Chapter 4

Rigid Body Motion

In this chapter we develop the dynamics of a rigid body, one in which all interparticle distances are fixed by internal forces of constraint. This is, of course, an idealization which ignores elastic and plastic deformations to which any real body is susceptible, but it is an excellent approximation for many situations, and vastly simplifies the dynamics of the very large number of constituent particles of which any macroscopic body is made. In fact, it reduces the problem to one with six degrees of freedom. While the ensuing motion can still be quite complex, it is tractible. In the process we will be dealing with a configuration space which is a group, and is not a Euclidean space. Degrees of freedom which lie on a group manifold rather than Euclidean space arise often in applications in quantum mechanics and quantum field theory, in addition to the classical problems we will consider such as gyroscopes and tops.

4.1 Configuration space for a rigid body

A macroscopic body is made up of a very large number of atoms. Describing the motion of such a system without some simplifications is clearly impossible. Many objects of interest, however, are very well approximated by the assumption that the distances between the atoms in the body are fixed[1],

$$|\vec{r}_\alpha - \vec{r}_\beta| = c_{\alpha\beta} = \text{ constant.} \tag{4.1}$$

[1]In this chapter we will use Greek letters as subscripts to represent the different particles within the body, reserving Latin subscripts to represent the three spatial directions.

This constitutes a set of holonomic constraints, but not independent ones, as we have here $\frac{1}{2}n(n-1)$ constraints on $3n$ coordinates. Rather than trying to solve the constraints, we can understand what are the generalized coordinates by recognizing that the possible motions which leave the interparticle lengths fixed are combinations of

- translations of the body as a whole, $\vec{r}_\alpha \to \vec{r}_\alpha + \vec{C}$,
- rotations of the body about some fixed, or "marked", point.

We will need to discuss how to represent the latter part of the configuration, (including what a rotation is), and how to reexpress the kinetic and potential energies in terms of this configuration space and its velocities.

The first part of the configuration, describing the translation, can be specified by giving the coordinates of the marked point fixed in the body, $\widetilde{R}(t)$. Often, but not always, we will choose this marked point to be the center of mass $\vec{R}(t)$ of the body. In order to discuss other points which are part of the body, we will use an orthonormal coordinate system fixed in the body, known as the **body coordinates**, with the origin at the fixed point $\widetilde{R}$. The constraints mean that the position of each particle of the body has fixed coordinates in terms of this coordinate system. Thus the dynamical configuration of the body is completely specified by giving the orientation of these coordinate axes in addition to $\widetilde{R}$. This orientation needs to be described relative to a fixed inertial coordinate system, or **inertial coordinates**, with orthonormal basis $\hat{e}_i$.

Let the three orthogonal unit vectors defining the body coordinates be $\hat{e}'_i$, for $i = 1, 2, 3$. Then the position of any particle α in the body which has coordinates $b'_{\alpha i}$ in the body coordinate system is at the position $\vec{r}_\alpha = \widetilde{R} + \sum_i b'_{\alpha i}\hat{e}'_i$. In order to know its components in the inertial frame $\vec{r}_\alpha = \sum_i r_{\alpha i}\hat{e}_i$ we need to know the coordinates of the three vectors $\hat{e}'_i$ in terms of the inertial coordinates,

$$\hat{e}'_i = \sum_j A_{ij}\hat{e}_j. \tag{4.2}$$

The nine quantities A_{ij}, together with the three components of $\widetilde{R} = \sum \widetilde{R}_i\hat{e}_i$, specify the position of every particle,

$$r_{\alpha i} = \tilde{R}_i + \sum_j b'_{\alpha j}A_{ji},$$

and the configuration of the system is completely specified by $\widetilde{R}_i(t)$ and $A_{ij}(t)$.

The nine real quantities in the matrix A_{ij} are not independent, for the basis vectors $\hat{e}'_i$ of the body-fixed coordinate system are orthonormal,

$$\hat{e}'_i \cdot \hat{e}'_k = \delta_{ik} = \sum_{j\ell} A_{ij} A_{k\ell} \hat{e}_j \cdot \hat{e}_\ell = \sum_{j\ell} A_{ij} A_{k\ell} \delta_{j\ell} = \sum_j A_{ij} A_{kj},$$

or in matrix languag $AA^T = 1\!\!1$. Such a matrix of real values, whose transpose is equal to its inverse, is called **orthogonal**, and is a transformation of basis vectors which preserves orthonormality of the basis vectors. Because they play such an important role in the study of rigid body motion, we need to explore the properties of orthogonal transformations in some detail.

4.1.1 Orthogonal Transformations

There are two ways of thinking about an orthogonal transformation A and its action on an orthonormal basis, (Eq. 4.2). One way is to consider that $\{\hat{e}_i\}$ and $\{\hat{e}'_i\}$ are simply different basis vectors used to describe the same physical vectors in the same vector space. A vector $\vec{V}$ is the same vector whether it is expanded in one basis $\vec{V} = \sum_j V_j \hat{e}_j$ or the other $\vec{V} = \sum_i V'_i \hat{e}'_i$. Thus

$$\vec{V} = \sum_j V_j \hat{e}_j = \sum_i V'_i \hat{e}'_i = \sum_{ij} V'_i A_{ij} \hat{e}_j,$$

and we may conclude from the fact that the $\hat{e}_j$ are linearly independent that $V_j = \sum_i V'_i A_{ij}$, or in matrix notation that $V = A^T V'$. Because A is orthogonal, multiplying by A (from the left) gives $V' = AV$, or

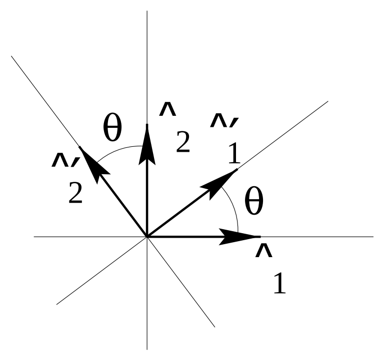
$$V'_i = \sum_j A_{ij} V_j. \tag{4.3}$$

Thus A is to be viewed as a rule for giving the primed basis vectors in terms of the unprimed ones (4.2), and also for giving the components of a vector in the primed coordinate system in terms of the components in the unprimed one (4.3). This picture of the role of A is called the **passive** interpretation.

One may also use matrices to represent a real physical transformation of an object or quantity. In particular, Eq. 4.2 gives A the interpretation of an operator that rotates each of the coordinate basis $\hat{e}_1, \hat{e}_2, \hat{e}_3$ into the corresponding new vector $\hat{e}'_1$, $\hat{e}'_2$, or $\hat{e}'_3$. For real rotation of the physical system, all the vectors describing the objects are changed by the rotation into new vectors $\vec{V} \to \vec{V}^{(R)}$, physically different from the original vector, but having the same coordinates in the primed basis as V has in the unprimed basis. This is called the **active** interpretation of the transformation. Both active and passive views of the transformation apply here, and this can easily lead to confusion. The transformation $A(t)$ is the physical transformation which rotated the body from some standard orientation, in which the body axes $\hat{e}'_i$ were parallel to the "lab frame" axes $\hat{e}_i$, to the configuration of the body at time t. But it also gives the relation of the components of the same position vectors (at time t) expressed in body fixed and lab frame coordinates.

If we first consider rotations in two dimensions, it is clear that they are generally described by the counterclockwise angle θ through which the basis is rotated,

$$\begin{aligned} \hat{e}'_1 &= \cos\theta \hat{e}_1 + \sin\theta \hat{e}_2 \\ \hat{e}'_2 &= -\sin\theta \hat{e}_1 + \cos\theta \hat{e}_2 \end{aligned}$$



corresponding to the matrix

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \qquad (4.4)$$

Clearly taking the transpose simply changes the sign of θ, which is just what is necessary to produce the inverse transformation. Thus each two dimensional rotation is an orthogonal transformation. The orthogonality equation $A \cdot A^T = 1$ has four matrix elements. It is straightforward to show that these four equations on the four elements of A determine A to be of the form (4.4) except that the sign of the bottom row is undetermined. For example, the transformation $\hat{e}'_1 = \hat{e}_1$, $\hat{e}'_2 = -\hat{e}_2$ is orthogonal but is not a rotation. Let us call this transformation P. Thus any two-dimensional orthogonal matrix is a rotation or is P followed by a rotation. The set of all real orthogonal matrices in two dimensions is called $O(2)$, and the subset consisting of rotations is called $SO(2)$.

In three dimensions we need to take some care with what we mean by a rotation. On the one hand, we might mean that the transformation has some fixed axis and is a rotation through some angle about that axis. Let us call that a **rotation about an axis**. On the other hand, we might mean all transformations we can produce by a sequence of rotations about various axes. Let us define **rotation** in this sense. Clearly if we consider the rotation R which rotates the basis $\{\hat{e}\}$ into the basis $\{\hat{e}'\}$, and if we have another rotation R' which rotates $\{\hat{e}'\}$ into $\{\hat{e}''\}$, then the transformation which first does R and then does R', called the **composition** of them, $\breve{R} = R' \circ R$, is also a rotation in this latter sense. As $\hat{e}''_i = \sum_j R'_{ij} \hat{e}'_j = \sum_{ij} R'_{ij} R_{jk} \hat{e}_k$, we see that $\breve{R}_{ik} = \sum_j R'_{ij} R_{jk}$ and $\hat{e}''_i = \sum_k \breve{R}_{ik} \hat{e}_k$. Thus the composition $\breve{R} = R'R$ is given by matrix multiplication. In two dimensions, straightforward evaluation will verify that if R and R' are of the form (4.4) with angles θ and θ' respectively, the product $\breve{R}$ is of the same form with angle $\breve{\theta} = \theta + \theta'$. Thus all rotations are rotations about an axis there. Rotations in three dimensions are a bit more complex, because they can take place in different directions as well as through different angles. We can still represent the composition of rotations with matrix multiplication, now of 3×3 matrices. In general, matrices do not commute, $AB \neq BA$, and this is indeed reflected in the fact that the effect of performing two rotations depends in the order in which they are done. A graphic illustration is worth trying. Let V be the process of rotating an object through $90°$ about the vertical z-axis, and H be a rotation through $90°$ about the x-axis, which goes goes off to our right. If we start with the book lying face up facing us on the table, and first apply V and then H, we wind up with the binding down and the front of the book facing us. If, however, we start from the same position but apply first H and then V, we wind up with the book standing upright on the table with the binding towards us. Clearly the operations H and V do not commute.

It is clear that any composition of rotations must be orthogonal, as any set of orthonormal basis vectors will remain orthonormal under each transformation. It is also clear that there is a three dimensional version of P, say $\hat{e}'_1 = \hat{e}_1$, $\hat{e}'_2 = \hat{e}_2$, $\hat{e}'_3 = -\hat{e}_3$, which is orthogonal but not a composition of rotations, for it changes a right-handed coordinate system (with $\hat{e}_1 \times \hat{e}_2 = \hat{e}_3$) to a left handed one, while rotations preserve the handedness. It is straightforward to show, in any dimension N, that any composition of orthogonal matrices is orthogonal, for if $AA^T = 1\!\!1$ and $BB^T = 1\!\!1$ and $C = AB$, then $CC^T = AB(AB)^T = ABB^TA^T = A\, 1\!\!1\, A^T = 1\!\!1$, and C is orthogonal as well. So the rotations are a subset of the set $O(N)$ of orthogonal matrices.

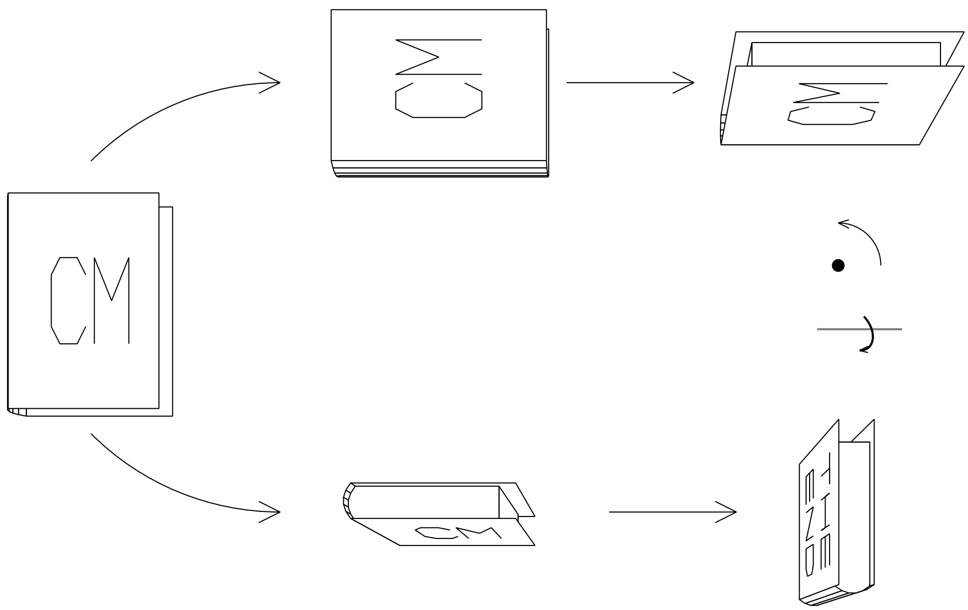


Figure 4.1: The results of applying the two rotations H and V to a book depends on which is done first. Thus rotations do not commute. Here we are looking down at a book which is originally lying face up on a table. V is a rotation about the vertical z-axis, and H is a rotation about a fixed axis pointing to the right, each through 90°.

4.1.2 Groups

This set of orthogonal matrices is a group, which means that the set $O(N)$ satisfies the following requirements, which we state for a general set G.

A set G of elements $A, B, C, ...$ together with a **group multiplication** rule ($\odot$) for combining two of them, is a **group** if

- *Given any two elements A and B in the group, the product $A \odot B$ is also in the group.* One then says that the set G is **closed under** $\odot$. In our case the group multiplication is ordinary matrix multiplication, the group consists of all $N \times N$ orthogonal real matrices, and we have just shown that it is closed.

- *The product rule is* **associative***; for every $A, B, C \in G$, we have $A \odot (B \odot C) = (A \odot B) \odot C$.* For matrix multiplication this is simply due to the commutivity of finite sums, $\sum_i \sum_j = \sum_j \sum_i$.

- *There is an element e in G, called the* **identity***, such that for every element $A \in G$, $e \odot A = A \odot e = A$.* In our case e is the unit matrix $\mathbb{1}$, $\mathbb{1}_{ij} = \delta_{ij}$.

- *Every element $A \in G$ has an element $A^{-1} \in G$ such that $A \odot A^{-1} = A^{-1} \odot A = e$.* This element is called the **inverse** of A, and in the case of orthogonal matrices is the inverse matrix, which always exists, because for orthogonal matrices the inverse is the transpose, which always exists for any matrix.

While the constraints (4.1) would permit $A(t)$ to be any orthogonal matrix, the nature of Newtonian mechanics of a rigid body requires it to vary continuously in time. If the system starts with $A = \mathbb{1}$, there must be a continuous path in the space of orthogonal matrices to the configuration $A(t)$ at any later time. But the set of matrices $O(3)$ is not connected in this fashion: there is no path from $A = \mathbb{1}$ to $A = P$. To see it is true, we look at the determinant of A. From $AA^T = \mathbb{1}$ we see that $\det(AA^T) = 1 = \det(A)\det(A^T) = (\det A)^2$ so $\det A = \pm 1$ for all orthogonal matrices A. But the determinant varies continuously as the matrix does, so no continuous variation of the matrix can lead to a jump in its determinant. Thus the matrices which represent rotations have unit determinant, $\det A = +1$, and are called **unimodular**.

The set of all unimodular orthogonal matrices in N dimensions is called $SO(N)$. It is a subset of $O(N)$, the set of all orthogonal matrices in N

dimensions. Clearly all rotations are in this subset. The subset is closed under multiplication, and the identity and the inverses of elements in $SO(N)$ are also in $SO(N)$, for their determinants are clearly 1. Thus $SO(N)$ is a **subgroup** of $O(N)$. It is actually the set of rotations, but we shall prove this statement only for the case $N = 3$, which is the immediately relevant one. Simultaneously we will show that every rotation in three dimensions is a rotation about an axis. We have already proven it for $N = 2$. We now show that every $A \in SO(3)$ has one vector it leaves unchanged or invariant, so that it is effectively a rotation in the plane perpendicular to this direction, or in other words a rotation about the axis it leaves invariant. The fact that every unimodular orthogonal matrix in three dimensions is a rotation about an axis is known as **Euler's Theorem**. To show that it is true, we note that if A is orthogonal and has determinant 1,

$$\begin{aligned}
\det\left\{(A - 1\!\!1)A^T\right\} &= \det(1\!\!1 - A^T) = \det(1\!\!1 - A) \\
&= \det(A - 1\!\!1)\det(A) = \det(-(1\!\!1 - A)) = (-1)^3\det(1\!\!1 - A) \\
&= -\det(1\!\!1 - A),
\end{aligned}$$

so $\det(1\!\!1 - A) = 0$ and $1\!\!1 - A$ is a singular matrix. Then there exists a vector $\vec{\omega}$ which is annihilated by it, $(1\!\!1 - A)\vec{\omega} = 0$, or $A\vec{\omega} = \vec{\omega}$, and $\vec{\omega}$ is invariant under A. Of course this determines only the direction of $\vec{\omega}$, and only up to sign. If we choose a new coordinate system in which the $\tilde{z}$-axis points along $\vec{\omega}$, we see that the elements $\tilde{A}_{i3} = (0, 0, 1)$, and orthogonality gives $\sum \tilde{A}_{3j}^2 = 1 = \tilde{A}_{33}^2$ so $\tilde{A}_{31} = \tilde{A}_{32} = 0$. Thus $\tilde{A}$ is of the form

$$\tilde{A} = \begin{pmatrix} \multicolumn{2}{c}{(\,B\,)} & \begin{matrix}0\\0\end{matrix} \\ 0 & 0 & 1 \end{pmatrix}$$

where B is an orthogonal unimodular 2×2 matrix, which is therefore a rotation about the z-axis through some angle ω, which we may choose to be in the range $\omega \in (-\pi, \pi]$. It is natural to define the vector $\vec{\omega}$, whose direction only was determined above, to be $\vec{\omega} = \omega\hat{e}_{\tilde{z}}$. Thus we see that the set of orthogonal unimodular matrices is the set of rotations, and elements of this set may be specified by a vector[2] of length $\leq \pi$.

[2] More precisely, we choose $\vec{\omega}$ along one of the two opposite directions left invariant by A, so that the the angle of rotation is non-negative and $\leq \pi$. This specifies a point in or on the surface of a three dimensional ball of radius π, but in the case when the angle is exactly π the two diametrically opposed points both describe the same rotation. Mathematicians say that the space of $SO(3)$ is three-dimensional real projective space $P_3(\mathbb{R})$[4].

Thus we see that the rotation which determines the orientation of a rigid body can be described by the three degrees of freedom $\vec{\omega}$. Together with the translational coordinates $\widetilde{R}$, this parameterizes the configuration space of the rigid body, which is six dimensional. It is important to recognize that this is not motion in a flat six dimensional configuration space, however. For example, the configurations with $\vec{\omega} = (0, 0, \pi - \epsilon)$ and $\vec{\omega} = (0, 0, -\pi + \epsilon)$ approach each other as $\epsilon \to 0$, so that motion need not even be continuous in $\vec{\omega}$. The composition of rotations is by multiplication of the matrices, *not* by addition of the $\vec{\omega}$'s. There are other ways of describing the configuration space, two of which are known as Euler angles and Cayley-Klein parameters, but none of these make describing the space very intuitive. For some purposes we do not need all of the complications involved in describing finite rotations, but only what is necessary to describe infinitesimal changes between the configuration at time t and at time $t + \Delta t$. We will discuss these applications first. Later, when we do need to discuss the configuration in section 4.4.2, we will define Euler angles.

4.2 Kinematics in a rotating coordinate system

We have seen that the rotations form a group. Let us describe the configuration of the body coordinate system by the position $\widetilde{R}(t)$ of a given point and the rotation matrix $A(t) : \hat{e}_i \to \hat{e}'_i$ which transforms the canonical fixed basis (inertial frame) into the body basis. A given particle of the body is fixed in the body coordinates, but this, of course, is not an inertial coordinate system, but a rotating and possibly accelerating one. We need to discuss the transformation of kinematics between these two frames. While our current interest is in rigid bodies, we will first derive a general formula for rotating (and accelerating) coordinate systems.

Suppose a particle has coordinates $\vec{b}(t) = \sum_i b'_i(t)\hat{e}'_i(t)$ in the body system. We are not assuming at the moment that the particle is part of the rigid body, in which case the $b'_i(t)$ would be independent of time. In the inertial coordinates the particle has its position given by $\vec{r}(t) = \widetilde{R}(t) + \vec{b}(t)$, but the coordinates of $\vec{b}(t)$ are different in the space and body coordinates. Thus

$$r_i(t) = \widetilde{R}_i(t) + b_i(t) = \widetilde{R}_i(t) + \sum_j \left(A^{-1}(t)\right)_{ij} b'_j(t).$$

The velocity is $\vec{v} = \sum_i \dot{r}_i \hat{e}_i$, because the $\hat{e}_i$ are inertial and therefore considered stationary, so

$$\vec{v} = \dot{\widetilde{R}} + \sum_{ij} \left[\left(\frac{d}{dt} A^{-1}(t) \right)_{ij} b'_j(t) + \left(A^{-1}(t) \right)_{ij} \frac{db'_j(t)}{dt} \right] \hat{e}_i,$$

and *not* $\dot{\widetilde{R}} + \sum_i (db'_i/dt)\hat{e}'_i$, because the $\hat{e}'_i$ are themselves changing with time. We might define a "body time derivative"

$$\left(\dot{\vec{b}}\right)_b := \left(\frac{d}{dt}\vec{b}\right)_b := \sum_i \left(\frac{db'_i}{dt}\right) \hat{e}'_i,$$

but it is *not* the velocity of the particle α, even with respect to $\widetilde{R}(t)$, in the sense that physically a vector is basis independent, and its derivative requires a notion of which basis vectors are considered time independent (inertial) and which are not. Converting the inertial evaluation to the body frame requires the velocity to include the dA^{-1}/dt term as well as the $\left(\dot{\vec{b}}\right)_b$ term.

What is the meaning of this extra term

$$\mathcal{V} = \sum_{ij} \left(\frac{d}{dt} A^{-1}(t) \right)_{ij} b'_j(t) \hat{e}_i \quad ?$$

The derivative is, of course,

$$\mathcal{V} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \sum_{ij} \left[A^{-1}(t + \Delta t)_{ij} - A^{-1}(t)_{ij} \right] b'_j(t) \hat{e}_i.$$

This expression has coordinates in the body frame with basis vectors from the inertial frame. It is better to describe it in terms of the body coordinates and body basis vectors by inserting $\hat{e}_i = \sum_k (A^{-1}(t)_{ik} \hat{e}'_k(t) = \sum_k A_{ki}(t) \hat{e}'_k(t)$. Then we have

$$\mathcal{V} = \sum_{kj} \hat{e}'_k \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[A(t) A^{-1}(t + \Delta t) - A(t) A^{-1}(t) \right]_{kj} b'_j(t).$$

The second term is easy enough to understand, as $A(t)A^{-1}(t) = 1\!\!\mathrm{I}$, so the full second term is just $\vec{b}$ expressed in the body frame. The interpretation of the first term is suggested by its matrix form: $A^{-1}(t + \Delta t)$ maps the body basis at $t + \Delta t$ to the inertial frame, and $A(t)$ maps this to the body basis at t. So together this is the infinitesimal rotation $\hat{e}'_i(t + \Delta t) \to \hat{e}'_i(t)$. This transformation must be close to an identity, as $\Delta t \to 0$. Let us expand it:

$$B := A(t)A^{-1}(t + \Delta t) = 1\!\!\mathrm{I} - \Omega' \Delta t + \mathcal{O}(\Delta t)^2. \tag{4.5}$$

Here Ω' is a matrix which has fixed (finite) elements as $\Delta t \to 0$, and is called the **generator** of the rotation. Note $B^{-1} = 1\!\!\mathrm{I} + \Omega' \Delta t$ to the order we are working, while the transpose $B^T = 1\!\!\mathrm{I} - \Omega'^T \Delta t$, so because we know B is orthogonal we must have that Ω' is **antisymmetric**, $\Omega' = -\Omega'^T$, $\Omega'_{ij} = -\Omega'_{ji}$.

Subtracting $1\!\!\mathrm{I}$ from both sides of (4.5) and taking the limit shows that the matrix

$$\Omega'(t) = -A(t) \cdot \frac{d}{dt} A^{-1}(t) = \left(\frac{d}{dt} A(t) \right) \cdot A^{-1}(t),$$

where the latter equality follows from differentiating $A \cdot A^{-1} = 1\!\!\mathrm{I}$. The antisymmetric 3×3 real matrix Ω' is determined by the three off-diagonal elements above the diagonal, $\Omega'_{23} = \omega'_1$, $\Omega'_{13} = -\omega'_2$, $\Omega'_{12} = \omega'_3$. as the others are given by antisymmetry. Thus it is effectively a vector. It is very useful to express this relationship by defining the **Levi-Civita** symbol ϵ_{ijk}, a totally antisymmetric rank 3 tensor specified by $\epsilon_{123} = 1$. Then the above expressions are given by $\Omega'_{ij} = \sum_k \epsilon_{ijk} \omega'_k$, and we also have

$$\frac{1}{2} \sum_{ij} \epsilon_{kij} \Omega'_{ij} = \frac{1}{2} \sum_{ij\ell} \epsilon_{kij} \epsilon_{ij\ell} \omega'_\ell = \omega'_k,$$

because, as explored in Appendix A.1,

$$\epsilon_{kij} = \epsilon_{ijk}, \qquad \sum_i \epsilon_{ijk} \epsilon_{ipq} = \delta_{jp} \delta_{kq} - \delta_{jq} \delta_{kp}, \qquad \text{so } \sum_{ij} \epsilon_{ijk} \epsilon_{ij\ell} = 2\delta_{k\ell}.$$

Thus ω'_k and Ω'_{ij} are essentially the same thing.

We have still not answered the question, "what is $\mathcal{V}$?"

$$\begin{aligned} \mathcal{V} &= \sum_{kj} \hat{e}'_k \lim_{\Delta t \to 0} \frac{1}{\Delta t} [B - 1\!\!\mathrm{I}]_{kj} b'_j = -\sum_{kj} \hat{e}'_k \Omega'_{kj} b'_j = -\sum_{kj\ell} \hat{e}'_k \epsilon_{kj\ell} \omega'_\ell b'_j \\ &= \vec{\omega} \times \vec{b}, \end{aligned}$$

where $\vec{\omega} = \sum_\ell \omega'_\ell \hat{e}'_\ell$. Note we have used Eq. A.4 for the cross-product. Thus we have shown that

$$\vec{v} = \dot{\tilde{R}} + \vec{\omega} \times \vec{b} + (\dot{\vec{b}})_b, \tag{4.6}$$

and the second term, coming from $\mathcal{V}$, represents the motion due to the rotating coordinate system.

When differentiating a true vector, which is independent of the origin of the coordinate system, rather than a position, the first term in (4.6) is absent, so in general for a vector $\vec{C}$,

$$\frac{d}{dt}\vec{C} = \left(\frac{d\vec{C}}{dt}\right)_b + \omega \times \vec{C}. \tag{4.7}$$

The velocity $\vec{v}$ is a vector, as are $\dot{\vec{R}}$ and $\vec{b}$, the latter because it is the difference of two positions. The angular velocity $\vec{\omega}$ is also a vector[3], and its derivative is particularly simple, because

$$\dot{\vec{\omega}} = \frac{d}{dt}\vec{\omega} = \left(\frac{d\vec{\omega}}{dt}\right)_b + \vec{\omega} \times \vec{\omega} = \left(\frac{d\vec{\omega}}{dt}\right)_b. \tag{4.8}$$

Another way to understand (4.7) is as a simple application of Leibnitz' rule to $\vec{C} = \sum C'_i \hat{e}'_i$, noting that

$$\frac{d}{dt}\hat{e}'_i(t) = \sum_j \frac{d}{dt}A_{ij}(t)\hat{e}_j = \sum_j \left(\Omega' A\right)_{ij}\hat{e}_j = \sum_k \Omega'_{ik}\hat{e}'_k,$$

which means that the second term from Leibnitz is

$$\sum C'_i \frac{d}{dt}\hat{e}'_i(t) = \sum_{ik} C'_i \Omega'_{ik}\hat{e}'_k = \sum_{ijk} C'_i \epsilon_{ikj}\omega'_j \hat{e}'_k = \vec{\omega} \times \vec{C},$$

as given in (4.7). This shows that even the peculiar object $(\dot{\vec{b}})_b$ obeys (4.7).

Applying this to the velocity itself (4.6), we find the acceleration

$$\begin{aligned}
\vec{a} &= \frac{d}{dt}\vec{v} = \frac{d}{dt}\dot{\vec{R}} + \frac{d\omega}{dt} \times \vec{b} + \omega \times \frac{d}{dt}\vec{b} + \frac{d}{dt}(\dot{\vec{b}})_b \\
&= \ddot{\vec{R}} + \dot{\vec{\omega}} \times \vec{b} + \omega \times \left[\left(\frac{d\vec{b}}{dt}\right)_b + \vec{\omega} \times \vec{b}\right] + \left(\frac{d^2\vec{b}}{dt^2}\right)_b + \omega \times \left(\frac{d\vec{b}}{dt}\right)_b \\
&= \ddot{\vec{R}} + \left(\frac{d^2\vec{b}}{dt^2}\right)_b + 2\omega \times \left(\frac{d\vec{b}}{dt}\right)_b + \dot{\vec{\omega}} \times \vec{b} + \vec{\omega} \times \left(\omega \times \vec{b}\right).
\end{aligned}$$

[3] Actually $\vec{\omega}$ is a **pseudovector**, which behaves like a vector under rotations but changes sign compared to what a vector does under reflection in a mirror.

This is a general relation between any orthonormal coordinate system and an inertial one, and in general can be used to describe physics in noninertial coordinates, regardless of whether that coordinate system is imbedded in a rigid body. The full force on the particle is $\vec{F} = m\vec{a}$, but if we use $\vec{r}$, $\vec{v}\,'$, and $\vec{a}\,'$ to represent $\vec{b}$, $(d\vec{b}/dt)_b$ and $(d^2\vec{b}/dt^2)_b$ respectively, we have an expression for the apparent force

$$m\vec{a}\,' = \vec{F} - m\ddot{\tilde{R}} - 2m\vec{\omega} \times \vec{v}\,' - m\dot{\vec{\omega}} \times \vec{r} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}).$$

The additions to the real force are the pseudoforce for an accelerating reference frame $-m\ddot{\tilde{R}}$, the Coriolus force $-2m\vec{\omega} \times \vec{v}\,'$, an unnamed force involving the angular acceleration of the coordinate system $-m\dot{\vec{\omega}} \times \vec{r}$, and the centrifugal force $-m\vec{\omega} \times (\vec{\omega} \times \vec{r})$ respectively.

4.3 The moment of inertia tensor

Let us return to a rigid body, where the particles are constrained to keep the distances between them constant. Then the coordinates $b'_{\alpha i}$ in the body frame are independant of time, and

$$\vec{v}_\alpha = \dot{\tilde{R}} + \omega \times \vec{b}_\alpha$$

so the individual momenta and the total momentum are

$$\begin{aligned} \vec{p}_\alpha &= m_\alpha \tilde{V} + m_\alpha \vec{\omega} \times \vec{b}_\alpha \\ \vec{P} &= M\tilde{V} + \vec{\omega} \times \sum_\alpha m_\alpha \vec{b}_\alpha \\ &= M\tilde{V} + M\vec{\omega} \times \vec{B} \end{aligned}$$

where $\vec{B}$ is the center of mass position relative to the marked point $\tilde{R}$.

4.3.1 Motion about a fixed point

Angular Momentum

We next evaluate the total angular momentum, $\vec{L} = \sum_\alpha \vec{r}_\alpha \times p_\alpha$. We will first consider the special case in which the body is rotating about the origin, so $\tilde{R} \equiv 0$, and then we will return to the general case. As $\vec{p}_\alpha = m_\alpha \vec{\omega} \times \vec{b}_\alpha$

already involves a cross product, we will find a triple product, and will use the reduction formula[4]

$$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \vec{B}\left(\vec{A} \cdot \vec{C}\right) - \vec{C}\left(\vec{A} \cdot \vec{B}\right).$$

Thus

$$\begin{aligned} \vec{L} &= \sum_\alpha m_\alpha \vec{b}_\alpha \times \left(\vec{\omega} \times \vec{b}_\alpha\right) && (4.9) \\ &= \vec{\omega} \sum_\alpha m_\alpha \vec{b}_\alpha^{\,2} - \sum_\alpha m_\alpha \vec{b}_\alpha \left(\vec{b}_\alpha \cdot \vec{\omega}\right). && (4.10) \end{aligned}$$

We see that, in general, $\vec{L}$ need not be parallel to the angular velocity $\vec{\omega}$, but it is always linear in $\vec{\omega}$. Thus it is possible to generalize the equation $\vec{L} = I\vec{\omega}$ of elementary physics courses, but we need to generalize I from a multiplicative number to a linear operator which maps vectors into vectors, not necessarily in the same direction. In component language this linear operation is clearly in the form $L_i = \sum_j I_{ij}\omega_j$, so $\mathbf{I}$ is a 3×3 matrix. Rewriting (4.10), we have

$$\begin{aligned} L_i &= \omega_i \sum_\alpha m_\alpha \vec{b}_\alpha^{\,2} - \sum_\alpha m_\alpha b_{\alpha i} \left(\vec{b}_\alpha \cdot \vec{\omega}\right). \\ &= \sum_j \sum_\alpha m_\alpha \left(\vec{b}_\alpha^{\,2}\delta_{ij} - b_{\alpha i} b_{\alpha j}\right)\omega_j \\ &\equiv \sum_j I_{ij}\omega_j, \end{aligned}$$

where

$$I_{ij} = \sum_\alpha m_\alpha \left(\vec{b}_\alpha^{\,2}\delta_{ij} - b_{\alpha i} b_{\alpha j}\right) \tag{4.11}$$

is the **inertia tensor** about the fixed point $\widetilde{R}$. In matrix form, we now have (4.10) as

$$\vec{L} = \mathbf{I} \cdot \vec{\omega}, \tag{4.12}$$

where $\mathbf{I} \cdot \vec{\omega}$ means a vector with components $(\mathbf{I} \cdot \vec{\omega})_i = \sum_j I_{ij}\omega_j$.

If we consider the rigid body in the continuum limit, the sum over particles becomes an integral over space times the density of matter,

$$I_{ij} = \int d^3b\, \rho(\vec{b}) \left(\vec{b}^{\,2}\delta_{ij} - b_i b_j\right). \tag{4.13}$$

[4]This formula is colloquially known as the **bac-cab** formula. It is proven in Appendix A.1.

Kinetic energy

For a body rotating about the origin

$$T = \frac{1}{2}\sum_\alpha m_\alpha \vec{v}_\alpha^{\,2} = \frac{1}{2}\sum_\alpha m_\alpha \left(\vec{\omega}\times\vec{b}_\alpha\right)\cdot\left(\vec{\omega}\times\vec{b}_\alpha\right).$$

From the general 3-dimensional identity[5]

$$\left(\vec{A}\times\vec{B}\right)\cdot\left(\vec{C}\times\vec{D}\right) = \vec{A}\cdot\vec{C}\vec{B}\cdot\vec{D} - \vec{A}\cdot\vec{D}\vec{B}\cdot\vec{C},$$

we have

$$\begin{aligned} T &= \frac{1}{2}\sum_\alpha m_\alpha \left[\vec{\omega}^{\,2}\vec{b}_\alpha^{\,2} - \left(\vec{\omega}\cdot\vec{b}_\alpha\right)^2\right] \\ &= \frac{1}{2}\sum_{ij}\omega_i\omega_j\sum_\alpha m_\alpha\left(\vec{b}_\alpha^{\,2}\delta_{ij} - \vec{b}_{\alpha i}\vec{b}_{\alpha j}\right) \\ &= \frac{1}{2}\sum_{ij}\omega_i I_{ij}\omega_j. \end{aligned} \tag{4.14}$$

or

$$T = \frac{1}{2}\vec{\omega}\cdot\mathbf{I}\cdot\vec{\omega}.$$

Noting that $\sum_j I_{ij}\omega_j = L_i$, $T = \frac{1}{2}\vec{\omega}\cdot\vec{L}$ for a rigid body rotating about the origin, with $\vec{L}$ measured from that origin.

4.3.2 More General Motion

When the marked point $\widetilde{R}$ is not fixed in space, there is nothing special about it, and we might ask whether it would be better to evaluate the moment of inertia about some other point. Working in the body-fixed coordinates, we may consider a given point $\vec{b}$ and evaluate the moment of inertia about that point, rather than about the origin. This means $\vec{b}_\alpha$ is replaced by $\vec{b}_\alpha - \vec{b}$, so

$$\begin{aligned} I_{ij}^{(\vec{b})} &= \sum_\alpha m_\alpha\left[\left(\vec{b}_\alpha - \vec{b}\right)^2\delta_{ij} - \left(b_{\alpha i} - b_i\right)\left(b_{\alpha j} - b_j\right)\right] \\ &= I_{ij}^{(0)} + M\left[\left(-2\vec{b}\cdot\vec{B} + b^2\right)\delta_{ij} + B_i b_j + b_i B_j - b_i b_j\right], \end{aligned} \tag{4.15}$$

where we recall that $\vec{B}$ is the position of the center of mass with respect to $\widetilde{R}$, the origin of the body fixed coordinates[6]. Subtracting the moment of inertia

[5] See Appendix A for a hint on how to derive this.

[6] $I^{(0)}$ is evaluated about the body-fixed position $\vec{b} = 0$, or about $\tilde{R}$, so it is given by Eq. 4.11.

about the center of mass, given by (4.15) with $b \to B$, we have

$$\begin{aligned} I_{ij}^{(\vec{b})} - I_{ij}^{(\vec{B})} &= M\left[\left(-2\vec{b}\cdot\vec{B} + b^2 + B^2\right)\delta_{ij} + B_i b_j + b_i B_j - b_i b_j - B_i B_j\right] \\ &= M\left[\left(\vec{b} - \vec{B}\right)^2 \delta_{ij} - (b_i - B_i)(b_j - B_j)\right]. \end{aligned} \tag{4.16}$$

Note the difference is independent of the origin of the coordinate system, depending only on the vector $\breve{b} = \vec{b} - \vec{B}$.

A possible axis of rotation can be specified by a point $\vec{b}$ through which it passes, together with a unit vector $\hat{n}$ in the direction of the axis[7]. The **moment of inertia** about the axis $(\vec{b}, \hat{n})$ is defined as $\hat{n}\cdot\mathbf{I}^{(\vec{b})}\cdot\hat{n}$. If we compare this to the moment about a parallel axis through the center of mass, we see that

$$\begin{aligned} \hat{n}\cdot\mathbf{I}^{(\vec{b})}\cdot\hat{n} - \hat{n}\cdot\mathbf{I}^{(\mathrm{cm})}\cdot\hat{n} &= M\left[\breve{b}^2\hat{n}^2 - (\breve{b}\cdot\hat{n})^2\right] \\ &= M(\hat{n}\times\breve{b})^2 = M\breve{b}_\perp^2, \end{aligned} \tag{4.17}$$

where $\breve{b}_\perp$ is the projection of the vector, from the center of mass to $\vec{b}$, onto the plane perpendicular to the axis. Thus the moment of inertia about any axis is the moment of inertia about a parallel axis through the center of mass, plus $M\ell^2$, where $\ell = \breve{b}_\perp$ is the distance between these two axes. This is known as the **parallel axis theorem**.

The general motion of a rigid body involves both a rotation and a translation of a given point $\widetilde{R}$. Then

$$\vec{r}_\alpha = \tilde{R} + \vec{b}_\alpha, \qquad \dot{\vec{r}}_\alpha = \widetilde{V} + \vec{\omega}\times\vec{b}_\alpha, \tag{4.18}$$

where $\widetilde{V}$ and $\vec{\omega}$ may be functions of time, but they are the same for all particles α. Then the angular momentum about the *origin* is

$$\begin{aligned} \vec{L} &= \sum_\alpha m_\alpha \vec{r}_\alpha \times \dot{\vec{r}}_\alpha = \sum_\alpha m_\alpha \vec{r}_\alpha \times \widetilde{V} + \sum_\alpha m_\alpha \left(\widetilde{R} + \vec{b}_\alpha\right)\times\left(\vec{\omega}\times\vec{b}_\alpha\right) \\ &= M\vec{R}\times\widetilde{V} + \mathbf{I}^{(0)}\cdot\vec{\omega} + M\widetilde{R}\times(\vec{\omega}\times\vec{B}), \end{aligned} \tag{4.19}$$

where the inertia tensor $\mathbf{I}^{(0)}$ is still measured[8] about $\widetilde{R}$, even though that is not a fixed point. Recall that $\vec{R}$ is the laboratory position of the center of

[7] Actually, this gives more information than is needed to specify an axis, as $\vec{b}$ and $\vec{b}'$ specify the same axis if $\vec{b} - \vec{b}' \propto \hat{n}$. In the expression for the moment of inertia about the axis, (4.17), we see that the component of $\vec{b}$ parallel to $\hat{n}$ does not affect the result.

[8] Recall the $(\vec{b})$ superscript in (4.15) refers to the body-fixed coordinate, so $\mathbf{I}^{(0)}$ is about $\vec{b} = 0$, not about the origin in inertial coordinates.

mass, while $\vec{B}$ is its position in the body-fixed system. The kinetic energy is now

$$\begin{aligned} T &= \sum_\alpha \frac{1}{2} m_\alpha \dot{\vec{r}}_\alpha^{\,2} = \frac{1}{2}\sum_\alpha m_\alpha \left(\widetilde{V} + \vec{\omega}\times\vec{b}_\alpha\right)\cdot\left(\widetilde{V} + \vec{\omega}\times\vec{b}_\alpha\right) \\ &= \frac{1}{2}\sum_\alpha m_\alpha \widetilde{V}^2 + \widetilde{V}\cdot\left(\vec{\omega}\times\sum_\alpha m_\alpha \vec{b}_\alpha\right) + \frac{1}{2}\sum_\alpha m_\alpha \left(\vec{\omega}\times\vec{b}_\alpha\right)^2 \\ &= \frac{1}{2}M\widetilde{V}^2 + M\widetilde{V}\cdot\left(\vec{\omega}\times\vec{B}\right) + \frac{1}{2}\vec{\omega}\cdot\mathbf{I}^{(0)}\cdot\vec{\omega} \end{aligned} \tag{4.20}$$

and again the inertia tensor $\mathbf{I}^{(0)}$ is calculated about the arbitrary point $\widetilde{R}$. We will see that it makes more sense to use the center of mass.

Simplification Using the Center of Mass

As each $\dot{\vec{r}}_\alpha = \widetilde{V} + \vec{\omega}\times\vec{b}_\alpha$, the center of mass velocity is given by

$$M\vec{V} = \sum_\alpha m_\alpha \dot{\vec{r}}_\alpha = \sum_\alpha m_\alpha \left(\widetilde{V} + \vec{\omega}\times\vec{b}_\alpha\right) = M\left(\widetilde{V} + \vec{\omega}\times\vec{B}\right), \tag{4.21}$$

so $\frac{1}{2}M\vec{V}^{\,2} = \frac{1}{2}M\widetilde{V}^2 + M\widetilde{V}\cdot(\vec{\omega}\times\vec{B}) + \frac{1}{2}M(\omega\times\vec{B})^2$. Comparing with 4.20, we see that

$$T = \frac{1}{2}M\vec{V}^{\,2} - \frac{1}{2}M(\vec{\omega}\times\vec{B})^2 + \frac{1}{2}\vec{\omega}\cdot\mathbf{I}^{(0)}\cdot\vec{\omega}.$$

The last two terms can be written in terms of the inertia tensor about the center of mass. From 4.16 with $\vec{b}=0$, as $\vec{B}$ is the center of mass,

$$I_{ij}^{(cm)} = I_{ij}^{(0)} - MB^2\delta_{ij} + MB_iB_j.$$

Using the formula for $\left(\vec{A}\times\vec{B}\right)\cdot\left(\vec{C}\times\vec{D}\right)$ again,

$$\begin{aligned} T &= \frac{1}{2}M\vec{V}^{\,2} - \frac{1}{2}M\left[\vec{\omega}^2\vec{B}^2 - \left(\vec{\omega}\cdot\vec{B}\right)^2\right] + \frac{1}{2}\vec{\omega}\cdot\mathbf{I}^{(0)}\cdot\vec{\omega} \\ &= \frac{1}{2}M\vec{V}^{\,2} + \frac{1}{2}\vec{\omega}\cdot\mathbf{I}^{(cm)}\cdot\vec{\omega}. \end{aligned} \tag{4.22}$$

A similar expression holds for the angular momentum. Inserting $\widetilde{V} = \vec{V} - \vec{\omega}\times\vec{B}$ into (4.19),

$$\vec{L} = M\widetilde{R}\times\left[\vec{V} - \vec{\omega}\times\vec{B}\right] + \mathbf{I}^{(0)}\cdot\vec{\omega} + M\widetilde{R}\times(\vec{\omega}\times\vec{B})$$

$$\begin{aligned}
&= M\vec{R}\times\vec{V} - M(\vec{R}-\widetilde{R})\times(\vec{\omega}\times\vec{B}) + \mathbf{I}^{(0)}\cdot\vec{\omega} \\
&= M\vec{R}\times\vec{V} - M\vec{B}\times(\vec{\omega}\times\vec{B}) + \mathbf{I}^{(0)}\cdot\vec{\omega} \\
&= M\vec{R}\times\vec{V} - M\vec{\omega}\,B^2 + M\vec{B}\,\vec{\omega}\cdot\vec{B} + \mathbf{I}^{(0)}\cdot\vec{\omega} \\
&= M\vec{R}\times\vec{V} + \mathbf{I}^{(cm)}\cdot\vec{\omega}. \qquad (4.23)
\end{aligned}$$

These two decompositions, (4.22) and (4.23), have a reasonable interpretation: the total angular momentum is the angular momentum about the center of mass, plus the angular momentum that a point particle of mass M and position $\vec{R}(t)$ would have. Similiarly, the total kinetic energy is the rotational kinetic energy of the body rotating about its center of mass, plus the kinetic energy of the fictious point particle moving with the center of mass.

Note that if we go back to the situation where the marked point $\widetilde{R}$ is stationary at the origin of the lab coordinates, $\widetilde{V}=0$, $\vec{L}=\mathbf{I}\cdot\vec{\omega}$, $T=\frac{1}{2}\vec{\omega}\cdot\mathbf{I}\cdot\vec{\omega}=\frac{1}{2}\vec{\omega}\cdot\vec{L}$.

The angular momentum in Eqs. 4.19 and 4.23 is the angular momentum measured about the origin of the lab coordinates, $\vec{L}=\sum_\alpha m_\alpha \vec{r}_\alpha\times v_\alpha$. It is useful to consider the angular momentum as measured about the center of mass,

$$\vec{L}^{\,\mathrm{cm}} = \sum_\alpha m_\alpha\left(\vec{r}_\alpha-\vec{R}\right)\times\left(\vec{v}_\alpha-\vec{V}\right) = \vec{L}-M\vec{R}\times\vec{V}, \qquad (4.24)$$

so we see that the angular momentum, measured about the center of mass, is just $\mathbf{I}^{(cm)}\cdot\vec{\omega}$.

The parallel axis theorem is also of the form of a decomposition. The inertia tensor about a given point $\vec{r}$ given by (4.16) is

$$I_{ij}^{(r)} = I_{ij}^{(cm)} + M\left[\left(\vec{r}-\vec{R}\right)^2\delta_{ij} - (r_i-R_i)\,(r_j-R_j)\right].$$

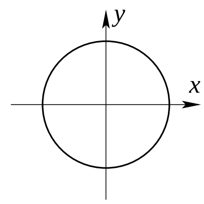
This is, once again, the sum of the quantity, here the inertia tensor, of the body about the center of mass, plus the value a particle of mass M at the center of mass $\vec{R}$ would have, evaluated about $\vec{r}$.

There is another theorem about moments of inertia, though much less general — it only applies to a planar object — let's say in the xy plane, so that $z_\alpha\approx 0$ for all the particles constituting the body. As

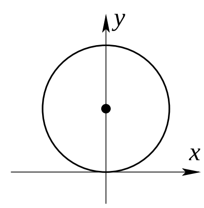
$$I_{zz} = \sum_\alpha m_\alpha\left(x_\alpha^2+y_\alpha^2\right)$$

$$\begin{aligned} I_{xx} &= \sum_\alpha m_\alpha \left(y_\alpha^2 + z_\alpha^2\right) = \sum_\alpha m_\alpha y_\alpha^2 \\ I_{yy} &= \sum_\alpha m_\alpha \left(x_\alpha^2 + z_\alpha^2\right) = \sum_\alpha m_\alpha x_\alpha^2, \end{aligned}$$

we see that $I_{zz} = I_{xx} + I_{yy}$, the moment of inertia about an axis perpendicular to the body is the sum of the moments about two perpendicular axes within the body, through the same point. This is known as the **perpendicular axis theorem**. As an example of its usefulness we calculate the moments for a thin uniform ring lying on the circle $x^2 + y^2 = R^2$, $z = 0$, about the origin. As every particle of the ring has the same distance R from the z-axis, the moment of inertia I_{zz} is simply MR^2. As $I_{xx} = I_{yy}$ by symmetry, and as the two must add up to I_{zz}, we have, by a simple indirect calculation, $I_{xx} = \frac{1}{2}MR^2$.



The parallel axis theorem (4.17) is also a useful calculational tool. Consider the moment of inertia of the ring about an axis parallel to its axis of symmetry but through a point on the ring. About the axis of symmetry, $I_{zz} = MR^2$, and $b_\perp = R$, so about a point on the ring, $I_{zz} = 2MR^2$. If instead, we want the moment about a tangent to the ring in the x direction, $I_{xx} = I_{xx}^{(cm)} + MR^2 = \frac{1}{2}MR^2 + MR^2 = 3MR^2/2$. Of course for I_{yy} the $b_\perp = 0$, so $I_{yy} = \frac{1}{2}MR^2$, and we may verify that $I_{zz} = I_{xx} + I_{yy}$ about this point as well.



For an object which has some thickness, with non-zero z components, the perpendicular axis theorem becomes an inequality, $I_{zz} \le I_{xx} + I_{yy}$.

Principal axes

If an object has an axial symmetry about z, we may use cylindrical polar coordinates (ρ, θ, z). Then its density $\mu(\rho, \theta, z)$ must be independent of θ, and

$$\begin{aligned} I_{ij} &= \int dz\, \rho\, d\rho\, d\theta\, \mu(\rho, z) \left[(\rho^2 + z^2)\delta_{ij} - r_i r_j\right], \\ \text{so} \qquad I_{xz} &= \int dz\, \rho\, d\rho\, d\theta\, \mu(\rho, z)(-z\rho\cos\theta) = 0 \\ I_{xy} &= \int dz\, \rho\, d\rho\, d\theta\, \mu(\rho, z)(\rho^2 \sin\theta\cos\theta) = 0 \end{aligned}$$

$$\begin{aligned} I_{xx} &= \int dz\,\rho\,d\rho\,d\theta\,\mu(\rho,z)\,\left[(\rho^2+z^2-\rho^2\cos^2\theta\right] \\ I_{yy} &= \int dz\,\rho\,d\rho\,d\theta\,\mu(\rho,z)\,\left[(\rho^2+z^2-\rho^2\sin^2\theta\right] \;=\; I_{xx} \end{aligned}$$

Thus the inertia tensor is diagonal and has two equal elements,

$$\mathbf{I} = \begin{pmatrix} I_{xx} & 0 & 0 \\ 0 & I_{xx} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix}.$$

In general, an object need not have an axis of symmetry, and even a diagonal inertia tensor need not have two equal "eigenvalues". Even if a body has no symmetry, however, there is always a choice of axes, a coordinate system, such that in this system the inertia tensor is diagonal. This is because I_{ij} is always a real symmetric tensor, and any such tensor can be brought to diagonal form by an orthogonal similiarity transformation[9]

$$\mathbf{I} = \mathcal{O}\mathbf{I}_D\mathcal{O}^{-1}, \quad \mathbf{I}_D = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{4.25}$$

An orthogonal matrix $\mathcal{O}$ is either a rotation or a rotation times P, and the P's can be commuted through $\mathbf{I}_D$ without changing its form, so there is a rotation $\mathcal{R}$ which brings the inertia tensor into diagonal form. The axes of this new coordinate system are known as the **principal axes**.

Tire balancing

Consider a rigid body rotating on an axle, and therefore about a fixed axis. What total force and torque will the axle exert? First, $\dot{\vec{R}} = \vec{\omega}\times\vec{R}$, so

$$\ddot{\vec{R}} = \dot{\vec{\omega}}\times\vec{R} + \vec{\omega}\times\dot{\vec{R}} = \dot{\vec{\omega}}\times\vec{R} + \vec{\omega}\times(\omega\times\vec{R}) = \dot{\vec{\omega}}\times\vec{R} + \vec{\omega}(\vec{\omega}\cdot\vec{R}) + \vec{R}\omega^2.$$

If the axis is fixed, $\vec{\omega}$ and $\dot{\vec{\omega}}$ are in the same direction, so the first term in the last expression is perpendicular to the other two. If we want the total force to be zero[10], $\ddot{\vec{R}} = 0$, so

$$\vec{R}\cdot\ddot{\vec{R}} = 0 = 0 + (\vec{\omega}\cdot\vec{R})^2 - R^2\omega^2.$$

[9] This should be proven in any linear algebra course. For example, see [1], Theorem 6 in Section 6.3.

[10] Here we are ignoring any constant force compensating the force exerted by the road which is holding the car up!

Thus the angle between $\vec{\omega}$ and $\vec{R}$ is 0 or π, and the center of mass must lie on the axis of rotation. This is the condition of static balance if the axis of rotation is horizontal in a gravitational field. Consider a car tire: to be stable at rest at any angle, $\vec{R}$ must lie on the axis or there will be a gravitational torque about the axis, causing rotation in the absense of friction. If the tire is not statically balanced, this force will rotate rapidly with the tire, leading to vibrations of the car.

Even if the net force is 0, there might be a torque. $\vec{\tau} = \dot{\vec{L}} = d(\mathbf{I} \cdot \vec{\omega})/dt$. If $\mathbf{I} \cdot \vec{\omega}$ is not parallel to $\vec{\omega}$ it will rotate with the wheel, and so $\dot{\vec{L}}$ will rapidly oscillate. This is also not good for your axle. If, however, $\vec{\omega}$ is parallel to one of the principal axes, $\mathbf{I} \cdot \vec{\omega}$ is parallel to $\vec{\omega}$, so if $\vec{\omega}$ is constant, so is $\vec{L}$, and $\vec{\tau} = 0$. The process of placing small weights around the tire to cause one of the principal axes to be aligned with the axle is called **dynamical balancing**.

Every rigid body has its principal axes; the problem of finding them and the moments of inertia about them, given the inertia tensor $\mathbf{I}$ in some coordiate system, is a mathematical question of finding a rotation $\mathcal{R}$ and "eigenvalues" I_1, I_2, I_3 (not components of a vector) such that equation 4.25 holds, with $\mathcal{R}$ in place of $\mathcal{O}$. The vector $\vec{v}_1 = \mathcal{R}\begin{pmatrix}1\\0\\0\end{pmatrix}$ is then an eigenvector, for

$$\mathbf{I} \cdot \vec{v}_1 = \mathcal{R}\mathbf{I}_D\mathcal{R}^{-1}\mathcal{R}\begin{pmatrix}1\\0\\0\end{pmatrix} = \mathcal{R}\mathbf{I}_D\begin{pmatrix}1\\0\\0\end{pmatrix} = I_1\mathcal{R}\begin{pmatrix}1\\0\\0\end{pmatrix} = I_1\vec{v}_1.$$

Similarly $\mathbf{I} \cdot \vec{v}_2 = I_2\vec{v}_2$ and $\mathbf{I} \cdot \vec{v}_3 = I_3\vec{v}_3$, where $\vec{v}_2$ and $\vec{v}_3$ are defined the same way, starting with $\hat{e}_2$ and $\hat{e}_3$ instead of $\hat{e}_1$. Note that, in general, $\mathbf{I}$ acts simply as a multiplier only for multiples of these three vectors individually, and not for sums of them. On a more general vector $\mathbf{I}$ will change the direction as well as the length of the vector it acts on.

Note that the I_i are all ≥ 0, for given any vector $\vec{n}$,

$$\vec{n} \cdot \mathbf{I} \cdot \vec{n} = \sum_\alpha m_\alpha[r_\alpha^2 n^2 - (\vec{r}_\alpha \cdot \vec{n})^2] = \sum_\alpha m_\alpha r_\alpha^2 n^2(1 - \cos^2\theta_\alpha) \geq 0,$$

so all the eigenvalues must be ≥ 0. It will be equal to zero only if all massive points of the body are in the $\pm\vec{n}$ directions, in which case the rigid body must be a thin line.

Finding the eigenvalues I_i is easier than finding the rotation $\mathcal{R}$. Consider the matrix $\mathbf{I}-\lambda\mathbb{1}$, which has the same eigenvectors as $\mathbf{I}$, but with eigenvalues $I_i - \lambda$. Then if λ is one of the eigenvalues I_i, this matrix will annihilate $\vec{v}_i$, so $\mathbf{I}-\lambda\mathbb{1}$ is a singular matrix with zero determinant. Thus the characteristic equation $\det(\mathbf{I}-\lambda\mathbb{1}) = 0$, which is a cubic equation in λ, gives as its roots the eigenvalues of $\mathbf{I}$.

4.4 Dynamics

4.4.1 Euler's Equations

So far, we have been working in an inertial coordinate system $\mathcal{O}$. In complicated situations this is rather unnatural; it is more natural to use a coordiate system $\mathcal{O}'$ fixed in the rigid body. In such a coordinate system, the vector one gets by differentiating the coefficients of a vector $\vec{b} = \sum b'_i \hat{e}'_i$ differs from the inertial derivative $\dot{\vec{b}}$ as given in Eq. 4.7. Consider two important special cases: either we have a system rotating about a fixed point $\widetilde{R}$, with $\vec{\tau}$, $\vec{L}$, and I'_{ij} all evaluated about that fixed point, or we are working about the center of mass, with $\vec{\tau}$, $\vec{L}$, and I'_{ij} all evaluated about the center of mass, even if it is in motion. In either case, we have $\vec{L} = \mathbf{I}' \cdot \vec{\omega}$, so for the time derivative of the angular momentum, we have

$$\begin{aligned} \vec{\tau} = \frac{d\vec{L}}{dt} &= \left(\frac{d\vec{L}}{dt}\right)_b + \vec{\omega} \times \vec{L} \\ &= \sum_{ij} \frac{d(I'_{ij}\omega'_j)}{dt} \hat{e}'_i + \vec{\omega} \times (I' \cdot \vec{\omega}), \end{aligned}$$

Now in the $\mathcal{O}'$ frame, all the masses are at fixed positions, so I'_{ij} is constant, and the first term is simply $\mathbf{I} \cdot (d\omega/dt)_b$, which by (4.8) is simply $\mathbf{I} \cdot \dot{\vec{\omega}}$. Thus we have (in the body coordinate system)

$$\vec{\tau} = \mathbf{I}' \cdot \dot{\vec{\omega}} + \vec{\omega} \times (\mathbf{I}' \cdot \omega). \tag{4.26}$$

We showed that there is always a choice of cartesian coordinates mounted on the body along the principal axes. For the rest of this section we will use this body-fixed coordinate system, so we will drop the primes.

The torque not only determines the rate of change of the angular momentum, but also does work in the system. For a system rotating about a fixed point, we see from the expression (4.14), $T = \frac{1}{2}\vec{\omega} \cdot \mathbf{I} \cdot \vec{\omega}$, that

$$\frac{dT}{dt} = \frac{1}{2}\dot{\vec{\omega}} \cdot \mathbf{I} \cdot \vec{\omega} + \frac{1}{2}\vec{\omega} \cdot \dot{\mathbf{I}} \cdot \vec{\omega} + \frac{1}{2}\vec{\omega} \cdot \mathbf{I} \cdot \dot{\vec{\omega}}.$$

The first and last terms are equal because the inertia tensor is symmetric, $I_{ij} = I_{ji}$, and the middle term vanishes in the body-fixed coordinate system because all particle positions are fixed. Thus $dT/dt = \vec{\omega} \cdot \mathbf{I} \cdot \dot{\vec{\omega}} = \vec{\omega} \cdot \dot{\vec{L}} = \vec{\omega} \cdot \vec{\tau}$. Thus the kinetic energy changes due to the work done by the external torque. Therefore, of course, if there is no torque the kinetic energy is constant.

We will write out explicitly the components of Eq. 4.26. In evaluating τ_1, we need the first component of the second term,

$$[(\omega_1, \omega_2, \omega_3) \times (I_1\omega_1, I_2\omega_2, I_3\omega_3)]_1 = (I_3 - I_2)\omega_2\omega_3.$$

Inserting this and the similar expressions for the other components into Eq. (4.26), we get **Euler's equations**

$$\begin{aligned} \tau_1 &= I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3, \\ \tau_2 &= I_2\dot{\omega}_2 + (I_1 - I_3)\omega_1\omega_3, \\ \tau_3 &= I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2. \end{aligned} \tag{4.27}$$

Using these equations we can address several situations of increasing difficulty.

First, let us ask under what circumstances the angular velocity will be fixed in the absense of a torque. As $\vec{\tau} = \dot{\vec{\omega}} = 0$, from the 1-component equation we conclude that $(I_2 - I_3)\omega_2\omega_3 = 0$. Then either the moments are equal ($I_2 = I_3$) or one of the two components ω_2 or ω_3 must vanish. Similarly, if $I_1 \neq I_2$, either ω_1 or ω_2 vanishes. So the only way more than one component of $\vec{\omega}$ can be nonzero is if two or more of the principal moments are equal. In this case, the principal axes are not uniquely determined. For example, if $I_1 = I_2 \neq I_3$, the third axes is unambiguously required as one of the principle axes, but any direction in the (12)-plane will serve as the second principal axis. In this case we see that $\vec{\tau} = \dot{\vec{\omega}} = 0$ implies either $\vec{\omega}$ is along the z-axis ($\omega_1 = \omega_2 = 0$) or it lies in the (12)-plane, ($\omega_3 = 0$). In any case, the angular velocity is constant in the absence of torques only if it lies along a principal axis of the body.

As our next example, consider an axially symmetric body with no external forces or torques acting on it. Then $\dot{\vec{R}}$ is a constant, and we will choose to work in an inertial frame where $\vec{R}$ is fixed at the origin. Choosing our body-fixed coordinates with z along the axis of symmetry, our axes are principal ones and $I_1 = I_2$, so we have

$$\begin{aligned} I_1\dot\omega_1 &= (I_1 - I_3)\omega_2\omega_3, \\ I_1\dot\omega_2 &= (I_3 - I_1)\omega_1\omega_3, \\ I_3\dot\omega_3 &= (I_1 - I_2)\omega_1\omega_2 = 0. \end{aligned}$$

We see that ω_3 is a constant. Let $\Omega = \omega_3(I_3 - I_1)/I_1$. Then we see that

$$\dot\omega_1 = -\Omega\omega_2, \qquad \dot\omega_2 = \Omega\omega_1.$$

Differentiating the first and plugging in the second, we find

$$\ddot\omega_1 = -\Omega\dot\omega_2 = -\Omega^2\omega_1,$$

which is just the harmonic oscillator equation. So $\omega_1 = A\cos(\Omega t + \phi)$ with some arbitrary amplitude A and constant phase ϕ, and $\omega_2 = -\dot\omega_1/\Omega = A\sin(\Omega t + \phi)$. We see that, in the body-fixed frame, the angular velocity rotates about the axis of symmetry in a circle, with arbitrary radius A, and a period $2\pi/\Omega$. The angular velocity vector $\vec\omega$ is therefore sweeping out a cone, called the **body cone** of precession with a half-angle $\phi_b = \tan^{-1} A/\omega_3$. Note the length of $\vec\omega$ is fixed.

What is happening in the lab frame? The kinetic energy $\frac{1}{2}\vec\omega\cdot\vec L$ is constant, as is the vector $\vec L$ itself. As the length of a vector is frame independent, $|\vec\omega|$ is fixed as well. Therefore the angle between them, called the **lab angle**, is constant,

$$\cos\phi_L = \frac{\vec\omega\cdot\vec L}{|\vec\omega||\vec L|} = \frac{2T}{|\vec\omega||\vec L|} = \text{constant}. \tag{4.28}$$

Thus $\vec\omega$ rotates about $\vec L$ in a cone, called the laboratory cone.

Note that ϕ_b is the angle between $\vec\omega$ and the z-axis of the body, while ϕ_L is the angle between $\vec\omega$ and $\vec L$, so they are not the same angle in two different coordinate systems.

The situation is a bit hard to picture. In the body frame it is hard to visualize $\vec\omega$, although that is the negative of the angular velocity of the universe in that system. In the lab frame the body is instantanously rotating about the axis $\vec{\omega}$, but this axis is not fixed in the body. At any instant, the points on this line are not moving, and we may think of the body rolling without slipping on the lab cone, with $\vec{\omega}$ the momentary line of contact. Thus the body cone rolls on the lab cone without slipping.

The Poinsot construction

This idea has an extension to the more general case where the body has no symmetry. The motion in this case can be quite complex, both for analytic solution, because Euler's equations are nonlinear, and to visualize, because the body is rotating and bobbing around in a complicated fashion. But as we are assuming there are no external forces or torques, the kinetic energy and total angular momentum vectors are constant, and this will help us understand the motion. To do so we construct an abstract object called the inertia ellipsoid. Working in the body frame, consider that the equation

$$2T = \sum_{ij} \omega_i I_{ij} \omega_j = f(\vec{\omega})$$

is a quadratic equation for $\vec{\omega}$, with constant coefficients, which therefore determines an ellipsoid[11] in the space of possible values of $\vec{\omega}$. This is called the **inertia ellipsoid**[12]. It is fixed in the body, and so if we were to scale it by some constant to change units from angular velocity to position, we could think of it as a fixed ellipsoid in the body itself, centered at the center of mass. At every moment the instantanous value of $\vec{\omega}$ must lie on this ellipsoid, so $\vec{\omega}(t)$ sweeps out a path on this ellipsoid called the **polhode**.

If we go to the lab frame, we see this ellipsoid fixed in and moving with the body. The instantaneous value of $\vec{\omega}$ still lies on it. In addition, the component of $\vec{\omega}$ in the (fixed) $\vec{L}$ direction is fixed, and as the center of mass is fixed, the point corresponding to $\vec{\omega}$ lies in a plane perpendicular to $\vec{L}$ a fixed distance from the center of mass, known as the **invariant plane**. Finally we note that the normal to the surface of the ellipsoid $f(\vec{\omega}) = 2T$ is parallel to $\nabla f = 2I \cdot \vec{\omega} = 2\vec{L}$, so the ellipsoid of inertia is tangent to the invariant plane

[11]We assume the body is not a thin line, so that I is a positive definite matrix (all its eigenvalues are strictly > 0), so the surface defined by this equation is bounded.

[12]Exactly which quantity forms the inertia ellipsoid varies by author. Goldstein scales $\vec{\omega}$ by a constant $1/\sqrt{2T}$ to form an object ρ whose ellipsoid he calls the inertia ellipsoid. Landau and Lifshitz discuss an ellipsoid of $\vec{L}$ values but don't give it a name. They then call the corresponding path swept out by $\vec{\omega}$ the polhode, as we do.

at the point $\vec{\omega}(t)$. The path that $\vec{\omega}(t)$ sweeps out on the invariant plane is called the **herpolhode**. At this particular moment, the point corresponding to $\vec{\omega}$ in the body is not moving, so the inertia ellipsoid is rolling, not slipping, on the invariant plane.

In general, if there is no special symmetry, the inertia ellipsoid will not be axially symmetric, so that in order to roll on the fixed plane and keep its center at a fixed point, it will need to bob up and down. But in the special case with axial symmetry, the inertia ellipsoid will also have this symmetry, so it can roll about a circle, with its symmetry axis at a fixed angle relative to the invariant plane. In the body frame, ω_3 is fixed and the polhode moves on a circle of radius $A = \omega \sin \phi_b$. In the lab frame, $\vec{\omega}$ rotates about $\vec{L}$, so it sweeps out a circle of radius $\omega \sin \phi_L$ in the invariant plane. One circle is rolling on the other, and the polhode rotates about its circle at the rate Ω in the body frame, so the angular rate at which the herpolhode rotates about $\vec{L}$, Ω_L, is

$$\Omega_L = \Omega \frac{\text{circumference of polhode circle}}{\text{circumference of herpolhode circle}} = \frac{I_3 - I_1}{I_1} \omega_3 \frac{\sin \phi_b}{\sin \phi_L}.$$

Stability of rotation about an axis

We have seen that the motion of a isolated rigid body is simple only if the angular velocity is along one of the principal axes, and can be very complex otherwise. However, it is worth considering what happens if $\vec{\omega}$ is very nearly, but not exactly, along one of the principal axes, say z. Then we may write $\vec{\omega} = \omega_3 \hat{e}_3 + \vec{\epsilon}$ in the body coordinates, and assume $\epsilon_3 = 0$ and the other components are small. We treat Euler's equations to first order in the small quantity $\vec{\epsilon}$. To this order, $\dot{\omega}_3 = (I_1 - I_2)\epsilon_1 \epsilon_2 / I_3 \approx 0$, so ω_3 may be considered a constant. The other two equations give

$$\begin{aligned} \dot{\omega}_1 = \dot{\epsilon}_1 &= \frac{I_2 - I_3}{I_1} \epsilon_2 \omega_3 \\ \dot{\omega}_2 = \dot{\epsilon}_2 &= \frac{I_3 - I_1}{I_2} \epsilon_1 \omega_3 \end{aligned}$$

so

$$\ddot{\epsilon}_1 = \frac{I_2 - I_3}{I_1} \frac{I_3 - I_1}{I_2} \omega_3^2 \epsilon_1.$$

What happens to $\vec{\epsilon}(t)$ depends on the sign of the coefficient, or the sign of $(I_2 - I_3)(I_3 - I_1)$. If it is negative, ϵ_1 oscillates, and indeed $\vec{\epsilon}$ rotates about z just as we found for the symmetric top. This will be the case if I_3 is either the largest or the smallest eigenvalue. If, however, it is the middle eigenvalue, the constant will be positive, and the equation is solved by exponentials, one damping out and one growing. Unless the initial conditions are perfectly fixed, the growing piece will have a nonzero coefficient and $\vec{\epsilon}$ will blow up. Thus a rotation about the intermediate principal axis is unstable, while motion about the axes with the largest and smallest moments are stable. For the case where two of the moments are equal, the motion will be stable about the third, and slightly unstable ($\vec{\epsilon}$ will grow linearly instead of exponentially with time) about the others.

An interesting way of understanding this stability or instability of rotation close to a principle axes involves another ellipsoid we can define for the free rigid body, an ellipsoid of possible angular momentum values. Of course in the inertial coordinates $\vec{L}$ is constant, but in body fixed language the coordinates vary with time, though the length of $\vec{L}$ is still constant. In addition, the conservation of kinetic energy

$$2T = \vec{L} \cdot \mathbf{I}^{-1} \cdot \vec{L}$$

(where $\mathbf{I}^{-1}$ is the inverse of the moment of inertia matrix) gives a quadratic equation for the three components of $\vec{L}$, just as we had for $\vec{\omega}$ and the ellipsoid of inertia. The path of $\vec{L}(t)$ on this ellipsoid is on the intersection of the ellisoid with a sphere of radius $|\vec{L}|$, for the length is fixed.

If $\vec{\omega}$ is near the principle axis with the largest moment of inertia, $\vec{L}$ lies near the major axis of the ellipsoid. The sphere is nearly circumscribing the ellipsoid, so the intersection consists only of two small loops surrounding each end of the major axis. Similiarly if $\vec{\omega}$ is near the smallest moment, the sphere is nearly inscribed in the ellipsoid, and again the possible values of $\vec{L}$ lie close to either end of the minor axis. Thus the subsequent motion is confined to one of these small loops. But if $\vec{\omega}$ starts near the intermediate principle axis, $\vec{L}$ does likewise, and the intersection consists of two loops which extend from near one end to near the other of the intermediate axis, and the possible continuous motion of $\vec{L}$ is not confined to a small region of the ellipsoid.

Because the rotation of the Earth flattens the poles, the Earth is approximately an **oblate** ellipsoid, with I_3 greater than $I_1 = I_2$ by about one part in 300. As ω_3 is 2π per siderial day, if $\vec{\omega}$ is not perfectly aligned with the axis, it will precess about the symmetry axis once every 10 months. This **Chandler wobble** is not of much significance, however, because the body angle $\phi_b \approx 10^{-6}$.

4.4.2 Euler angles

Up to this point we have managed to describe the motion of a rigid body without specifying its coordinates. This is not possible for most problems with external forces, for which the torque will generally depend on the orientation of the body. It is time to face up to the problem of using three generalized coordinates to describe the orientation.

In section 4.1.1 we described the orientation of a rigid body in terms of a rotation through a finite angle in a given direction, specified by ω. This does not give a simple parameterization of the matrix A, and it is more common to use an alternate description known as Euler angles. Here we describe the rotation A as a composition of three simpler rotations about specified coordinates, so that we are making a sequence of changes of coordinates

$$(x, y, z) \xrightarrow{R_z(\phi)} (x_1, y_1, z_1) \xrightarrow{R_{y_1}(\theta)} (x_2, y_2, z_2) \xrightarrow{R_{z_2}(\psi)} (x', y', z').$$

We have chosen three specific directions about which to make the three rotations, namely the original z-axis, the next y-axis, y_1, and then the new z-axis, which is both z_2 and z'. This choice is not universal, but is the one generally used in quantum mechanics. Many of the standard classical mechanics texts[13] take the second rotation to be about the x_1-axis instead of y_1, but quantum mechanics texts[14] avoid this because the action of R_y on a spinor is real, while the action of R_x is not. While this does not concern us here, we prefer to be compatible with quantum mechanics discussions.

This procedure is pictured in Figure 4.2. To see that any rotation can be written in this form, and to determine the range of the angles, we first discuss what fixes the y_1 axis. Notice that the rotation about the z-axis leaves z uneffected, so $z_1 = z$, Similiarly, the last rotation leaves the z_2 axis unchanged, so it is also the z' axis. The planes orthogonal to these axes are also left invariant[15]. These planes, the xy-plane and the $x'y'$-plane respectively, intersect in a line called the **line of nodes**[16]. These planes are also the x_1y_1 and x_2y_2 planes respectively, and as the second rotation

[13] See [2], [6], [9], [10], [11] and [17].

[14] For example [13] and [20].

[15] although the points in the planes are rotated by 4.4.

[16] The case where the xy and $x'y'$ are identical, rather than intersecting in a line, is exceptional, corresponding to $\theta = 0$ or $\theta = \pi$. Then the two rotations about the z-axis add or subtract, and many choices for the Euler angles (ϕ, ψ) will give the same full rotation.

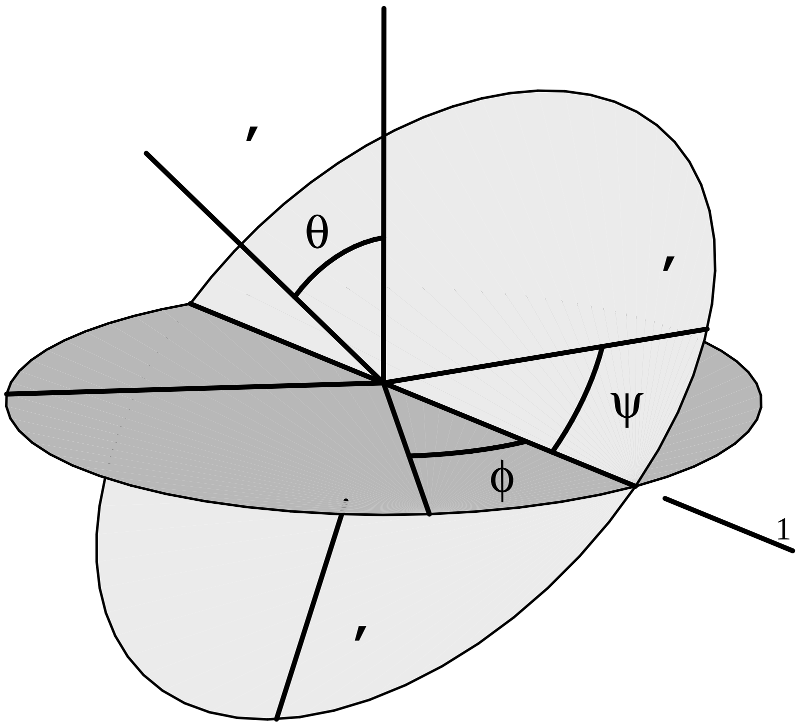


Figure 4.2: The Euler angles as rotations through ϕ, θ, ψ, about the z, y_1, and z_2 axes sequentially

$R_{y_1}(\theta)$ must map the first into the second plane, we see that y_1, which is unaffected by R_{y_1}, must be along the line of nodes. We choose between the two possible orientations of y_1 to keep the necessary θ angle in $[0, \pi]$. The angles ϕ and ψ are then chosen $\in [0, 2\pi)$ as necessary to map $y \to y_1$ and $y_1 \to y'$ respectively.

While the rotation about the z-axis leaves z uneffected, it rotates the x and y components by the matrix (4.4). Thus in three dimensions, a rotation about the z axis is represented by

$$R_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.29}$$

Similarly a rotation through an angle θ about the current y axis has a similar form

$$R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{4.30}$$

The reader needs to assure himself, by thinking of the rotations as active transformations, that the action of the matrix R_y after having applied R_z produces a rotation about the y_1-axis, not the original y-axis.

The full rotation $A = R_z(\psi) \cdot R_y(\theta) \cdot R_z(\phi)$ can then be found simply by matrix multiplication:

$$\begin{aligned} A(\phi,\theta,\psi) &= \\ &\begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \\ &\begin{pmatrix} -\sin\phi\sin\psi + \cos\theta\cos\phi\cos\psi & \cos\phi\sin\psi + \cos\theta\sin\phi\cos\psi & -\sin\theta\cos\psi \\ -\sin\phi\cos\psi - \cos\theta\cos\phi\sin\psi & \cos\phi\cos\psi - \cos\theta\sin\phi\sin\psi & \sin\theta\sin\psi \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix}. \end{aligned} \tag{4.31}$$

We need to reexpress the kinetic energy in terms of the Euler angles and their time derivatives. From the discussion of section 4.2, we have

$$\Omega' = -A(t) \cdot \frac{d}{dt} A^{-1}(t)$$

The inverse matrix is simply the transpose, so finding Ω' can be done by straightforward differentiation and matrix multiplication[17]. The result is

$$\begin{aligned} \Omega' &= \\ &\begin{pmatrix} 0 & \dot\psi + \dot\phi\cos\theta & -\dot\theta\cos\psi - \dot\phi\sin\theta\sin\psi \\ -\dot\psi - \dot\phi\cos\theta & 0 & \dot\theta\sin\psi - \dot\phi\sin\theta\cos\psi \\ \dot\theta\cos\psi + \dot\phi\sin\theta\sin\psi & -\dot\theta\sin\psi + \dot\phi\sin\theta\cos\psi & 0 \end{pmatrix}. \end{aligned} \tag{4.32}$$

Note Ω' is antisymmetric as expected, so it can be recast into the axial vector ω

$$\begin{aligned} \omega'_1 &= \Omega'_{23} = \dot\theta\sin\psi - \dot\phi\sin\theta\cos\psi, \\ \omega'_2 &= \Omega'_{31} = \dot\theta\cos\psi + \dot\phi\sin\theta\sin\psi, \\ \omega'_3 &= \Omega'_{12} = \dot\psi + \dot\phi\cos\theta. \end{aligned} \tag{4.33}$$

[17] Verifying the above expression for A and the following one for Ω' is a good application for a student having access to a good symbolic algebra computer program. Both Mathematica and Maple handle the problem nicely.

This expression for $\vec{\omega}$ gives the necessary velocities for the kinetic energy term (4.20 or 4.22) in the Lagrangian, which becomes

$$L = \frac{1}{2}M\tilde{V}^2 + M\tilde{V}\cdot\left(\vec{\omega}\times\vec{B}\right) + \frac{1}{2}\vec{\omega}\cdot I^{(\tilde{R})}\cdot\vec{\omega} - U(\tilde{R},\theta,\psi,\phi), \tag{4.34}$$

or

$$L = \frac{1}{2}M\vec{V}^{\,2} + \frac{1}{2}\vec{\omega}\cdot I^{(cm)}\cdot\vec{\omega} - U(\vec{R},\theta,\psi,\phi), \tag{4.35}$$

with $\vec{\omega} = \sum_i \omega_i'\hat{e}_i'$ given by (4.33).

4.4.3 The symmetric top

Now let us consider an example with external forces which constrain one point of a symmetrical top to be stationary. Then we choose this to be the fixed point, at the origin $\tilde{R} = 0$, and we choose the body-fixed z'-axis to be along the axis of symmetry. Of course the center of mass in on this axis, so $\vec{R} = (0,0,\ell)$ in body-fixed coordinates. We will set up the motion by writing the Lagrangian from the forms for the kinetic and potential energy, due entirely to the gravitational field[18].

$$\begin{aligned}
T &= \frac{1}{2}(\omega_1^2+\omega_2^2)I_1 + \frac{1}{2}\omega_3^2 I_3 \\
&= \frac{1}{2}\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right)I_1 + \frac{1}{2}\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 I_3, \qquad (4.36)\\
U &= Mgz_{\rm cm} = Mg\ell\left(A^{-1}\right)_{zz} = Mg\ell\cos\theta. \qquad (4.37)
\end{aligned}$$

So $L = T - U$ is independent of ϕ, ψ, and the corresponding momenta

$$\begin{aligned}
p_\phi &= \dot{\phi}\sin^2\theta I_1 + \left(\dot{\phi}\cos\theta + \dot{\psi}\right)\cos\theta I_3 \\
&= \dot{\phi}\sin^2\theta I_1 + \cos\theta\omega_3 I_3, \\
p_\psi &= \left(\dot{\phi}\cos\theta + \dot{\psi}\right)I_3 = \omega_3 I_3
\end{aligned}$$

are constants of the motion. Let us use parameters $a = p_\psi/I_1$ and $b = p_\phi/I_1$, which are more convenient, to parameterize the motion, instead of p_ϕ, p_ψ, or

[18] As we did in discussing Euler's equations, we drop the primes on ω_i and on I_{ij} even though we are evaluating these components in the body fixed coordinate system. The coordinate z, however, is still a lab coordinate, with $\hat{e}_z$ pointing upward.

even ω_3, which is also a constant of the motion and might seem physically a more natural choice. A third constant of the motion is the energy,

$$E = T + U = \frac{1}{2} I_1 \left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} \omega_3^2 I_3 + Mg\ell \cos \theta.$$

Solving for $\dot{\phi}$ from $p_\phi = I_1 b = \dot{\phi} \sin^2 \theta I_1 + I_1 a \cos \theta$,

$$\dot{\phi} = \frac{b - a\cos\theta}{\sin^2\theta}, \tag{4.38}$$

$$\dot{\psi} = \omega_3 - \dot{\phi}\cos\theta = \frac{I_1 a}{I_3} - \frac{b - a\cos\theta}{\sin^2\theta}\cos\theta, \tag{4.39}$$

Then E becomes

$$E = \frac{1}{2} I_1 \dot{\theta}^2 + U'(\theta) + \frac{1}{2} I_3 \omega_3^2,$$

where

$$U'(\theta) := \frac{1}{2} I_1 \frac{(b - a\cos\theta)^2}{\sin^2\theta} + Mg\ell\cos\theta.$$

The term $\frac{1}{2} I_3 \omega_3^2$ is an ignorable constant, so we consider $E' := E - \frac{1}{2} I_3 \omega_3^2$ as the third constant of the motion, and we now have a one dimensional problem for $\theta(t)$, with a first integral of the motion. Once we solve for $\theta(t)$, we can plug back in to find $\dot{\phi}$ and $\dot{\psi}$.

Substitute $u = \cos\theta$, $\dot{u} = -\sin\theta\dot{\theta}$, so

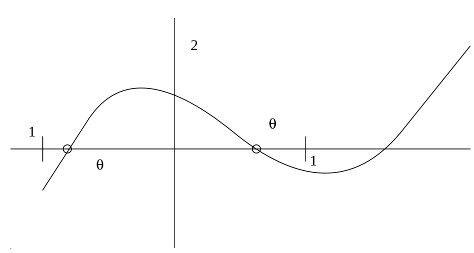
$$E' = \frac{I_1 \dot{u}^2}{2(1-u^2)} + \frac{1}{2} I_1 \frac{(b - au)^2}{1 - u^2} + Mg\ell u,$$

or

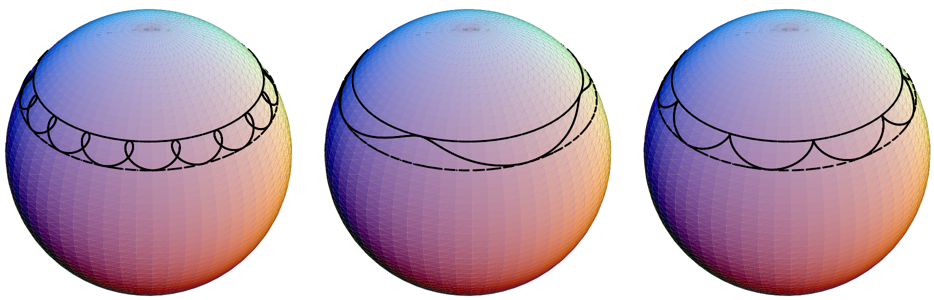
$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2 =: f(u), \tag{4.40}$$

with $\alpha = 2E'/I_1$, $\beta = 2Mg\ell/I_1$.

$f(u)$ is a cubic with a positive u^3 term, and is negative at $u = \pm 1$, where the first term vanishes, and which are also the limits of the physical range of values of u. If there are to be any allowed values for $\dot{u}^2$, $f(u)$ must be nonnegative somewhere in $u \in [-1, 1]$, so f must look very much like what is shown.



To visualize what is happening, note that a point on the symmetry axis moves on a sphere, with θ and ϕ representing the usual spherical coordinates, as can be seen by examining what A^{-1} does to $(0, 0, z')$. So as θ moves back and forth between $\theta_{\min}$ and $\theta_{\max}$, the top is wobbling closer and further from the vertical, called **nutation**. At the same time, the symmetry axis



$\theta' = 52°$ $\theta' = 44°$ $\theta' = \theta_{\min}$

Figure 4.3: Possible loci for a point on the symmetry axis of the top. The axis nutates between $\theta_{\min} = 50°$ and $\theta_{\max} = 60°$

is **precessing**, rotating about the vertical axis, at a rate $\dot{\phi}$ which is not constant but a function of θ (Eq. 4.38). Qualitatively we may distinguish three kinds of motion, depending on the values of $\dot{\phi}$ at the turning points in θ. These in turn depend on the initial conditions and the parameters of the top, expressed in a, b, and $\theta_{\min}$, $\theta_{\max}$. If the value of $u' = \cos\theta'$ at which $\dot{\phi}$ vanishes is within the range of nutation, then the precession will be in different directions at $\theta_{\min}$ and $\theta_{\max}$, and the motion is as in Fig. 4.3a. On the other hand, if $\theta' = \cos^{-1}(b/a) \notin [\theta_{\min}, \theta_{\max}]$, the precession will always be in the same direction, although it will speed up and slow down. We then get a motion as in Fig. 4.3b. Finally, it is possible that $\cos\theta_{\min} = b/a$, so that the precession stops at the top, as in Fig. 4.3c. This special case is of interest, because if the top's axis is held still at an angle to the vertical, and then released, this is the motion we will get.

Exercises

4.1 Prove the following properties of matrix algebra:
(a) Matrix multiplication is associative: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

(b) $(A \cdot B)^T = B^T \cdot A^T$, where A^T is the **transpose** of A, that is $(A^T)_{ij} := A_{ji}$.
(c) If A^{-1} and B^{-1} exist, $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.
(d) The complex conjugate of a matrix $(A^*)_{ij} = A^*_{ij}$ is the matrix with every element complex conjugated. The **hermitean conjugate** $A^\dagger$ is the transpose of that, $A^\dagger := (A^*)^T = (A^T)^*$, with $(A^\dagger)_{ij} := A^*_{ji}$. Show that $(A \cdot B)^* = A^* \cdot B^*$ and $(A \cdot B)^\dagger = B^\dagger \cdot A^\dagger$.

4.2 In section (4.1) we considered reexpressing a vector $\vec{V} = \sum_i V_i \hat{e}_i$ in terms of new orthogonal basis vectors. If the new vectors are $\vec{e}\,'_i = \sum_j A_{ij}\hat{e}_j$, we can also write $\hat{e}_i = \sum_j A_{ji}\vec{e}\,'_j$, because $A^T = A^{-1}$ for an orthogonal transformation.
Consider now using a new basis $\vec{e}\,'_i$ which are *not* orthonormal. Then we must choose which of the two above expressions to generalize. Let $\hat{e}_i = \sum_j A_{ji}\vec{e}\,'_j$, and find the expressions for (a) $\vec{e}\,'_j$ in terms of $\hat{e}_i$; (b) V'_i in terms of V_j; and (c) V_i in terms of V'_j. Then show (d) that if a linear tranformation **T** which maps vectors $\vec{V} \to \vec{W}$ is given in the $\hat{e}_i$ basis by a matrix B_{ij}, in that $W_i = \sum B_{ij} V_j$, then the same transformation **T** in the $\vec{e}\,'_i$ basis is given by $C = A \cdot B \cdot A^{-1}$. This transformation of matrices, $B \to C = A \cdot B \cdot A^{-1}$, for an arbitrary invertible matrix A, is called a **similarity transformation**.

4.3 Two matrices B and C are called **similar** if there exists an invertible matrix A such that $C = A \cdot B \cdot A^{-1}$, and this transformation of B into C is called a similarity transformation, as in the last problem. Show that, if B and C are similar, (a) $\operatorname{Tr} B = \operatorname{Tr} C$; (b) $\det B = \det C$; (c) B and C have the same eigenvalues; (d) If A is orthogonal and B is symmetric (or antisymmetric), then C is symmetric (or antisymmetric).

4.4 From the fact that $AA^{-1} = 1$ for any invertible matrix, show that if $A(t)$ is a differentiable matrix-valued function of time,

$$\dot{A}\, A^{-1} = -A \frac{dA^{-1}}{dt}.$$

4.5 Show that a counterclockwise rotation through an angle θ about an axis in the direction of a unit vector $\hat{n}$ passing through the origin is given by the matrix

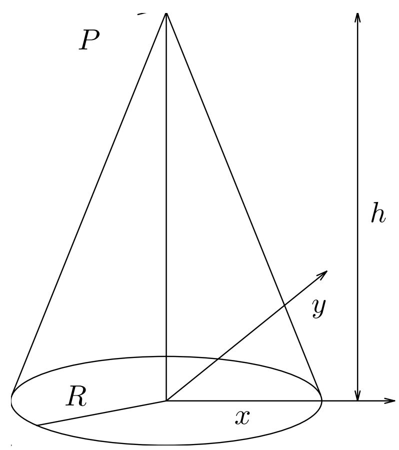
$$A_{ij} = \delta_{ij} \cos\theta + n_i n_j (1 - \cos\theta) - \epsilon_{ijk} n_k \sin\theta.$$

4.6 Consider a rigid body in the shape of a right circular cone of height h and a base which is a circle of radius R, made of matter with a uniform density ρ.

a) Find the position of the center of mass. Be sure to specify with respect to what.

b) Find the moment of inertia tensor in some suitable, well specified coordinate system about the center of mass.

c) Initially the cone is spinning about its symmetry axis, which is in the z direction, with angular velocity ω_0, and with no external forces or torques acting on it. At time $t = 0$ it is hit with a momentary laser pulse which imparts an impulse P in the x direction at the apex of the cone, as shown.

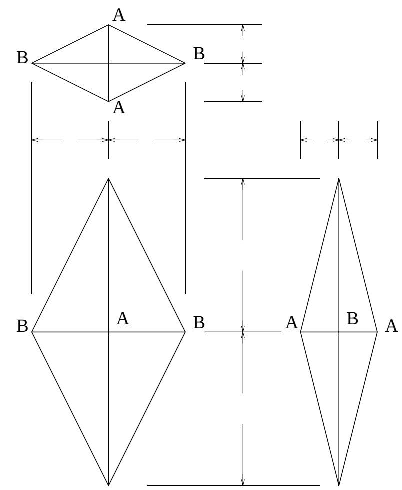


Describe the subsequent force-free motion, including, as a function of time, the angular velocity, angular momentum, and the position of the apex, in any inertial coordinate system you choose, provided you spell out the relation to the initial inertial coordinate system.

4.7 We defined the general rotation as $A = R_z(\psi) \cdot R_y(\theta) \cdot R_z(\phi)$. Work out the full expression for $A(\phi, \theta, \psi)$, and verify the last expression in (4.31). [For this and exercise 4.8, you might want to use a computer algebra program such as mathematica or maple, if one is available.]

4.8 Find the expression for $\vec{\omega}$ in terms of $\phi, \theta, \psi, \dot{\phi}, \dot{\theta}, \dot{\psi}$. [This can be done simply with computer algebra programs. If you want to do this by hand, you might find it easier to use the product form $A = R_3 R_2 R_1$, and the rather simpler expressions for $R\dot{R}^T$. You will still need to bring the result (for $R_1 \dot{R}_1^T$, for example) through the other rotations, which is somewhat messy.]

4.9 A diamond shaped object is shown in top, front, and side views. It is an octahedron, with 8 triangular flat faces.



It is made of solid aluminum of uniform density, with a total mass M. The dimensions, as shown, satisfy $h > b > a$.
(a) Find the moment of inertia tensor about the center of mass, clearly specifying the coordinate system chosen.
(b) About which lines can a stable spinning motion, with fixed $\vec{\omega}$, take place, assuming no external forces act on the body?

4.10 From the expression 4.40 for $u = \cos\theta$ for the motion of the symmetric top, we can derive a function for the time $t(u)$ as an indefinite integral

$$t(u) = \int^u f^{-1/2}(z)\, dz.$$

For values which are physically realizable, the function f has two (generically distinct) roots, $u_X \leq u_N$ in the interval $u \in [-1, 1]$, and one root $u_U \in [1, \infty)$, which does not correspond to a physical value of θ. The integrand is then generically an analytic function of z with square root branch points at u_N, u_X, u_U, and ∞, which we can represent on a cut Riemann sheet with cuts on the real axis, $[-\infty, u_X]$ and $[u_N, u_U]$, and $f(u) > 0$ for $u \in (u_X, u_N)$. Taking $t = 0$ at the time the top is at the bottom of a wobble, $\theta = \theta_{\rm max}, u = u_X$, we can find the time at which it first reaches another $u \in [u_X, u_N]$ by integrating along the real axis. But we could also use any other path in the upper half plane, as the integral of a complex function is independent of deformations of the path through regions where the function is analytic.
(a) Extend this definition to a function $t(u)$ defined for $\text{Im } u \geq 0$, with u not on a cut, and show that the image of this function is a rectangle in the complex t plane, and identify the pre-images of the sides. Call the width $T/2$ and the height $\tau/2$
(b) Extend this function to the lower half of the same Riemann sheet by allowing contour integrals passing through $[u_X, u_N]$, and show that this extends the image in t to the rectangle $(0, T/2) \times (-i\tau/2, i\tau/2)$.
(c) If the coutour passes through the cut $(-\infty, u_X]$ onto the second Riemann sheet, the integrand has the opposite sign from what it would have at the corresponding point of the first sheet. Show that if the path takes this path onto the second sheet

and reaches the point u, the value $t_1(u)$ thus obtained is $t_1(u) = -t_0(u)$, where $t_0(u)$ is the value obtained in (a) or (b) for the same u on the first Riemann sheet.
(d) Show that passing to the second Riemann sheet by going through the cut $[u_N, u_U]$ instead, produces a $t_2(u) = t_1 + T$.
(e) Show that evaluating the integral along two contours, Γ_1 and Γ_2, which differ only by Γ_1 circling the $[u_N, u_U]$ cut clockwise once more than Γ_2 does, gives $t_1 = t_2 + i\tau$.
(f) Show that any value of t can be reached by some path, by circling the $[u_N, u_U]$ as many times as necessary, and also by passing downwards through it and upwards through the $[-\infty, u_X]$ cut as often as necessary (perhaps reversed).
(g) Argue that thus means the function $u(t)$ is an analytic function from the complex t plane into the u complex plane, analytic except at the points $t = nT + i(m+\frac{1}{2})\tau$, where $u(t)$ has double poles. Note this function is doubly periodic, with $u(t) = u(t + nT + im\tau)$.
(g) Show that the function is then given by $u = \beta\,\wp(t - i\tau/2) + c$, where c is a constant, β is the constant from (4.40), and

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{m,n\in\mathbf{Z}\\(m,n)\neq 0}} \left(\frac{1}{(z - nT - mi\tau)^2} - \frac{1}{(nT + mi\tau)^2} \right)$$

is the Weierstrass' $\wp$-Function.
(h) Show that $\wp$ satisfies the differential equation

$$\wp'^2 = 4\wp^3 - g_2\wp - g_3,$$

where

$$g_2 = \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (mT + in\tau)^{-4}, \qquad g_3 = \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (mT + in\tau)^{-6}.$$

[Note that the Weierstrass function is defined more generally, using parameters $\omega_1 = T/2$, $\omega_2 = i\tau/2$, with the ω's permitted to be arbitrary complex numbers with differing phases.]

4.11 As a rotation about the origin maps the unit sphere into itself, one way to describe rotations is as a subset of maps $f : S^2 \to S^2$ of the (surface of the) unit sphere into itself. Those which correspond to rotations are clearly one-to-one, continuous, and preserve the angle between any two paths which intersect at a point. This is called a conformal map. In addition, rotations preserve the distances between points. In this problem we show how to describe such mappings, and therefore give a representation for the rotations in three dimensions.

(a) Let N be the north pole $(0,0,1)$ of the unit sphere $\Sigma = \{(x,y,z), x^2+y^2+z^2 = 1\}$. Define the map from the rest of the sphere $s : \Sigma - \{N\} \to \mathbb{R}^2$ given by a stereographic projection, which maps each point on the unit sphere, other than the north pole, into the point (u,v) in the equatorial plane $(x,y,0)$ by giving the intersection with this plane of the straight line which joins the point $(x,y,z) \in \Sigma$ to the north pole. Find (u,v) as a function of (x,y,z), and show that the lengths of infinitesimal paths in the vicinity of a point are scaled by a factor $1/(1-z)$ independent of direction, and therefore that the map s preserves the angles between intersecting curves (*i.e.* is **conformal**).
(b) Show that the map $f((u,v)) \to (u',v')$ which results from first applying s^{-1}, then a rotation, and then s, is a conformal map from $\mathbb{R}^2$ into $\mathbb{R}^2$, except for the pre-image of the point which gets mapped into the north pole by the rotation.
By a general theorem of complex variables, any such map is analytic, so $f : u+iv \to u'+iv'$ is an analytic function except at the point $\xi_0 = u_0 + iv_0$ which is mapped to infinity, and ξ_0 is a simple pole of f. Show that $f(\xi) = (a\xi + b)/(\xi - \xi_0)$, for some complex a and b. This is the set of complex Mobius transformations, which are usually rewritten as

$$f(\xi) = \frac{\alpha\xi + \beta}{\gamma\xi + \delta},$$

where $\alpha, \beta, \gamma, \delta$ are complex constants. An overall complex scale change does not affect f, so the scale of these four complex constants is generally fixed by imposing a normalizing condition $\alpha\delta - \beta\gamma = 1$.
(c) Show that composition of Mobius transformations $f'' = f' \circ f : \xi \underset{f}{\longrightarrow} \xi' \underset{f'}{\longrightarrow} \xi''$ is given by matrix multiplication,

$$\begin{pmatrix} \alpha'' & \beta'' \\ \gamma'' & \delta'' \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

(d) Not every mapping $s^{-1} \circ f \circ s$ is a rotation, for rotations need to preserve distances as well. We saw that an infinitesimal distance $d\ell$ on Σ is mapped by s to a distance $|d\xi| = d\ell/(1-z)$. Argue that the condition that $f : \xi \to \tilde{\xi}$ correspond to a rotation is that $d\tilde{\ell} \equiv (1-\tilde{z})|df/d\xi||d\xi| = d\ell$. Express this change of scale in terms of ξ and $\tilde{\xi}$ rather than z and $\tilde{z}$, and find the conditions on $\alpha, \beta, \gamma, \delta$ that insure this is true for all ξ. Together with the normalizing condition, show that this requires the matrix for f to be a unitary matrix with determinant 1, so that the set of rotations corresponds to the group $SU(2)$. The matrix elements are called Cayley-Klein parameters, and the real and imaginary parts of them are called the Euler parameters.