obtained from equation (4.8). This leads to the conditions obtained earlier for static stability.

Using equation (4.40) as well as the assumption that the top is fast and therefore $\mathbf{J}$ is parallel to $\boldsymbol{\mu}$, we are able to write the geometric component of magnetic force, $\mathbf{F}_G$ (which is a post-adiabatic force) as

$$\begin{aligned}\mathbf{F}_G &= -\frac{\mu_B}{\Omega}(\mathbf{b} \times \dot{\mathbf{b}}) \cdot \nabla\mathbf{B}(\mathbf{r}) \\ &= \frac{J_B}{B}(\mathbf{b} \times \dot{\mathbf{b}}) \cdot \nabla\mathbf{B}(\mathbf{r})\end{aligned} \tag{4.41}$$

where $J_B \equiv \mathbf{J} \cdot \mathbf{b}$. As the top moves through the inhomogeneous field $\mathbf{B}(\mathbf{r})$, it causes a change $\dot{\mathbf{b}}$:

$$\dot{\mathbf{b}} = (\mathbf{v} \cdot \nabla)\mathbf{b}(\mathbf{r}). \tag{4.42}$$

The force $\mathbf{F}_G$ is dependent on both the position of the top and its velocity:

$$\mathbf{F}_G = -\frac{J_B}{B^3}\left[\mathbf{B} \times (\mathbf{v} \cdot \nabla)\mathbf{B}\right] \cdot \nabla\mathbf{B}(\mathbf{r}). \tag{4.43}$$

It is possible to use vector algebra to rewrite equation (4.43) as

$$\mathbf{F}_G = \mathbf{v} \times \mathbf{B}_{eff}(\mathbf{r}) \tag{4.44}$$

with the vector $\mathbf{v} = \dot{\mathbf{r}}$, and where $\mathbf{B}_{eff}$ is dependent on the components of $\mathbf{B}$, as dictated by

$$\mathbf{B}_{eff} = -\frac{J_B}{B^3}(B_x\nabla B_y \times \nabla B_Z + B_y\nabla B_Z \times \nabla B_x + B_z\nabla B_x \times \nabla B_y \tag{4.45}$$

The geometric magnetism force, $\mathbf{F}_M$ has the same dependence on velocity as the Lorentz force, it seems as if the top carries a unit charge in response to the effective field $\mathbf{B}_{eff}$. The geometric post-adiabatic magnetism force acts as a reaction to the fast spin $\mathbf{J}(t)$ of the top on the slow motion $\mathbf{r}(t)$ of its centre of mass.

In all the calculations, we have regarded the precession to be fast, so that it is slaved to the slow variable $\mathbf{r}$ and $\mathbf{J}$ then reacts magnetically on $\mathbf{r}$. We can instead show that the precession can be considered to be geometric because $\mathbf{J}$ is actually slaved to the spin of the top - the motion of the axis of the top is slow when compared to the spin. When we average the precession over the nutation of the body axes, we can consider it to be a geometric reaction which results from a monopole source of magnetism fixed at the point of precession. For the Levitron, this fixed point is simply the centre of mass of the top, but in the regular spinning top it would be the point where the top is in contact with the surface while spinning.

**4.6 The Effect of Geometric Magnetism on Stability**

We can use the previously seen equations (4.33), (4.37) and (4.44) to write an equation of motion for the top when it is subjected to the force of gravity as well as the two magnetic forces. We obtain

$$m\ddot{\mathbf{r}} = -mg\mathbf{k} + \mu_B \nabla B(\mathbf{r}) + \mathbf{v} \times \mathbf{B}_{eff}(\mathbf{r}). \tag{4.46}$$

Using the adiabatic force $\mathbf{F}_A$ from equation (4.63) with earlier obtained equations (4.12), (4.13) and (4.17), we get

$$\nabla B = sgn\phi_1 \left[\tfrac{1}{4}\left(\tfrac{\phi_2^2}{\phi_1} - 2\phi_3\right)\mathbf{R} + \phi_2\mathbf{k}\right]. \tag{4.47}$$

The effective magnetic field on the axis in equation (4.45) can be used with equations (4.10) and (4.16) to give

$$\mathbf{B}_{eff} = J_B sgn\phi_1 \tfrac{\phi_2^2}{4\phi_1^2}\mathbf{k}. \tag{4.48}$$

This tells us that geometric magnetism does not have an effect on vertical motion; only the static gravitational force and the adiabatic magnetism force will affect the motion vertically. The geometric magnetism is instead responsible for affecting the horizontal motion of the top, and we can write a linear equation for the horizontal acceleration:

$$\ddot{\mathbf{R}} = \tfrac{g}{4}\left(\tfrac{\phi_2}{\phi_1} - 2\tfrac{\phi_3}{\phi_2}\right)\mathbf{R} + J_B sgn\phi_1 \tfrac{\phi_2^2}{4m\phi_1^2}\dot{\mathbf{R}} \times \mathbf{k} \tag{4.49}$$

where there are coefficients which are dependent on the height of the top, and we have made use of equation (4.18) to eliminate $\mu_B$. Defining $u(t)$ as

$$u(t) \equiv x(t) + iy(t) \tag{4.50}$$

will allow us to write equation (4.49) as

$$\ddot{u} = A(z)u + iB(z)\dot{u}, \tag{4.51}$$

in scalar form. This has the general solution:

$$u(t) = u_+ e^{i\nu_+(z)t} + u_- e^{i\nu_-(z)t} \tag{4.52}$$

where

$$\nu_\pm(z) = \tfrac{1}{2}(B \pm \sqrt{B^2 - 4A}). \tag{4.53}$$

We need $\nu_\pm$ to be real for the top to be stable horizontally, which in turn requires $B^2 > 4A$. With equation (4.49), this gives the condition

$$\tfrac{J_B^2}{m^2 g} > G(z) \equiv 32 \mid \tfrac{\phi_1}{\phi_3} \mid^3 \left(\tfrac{\phi_1\phi_3}{\phi_2^2} - \tfrac{1}{2}\right). \tag{4.54}$$

If the condition (4.20(iii)) is satisfied, then the condition (4.54) is also satisfied, since $G(z)$ will be negative. If $G(z)$ is positive then condition (4.20(iii)) will not be satisfied and horizontal stability cannot be achieved. Condition (4.54) shows that if the top is spun fast enough, geometric magnetism can indeed provide post-adiabatic stabilization. If we set the top to have a vertical axis with radius of gyration $d$ and spin frequency $\omega$, we are able to evaluate the last part of equation (4.54) with the potential of a disc of radius $a$ from equation (4.22). The horizontal motion will then be stable if $\gamma = \frac{z}{a}$ satisfies the condition:

$$\frac{J_B^2}{m^2 g a^3} = \frac{4\pi^2\omega^2 d^4}{g a^3} > \frac{G(z)}{a^3} = \frac{16(5\gamma^2-2)(\gamma^2+1)^3}{\gamma^5}. \tag{4.55}$$

This is equivalent to the condition:

$$\gamma - \sqrt{\frac{2}{5}} < \frac{81\pi^2\omega^2 d^4}{686 g a^3}.. \tag{4.56}$$

Physical measurements for the Levitron are $a \approx 5$cm, $d = 1.13$cm, and hand spinning gives a spin frequency $\omega \sim 20$Hz. Stability requires $\gamma - \sqrt{\frac{2}{5}} < 0.0062$. The total range of stable spin rates is approximately $18\text{Hz} \leq \omega \leq 40\text{Hz}$, and is achievable by hand-spinning. The lower spin limit corresponds to the *sleeping top* condition. Geometric magnetism only contributes a small (about 5%) increase to the stable interval of $\gamma$ when spun at $\omega = 20$Hz, but would contribute up to a 20% increase if the top were spun at $\omega = 40$Hz.

### 4.7 Adiabatic Conditions

Practically, the ratio of the moments of inertia is somewhere between $\frac{1}{2}$ and 1; we can assume that all of the principal moments of inertia of the top are approximately the same size, and we then have the following condition for the top to be fast:

$$2\pi\omega \gg \mid \Omega \mid$$

where $2\pi\omega$ is the spin angular velocity and $\mid \Omega \mid$ is the precession angular velocity. We can use equations (4.6) and (4.18) to express the above inequality in terms of the magnetic potential. We get

$$\omega \gg \omega_{min} \tag{4.57}$$

where

$$\omega_{min} = \frac{1}{2\pi d}\sqrt{g\left|\frac{\phi_1}{\phi_2}\right|}. \tag{4.58}$$

We can then write the precession frequency in terms of $\omega_{min}$ :

$$\frac{\Omega}{2\pi} = \frac{\omega_{min}^2}{\omega}. \tag{4.59}$$

By the adiabaticity, the rate at which the field $\mathbf{b}$ is changing (equation (4.42)) must be much slower than precession. This condition is written as

$$\mid \Omega \mid \gg \mid (\mathbf{v} \cdot \nabla)\mathbf{b} \mid . \tag{4.60}$$

We can use equations (4.10), (4.11) and (4.16) to give the result

$$\mathbf{b} = \left\{ x\frac{\phi_2}{2|\phi_1|}, y\frac{\phi_2}{2|\phi_1|}, -sgn\phi_1 \right\} + ... \tag{4.61}$$

and this means that equation (4.59) is dependent only on the transverse speed of the top, $\mathbf{v}_\perp$. We can apply some algebra to the adiabatic condition (4.60) to give the new condition

$$\mathbf{v}_\perp \gg \frac{2\sqrt{g}}{d} \left( \left| \frac{\phi_1}{\phi_2} \right| \right)^{\frac{3}{2}} \frac{\omega_{min}}{\omega}. \tag{4.62}$$

The nutation frequency is related to the transverse speed of the top. If we consider the frequency of this vertical motion, equation (4.46) gives

$$\omega_z = \frac{1}{2\pi} \sqrt{g \left| \frac{\phi_3}{\phi_2} \right|} = \frac{1}{2\pi} \sqrt{\frac{g}{a} \frac{(4\gamma^2 - 1)}{\gamma(\gamma^2+1)}} \tag{4.63}$$

where we are referring to the circular disc base.

Using the previously stated measurements of $a$ and $d$ for the Levitron, and using an approximate spin frequency of 20Hz, the potential from equation (4.22) implies that $0.73a < \mid \frac{\phi_1}{\phi_2} \mid < 0.83a$ and

$$\omega_{min} \approx \frac{0.88}{2\pi} \sqrt{\frac{ga}{d^2}} \sim 8.7\text{Hz} \tag{4.64}$$

over the interval of stability, $\frac{1}{2} < \gamma < \sqrt{\frac{2}{5}}$. This frequency is much less than the lowest speed achievable by hand-spinning. Equation (4.59) then tells us that the corresponding precession frequency here would be $\frac{\Omega}{2\pi} = 3.8\text{Hz}$ , and the Levitron is indeed a fast top.

The inequality (4.62) gives an upper limit for the transverse speed of the top. When the top is spun by hand, its horizontal speed is much slower than this limit and so we have satisfied the adiabatic condition for motion.

Earlier we discussed $M_S$, the mass for which the most stable motion can occur. Putting this together with equation (4.63), we obtain

$$\omega_z = \frac{0.61106}{2\pi} \sqrt{\frac{g}{a}} \approx 0.69\omega_{min} \frac{d}{a} \sim 1.4\text{Hz}. \tag{4.65}$$

This nutation frequency is much less than the precession frequency.

If we spin the top too fast, the precession frequency will be too slow to prevent the top from tipping over so much that it can no longer be supported by the magnetic field, and the top will fall over.

### 4.8 Analogy with Microscopic Particle Traps

While I will not look into it here, it is worth noting that the trap mechanism in the Levitron is analogous to magnetic gradient traps for neutral particles with a quantum magnetic dipole moment. These traps were first used to trap cold neutrons and currently are being used to trap atoms.

# Chapter 5

## Conclusion

Rigid body dynamics is a broad and interesting topic, with many useful applications. In this project, we began by defining a rigid body and went on to look at the basics of rigid bodies and their motion, describing Euler angles and considering infinitesimal rotations. We then went on to look at angular momentum and the moment of inertia tensor, as well as the principal axis, combining these ideas to find the equations of motion for a rigid body to be Euler's equations. We looked into the motion of symmetric rigid bodies and finally were able to consider the motion of a heavy symmetrical top.

We have seen that precession is an exciting phenomenon and although it may imply some counter-intuitive ideas, fully understood it can be used to describe the motion of the spinning top, and further the motion of the Levitron, an exceptional toy which had scientists baffled for a long time, since it seems to violate Earnshaw's theorem.

In examining the principles behind the Levitron, we have seen that it is the unique coupling of the magnetic forces and torques with the gyroscopic motion of the top that is responsible for permitting stable levitation to occur.

There are many areas where I could have extended the study of precession. An obvious one is to study the precession of the Earth. This would involve considering local torque-induced precession due to the gravitational effects of the sun and moon acting upon the axis of the Earth. There is also a minor amount of non-local torque-free precession due to the motion of the solar system. The Earth has a central bulge at the equator, so it is in fact not a sphere, but a symmetrical top. We could also look into the effects of the Earth's precession on astronomical observations and the precession of orbital objects. Clearly, this topic is very rich and there is wide scope for further study here.

Very recently, I discovered a special type of spinning top called the Rattleback, or Celt, which is a semi-ellipsoidal top which can be initially spun in any direction, but if not spun in its preferred direction, it will become unstable, "rattle", stop and reverse its spin. This spin-reversal seems to violate angular momentum conservation laws, and I think it would be very interesting to study the mathematical principles behind the Rattleback more closely.

There are many other areas of the project which I haven't touched on at all. We could have considered Thomas precession, which is a special correction to gyroscopic precession in a rotating non-inertial frame. It has many applications, such as in quantum mechanics where it is a correction to the spin-orbit interaction, and takes into account the relativistic time dilation between the electron and the nucleus in Hydrogen atoms.

We could further extend the study of the Levitron to consider more closely the analogy between it and microscopic particle traps.

It is astounding to think that the simple spinning top, one of the oldest discovered toys has such stimulating mathematics behind it, and similar mathematics can help explain such a wide range of different phenomena in the areas of astrology, electromagnetism and quantum mechanics.

# Bibliography


[1] A. Modi, Dynamics of a Spinning Top:
http://www.anirudh.net/courses/emch520/html/node3.html

[2] D. Kleppner and R. J. Kolenkow: *An Introduction to Mechanics* (McGraw-Hill)

[3] D. Tong, Cambridge University: Dynamics Lectures:
http://www.damtp.cam.ac.uk/user/tong/dynamics/three.pdf

[4] G. R. Fowles and G. L. Cassiday: *Analytical Mechanics* (Harcourt Brace)

[5] H. Goldstein: *Classical Mechanics* (Addison Wesley)

[6] Hyperphyics, Gyroscopes:
http://hyperphysics.phy-astr.gsu.edu/hbase/gyr.html

[7] J. B. Marion and S. T. Thornton: *Classical Dynamics of Particles and Systems* (Saunders College)

[8] J. Huang, Texas Tech University:
http://www.phys.ttu.edu/~huang24/Teaching/Phys5306/

[9] M. Burske, University of Colarado, Advanced Dynamics:
http://theory.ph.man.ac.uk/~mikeb/lecture/pc167/rigidbody/contents.html

[10] Morehouse, Orthogonal Transformations:
http://www.morehouse.edu/facstaff/cmoore/Orthogonal%20Transformations.htm

[11] M. Simon, L. Heflinger & S. Ridgway: *Spin Stabilized Magnetic Levitation*: Am. J. Phys., Vol. 65, No. 4 (1997)

[12] M. V. Berry: *The Levitron: An Adiabatic Trap for Spins:* The Royal Society (1996)

[13] Queen's University Mechanics Lectures:
http://me.queensu.ca/courses/MECH494/documents/06-EulersTheorem.pdf

[14] Tabitha Rigid Bodies:
http://tabitha.phas.ubc.ca/wiki/index.php/Rigid_Bodies

[15] The Levitron: http://www.levitron.com/physics.html

[16] T. W. B. Kibble and F. H. Berkshire: *Classical Mechanics* (Longman)

[17] V. D. Barger and M. G. Olsson: *Classical Mechanics* A Modern Perspective (McGraw-Hill)

[18] Wikipedia, Gyroscopes: http://en.wikipedia.org/wiki/Gyroscope

[19] Wolfram MathWorld, Infinitesimal Rotations:
http://mathworld.wolfram.com/InfinitesimalRotation.html