

# 2

## GYRODYNAMICS

### *Introduction to the dynamics of rigid bodies*

**Introduction.** Though Newton wrote on many topics—and may well have given thought to the odd behavior of tops—I am not aware that he committed any of that thought to writing. But by 1750 Euler was active in the field, and it has continued to bedevil the thought of mathematical physicists. “Extended rigid bodies” are classical abstractions—alien both to relativity and to quantum mechanics—which are revealed to our dynamical imaginations not so much by commonplace Nature as by, in Maxwell’s phrase, the “toys of Youth.” That such toys behave “strangely” is evident to the most casual observer, but the detailed theory of their behavior has become notorious for being elusive, surprising and difficult at every turn. Its formulation has required and inspired work of wonderful genius: it has taught us much of great worth, and clearly has much to teach us still.

Early in my own education as a physicist I discovered that I could not understand—or, when I could understand, remained unpersuaded by—the “elementary explanations” of the behavior of tops & gyros which are abundant in the literature. So I fell into the habit of avoiding the field, waiting for the day when I could give to it the time and attention it obviously required and deserved. I became aware that my experience was far from atypical: according to Goldstein it was in fact a similar experience that motivated Klein & Sommerfeld to write their 4-volume classic, *Theorie des Kreisels* (1897–1910).

In November 1976 I had occasion to consult my Classical Mechanics II students concerning what topic we should take up to finish out the term. It was with feelings of mixed excitement and dread that I heard their suggestion that we turn our attention to the theory of tops. The following material takes as its point of departure the class notes that were written on that occasion.

**1. Kinematic preliminaries.** Let  $\mathbf{x}_i$  be the inertial Cartesian coordinates of an arbitrary—and, for the moment, *not necessarily rigid*—assemblage of point masses  $m_i$ . Writing

$$\mathbf{x}_i = \mathbf{X} + \mathbf{r}_i \quad (1)$$

the *total kinetic energy* of the system becomes

$$\begin{aligned} T &= \frac{1}{2} \sum_i m_i \dot{\mathbf{x}}_i^2 \\ &= \frac{1}{2} M \dot{\mathbf{X}}^2 + \dot{\mathbf{X}} \cdot \sum_i m_i \dot{\mathbf{r}}_i + \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2 \end{aligned} \quad (2)$$

Noting that the variables  $\{\mathbf{X}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$  cannot, on numerological grounds, be independent, but must be subject to a single vectorial constraint, it is “to kill the cross term” in (2) that we require

$$\sum_i m_i \mathbf{r}_i = \mathbf{0} \quad (3)$$

It follows then from (1) that

$$\sum_i m_i \mathbf{x}_i = M \mathbf{X} \quad \text{with} \quad M \equiv \sum_i m_i \quad (4)$$

$\mathbf{X} \equiv$  coordinates of the **center of mass**

Time-differentiation of (4) gives

$$\mathbf{P} = \sum_i \mathbf{p}_i$$

where  $\mathbf{p}_i = m_i \dot{\mathbf{x}}_i$ .  $\mathbf{P} = M \dot{\mathbf{X}}$  is the *total linear momentum* of the system and is for isolated systems conserved, whatever may be the nature of the intra-system interactions. In this notation (2) has become

$$T = T_0 + \mathcal{T} \quad (6)$$

$$\mathcal{T} = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i^2 = \text{kinetic energy relative to the center of mass}$$

$$T_0 = \frac{1}{2} M \dot{\mathbf{X}}^2 = \frac{1}{2M} P^2 = \text{kinetic energy of the center of mass}$$

To impose **rigidity** upon the  $N$ -particle system amounts formally to imposing the stipulation that the numbers

$$a_{ij} \equiv |\mathbf{x}_i - \mathbf{x}_j| = |\mathbf{r}_i - \mathbf{r}_j|$$

shall be time-independent *constants*:

$$\dot{a}_{ij} = 0 \quad (7)$$

Equivalently, we might stipulate that the numbers  $\mathbf{r}_i \cdot \mathbf{r}_j$  be constant. It is, however, intuitively obvious that the conditions (7) are highly redundant:

number $N$ of particles	number $n$ of degrees of freedom
1	3
2	5
3	<b>6</b>
4	<b>6</b>
5	<b>6</b>
$\vdots$	$\vdots$

We on this basis expect a rigid blob of material to have six degrees of freedom. The point is sharpened by *Chasle's theorem*,<sup>1</sup> according to which the constituent parts  $m_i$  of a rigid assemblage have instantaneous positions that can be described

$$\mathbf{x}_i(t) = \mathbf{X}(t) + \mathbb{R}(t)\mathbf{r}_i^0 \quad (8)$$

where  $\mathbf{X}(t)$  locates the moving center of mass, where  $\mathbb{R}(t)$  is a time-dependent (proper) rotation matrix

$$\mathbb{R}(0) = \mathbb{I} \quad \text{and} \quad \mathbb{R}(t)\mathbb{R}^\top(t) = \mathbb{I} \quad : \quad \text{all } t \geq 0 \quad (9)$$

and where the constant vectors  $\mathbf{r}_i^0$  record the initial and enduring *design* of the rigid assemblage:  $\mathbf{r}_i^0 = \mathbf{x}_i(0) - \mathbf{X}(0)$ . Three degrees of freedom enter into the specification of  $\mathbf{X}$ , and (in 3-dimensional space) three more into the specification of  $\mathbb{R}$ .

Chasle's equation (8) amounts to the assertion that in a *rigid* assemblage

$$\mathbf{r}_i = \mathbb{R}\mathbf{r}_i^0 \quad (10)$$

The basic decomposition (1) assumes therefore the sharpened form

$$\mathbf{x}_i = \mathbf{X} + \mathbb{R}\mathbf{r}_i^0$$

which after  $t$ -differentiation becomes

$$\begin{aligned} \dot{\mathbf{x}}_i &= \dot{\mathbf{X}} + \dot{\mathbb{R}}\mathbf{r}_i^0 \\ &= \dot{\mathbf{X}} + \dot{\mathbb{R}}\mathbb{R}^{-1}\mathbf{r}_i \end{aligned}$$

But it is a familiar implication of (9) that  $\mathbb{A} \equiv \dot{\mathbb{R}}\mathbb{R}^{-1}$  is invariably/necessarily antisymmetric. Writing

$$\mathbb{A} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \quad (11)$$

we obtain finally

$$\dot{\mathbf{x}}_i = \dot{\mathbf{X}} + \mathbb{A}\mathbf{r}_i = \dot{\mathbf{X}} + (\boldsymbol{\omega} \times \mathbf{r}_i) \quad (12)$$

---

<sup>1</sup> See E. T. Whittaker, *Analytical Dynamics* (4<sup>th</sup> edition 1937), page 4.

where  $\mathbb{A}$  and (equivalently)  $\boldsymbol{\omega}$  will, in general, be  $t$ -dependent.

Returning with this information to (6) we find that for *rigid* bodies the kinetic energy relative to the center of mass—the “intrinsic” kinetic energy—can be described

$$\begin{aligned}
 \mathcal{T} &= \frac{1}{2} \sum_i m_i \mathbf{r}_i^T \mathbb{A}^T \mathbb{A} \mathbf{r}_i = \frac{1}{2} \sum_i m_i (\boldsymbol{\omega} \times \mathbf{r}_i) \cdot (\boldsymbol{\omega} \times \mathbf{r}_i) \\
 &= \frac{1}{2} \sum_i m_i (\mathbf{r}_i \times \boldsymbol{\omega}) \cdot (\mathbf{r}_i \times \boldsymbol{\omega}) \\
 &= \frac{1}{2} \sum_i m_i \boldsymbol{\omega}^T \mathbb{B}_i^T \mathbb{B}_i \boldsymbol{\omega} \quad \text{with} \quad \mathbb{B}_i \equiv \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} \\
 &= \frac{1}{2} \boldsymbol{\omega}^T \mathbb{I} \boldsymbol{\omega}
 \end{aligned} \tag{13}$$

where

$$\begin{aligned}
 \mathbb{I} &\equiv \sum_i m_i \mathbb{B}_i^T \mathbb{B}_i \\
 &= \sum_i m_i \begin{pmatrix} r^2 - r_1 r_1 & -r_1 r_2 & -r_1 r_3 \\ -r_2 r_1 & r^2 - r_2 r_2 & -r_2 r_3 \\ -r_3 r_1 & -r_3 r_2 & r^2 - r_3 r_3 \end{pmatrix}_i
 \end{aligned} \tag{14}$$

serves to define the “moment of inertia matrix”—also called the “moment of inertia tensor” or simply the “inertia matrix.” Here  $r_i^2 \equiv \mathbf{r}_i \cdot \mathbf{r}_i$ , and in the continuous limit we expect to have

$$= \iiint \begin{pmatrix} r^2 - r_1 r_1 & -r_1 r_2 & -r_1 r_3 \\ -r_2 r_1 & r^2 - r_2 r_2 & -r_2 r_3 \\ -r_3 r_1 & -r_3 r_2 & r^2 - r_3 r_3 \end{pmatrix} \rho(\mathbf{r}) dr_1 dr_2 dr_3$$

**REMARK:** In the little argument that led to (13)—whence to the invention of the moment of inertia matrix  $\mathbb{I}$ —we made essential use of a property of the “cross product:”  $\mathbf{r} \times \boldsymbol{\omega} = -\boldsymbol{\omega} \times \mathbf{r}$ . But the cross product is a peculiarly 3-dimensional construct, and so also, therefore, are the results reported above. In a one-dimensional world smoothly graded rotation is impossible, and the theory of rigid bodies trivializes. It is established in introductory physics courses that for rigid bodies that are constrained to move in what is, in effect, two dimensions one has

$$T = \frac{1}{2} M \dot{X}^2 + \frac{1}{2} I \omega^2 \quad \text{with} \quad I \equiv \iint r^2 \rho(\mathbf{r}) dr_1 dr_2$$

It is in the cases of spatial dimension  $N \geq 4$  that things become mathematically interesting, but in those cases we have no physical interest.

Look finally to the *total angular momentum* of a rigid body. For *any* multi-particle system one has

$$\mathbf{J} = \sum_i \mathbf{x}_i \times \mathbf{p}_i = \sum_i m_i \mathbf{x}_i \times \dot{\mathbf{x}}_i$$

which by (1) becomes

$$= \sum_i m_i (\mathbf{X} + \mathbf{r}_i) \times (\dot{\mathbf{X}} + \dot{\mathbf{r}}_i)$$

The expression on the right develops into a sum of four terms, of which two—the “cross terms”—vanish in consequence of the fundamental constraint (3). We are left with

$$= \mathbf{L}_{\text{orbital}} + \mathbf{L}_{\text{intrinsic}} \quad (15)$$

with

$$\begin{aligned} \mathbf{L}_{\text{orbital}} &= \mathbf{X} \times M \dot{\mathbf{X}} = \mathbf{X} \times \mathbf{P} = \begin{cases} \text{angular momentum of the} \\ \text{center of mass, relative to} \\ \text{the coordinate origin} \end{cases} \\ \mathbf{L}_{\text{intrinsic}} &= \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i = \begin{cases} \text{angular momentum relative to} \\ \text{the center of mass} \end{cases} \end{aligned}$$

The intrinsic angular momentum of *rigid* multi-particle systems is called *spin*: one has

$$\begin{aligned} \mathbf{L}_{\text{intrinsic}} &\longrightarrow \mathbf{S} = \sum_i m_i \mathbf{r}_i \times \mathbb{A} \mathbf{r}_i \\ &= \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i) \\ &= - \sum_i m_i \mathbf{r}_i \times (\mathbf{r}_i \times \boldsymbol{\omega}) \\ &= \sum_i m_i \mathbb{B}_i^T \mathbb{B}_i \boldsymbol{\omega} \\ &= \mathbb{I} \boldsymbol{\omega} \end{aligned} \quad (16)$$

and agrees, in place of (15), to write

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (17)$$

Several concluding remarks are now in order:

- Equations (6) and (15) share a structure

$$\text{total} = \text{orbital} + \text{intrinsic}$$

which they owe to the disappearance of certain cross terms. That disappearance is a consequence of the way the center of mass was defined: a consequence, that is to say, of (3):  $\sum m_i \mathbf{r}_i = \mathbf{0}$

- One does not write  $\mathbf{P}_{\text{total}} = \mathbf{P}_{\text{orbital}} + \mathbf{P}_{\text{intrinsic}}$  because, by (3),  $\mathbf{P}_{\text{intrinsic}}$  vanishes identically.
- Familiar analogs of the conditions that yield the conservation law  $\dot{\mathbf{P}} = \mathbf{0}$  also yield  $\dot{\mathbf{J}} = \mathbf{0}$ , which *may but need not* arise from conservation separately of  $\mathbf{L}$  and  $\mathbf{S}$ .
- We write  $\mathbf{p} = m\mathbf{v}$ , and from the fact that  $m$  is a scalar conclude that the momentum  $\mathbf{p}$  of a particle is always parallel to (in fact: a fixed multiple of) its velocity  $\mathbf{v}$ . But from  $\mathbf{S} = \mathbb{I}\boldsymbol{\omega}$  and the fact that  $\mathbb{I}$  is a  $3 \times 3$  *matrix* we see that the spin angular momentum  $\mathbf{S}$  and angular velocity  $\boldsymbol{\omega}$  need not—and typically will not—be parallel.
- Returning with  $\boldsymbol{\omega} = \mathbb{I}^{-1}\mathbf{S}$  to (13), we find

$$\mathcal{T} = \frac{1}{2}\mathbf{S}^T \mathbb{I}^{-1} \mathbf{S} \quad \text{provided} \quad \det \mathbb{I} \neq 0 \quad (18.1)$$

which is of interest as the formal analog of  $T = \frac{1}{2}\mathbf{p}^T m^{-1} \mathbf{p}$ . Equivalently

$$= \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{S} \quad (18.2)$$

**2. Nature & properties of the moment of inertia matrix.** Let the distribution function  $\rho(x)$  describe how some material of interest (mass, let us say) is distributed along the real line. One writes

$$\begin{aligned} M &= m^{(0)} = \int x^0 \rho(x) dx & : & \quad 0^{\text{th}} \text{ moment} \\ MX &= m^{(1)} = \int x^1 \rho(x) dx & : & \quad 1^{\text{st}} \text{ moment} \\ m^{(2)} &= \int x^2 \rho(x) dx & : & \quad 2^{\text{nd}} \text{ moment} \\ & \vdots & & \end{aligned}$$

to define the “moments” of the distribution.<sup>2</sup> From the set of *all* moments one can construct the “moment generating function”

$$\varphi(k) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} (ik)^n m^{(n)} = \int e^{ikx} \rho(x) dx$$

and by Fourier transformation recover the distribution itself:

$$\rho(x) = \frac{1}{2\pi} \int e^{-ikx} \varphi(k) dk$$

If we translate the coordinate origin to the center of mass—which is to say: if we introduce new coordinates  $r = x - X$  and proceed as before we obtain the

---

<sup>2</sup> In conventional terminology and notation,  $M$  is the “total mass” and  $X$  defines the position of the “center of mass.”

so-called “centered moments”<sup>3</sup>

$$\mu^{(n)} = \int r^n \rho(r) dr$$

Evidently

$$\mu^{(0)} = M$$

$$\mu^{(1)} = m^{(1)} - X m^{(0)} = 0 : \begin{cases} \text{the center of mass relative to the} \\ \text{center of mass resides at the origin} \end{cases}$$

All of which carries over straightforwardly to higher-dimensional situations. In three dimensions we have

$$\begin{aligned} M &= \iiint \rho(\mathbf{x}) d^3x && : \text{solitary } 0^{\text{th}} \text{ moment} \\ M\mathbf{X} &= \iiint \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rho(\mathbf{x}) d^3x && : \text{vector of } 1^{\text{st}} \text{ moments} \\ \mathbb{M} &= \iiint \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix} \rho(\mathbf{r}) d^3r && : \text{matrix of centered } 2^{\text{nd}} \text{ moments} \end{aligned}$$

and are placed now in position to recognize that the moment of inertia matrix is an object assembled from centered second moment data:<sup>4</sup>

$$\mathbb{I} = (\text{trace}\mathbb{M}) \cdot \mathbb{U} - \mathbb{M} \quad (19)$$

Remarkably, the low-order moment data built into the designs of  $M$ ,  $\mathbf{X}$  and  $\mathbb{I}$  is central to the dynamical theory of rigid bodies, but the moments of higher order are (in most contexts) utterly irrelevant: distinct rigid bodies can be expected to move identically if they have identical  $0^{\text{th}}$ ,  $1^{\text{st}}$  and  $2^{\text{nd}}$  order moments.

The moment of inertia matrix  $\mathbb{I}$  is manifestly *real* and *symmetric*:  $\mathbb{I}^T = \mathbb{I}$ . We are assured, therefore, that the eigenvalues of  $\mathbb{I}$  (call them  $\{I_1, I_2, I_3\}$  or  $\{A, B, C\}$ ) are *real*, and the associated eigenvectors (call them  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  or  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ ) are—or can always be taken to be—*orthogonal*:  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ . We are assured, moreover, that  $\mathbb{I}$  can in every case be *diagonalized by a suitably chosen rotation matrix*:

$$\mathbb{R}^T \mathbb{I} \mathbb{R} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad \text{with} \quad \mathbb{R}^T \mathbb{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

<sup>3</sup> I have here allowed myself to write  $\rho(r)$  where I should more properly have written something like  $\tilde{\rho}(r) \equiv \rho(X + r)$ . Similarly, I will later write  $\rho(\mathbf{r})$  when actually it was  $\rho(\mathbf{x})$  that was given and  $\rho(\mathbf{X} + \mathbf{r})$  that is intended.

<sup>4</sup> Since the symbol  $\mathbb{I}$  is busy, I have here had to use  $\mathbb{U}$  to represent the  $3 \times 3$  identity matrix.

The orthonormal triple  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  serves to define the “principal axes” of the rigid body. With respect to the “principal coordinate frame” that has its origin at the center of mass and coordinate axes parallel to the principal axes one has

$$\begin{aligned} \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= I_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} &= I_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} &= I_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

If  $\{r_1, r_2, r_3\}$  refer to the principal frame, then the continuous version of (14) supplies

$$\begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} = \iiint \begin{pmatrix} r_2^2 + r_3^2 & 0 & 0 \\ 0 & r_1^2 + r_3^2 & 0 \\ 0 & 0 & r_1^2 + r_2^2 \end{pmatrix} \rho(\mathbf{r}) d^3r \quad (20)$$

from which it follows trivially that

$$\text{All the eigenvalues of } \mathbb{I} \text{ are positive.} \quad (21.1)$$

More interestingly,

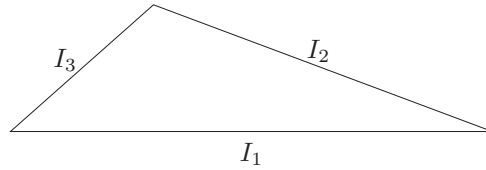
$$\begin{aligned} I_1 + I_2 &= \iiint (r_1^2 + r_2^2 + 2r_3^2) \rho(\mathbf{r}) d^3r \geq \iiint (r_1^2 + r_2^2) \rho(\mathbf{r}) d^3r = I_3 \\ &\vdots \\ &\text{etc.} \end{aligned}$$

—the implication being that

$$\text{No eigenvalue can exceed the sum of the other two,} \quad (21.2)$$

which is to say:

$$\text{The eigenvalues of } \mathbb{I} \text{ satisfy the triangle inequality:} \quad (21.3)$$





This occurrence of the triangle inequality is, in view of (20), not at all surprising, for if  $\alpha$ ,  $\beta$  and  $\gamma$  are *any* positive numbers then

$$i = \alpha + \beta$$

$$j = \beta + \gamma$$

$$k = \gamma + \alpha$$

*invariably and automatically* satisfy the  $\triangle$  inequality.<sup>5</sup>

We have been brought to the conclusion that to every rigid blob can be associated

- a naturally preferred point (the center of mass);
- a naturally preferred “principal axis frame,” with origin at that point (defined by the eigenvectors of the moment of inertia matrix  $\mathbb{I}$ );
- non-negative numbers  $\{A, B, C\}$  associated with the respective legs of the principal axis frame. Those numbers (eigenvalues of  $\mathbb{I}$ ) can in all cases be identified with the sides of a triangle, or alternatively: with the semi-axes of an ellipsoid

$$\frac{r_1^2}{A^2} + \frac{r_2^2}{B^2} + \frac{r_3^2}{C^2} = 1$$

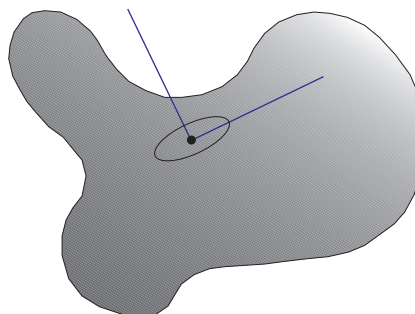


FIGURE 1: *Two-dimensional representation of a rigid body with preferred point and eigenvalue-weighted principal axes. Note that those attributes attach instantaneously even to non-rigid blobs, but it is only in the presence of rigidity that they acquire importance.*

**3. Moment of inertia with respect to an axis.** In introductory physics one learns to write

$$I = \iiint r^2 dm$$

to describe the moment of inertia of a rigid body with *respect to a prescribed axis*: here  $r$  is understood to denote the normal distance from the mass element

---

<sup>5</sup> I am indebted to Tom Wieting for this observation.

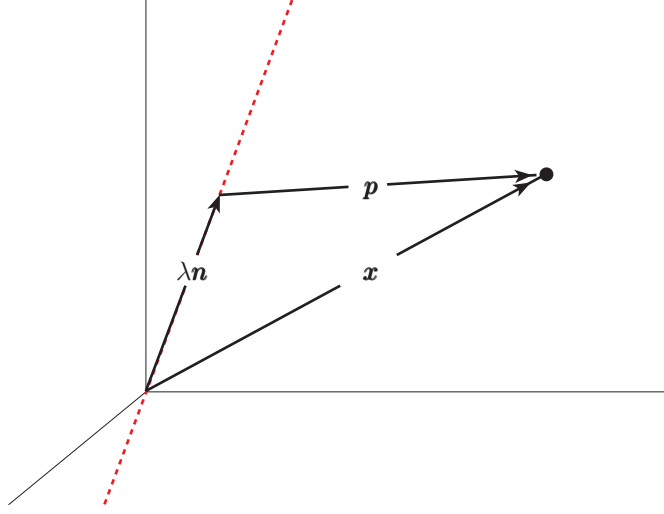


FIGURE 2: In the text we compute the shortest distance from a point to a line. The figure explains the notation used in that argument.

$dm$  to the axial line. The question now before us: how does that *scalar* moment of inertia relate to the matrix-valued construction  $\mathbb{I}$ ?

We confront first a simple geometrical problem: how to describe the length of the normal dropped from a point to a line? To describe a line through the origin we write  $\lambda \mathbf{n}$  ( $\lambda$  variable,  $\mathbf{n}$  a fixed unit vector). The condition that the vector  $\mathbf{p}(\lambda) \equiv \mathbf{x} - \lambda \mathbf{n}$  be normal to the line ( $\mathbf{n} \cdot \mathbf{p}(\lambda) = 0$ ) enforces  $\lambda = \mathbf{n} \cdot \mathbf{x}$ . The *length* of the normal dropped from  $\mathbf{x}$  to the line can therefore be described

$$\begin{aligned} r^2(\mathbf{x}) &= [\mathbf{x} - (\mathbf{n} \cdot \mathbf{x}) \mathbf{n}] \cdot [\mathbf{x} - (\mathbf{n} \cdot \mathbf{x}) \mathbf{n}] \\ &= \mathbf{x} \cdot \mathbf{x} - (\mathbf{x} \cdot \mathbf{n})(\mathbf{n} \cdot \mathbf{x}) \\ &= \mathbf{x} \cdot (\mathbb{U} - \mathbb{N}) \mathbf{x} \end{aligned} \tag{22.1}$$

$$= \mathbf{n} \cdot [(\text{trace } \mathbb{X}) \cdot \mathbb{U} - \mathbb{X}] \mathbf{n} \tag{22.2}$$

where, as before,  $\mathbb{U}$  is the identity matrix<sup>4</sup>, where

$$\mathbb{N} \equiv \begin{pmatrix} n_1 n_1 & n_1 n_2 & n_1 n_3 \\ n_2 n_1 & n_2 n_2 & n_2 n_3 \\ n_3 n_1 & n_3 n_2 & n_3 n_3 \end{pmatrix} \quad \text{projects onto } \mathbf{n}$$

and where  $\mathbb{X}$  is constructed similarly:

$$\mathbb{X} \equiv \begin{pmatrix} x_1 x_1 & x_1 x_2 & x_1 x_3 \\ x_2 x_1 & x_2 x_2 & x_2 x_3 \\ x_3 x_1 & x_3 x_2 & x_3 x_3 \end{pmatrix}$$

It now follows that

$$I_0 = \iiint r^2(\mathbf{x})\rho(\mathbf{x})d^3x = \begin{cases} \text{moment of inertial about an axis through} \\ \text{the center of mass, in the direction defined} \\ \text{by the unit vector } \mathbf{n} \end{cases}$$

can, by (19), be written

$$I_0 = \mathbf{n} \cdot \mathbb{I} \mathbf{n} \quad (23)$$

which could hardly be simpler or prettier.

What can one say about the moment of inertia about an axis that does *not* pass through the center of mass? To describe points on such an axis we write  $\mathbf{a} + \lambda \mathbf{n}$  and assume without loss of generality that  $\mathbf{a} \perp \mathbf{n}$ . Proceeding as before, we introduce  $\mathbf{p}(\lambda) \equiv \mathbf{x} - (\mathbf{a} + \lambda \mathbf{n})$ , insist that  $\mathbf{p}(\lambda) \perp (\mathbf{a} + \lambda \mathbf{n})$  and are led to the conclusion that the line dropped normally from  $\mathbf{x}$  to the displaced axis has squared length

$$r^2(\mathbf{x}) = (\mathbf{x} - \mathbf{a}) \cdot (\mathbb{U} - \mathbb{N})(\mathbf{x} - \mathbf{a})$$

which, we note, does give back (22.1) at  $\mathbf{a} = \mathbf{0}$ . More particularly, we have

$$= \mathbf{x} \cdot (\mathbb{U} - \mathbb{N})\mathbf{x} - 2\mathbf{x} \cdot (\mathbb{U} - \mathbb{N})\mathbf{a} + \mathbf{a} \cdot (\mathbb{U} - \mathbb{N})\mathbf{a}$$

which we introduce into

$$I = \iiint r^2(\mathbf{x})\rho(\mathbf{x})d^3x$$

and by quick easy steps recover the “parallel axis theorem:”

$$I = I_0 + Ma^2 \quad (24)$$

Equation (23) indicates how moments with respect to axes (numbers of the type  $I_0$ ) can be extracted from the data written into  $\mathbb{I}$ . One can also proceed in the reverse direction. Suppose, for example, we were to set

$$\mathbf{n} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

We would then have

$$\text{measured value of } I_0 = I_{11}$$

More generally, we have, for each selected  $\mathbf{n}$ ,

$$\begin{aligned} \text{measured value of } I_0(\mathbf{n}) = & n_1 n_1 I_{11} + 2n_1 n_2 I_{12} + 2n_1 n_3 I_{13} \\ & + n_2 n_2 I_{22} + 2n_2 n_3 I_{23} \\ & + n_3 n_3 I_{33} \end{aligned}$$

Given six well-chosen instances of that equation, we would have enough information to compute all the  $I_{ij}$  by straightforward linear algebra.<sup>6</sup>

---

<sup>6</sup> Is there an computationally optimal way to select  $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5, \mathbf{n}_6$ ?

**4. How the moment of inertia matrix responds to rotations.** Shown below is a representation of a rigid body that has rotated about its fixed center of mass. The component parts of the body have retained their *relative* positions, but their positions relative to fixed exterior reference frames have, in general,

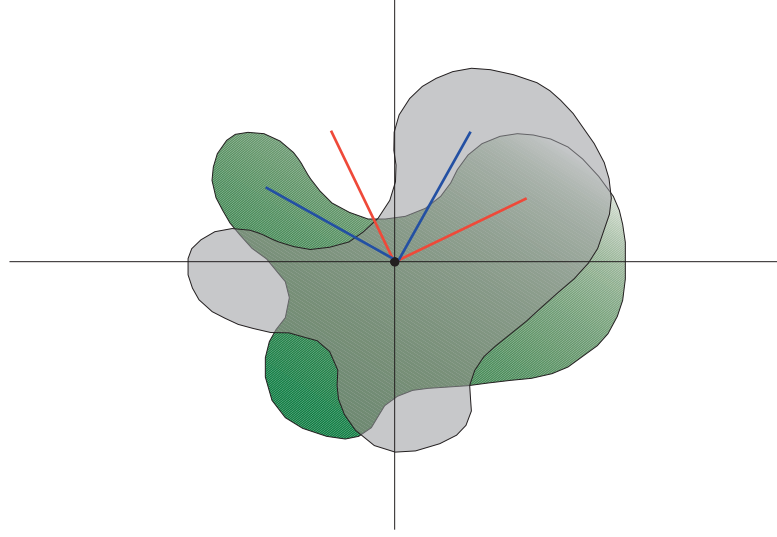


FIGURE 3: *Schematic representation of a rigid body has been rotated about its immobile center of mass.*

changed. The issue before us: How is  $\mathbb{I}_{\text{final}}$  related to  $\mathbb{I}_{\text{initial}}$ ? Working from (14), *i.e.*, from (compare (19))

$$\mathbb{I} = \sum_i m_i [r^2 \cdot \mathbb{U} - \mathbb{W}]_i \quad \text{with} \quad \mathbb{W} \equiv \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}$$

The effect of body rotation—as was remarked already at (10)—can be described

$$\mathbf{r}_i^0 \mapsto \mathbf{r}_i = \mathbb{R} \mathbf{r}_i^0 \quad : \quad \text{all } i$$

which induces

$$W_{\alpha\beta}^0 \equiv r_\alpha^0 r_\beta^0 \mapsto W_{\alpha\beta} \equiv r_\alpha r_\beta = R_{\alpha\mu} r_\mu^0 R_{\beta\nu} r_\nu^0 = R_{\alpha\mu} W_{\mu\nu}^0 R_{\beta\nu}$$

giving  $\mathbb{W} = \mathbb{R} \mathbb{W}^0 \mathbb{R}^\top = \mathbb{R} \mathbb{W}^0 \mathbb{R}^{-1}$ . Immediately  $r^2 = \text{trace} \mathbb{W} = \text{trace} \mathbb{W}^0 = (r^0)^2$  so we have

$$\mathbb{I}^0 \mapsto \mathbb{I} = \mathbb{R} \mathbb{I}^0 \mathbb{R}^{-1} \quad (25)$$

The importance of this result emerges when one supposes  $\mathbb{R}$  to be a function of time, writing

$$\mathbb{I}(t) = \mathbb{R}(t) \mathbb{I}(0) \mathbb{R}^{-1}(t)$$

Differentiation of  $\mathbb{R} \mathbb{R}^\top = \mathbb{U}$  leads quickly to the important conclusion (remarked already on page 3) that

$$\dot{\mathbb{R}} = \mathbb{A} \mathbb{R} \quad \text{with} \quad \mathbb{A}^\top = -\mathbb{A}$$

so we have

$$\dot{\mathbb{I}} = \mathbb{A} \mathbb{I} - \mathbb{I} \mathbb{A} \quad (26)$$

Look in this light back to the equation

$$\mathbf{S} = \mathbb{I} \boldsymbol{\omega}$$

that at (16) was seen to relate spin to angular velocity. Time-differentiation gives

$$\dot{\mathbf{S}} = \mathbb{I} \dot{\boldsymbol{\omega}} + (\mathbb{A} \mathbb{I} - \mathbb{I} \mathbb{A}) \boldsymbol{\omega}$$

But  $\mathbb{A}$  was seen at (11) to be just another name for the operation  $\boldsymbol{\omega} \times$ , so we have

$$\begin{aligned} \dot{\mathbf{S}} &= \mathbb{I} \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbb{I} \boldsymbol{\omega} - \mathbb{I} \boldsymbol{\omega} \times \boldsymbol{\omega} \\ &= \mathbb{I} \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbf{S} \end{aligned} \quad (27)$$

**5. Equations of rigid body motion: Newtonian approach.** Look now to the Newtonian *dynamics* of the  $N$ -body system contemplated at the beginning of §1, again suspending for the moment any assumption concerning the rigidity of the system. Immediately

$$\begin{aligned} \mathbf{F}_i &\equiv \mathbf{F}_i^{\text{impressed}} + \sum_j \mathbf{F}_{ij}^{\text{interactive}} = m_i \ddot{\mathbf{x}}_i \\ &= m_i \ddot{\mathbf{X}} + m_i \ddot{\mathbf{r}}_i \end{aligned} \quad (28)$$

where evidently  $\mathbf{F}_{ii}^{\text{interactive}} = \mathbf{0}$  while by the 3<sup>rd</sup> law  $\mathbf{F}_{ij}^{\text{interactive}} = -\mathbf{F}_{ji}^{\text{interactive}}$ . Summation on  $i$  supplies

$$\begin{aligned} \mathbf{F}^{\text{total impressed}} &\equiv \sum_i \mathbf{F}_i^{\text{impressed}} = M \ddot{\mathbf{X}} \\ &= \dot{\mathbf{P}} \end{aligned} \quad (29)$$

where the 3<sup>rd</sup> law has killed the interactive force terms, and the constraint (3) has served to kill the relative acceleration terms. Application of

$$\sum_i \mathbf{x}_i \times = \sum_i (\mathbf{X} + \mathbf{r}_i) \times \quad (30)$$

to the left side of (28) gives

$$\mathbf{X} \times \mathbf{F}^{\text{total impressed}} + \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{impressed}} + \sum_{ij} \mathbf{r}_i \times \mathbf{F}_{ij}^{\text{interactive}}$$

where the final term actually vanishes if—as we will assume—the interactive forces are central:  $\mathbf{F}_{ij}^{\text{interactive}} \parallel (\mathbf{r}_i - \mathbf{r}_j)$ . On the other hand, application of (30) to the right side of (28) was already seen at (15) to give  $\mathbf{L}_{\text{orbital}} + \mathbf{L}_{\text{intrinsic}}$ . So we have, for any centrally interactive  $N$ -body system,

$$\mathbf{N}_{\text{orbital}} + \mathbf{N}_{\text{intrinsic}} = \dot{\mathbf{L}}_{\text{orbital}} + \dot{\mathbf{L}}_{\text{intrinsic}}$$

where

$$\begin{aligned} \mathbf{N}_{\text{orbital}} &\equiv \mathbf{X} \times \mathbf{F}^{\text{total impressed}} \\ \mathbf{N}_{\text{intrinsic}} &\equiv \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{impressed}} \end{aligned}$$

But it follows already from (29) that  $\mathbf{N}_{\text{orbital}} = \dot{\mathbf{L}}_{\text{orbital}}$  so, collecting the results now in hand, we have

$$\left. \begin{aligned} \mathbf{F}^{\text{total impressed}} &= \dot{\mathbf{P}} \\ \mathbf{N}_{\text{orbital}} &= \dot{\mathbf{L}}_{\text{orbital}} \end{aligned} \right\} : \text{refer to motion of the center of mass}$$

$$\mathbf{N}_{\text{intrinsic}} = \dot{\mathbf{L}}_{\text{intrinsic}} : \text{refers to motion relative to the center of mass}$$

Now impose the assumption of rigidity upon our  $N$ -body system, and emphasize that we have done so by notational adjustment:  $\mathbf{L}_{\text{intrinsic}} \mapsto \mathbf{S} = \mathbb{I}\boldsymbol{\omega}$ . Drawing upon (27) we then have

$$\begin{aligned} \mathbf{N}_{\text{intrinsic}} &= \dot{\mathbf{S}} \\ &= \mathbb{I}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega} \end{aligned} \tag{31}$$

Here  $\mathbf{N}_{\text{intrinsic}}$ ,  $\boldsymbol{\omega}$  and the integrals that assign instantaneous meaning to  $\mathbb{I}$  all refer to the **space frame**, a (generally non-inertial) translated copy of the inertial **lab frame**. Several circumstances limit the utility of this result:

- The value of  $\mathbf{N}_{\text{intrinsic}}$  will, in the general case, depend upon both the location and the orientation of the rigid body, and may even depend upon its instantaneous state of motion (as would happen if the body had been assembled from charged particles and were placed in a magnetic field). This circumstance introduces an element of circularity into the discussion: one must know the net effect of all past motion to understand what (31) has to say about present motion.
- Even in the simplest case  $\mathbf{N}_{\text{intrinsic}} = \mathbf{0}$  an awkward time-dependence lurks in the design of  $\mathbb{I}$ , which changes moment to moment as the body rotates.

The latter difficulty can be circumvented by a strategy introduced by Euler (1758), the essential idea being to kill the time-dependence of  $\mathbb{I}$  by passing to

a frame that is fixed in the body. If  $\mathbb{I}^0 \equiv \mathbb{I}(0)$  refers to the initial orientation of the body, and  $\mathbb{I} \equiv \mathbb{I}(t)$  to its evolved orientation, then by (25) we have

$$\mathbb{I} = \mathbb{R} \mathbb{I}^0 \mathbb{R}^{-1} \quad (32)$$

where  $\mathbb{R} \equiv \mathbb{R}(t)$  refers to the rotational that has been accomplished during the interval in question. In this notation (31)—after multiplication on the left by  $\mathbb{R}^{-1}$ —becomes

$$\begin{aligned} \mathbf{N}^0 &= \mathbb{I}^0 \mathbb{R}^{-1} \dot{\boldsymbol{\omega}} + \mathbb{R}^{-1} \boldsymbol{\omega} \times \mathbb{R} \mathbb{I}^0 \mathbb{R}^{-1} \boldsymbol{\omega} \quad \text{with} \quad \mathbf{N}^0 \equiv \mathbb{R}^{-1} \mathbf{N} \\ &= \mathbb{I}^0 \cdot \mathbb{R}^{-1} \dot{\boldsymbol{\omega}} + \mathbb{R}^{-1} \boldsymbol{\omega} \times \mathbb{R} \cdot \mathbb{I}^0 \boldsymbol{\omega}^0 \quad \text{with} \quad \boldsymbol{\omega}^0 \equiv \mathbb{R}^{-1} \boldsymbol{\omega} \end{aligned}$$

Our further progress hinges on the following

**LEMMA:** Notice that if  $\mathbb{A}$  is a  $3 \times 3$  antisymmetric matrix

$$\mathbb{A} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

then  $\mathbb{A}\mathbf{x} = \mathbf{a} \times \mathbf{x}$  establishes the sense in which “ $\mathbb{A} = \mathbf{a} \times$ ”. Let  $\mathbb{R}$  be a proper  $3 \times 3$  rotation matrix:  $\mathbb{R}^{-1} = \mathbb{R}^\top$  and  $\det \mathbb{R} = 1$ . Then

$$\mathbb{R} \mathbb{A} \mathbb{R}^{-1} = (\mathbb{R} \mathbf{a}) \times \quad (33)$$

We can on this basis write  $\mathbb{R}^{-1} \boldsymbol{\omega} \times \mathbb{R} = \boldsymbol{\omega}^0 \times$ . Differentiation of  $\boldsymbol{\omega}^0 = \mathbb{R}^{-1} \boldsymbol{\omega}$  leads moreover to the conclusion that  $\dot{\boldsymbol{\omega}}^0 = \mathbb{R}^{-1} \dot{\boldsymbol{\omega}} + \dot{\mathbb{R}}^\top \boldsymbol{\omega}$ . But transposition of<sup>7</sup>  $\dot{\mathbb{R}} = \mathbb{A} \mathbb{R}$  gives  $\dot{\mathbb{R}}^\top = -\mathbb{R}^\top \mathbb{A} = -\mathbb{R}^\top (\boldsymbol{\omega} \times)$  so we have  $\dot{\boldsymbol{\omega}}^0 = \mathbb{R}^{-1} \dot{\boldsymbol{\omega}} - \mathbb{R}^\top \boldsymbol{\omega} \times \boldsymbol{\omega} = \mathbb{R}^{-1} \dot{\boldsymbol{\omega}}$ , so we have

$$\mathbf{N}^0 = \mathbb{I}^0 \dot{\boldsymbol{\omega}}^0 + \boldsymbol{\omega}^0 \times \mathbb{I}^0 \boldsymbol{\omega}^0 \quad (34)$$

The preceding equation describes the motion of  $\boldsymbol{\omega}$  as viewed by a non-inertial observer who is at rest with respect to the gyrating rigid body—difficult to imagine when you contemplate a spinning/precessing top, yet an entirely commonplace notion: you sit comfortably on the non-inertial earth, look up into the night sky and see  $\boldsymbol{\omega}^0$  as the vector about which the stars appear to revolve. And if you wait long enough (thousands of years) will notice that  $\boldsymbol{\omega}^0$  traces a curve in the patterned fixed stars. What is remarkable is that (34) is *structurally identical* to (31).<sup>8</sup>

<sup>7</sup> See again page 3.

<sup>8</sup> On pages 27–35 of *GYRODYNAMICS* (1976/77) and again on page 92 below I discuss in detail how Coriolis and centrifugal forces—universal symptoms of non-inertiality—conspire to achieve this remarkable result.

Computation<sup>9</sup> shows the vector  $\boldsymbol{\omega}^0 \times \mathbb{I}^0 \boldsymbol{\omega}^0$  to be, in the general case, a fairly intricate object. It is, however, entirely natural to identify the body frame with the principal axis frame

$$\begin{pmatrix} I_{11} & I_{12} & I_{13} \\ \bullet & I_{22} & I_{23} \\ \bullet & \bullet & I_{33} \end{pmatrix} \xrightarrow{\text{passage to principal axis frame}} \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}$$

Major simplifications are then achieved: one is led from (34) by quick calculation to the so-called **Euler equations**

$$\left. \begin{aligned} N_1^0 &= I_1^0 \dot{\omega}_1^0 + (I_3^0 - I_2^0) \omega_3^0 \omega_2^0 \\ N_2^0 &= I_2^0 \dot{\omega}_2^0 + (I_1^0 - I_3^0) \omega_1^0 \omega_3^0 \\ N_3^0 &= I_3^0 \dot{\omega}_3^0 + (I_2^0 - I_1^0) \omega_2^0 \omega_1^0 \end{aligned} \right\} \quad (35)$$

We have here a coupled system of three non-linear first-order differential conditions on three unknown functions  $\boldsymbol{\omega}^0(t)$ .

Suppose for the moment that equations (35) have been *solved*. How does knowledge of  $\boldsymbol{\omega}^0(t)$  determine the rotation matrix  $\mathbb{R}(t)$  by means of which we—if not riding on the body but watching it from a position at rest with respect to the space frame—propose to understand the perceived motion of the rigid body? We have

$$\begin{aligned} \dot{\mathbb{R}} \mathbb{R}^{-1} &= \mathbb{A} = \boldsymbol{\omega} \times \\ &= (\mathbb{R} \boldsymbol{\omega}^0) \times \\ &= \mathbb{R}(\boldsymbol{\omega}^0 \times) \mathbb{R}^{-1} \quad \text{by LEMMA (33)} \end{aligned}$$

giving

$$\dot{\mathbb{R}} = \mathbb{R} \mathbb{A}^0 \quad \text{with (we may assume) } \mathbb{R}(0) = \mathbb{U} \quad (36.1)$$

Equivalently

$$\mathbb{R}(t) = \mathbb{U} + \int_0^t \mathbb{R}(\tau) \mathbb{A}^0(\tau) d\tau \quad (36.2)$$

which can, in principle, be solved by iteration. But except in the simplest of cases we can expect the solution of (35) to be very difficult, and the solution of (36.1) to be also very difficult.

**6. Equations of rigid body motion: Lagrangian approach.** The idea here is to look upon the elements  $R_{ij}$  of  $\mathbb{R}$  as “generalized coordinates,” to construct a Lagrangian of the form  $\mathcal{L}(\dot{\mathbb{R}}, \mathbb{R})$ , and then to write

$$\left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{\mathbb{R}}} - \frac{\partial}{\partial \mathbb{R}} \right\} \mathcal{L}(\dot{\mathbb{R}}, \mathbb{R}) = 0$$

We confront, however, the fundamental difficulty that the nine variables  $R_{ij}$

---

<sup>9</sup> Which might be entrusted to *Mathematica*. Or see pages 24–25 in the notes just cited.





first equation, and noting that  $\mathbb{A} - \mathbb{A}^\top = \mathbb{O}$ , we obtain

$$(\partial\mathcal{U}/\partial\mathbb{R})^\top\mathbb{R} - \mathbb{R}^\top(\partial\mathcal{U}/\partial\mathbb{R}) = \mathbb{R}^\top\ddot{\mathbb{R}}\mathbb{M}^0 - \mathbb{M}^0\ddot{\mathbb{R}}^\top\mathbb{R} \quad (37)$$

from which all reference to the Lagrange multipliers has been eliminated. This second-order matrix equation is (by antisymmetry) equivalent to a system of three scalar equations, and will now be shown to comprise no more nor less than a *matrix formulation of the Euler equations* (35).

Let us first of all agree to write

$$(\partial\mathcal{U}/\partial\mathbb{R})^\top\mathbb{R} - \mathbb{R}^\top(\partial\mathcal{U}/\partial\mathbb{R}) \equiv \mathbb{N}^0$$

and to note that  $\mathbb{N}^0$  is antisymmetric. Turning our attention to the expression on the right side of (37), we notice that differentiation of  $\dot{\mathbb{R}} = \mathbb{A}\mathbb{R}$  supplies  $\ddot{\mathbb{R}} = (\dot{\mathbb{A}} + \mathbb{A}^2)\mathbb{R}$  whence

$$\mathbb{R}^\top\ddot{\mathbb{R}} = \mathbb{R}^{-1}(\dot{\mathbb{A}} + \mathbb{A}^2)\mathbb{R}$$

But it was the upshot of our LEMMA (33) that if  $\mathbb{A} = \boldsymbol{\omega} \times$  then

$$\text{If } \mathbb{A} = \boldsymbol{\omega} \times \text{ then } \mathbb{R}^{-1}\mathbb{A}\mathbb{R} = \mathbb{A}^0 \text{ with } \mathbb{A}^0 = \boldsymbol{\omega}^0 \times \text{ and } \boldsymbol{\omega}^0 = \mathbb{R}^{-1}\boldsymbol{\omega}$$

Moreover, we by differentiation of  $\mathbb{A}^0$  have

$$\begin{aligned} \dot{\mathbb{A}}^0 &= \dot{\mathbb{R}}^\top\mathbb{A}\mathbb{R} + \mathbb{R}^\top\dot{\mathbb{A}}\mathbb{R} + \mathbb{R}^\top\mathbb{A}\dot{\mathbb{R}} \\ &= \mathbb{R}^\top(-\mathbb{A}^2 + \dot{\mathbb{A}} + \mathbb{A}^2)\mathbb{R} \\ &= \mathbb{R}^{-1}\dot{\mathbb{A}}\mathbb{R} \end{aligned}$$

according to which *the derivative of the transform is the transform of the derivative* of  $\mathbb{A}$ . What we have established is that

$$\mathbb{R}^\top\ddot{\mathbb{R}} = \dot{\mathbb{A}}^0 + \mathbb{A}^0\mathbb{A}^0$$

from which information it follows that (37) can be written

$$\begin{aligned} \mathbb{N}^0 &= (\dot{\mathbb{A}}^0 + \mathbb{A}^0\mathbb{A}^0)\mathbb{M}^0 - \mathbb{M}^0(-\dot{\mathbb{A}}^0 + \mathbb{A}^0\mathbb{A}^0) \\ &= (\dot{\mathbb{A}}^0\mathbb{M}^0 + \mathbb{M}^0\dot{\mathbb{A}}^0) + (\mathbb{A}^0\mathbb{A}^0\mathbb{M}^0 - \mathbb{M}^0\mathbb{A}^0\mathbb{A}^0) \end{aligned} \quad (38)$$

But<sup>11</sup>  $\mathbb{I} = (\text{tr}\mathbb{M})\cdot\mathbb{U} - \mathbb{M}$  entails  $\text{tr}\mathbb{I} = 2\text{tr}\mathbb{M}$  whence

$$\mathbb{M}^0 = \frac{1}{2}(\text{tr}\mathbb{I}^0)\cdot\mathbb{U} - \mathbb{I}^0$$

giving finally

$$\mathbb{N}^0 = [(\text{tr}\mathbb{I}^0)\cdot\dot{\mathbb{A}}^0 - (\dot{\mathbb{A}}^0\mathbb{I}^0 + \mathbb{I}^0\dot{\mathbb{A}}^0)] - (\mathbb{A}^0\mathbb{A}^0\mathbb{I}^0 - \mathbb{I}^0\mathbb{A}^0\mathbb{A}^0) \quad (39)$$

The claim now is that (39) stands to Euler's equations (35) in precisely the

---

<sup>11</sup> See again (19) on page 7.

relation that  $\mathbb{A}$  stands to  $\boldsymbol{\omega}$ . The point is most easily established by direct (*Mathematica*-assisted) calculation: set

$$\mathbb{A}^0 = \begin{pmatrix} 0 & -\omega_3^0 & \omega_2^0 \\ \omega_3^0 & 0 & -\omega_1^0 \\ -\omega_2^0 & \omega_1^0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbb{I}^0 = \begin{pmatrix} I_1^0 & 0 & 0 \\ 0 & I_2^0 & 0 \\ 0 & 0 & I_3^0 \end{pmatrix}$$

and discover that the expression on the right side of (39) can be written

$$\begin{pmatrix} 0 & -[I_3^0 \dot{\omega}_3^0 + (I_2^0 - I_1^0) \omega_2^0 \omega_1^0] & [I_2^0 \dot{\omega}_2^0 + (I_1^0 - I_3^0) \omega_1^0 \omega_3^0] \\ [I_3^0 \dot{\omega}_3^0 + (I_2^0 - I_1^0) \omega_2^0 \omega_1^0] & 0 & -[I_1^0 \dot{\omega}_1^0 + (I_3^0 - I_2^0) \omega_3^0 \omega_2^0] \\ -[I_2^0 \dot{\omega}_2^0 + (I_1^0 - I_3^0) \omega_1^0 \omega_3^0] & [I_1^0 \dot{\omega}_1^0 + (I_3^0 - I_2^0) \omega_3^0 \omega_2^0] & 0 \end{pmatrix}$$

Taking the antisymmetric matrix on the left side of (39) to mean

$$\begin{pmatrix} 0 & -N_3^0 & N_2^0 \\ N_3^0 & 0 & -N_1^0 \\ -N_2^0 & N_1^0 & 0 \end{pmatrix}$$

we see that the Lagrangian formalism has led us to what is in effect the *dual* of Euler's system of equations.

**7. Euler angles.** We could sidestep the constraint problem altogether if we could produce a parameterized description of the elements of  $O(3)$ , analogous to the description which

$$\mathbb{R}(\theta) \equiv \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

provides of the elements of  $O(2)$ . This was first accomplished by Euler,<sup>12</sup> who observed (see the following figure) that if one

- rotates through an appropriate angle  $\phi$  about the 3-axis, then
- rotates through an appropriate angle  $\theta$  about the new 1-axis, then
- rotates through an appropriate angle  $\psi$  about the newest 3-axis

one can bring any **frame** into coincidence with any other **frame**.<sup>13</sup> Reading from

---

<sup>12</sup> L. Euler (1707–1783) retained an interest in dynamics—particularly the dynamics of rigid bodies—throughout his professional career, but his papers on the subject were written mainly between 1749 and 1760, when he was attached to the court of Frederick the Great, in Berlin. I suspect it was his physical work which stimulated the *invention* of the Euler angles, though they are of independent mathematical interest and importance.

<sup>13</sup> We assume, of course, that the two frames share the same origin and are similarly handed.

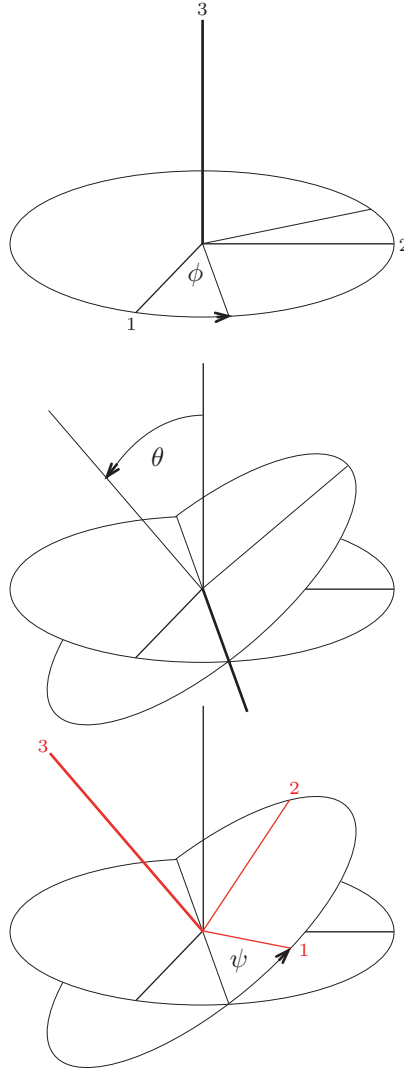


FIGURE 5: *The sequence of axial rotations that give rise to Euler's description of the elements of the group  $O(3)$  of rotations in 3-space. The angles  $\{\phi, \theta, \psi\}$  are called "Euler angles" and the two planes intersect in what is called the "line of nodes."*

the figure we have

$$\begin{aligned} \mathbf{r} &= \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{r} \quad (40.1) \\ &\equiv \mathbb{E}(\phi, \theta, \psi) \mathbf{r} \end{aligned}$$

serves to describe how the coordinates  $\mathbf{r}$  relative to the rotated **red frame** of a fixed point  $P$  are related to its coordinates  $\mathbf{r}$  relative to the **black frame**:

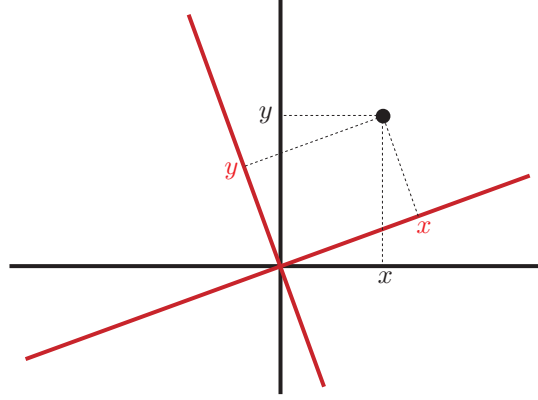


FIGURE 6: *The passive (or “alias”) interpretation that at (40.1) has been used to introduce the Euler angles.*

By *Mathematica*-assisted calculation we find

$$\begin{aligned} \mathbb{E}(\phi, \theta, \psi) &= \begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \psi \sin \phi & \cos \psi \sin \phi + \cos \theta \sin \psi \cos \phi & \sin \theta \sin \psi \\ -\sin \psi \cos \phi - \cos \theta \cos \psi \sin \phi & -\sin \psi \sin \phi + \cos \theta \cos \psi \cos \phi & \sin \theta \cos \psi \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix} \\ &= \mathbb{E}(\phi + \pi, -\theta, \psi + \pi) \end{aligned} \quad (40.2)$$

So archetypically symmetric would  $O(3)$  appear to be that it seems distinctly odd that Euler would have adopted such an asymmetric procedure to assign names to the elements of  $O(3)$ . It becomes in this light natural to ask: “Can a more symmetrical variant of Euler’s procedure be devised—a procedure that extends straightforwardly to the general case  $O(n)$ ?” Whatever may be the answers to those questions, it will be noticed (see Figure 7) that Euler’s defining procedure does relate very naturally/directly to the casually observed *behavior* of tops.

From our interest in the motion of rigid bodies we acquire interest in the rotation-induced adjustments  $\mathbf{x}^0 \mapsto \mathbf{x}$  of the coordinates relative to the **space frame** of points that are fixed in the body (*i.e.*, with respect to the **body frame**). The transformation of interest to us is therefore not passive but *active* (not an “alias” but an “alibi transformation” in Wigner’s language): compare Figure 8 with Figure 6. We write, as has been our practice since page 2,

$$\mathbf{x} = \mathbb{R} \mathbf{x}^0$$

where now

$$\mathbb{R} = \mathbb{E}^{-1}(\phi, \theta, \psi)$$

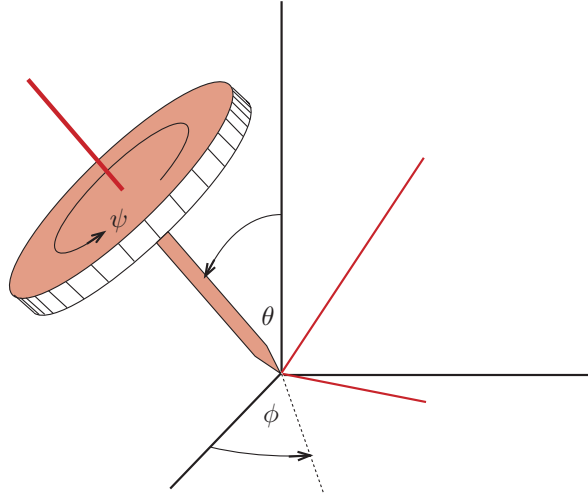


FIGURE 7: *It is, I suspect, not by accident that Euler's angles equip one to describe very simply and directly the motion executed by tops:*

$\phi$  measures **precession**,  
 $\theta$  measures **tilt** (*nutation*),  
 $\psi$  measures **spin**.

*The line of nodes is shown here as a dotted line.*

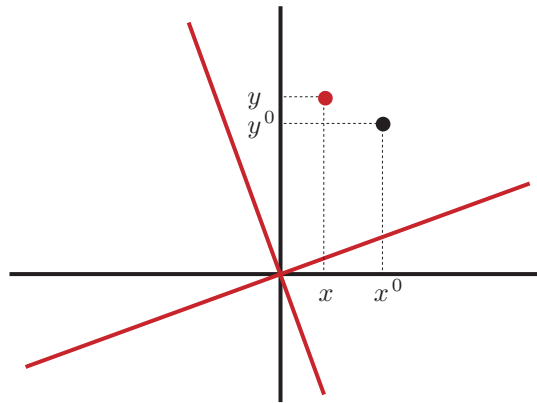


FIGURE 8: *The active (or "alibi") interpretation that at (41) is used to introduce Euler angles into the theory of tops.*

Explicitly

$$\begin{aligned}\mathbb{R} &= \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \mathbb{R}_3(\phi) \cdot \mathbb{R}_1(\theta) \cdot \mathbb{R}_3(\psi)\end{aligned}\quad (41)$$

To compute  $\boldsymbol{\omega}$  we proceed from

$$\begin{aligned}\begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} &= \dot{\mathbb{R}} \mathbb{R}^{-1} \\ &= \{ \dot{\mathbb{R}}_3(\phi) \cdot \mathbb{R}_1(\theta) \cdot \mathbb{R}_3(\psi) + \mathbb{R}_3(\phi) \cdot \dot{\mathbb{R}}_1(\theta) \cdot \mathbb{R}_3(\psi) + \mathbb{R}_3(\phi) \cdot \mathbb{R}_1(\theta) \cdot \dot{\mathbb{R}}_3(\psi) \} \\ &\quad \cdot \mathbb{R}_3^T(\psi) \cdot \mathbb{R}_1^T(\theta) \cdot \mathbb{R}_3^T(\phi) \\ &= \dot{\mathbb{R}}_3(\phi) \cdot \mathbb{R}_3^T(\phi) + \mathbb{R}_3(\phi) \cdot \dot{\mathbb{R}}_1(\theta) \cdot \mathbb{R}_1^T(\theta) \cdot \mathbb{R}_3^T(\phi) \\ &\quad + \mathbb{R}_3(\phi) \cdot \mathbb{R}_1(\theta) \cdot \dot{\mathbb{R}}_3(\psi) \cdot \mathbb{R}_3^T(\psi) \cdot \mathbb{R}_1^T(\theta) \cdot \mathbb{R}_3^T(\phi)\end{aligned}$$

But by quick calculation

$$\dot{\mathbb{R}}_3(\phi) \cdot \mathbb{R}_3^T(\phi) = \dot{\phi} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \dot{\phi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times$$

and similarly

$$\begin{aligned}\dot{\mathbb{R}}_1(\theta) \cdot \mathbb{R}_1^T(\theta) &= \dot{\theta} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \\ \dot{\mathbb{R}}_3(\psi) \cdot \mathbb{R}_3^T(\psi) &= \dot{\psi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times\end{aligned}$$

Drawing upon LEMMA (33) we therefore have the deceptively suggestive formal statement

$$\boldsymbol{\omega} = \dot{\boldsymbol{\phi}} + \dot{\boldsymbol{\theta}} + \dot{\boldsymbol{\psi}} \quad (41.1)$$

where

$$\dot{\boldsymbol{\phi}} \equiv \dot{\phi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \dot{\boldsymbol{\theta}} \equiv \dot{\theta} \mathbb{R}_3(\phi) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \dot{\boldsymbol{\psi}} \equiv \dot{\psi} \mathbb{R}_3(\phi) \mathbb{R}_1(\theta) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

give

$$\dot{\boldsymbol{\phi}} = \dot{\phi} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \dot{\boldsymbol{\theta}} = \dot{\theta} \begin{pmatrix} \cos \phi \\ \sin \phi \\ 0 \end{pmatrix}, \quad \dot{\boldsymbol{\psi}} = \dot{\psi} \begin{pmatrix} \sin \theta \sin \phi \\ -\sin \theta \cos \phi \\ \cos \theta \end{pmatrix} \quad (41.2)$$

whence finally

$$\boldsymbol{\omega} = \begin{pmatrix} \dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi \\ \dot{\theta} \sin \phi - \dot{\psi} \sin \theta \cos \phi \\ \dot{\phi} + \dot{\psi} \cos \theta \end{pmatrix} \quad (42)$$

The symbols  $\dot{\boldsymbol{\phi}}$ ,  $\dot{\boldsymbol{\theta}}$  and  $\dot{\boldsymbol{\psi}}$  are “deceptively suggestive” in that they are *intended to be read wholistically*: they are *not* intended to be read as references to the time derivatives of vectors  $\boldsymbol{\phi}$ ,  $\boldsymbol{\theta}$  and  $\boldsymbol{\psi}$ . Suppose it *were* possible to write

$$\boldsymbol{\omega} = \dot{\boldsymbol{\alpha}} \quad \text{with} \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_1(\phi, \theta, \psi) \\ \alpha_2(\phi, \theta, \psi) \\ \alpha_3(\phi, \theta, \psi) \end{pmatrix}$$

We would then have

$$\boldsymbol{\omega} dt = \begin{pmatrix} \alpha_{1,\phi}(\phi, \theta, \psi) d\phi + \alpha_{1,\theta}(\phi, \theta, \psi) d\theta + \alpha_{1,\psi}(\phi, \theta, \psi) d\psi \\ \alpha_{2,\phi}(\phi, \theta, \psi) d\phi + \alpha_{2,\theta}(\phi, \theta, \psi) d\theta + \alpha_{2,\psi}(\phi, \theta, \psi) d\psi \\ \alpha_{3,\phi}(\phi, \theta, \psi) d\phi + \alpha_{3,\theta}(\phi, \theta, \psi) d\theta + \alpha_{3,\psi}(\phi, \theta, \psi) d\psi \end{pmatrix}$$

and the functions  $\alpha_{k,\lambda}(\phi, \theta, \psi) : k \in \{1, 2, 3\}, \lambda \in \{\phi, \theta, \psi\}$  would assuredly satisfy the integrability conditions

$$\alpha_{k,\phi\theta} = \alpha_{k,\theta\phi}, \quad \alpha_{k,\phi\psi} = \alpha_{k,\psi\phi}, \quad \alpha_{k,\theta\psi} = \alpha_{k,\psi\theta}$$

Which the  $\alpha_{k\lambda}$ -functions latent in (42) obviously do *not* satisfy: it is *not possible* to write  $\boldsymbol{\omega} = \dot{\boldsymbol{\alpha}}$ . Right at the heart of 3-dimensional rotational kinematics lives a **non-integrability condition**. Contrast this with the 2-dimensional situation, where if

$$\mathbb{R} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

then

$$\dot{\mathbb{R}} \mathbb{R}^\top = \frac{d}{dt} \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}, \quad \text{which by dualization becomes simply } \boldsymbol{\omega} = \frac{d}{dt} \theta$$

We found at (13) that the intrinsic (or rotational) kinetic energy can be described

$$\mathcal{T} = \frac{1}{2} \boldsymbol{\omega}^\top \mathbb{I} \boldsymbol{\omega}$$

which by (25) becomes

$$\begin{aligned} \mathcal{T} &= \frac{1}{2} \boldsymbol{\omega}^\top \mathbb{R} \mathbb{I}^0 \mathbb{R}^{-1} \boldsymbol{\omega} \\ &= \frac{1}{2} \boldsymbol{\omega}^{0\top} \mathbb{I}^0 \boldsymbol{\omega}^0 \quad \text{with} \quad \boldsymbol{\omega}^0 \equiv \mathbb{R}^{-1} \boldsymbol{\omega} \end{aligned} \quad (43)$$

One could evaluate  $\boldsymbol{\omega}^0$  by *Mathematica*-assisted brute force. Or one could work



from the dual of (43); *i.e.*, from

$$\begin{aligned}\mathbf{A}^0 &= \mathbb{R}^{-1} \mathbf{A} \mathbb{R} \\ &= \mathbb{R}^{-1} \dot{\mathbb{R}} \mathbb{R}^{-1} \mathbb{R} = \mathbb{R}^{-1} \dot{\mathbb{R}}\end{aligned}$$

by the methods that led to (42). By either procedure one is led to

$$\boldsymbol{\omega}^0 = \begin{pmatrix} \dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi \\ \dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi \\ \dot{\phi} \cos \theta \qquad \qquad \qquad + \dot{\psi} \end{pmatrix} \quad (44)$$

**8. Lagrangian formalism using Euler angles.** It follows clearly from results now in hand that the rotational dynamics of a rigid body can be considered to devolve from a Lagrangian of—if we exercise our option to identify the body frame with the principal axis frame—the form

$$\begin{aligned}\mathcal{L} &= \mathcal{L}(\dot{\phi}, \dot{\theta}, \dot{\psi}, \phi, \theta, \psi) \\ &= \frac{1}{2} I_1^0 (\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)^2 + \frac{1}{2} I_2^0 (\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)^2 \\ &\quad + \frac{1}{2} I_3^0 (\dot{\phi} \cos \theta + \dot{\psi})^2 - \mathcal{U}(\phi, \theta, \psi)\end{aligned} \quad (45)$$

The resulting Lagrange equations are coupled differential equations of *second* order, and are of much more complicated appearance than the Euler equations, which were seen at (35) to be a symmetric set of *first* order equations. Note, however, that

$$\left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{\psi}} - \frac{\partial}{\partial \psi} \right\} \mathcal{L} = 0 \quad (46.1)$$

becomes

$$\underbrace{-\partial \mathcal{U} / \partial \psi}_{N_3^0} = I_3^0 \underbrace{\frac{d}{dt} (\dot{\phi} \cos \theta + \dot{\psi})}_{\omega_3^0} + (I_2^0 - I_1^0) \underbrace{(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)}_{\omega_2^0} \underbrace{(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)}_{\omega_1^0}$$

which precisely reproduces the third of the Euler equations (35). Certain *linear combinations* of the remaining Lagrange equations

$$\left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{\theta}} - \frac{\partial}{\partial \theta} \right\} \mathcal{L} = 0 \quad (46.2)$$

$$\left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{\phi}} - \frac{\partial}{\partial \phi} \right\} \mathcal{L} = 0 \quad (46.3)$$

serve to reproduce first and second of the Euler equations.

One can understand the relative complexity the Lagrange equations (46) on grounds that they undertake to accomplish more than Euler equations. Solutions  $\{\phi(t), \theta(t), \psi(t)\}$  of (46) serve in themselves to describe *how the body gyrates*, while solutions  $\boldsymbol{\omega}(t)$  of (35) leave us—as we saw at (36)—one awkward integration away from such explicit information.

If  $I_1^0 = I_2^0$  then the rigid body (or “top”) is said to be **symmetrical**,<sup>14</sup> and the Lagrangian (45) assumes the form

$$\mathcal{L} = \frac{1}{2}I_1^0(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}I_3^0(\dot{\phi} \cos \theta + \dot{\psi})^2 - \mathcal{U}(\phi, \theta, \psi) \quad (47)$$

while the Euler equations (35) become

$$\left. \begin{aligned} N_1^0 &= I_1^0 \dot{\omega}_1^0 - (I_1^0 - I_3^0) \omega_3^0 \omega_2^0 \\ N_2^0 &= I_2^0 \dot{\omega}_2^0 + (I_1^0 - I_3^0) \omega_1^0 \omega_3^0 \\ N_3^0 &= I_3^0 \dot{\omega}_3^0 \end{aligned} \right\} \quad (48)$$

Had we instead set (not  $I_1^0 = I_2^0$  but, say)  $I_1^0 = I_3^0$  then the simplification of (45) would have been masked or disguised, while the simplified Euler equations would be precisely similar to (48). It is, within the Lagrangian formalism, as a mere convenience, and without real loss of generality, that one identifies the 3-axis of the principal axis frame with the symmetry axis of a symmetrical top.

If  $I_1^0 = I_2^0 = I_3^0$  then the Lagrangian simplifies still further, to become

$$\mathcal{L} = \frac{1}{2}I_1^0(\dot{\phi}^2 + 2\dot{\phi}\dot{\psi} \cos \theta + \dot{\psi}^2) - \mathcal{U}(\phi, \theta, \psi) \quad (49)$$

while the Euler equations assume the trivial form

$$\mathbf{N}^0 = I_1^0 \dot{\boldsymbol{\omega}}^0 \quad (50)$$

We will return to discussion of some of the remarkably rich physics that arises in these important special cases; *i.e.*, to a discussion of the *solutions* of the associated equations of motion.

**9. Free motion of a rigid body.** Working first in the *space frame* (since it is from a position at rest with respect to the almost-inertial laboratory that we expect to view our gyrating rigid objects), we return to (31) and, setting  $\mathbf{N} = \mathbf{0}$ , obtain

$$\dot{\mathbf{S}} = \mathbb{I} \dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbb{I} \boldsymbol{\omega} = \mathbf{0} \quad (51)$$

according to which

$$\text{Spin } \mathbf{S} \text{ is a constant of the free motion of a rigid body} \quad (52)$$

Trivially,

$$\text{So also is } S^2 \equiv \mathbf{S} \cdot \mathbf{S} \text{ a constant of the free motion} \quad (53)$$

---

<sup>14</sup> The term conveys no information about the actual *shape* of the top.

But the general constancy of  $\mathbf{S}$  does, in general, *not* imply “uniformity” of the rotation: rotation typically causes  $\mathbb{I}$  to become time-dependent, which by  $\boldsymbol{\omega} = \mathbb{I}^{-1}(t)\mathbf{S}$  forces  $\boldsymbol{\omega}$  to be time-dependent. Rotational uniformity  $\dot{\boldsymbol{\omega}} = \mathbf{0}$  is seen by (51) to entail that  $\boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega} = \mathbf{0}$ ; *i.e.*, that  $\boldsymbol{\omega}$  and  $\mathbb{I}\boldsymbol{\omega}$  be *parallel*:

$$\text{Rotational uniformity } \dot{\boldsymbol{\omega}} = \mathbf{0} \text{ requires that } \boldsymbol{\omega} \text{ be an eigenvector of } \mathbb{I}: \mathbb{I}\boldsymbol{\omega} = \lambda\boldsymbol{\omega} \text{ with } \lambda \in \{I_1, I_2, I_3\} \quad (54)$$

For a *free* rigid body the rotational energy is all kinetic, and as we saw at (18.2) can be described  $\mathcal{T} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{S}$ . To prepare for a proof that  $\dot{\mathcal{T}} = 0$  I digress to establish the following

**LEMMA:** The objects

$$\mathbb{A} = \|A_{ij}\| = \begin{pmatrix} 0 & -a^3 & +a^2 \\ +a^3 & 0 & -a^1 \\ -a^2 & +a^1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{a} = \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix}$$

encountered on page 15 are “dual” in the sense that

$$A_{ij} = \epsilon_{ikj} a^k \quad (55)$$

The Levi-Civita symbol is known<sup>15</sup> to assume the *same numerical values in every coordinate system* if transformed as a tensor density of weight  $w = -1$ :

$$\begin{aligned} \tilde{\epsilon}_{ikj} &\equiv W^{-1} \cdot W^a{}_i W^c{}_k W^b{}_j \epsilon_{acb} \quad \text{where} \quad W \equiv \det \mathbb{W} \\ &= \epsilon_{ikj} \end{aligned}$$

Let it be assumed that the  $a^k$  transform as components of a weightless contravariant vector, and that the  $A_{ij}$  transform as components of a covariant tensor density of second rank and negative unit weight. The assertion (55) then preserves its design in all coordinate systems, and in

$$\tilde{A}_{ij} = \epsilon_{ikj} \tilde{a}^k \quad (56.1)$$

we have simply the statement that

$$\text{transform of dual} = \text{dual of transform} \quad (56.2)$$

Explicitly

$$W^{-1} \cdot W^m{}_i W^n{}_j A_{mn} = \epsilon_{ikj} M^k{}_p a^p$$

which in index-free notation becomes

$$(\det \mathbb{W})^{-1} \cdot \mathbb{W}^T \mathbb{A} \mathbb{W} = (\mathbb{M} \mathbf{a}) \times \quad (56.3)$$

where  $\mathbb{W} \equiv \mathbb{M}^{-1}$ . Notice that (56) gives back (33) as a special case.

---

<sup>15</sup> See CLASSICAL ELECTRODYNAMICS (1980/81), pages 172/3.

Returning now to the demonstration that  $\mathcal{T}$  is conserved, we have

$$\begin{aligned}\dot{\mathcal{T}} &= \frac{1}{2}(\dot{\boldsymbol{\omega}} \cdot \mathbf{S} + \boldsymbol{\omega} \cdot \dot{\mathbf{S}}) \\ &= \frac{1}{2}\dot{\boldsymbol{\omega}} \cdot \mathbf{S} \quad \text{by (51): } \dot{\mathbf{S}} = \mathbf{0}\end{aligned}$$

But—again by (51)—

$$\begin{aligned}\dot{\boldsymbol{\omega}} &= -\mathbb{I}^{-1}(\boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega}) \\ &= -\mathbb{I}^{-1}(\boldsymbol{\omega} \times) \mathbb{I}^{-1} \cdot \mathbb{I}^2 \boldsymbol{\omega}\end{aligned}$$

But the symmetry of  $\mathbb{I}$  implies that of  $\mathbb{I}^{-1}$ , so we can use LEMMA (56.3) to obtain

$$= -(\det \mathbb{I})^{-1}(\mathbb{I}\boldsymbol{\omega}) \times \mathbb{I}^2 \boldsymbol{\omega} \quad (57)$$

giving

$$\dot{\mathcal{T}} = \frac{1}{2}(\det \mathbb{I})^{-1}(\mathbb{I}^2 \boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega}) \cdot \mathbb{I}\boldsymbol{\omega}$$

But the triple scalar product

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0 \text{ unless } \mathbf{a}, \mathbf{b} \text{ and } \mathbf{c} \text{ are linearly independent}$$

Which in the preceding equation is clearly not the case. So we have

$$\dot{\mathcal{T}} = 0: \mathcal{T} \text{ is a constant of the free motion} \quad (58)$$

It is interesting to notice that while

$$\boldsymbol{\omega} \cdot \mathbb{I}^2 \boldsymbol{\omega} = S^2$$

and

$$\boldsymbol{\omega} \cdot \mathbb{I} \boldsymbol{\omega} = 2\mathcal{T}$$

are constants of free rigid body motion,  $\boldsymbol{\omega} \cdot \boldsymbol{\omega} = \omega^2$  is typically *not* constant, for  $\frac{1}{2} \frac{d}{dt} \omega^2 = \dot{\boldsymbol{\omega}} \cdot \boldsymbol{\omega} = \dot{\boldsymbol{\omega}} \cdot \boldsymbol{\omega}$  which by (57) supplies

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \omega^2 &= (\det \mathbb{I})^{-1}(\mathbb{I}^2 \boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega}) \cdot \boldsymbol{\omega} \\ &= (\det \mathbb{I})^{-1}[\mathbb{I}^2 \boldsymbol{\omega}, \mathbb{I}\boldsymbol{\omega}, \boldsymbol{\omega}] \text{ in a standard triple scalar product notation} \\ &= (\det \mathbb{I})^{-1}(I_1 - I_2)(I_1 - I_3)(I_2 - I_3)\omega_1\omega_2\omega_3\end{aligned}$$

Drawing upon  $\boldsymbol{\omega} = \mathbb{R}\boldsymbol{\omega}^0$ ,  $\mathbf{S} = \mathbb{R}\mathbf{S}^0$ ,  $\mathbb{I} = \mathbb{R}\mathbb{I}^0\mathbb{R}^{-1}$  we see by easy arguments

- That  $\mathbf{S}^0$  is *not* constant unless  $\mathbb{R}$  happens to describe spin about the (invariable)  $\mathbf{S}$ -axis;
- That therefore  $\boldsymbol{\omega}^0 = (\mathbb{I}^0)^{-1}\mathbf{S}^0$  is generally not constant either;
- That  $S^2 = \mathbf{S} \cdot \mathbf{S} = \mathbf{S}^0 \cdot \mathbf{S}^0$  provide alternative descriptions of  $S^2$  (conserved);
- That  $\mathcal{T} = \frac{1}{2}\mathbf{S}^T \mathbb{I}^{-1} \mathbf{S} = \frac{1}{2}\mathbf{S}^{0T} (\mathbb{I}^0)^{-1} \mathbf{S}^0$  provide alternative descriptions of  $\mathcal{T}$  (conserved);
- That  $\omega^2 = \boldsymbol{\omega} \cdot \boldsymbol{\omega} = \boldsymbol{\omega}^0 \cdot \boldsymbol{\omega}^0$  provide alternative descriptions of  $\omega^2$  (not conserved).

We look now to the time-dependence of  $\omega^2$ . Drawing upon (57) we have

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \omega^2 &= \dot{\boldsymbol{\omega}}^0 \cdot \boldsymbol{\omega}^0 \\ &= (\det \mathbb{I}^0)^{-1} (\mathbb{I}^{02} \boldsymbol{\omega}^0 \times \mathbb{I}^0 \boldsymbol{\omega}^0) \cdot \boldsymbol{\omega}^0\end{aligned}$$

Electing to work in the principal axis frame, where  $\mathbb{I}^0$  is diagonal, and agreeing to omit all <sup>0</sup> superscripts for the duration of this argument, we therefore have

$$\begin{aligned}&= \frac{1}{I_1 I_2 I_3} \begin{vmatrix} I_1^2 \omega_1 & I_1 \omega_1 & \omega_1 \\ I_2^2 \omega_2 & I_2 \omega_2 & \omega_2 \\ I_3^2 \omega_3 & I_3 \omega_3 & \omega_3 \end{vmatrix} \\ &= \frac{\det \mathbb{J}}{I_1 I_2 I_3} \omega_1 \omega_2 \omega_3 \quad \text{with} \quad \mathbb{J} \equiv \begin{pmatrix} I_1^2 & I_2^2 & I_3^2 \\ I_1 & I_2 & I_3 \\ 1 & 1 & 1 \end{pmatrix}\end{aligned}\quad (59)$$

To bring our conservation laws into play we write

$$\mathbb{J} \begin{pmatrix} \omega_1^2 \\ \omega_2^2 \\ \omega_3^2 \end{pmatrix} = \begin{pmatrix} S^2 \\ 2\mathcal{T} \\ \omega^2 \end{pmatrix} \equiv \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

which gives

$$\begin{aligned}\begin{pmatrix} \omega_1^2 \\ \omega_2^2 \\ \omega_3^2 \end{pmatrix} &= \mathbb{J}^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \\ \mathbb{J}^{-1} &= (\det \mathbb{J})^{-1} \cdot \begin{pmatrix} (I_2 - I_3) & -(I_2 + I_3)(I_2 - I_3) & I_2 I_3 (I_2 - I_3) \\ (I_3 - I_1) & -(I_3 + I_1)(I_3 - I_1) & I_3 I_1 (I_3 - I_1) \\ (I_1 - I_2) & -(I_1 + I_2)(I_1 - I_2) & I_1 I_2 (I_1 - I_2) \end{pmatrix}\end{aligned}$$

From this equation it follows in particular that

$$(\omega_1 \omega_2 \omega_3)^2 = \frac{(I_2 - I_3)(I_3 - I_1)(I_1 - I_2)}{(\det \mathbb{J})^3} \{\text{etc.}\}$$

with

$$\{\text{etc.}\} \equiv [\alpha_1 - (I_2 + I_3)\alpha_2 + I_2 I_3 \alpha_3] [\alpha_1 - (I_2 + I_3)\alpha_2 + I_2 I_3 \alpha_3] [\alpha_1 - (I_2 + I_3)\alpha_2 + I_2 I_3 \alpha_3]$$

But  $\det \mathbb{J} = -(I_2 - I_3)(I_3 - I_1)(I_1 - I_2)$ , so we have

$$\begin{aligned}\left( \frac{\det \mathbb{J}}{I_1 I_2 I_3} \omega_1 \omega_2 \omega_3 \right)^2 &= -\frac{1}{(I_1 I_2 I_3)^2} \{\text{etc.}\} \\ &= (\lambda_1 - \omega^2)(\lambda_2 - \omega^2)(\lambda_3 - \omega^2)\end{aligned}$$

with

$$\begin{aligned}\lambda_1 &= \frac{2(I_2 + I_3)\mathcal{T} - S^2}{I_2 I_3} \\ \lambda_2 &= \frac{2(I_3 + I_1)\mathcal{T} - S^2}{I_3 I_1} \\ \lambda_3 &= \frac{2(I_1 + I_2)\mathcal{T} - S^2}{I_1 I_2}\end{aligned}$$

which, it will be noticed, are assembled from frame-independent

- *system parameters* and
- *constants of the free motion*

Returning with this information to (59) we come at last to the statement

$$\omega \frac{d\omega}{dt} = \sqrt{(\lambda_1 - \omega^2)(\lambda_2 - \omega^2)(\lambda_3 - \omega^2)}$$

which can be used to compute

$$\begin{aligned}\tau &\equiv \text{transit time: } \omega_{\text{initial}} \rightarrow \omega_{\text{final}} \\ &= \int_{\omega_{\text{initial}}}^{\omega_{\text{final}}} \frac{\omega}{\sqrt{(\lambda_1 - \omega^2)(\lambda_2 - \omega^2)(\lambda_3 - \omega^2)}} d\omega\end{aligned}\tag{60}$$

The integral leads to elliptic functions with complicated arguments, but is an integral with which *Mathematica* appears to be quite comfortable.<sup>16</sup>

Looking again to points enumerated at the bottom of page 28, we see that  $\mathbf{S}^0$  ranges simultaneously on a sphere of radius  $S$  in spin space

$$(S_1^0)^2 + (S_2^0)^2 + (S_3^0)^2 = S^2$$

and on an energy ellipsoid

$$\left(\frac{S_1^0}{\sqrt{2I_1}}\right)^2 + \left(\frac{S_2^0}{\sqrt{2I_2}}\right)^2 + \left(\frac{S_3^0}{\sqrt{2I_3}}\right)^2 = \mathcal{T}$$

with semi-axes that we may, as a handy convention, assume to have been indexed in ordered sequence

$$\sqrt{2\mathcal{T}I_1} \geq \sqrt{2\mathcal{T}I_2} \geq \sqrt{2\mathcal{T}I_3}$$

It is clear therefore on simple geometrical grounds (see Figure 9) that if  $S^2$  is given/fixed then

$$\mathcal{T}_{\text{least}} \leq \mathcal{T} \leq \mathcal{T}_{\text{most}}$$

---

<sup>16</sup> My derivation of (60) has been freely adapted from the discussion that can be found in §137 of E. J. Routh, *Advanced Dynamics of a System of Rigid Bodies* (6<sup>th</sup> edition 1905).

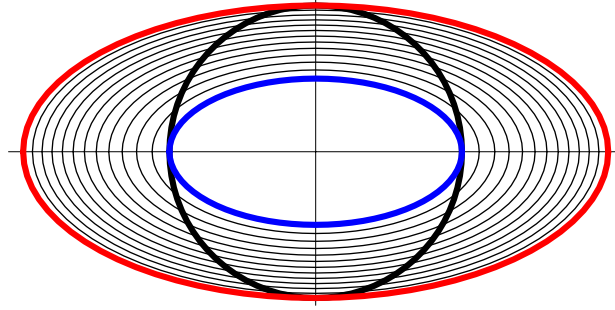


FIGURE 9: The **black** circle represents the sphere of radius  $S^2$  in 3-dimensional  $\mathcal{S}^0$ -space. From the relationship of the small **blue** energy ellipse to the sphere one deduces on purely geometrical grounds that

$$\mathcal{T}_{\text{least}} = \frac{S^2}{2I_{\text{largest}}} \quad (61.1)$$

while from the **red** ellipse one obtains

$$\mathcal{T}_{\text{most}} = \frac{S^2}{2I_{\text{smallest}}} \quad (61.2)$$

The intermediate ellipses were produced by incrementing the energy in equal steps.

Figures 10 are taken from a filmstrip,<sup>17</sup> and reveal a physically important new aspect of the situation that becomes evident only when advances from two to three dimensions. The free gyration of a rigid body causes  $\mathcal{S}^0$  to wander (precisely how?) along the *intersection* of the  $S$ -sphere and  $\mathcal{T}$ -ellipsoid. We learn from the figures to expect

- stable rotation about the major axis if  $\mathcal{T} = \mathcal{T}_{\text{least}}$ ;
- stable rotation about the minor axis if  $\mathcal{T} = \mathcal{T}_{\text{most}}$ ; but
- rotation about the intermediate axis (energy  $\mathcal{T} = \mathcal{T}_{\text{critical}}$ ) to be *unstable*.

---

<sup>17</sup> I have written

$$x^2 + y^2 + z^2 = 2^2$$

to describe the angular momentum sphere, and

$$(x/4)^2 + (3y/8)^2 + (z/2)^2 = (\frac{1}{4} + n\frac{1}{16}) \quad : \quad n = 0, 1, 2, \dots, 12$$

to describe a sequence of progressively more energetic ellipsoids. We then have  $x$ -puncture at  $n = 0$ ,  $y$ -puncture at  $n = 5$ ,  $z$ -puncture at  $n = 12$ . For relevant discussion see P. L. Lamy & J. A. Burns, "Geometrical approach to torque free motion of a rigid body having internal energy dissipation," AJP **40**, 441 (1972) and W. G. Harter & C. C. Kim, "Singular motions of asymmetric rotators," AJP **44**, 1080 (1976).

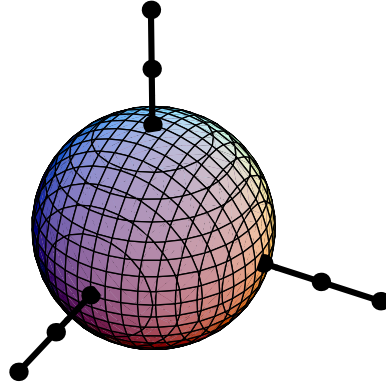


FIGURE 10a: *The ellipsoid of least energy is entirely interior to the  $S$ -sphere, to which it is tangent at only two points—the puncture points of the major principal axis.*

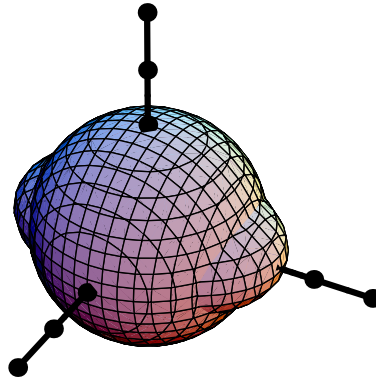


FIGURE 10b: *At an energy  $\mathcal{J}$  somewhat greater than  $\mathcal{J}_{\text{least}}$  the energy ellipsoid has become visible in the neighborhood of the major axis. Note that sphere and ellipsoid intersect on a roughly elliptical curve that envelops the major axis.*





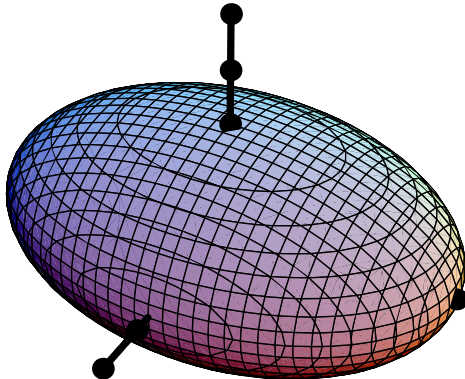


FIGURE 10e: *The ellipsoid of greatest energy is entirely exterior to the  $S$ -sphere, to which it is tangent at only two points—the puncture points of the minor axis.*

**10. Poinso's construction.** The preceding constructions live in  $\mathbf{S}^0$ -space: they tell us—on the assumption that the values of  $I_1, I_2, I_3$  are known, and that the values of  $S$  and  $\mathcal{T}$  have been prescribed—where  $\mathbf{S}^0$  is free to roam, but not how in time it elects to do so. We have

$$\omega_1^0 = S_1^0/I_1, \quad \omega_2^0 = S_2^0/I_2, \quad \omega_3^0 = S_3^0/I_3$$

but have, as yet, no *diagramatic* interpretation of the *motion* of  $\boldsymbol{\omega}^0$ . And even if we did possess  $\boldsymbol{\omega}^0(t)$ , it would be a long and arduous journey back to  $\mathbb{R}(t)$ , to understanding of how the rigid body itself moved.

These limitations were neatly circumvented by Louis Poinso (1777–1859), who devised a construction that owes its striking success mainly to the fact that it employs variables that refer not to the body frame but to the **space frame**. The equation

$$\mathcal{T}(\boldsymbol{\omega}) \equiv \frac{1}{2} \boldsymbol{\omega}^\top \mathbb{I}(t) \boldsymbol{\omega} = \mathcal{T}$$

defines what I will call the  $\mathcal{T}$ -ellipse: its center is pinned to the origin of  $\boldsymbol{\omega}$ -space, and its axes—of lengths

$$\sqrt{2\mathcal{T}/I_1} \leq \sqrt{2\mathcal{T}/I_2} \leq \sqrt{2\mathcal{T}/I_3}$$

—wobble about, reflecting the  $t$ -dependence of  $\mathbb{I}$  and remaining always in coincidence with the principal axes of the body (though the latter live not in  $\boldsymbol{\omega}$ -space but in  $\mathbf{r}$ -space).

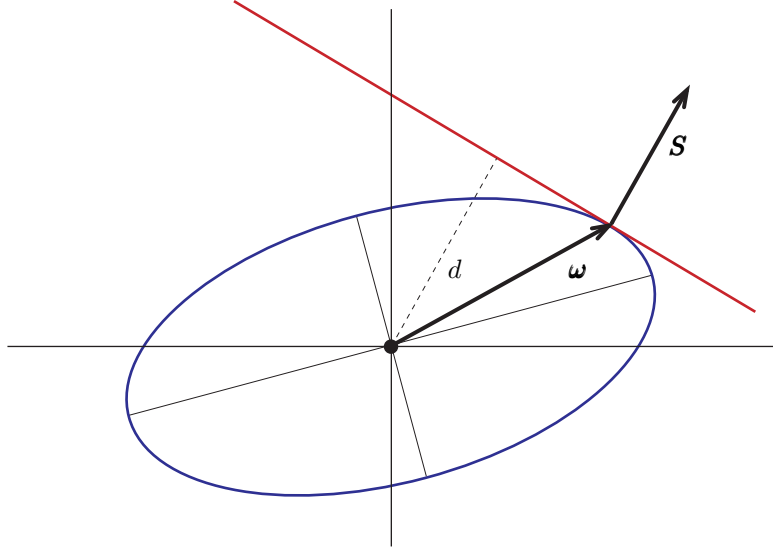


FIGURE 11: Physical variables  $\mathbf{S}$  and  $\omega_{\perp} \equiv 2\mathcal{T}/S$  determine the placement and orientation of Poincot's **invariant plane**, to which the  $\mathcal{T}$ -ellipse remains ever tangent

Poincot's construction proceeds now from two key observations, the first of which is that

$$\nabla_{\omega}\mathcal{T}(\omega) = \mathbb{I}\omega = \mathbf{S}$$

In words:  $\mathbf{S}$  stands normal to the plane which tangentially kisses the  $\mathcal{T}$ -ellipsoid at  $\omega$ . The planes thus constructed at various times  $t$  will, by the invariance of  $\mathbf{S}$ , be *parallel*. Poincot observed further that those planes are in fact *identical*, for the normal distance from origin to plane

$$\omega_{\perp} = \omega \cdot \hat{\mathbf{S}} = 2\mathcal{T}/S = \text{constant of free gyro motion}$$

In short:  $\hat{\mathbf{S}} \equiv \mathbf{S}/S$  and  $\omega_{\perp} = 2\mathcal{T}/S$  serve conjointly to identify a unique plane—call it  $\Pi(\mathbf{S}, \mathcal{T})$ —and since they are constants of the motion the plane is *invariable*. Evidently the  $\mathcal{T}$ -ellipsoid is forced by the Euler equations to move in such a way that

- while the center of the ellipsoid remains pinned at the origin
- the ellipsoid remains at every instant tangent to the invariable plane;
- the point  $\omega$  of tangency announces the instantaneous angular velocity of the ellipsoid (and evidently traces on the surface of the ellipsoid a closed curve).

To complete Poinso's construction we return to page 11, where it was reported that the moment of inertia *about the axis (through the center of mass) defined by the unit vector  $\mathbf{n}$*  can be described

$$\mathbf{n} \cdot \mathbb{I} \mathbf{n} = MR^2(\mathbf{n}) \quad (62)$$

where  $M$  refers to the *total mass* of the rigid body, and  $R(\mathbf{n})$  is the so-called *radius of gyration*. Introducing

$$\boldsymbol{\rho} \equiv \mathbf{n} / \sqrt{MR^2(\mathbf{n})} \quad (63)$$

we find that (62) can be written

$$\boldsymbol{\rho} \cdot \mathbb{I} \boldsymbol{\rho} = 1 \quad (64)$$

which serves to define the “inertia ellipsoid” in  $\boldsymbol{\rho}$ -space. The point to notice is that (64) follows also from  $\boldsymbol{\omega}^\top \mathbb{I}(t) \boldsymbol{\omega} = 2\mathcal{T}$  upon setting

$$\boldsymbol{\rho} = \boldsymbol{\omega} / \sqrt{2\mathcal{T}} \quad (65.1)$$

Evidently the inertia ellipsoid is—though it lives not in  $\boldsymbol{\omega}$ -space but in  $\boldsymbol{\rho}$ -space—a similarly oriented but rescaled copy of the  $\mathcal{T}$ -ellipsoid. As such, it moves in such a way as to be ever tangent to an invariable plane, from which its center maintains a distance

$$\rho_\perp = \omega_\perp / \sqrt{2\mathcal{T}} \quad (65.2)$$

While it would border on deceptive absurdity to attempt to “unpin the center of the  $\mathcal{T}$ -ellipsoid and to transport it from point to point in  $\boldsymbol{\omega}$ -space,” it *is* meaningful to do such a thing with the inertia ellipsoid in  $\boldsymbol{\rho}$ -space.<sup>18</sup> With Poinso, we observe that the vector

$$\boldsymbol{\rho} \quad : \quad \text{directed center} \longrightarrow \text{contact point}$$

is, by (65.1), momentarily axial: points on that line are therefore momentarily at rest. Which is to say: the inertial ellipsoid rolls without slipping on the invariant plane, and at an instantaneous rate proportional to the length of the  $\boldsymbol{\rho}$  vector (see FIGURE 12.) Since the principal axes of the inertial ellipsoid remain ever parallel to those of the rigid body itself, Poinso has given us what is, in effect, an ingenious special-purpose analog computer—its only limitation being that it lives in a fairly abstract space.

It should, perhaps, be noted that the figure of Poinso's inertia ellipsoid—with its semi-axes of lengths

$$\sqrt{1/I_1} \leq \sqrt{1/I_2} \leq \sqrt{1/I_3} \quad (66.1)$$

—is in an obvious sense “reciprocal” to that of the energy ellipsoid encountered on page 30 and in Figures 10: the former lives in  $\boldsymbol{\rho}$ -space, the latter in  $\boldsymbol{S}$ -space, and has semi-axes of lengths

$$\sqrt{2\mathcal{T}I_1} \geq \sqrt{2\mathcal{T}I_2} \geq \sqrt{2\mathcal{T}I_3} \quad (66.2)$$

---

<sup>18</sup> That is *why we took the trouble to introduce  $\boldsymbol{\rho}$ -space!*

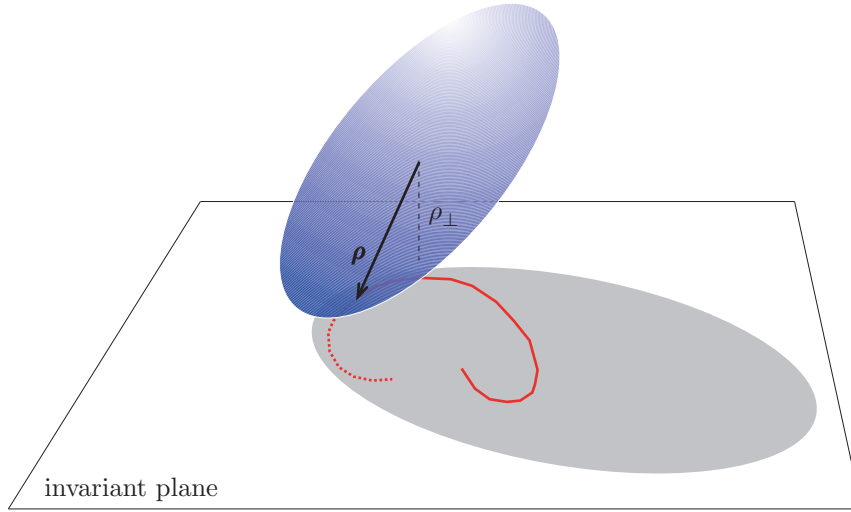


FIGURE 12: **Poinso's construction:** *The inertial ellipsoid, which lives in  $\boldsymbol{\rho}$ -space, rolls without slipping on the invariant plane, with  $\rho_{\perp}$  held constant.. The moving contact point traces a closed curve, called the “polhode” (from  $\pi\acute{o}\lambda o\varsigma$  = axis +  $\acute{o}\delta\acute{o}\varsigma$  = path), on the surface of the ellipsoid, and a typically more complicated curve called the “herpolhode” on the plane. In the figure the invariant plane has—for clarity—been laid flat, and only the herpolhode is shown.*

It will prove useful to observe in this connection that if the physical ellipsoid

$$(x_1/a_1)^2 + (x_2/a_2)^2 + (x_3/a_3)^2 = 1$$

is filled with material of uniform density  $d$  then, by quick calculation,<sup>19</sup> the total mass of the object is given by

$$M = d \cdot \frac{4}{3} \pi a_1 a_2 a_3$$

and the moment of inertia matrix becomes

$$\mathbb{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad \text{with} \quad \begin{cases} I_1 = \frac{1}{5} M (a_2^2 + a_3^2) \\ I_2 = \frac{1}{5} M (a_1^2 + a_3^2) \\ I_3 = \frac{1}{5} M (a_1^2 + a_2^2) \end{cases}$$

(which gives back a very familiar result when  $a_1 = a_2 = a_3 = r$ ). Let it be assumed that  $a_1 \leq a_2 \leq a_3$ . We then have

$$I_1 \geq I_2 \geq I_3 \quad \text{whence} \quad \sqrt{1/I_1} \leq \sqrt{1/I_2} \leq \sqrt{1/I_3}$$

<sup>19</sup> See Problem 2-1.

—the implication being that *the physical ellipsoid and the associated Poinso<sup>t</sup> ellipsoid have distinct but qualitatively similar figures*. During the first half of the 19<sup>th</sup> Century close study of the geometry of polhode and herpolhode appears to have been a flourishing industry. Many wonderful facts were discovered, most of which are now forgotten, but some of which can be recovered from the old textbooks. Webster reports, for example,<sup>20</sup> that while polhodes are invariably reentrant, herpolhodes are usually not, and never possess inflection points: their name (from  $\epsilon\rho\pi\epsilon\iota\nu$  = to creep like a snake) is therefore somewhat misleading. The old literature provides elaborate figures produced by laborious hand calculation. It would be amusing—possibly instructive—to use modern computer resources to recreate some of that material, to produce animated images of rolling Poinso<sup>t</sup> ellipsoids, *etc.*

I must emphasize that Poinso<sup>t</sup>'s construction pertains to the gyrodynamics of *free* bodies. The application of torques would, in general, cause both  $\mathbf{S}$  and  $\mathcal{T}$  to become time-dependent. The formerly “invariant plane” would begin to move, to wobble, and the center of the inertial ellipsoid to rise and fall with respect to that plane: the whole construction would become “seasick,” and rapidly lose its utility.

**11. First look at the free gyration of a symmetric top.** By “symmetric” we refer here not to the shape of the body itself, but to the shape of its only dynamically relevant feature—the inertia ellipsoid (or—reciprocally—the energy ellipsoid in spin space) . . . though in practice most rigid bodies that are symmetric in the above sense are axially symmetric also in their spatial form. I will occasionally allow myself to call such bodies “tops.”

We are obliged at the outset to distinguish (see FIGURE 13) two principal classes of axially symmetric tops:

$$\begin{array}{ll} \text{OBLATE} & \text{PROLATE} \\ A \equiv I_1 > I_2 = I_3 \equiv B & B \equiv I_1 = I_2 > I_3 \equiv A \end{array}$$

Standing at the interface between those two classes is the essentially trivial class

$$\begin{array}{c} \text{SPHERICAL} \\ A \equiv I_1 = I_2 = I_3 \end{array}$$

of *fully* symmetric tops. In the presence of symmetry the Euler equations (35) simplify: we have

$$\left[ \begin{array}{l} I_1 \dot{\omega}_1 = 0 \\ I_3 \dot{\omega}_2 + (I_1 - I_3) \omega_1 \omega_3 = 0 \\ I_3 \dot{\omega}_3 - (I_1 - I_3) \omega_1 \omega_2 = 0 \end{array} \right]^0 : \quad \text{oblate case}$$

---

<sup>20</sup> A. G. Webster, *The Dynamics of Particles and of Rigid Bodies* (2<sup>nd</sup> edition 1912; Dover reprint 1959), page 264; see also J. B. Hart, “Incorrect herpolhodes in textbooks,” *AJP* **37**, 1064 (1969).

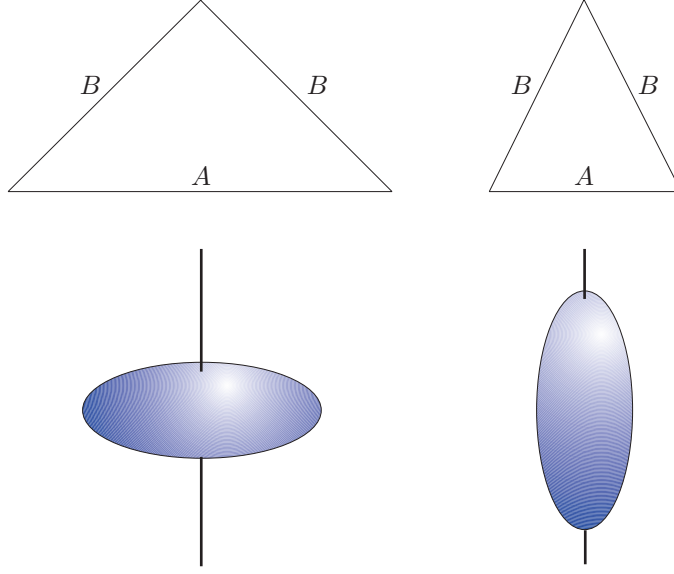


FIGURE 13: *Symmetric tops come in two flavors. At the top are triangle diagrams of the sort first encountered on page 8, and below are oblate/prolate Poinsot ellipsoids (figures of revolution) typical of that symmetry class.*

and

$$\begin{bmatrix} I_1 \dot{\omega}_1 - (I_1 - I_3) \omega_3 \omega_2 = 0 \\ I_1 \dot{\omega}_2 + (I_1 - I_3) \omega_3 \omega_1 = 0 \\ I_3 \dot{\omega}_3 = 0 \end{bmatrix}^0 \quad : \quad \text{prolate case}$$

which share the same abstract structure, together with

$$\begin{bmatrix} I_1 \dot{\omega}_1 = 0 \\ I_1 \dot{\omega}_2 = 0 \\ I_1 \dot{\omega}_3 = 0 \end{bmatrix}^0 \quad : \quad \text{spherical case}$$

which is trivial. Looking first to the former, we have

$$\begin{aligned} \omega_1^0(t) &= \lambda \quad : \quad \text{constant} \\ \begin{pmatrix} \dot{\omega}_2^0 \\ \dot{\omega}_3^0 \end{pmatrix} &= \begin{pmatrix} 0 & -\Omega \\ \Omega & 0 \end{pmatrix} \begin{pmatrix} \omega_2^0 \\ \omega_3^0 \end{pmatrix} \quad \text{with} \quad \Omega \equiv \frac{I_1 - I_3}{I_3} \lambda \end{aligned}$$

so

$$\dot{\omega}^0 = \mathbf{\Omega} \times \boldsymbol{\omega} \quad \text{with} \quad \mathbf{\Omega} \equiv \begin{pmatrix} \Omega \\ 0 \\ 0 \end{pmatrix}$$

of which the immediate solution is

$$\boldsymbol{\omega}^0(t) = e^{t\mathbb{W}}\boldsymbol{\omega}^0(0) \quad \text{with} \quad \mathbb{W} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\Omega \\ 0 & \Omega & 0 \end{pmatrix} = \boldsymbol{\Omega} \times \quad (67)$$

We conclude (in a phrase that describes not the motion of the top itself, but the motion of the vector  $\boldsymbol{\omega}^0(t)$  in  $\boldsymbol{\omega}^0$ -space) that

$\boldsymbol{\omega}^0(t)$  precesses about the  $\boldsymbol{\Omega}$ -vector with angular velocity  $\Omega$

For a prolate top the same line of argument gives

$$\begin{aligned} \dot{\boldsymbol{\omega}}^0 &= -\boldsymbol{\Omega} \times \boldsymbol{\omega} \quad \text{with} \quad \boldsymbol{\Omega} \equiv \begin{pmatrix} 0 \\ 0 \\ \Omega \end{pmatrix} \\ \Omega &\equiv \frac{I_1 - I_3}{I_1} \lambda \\ \lambda &\equiv \text{constant value of } \omega_3^0(t) \end{aligned}$$

—the conclusion being *similar to within a sign*. Reverting to the notation introduced at FIGURE 13, we have

$$\Omega = \begin{cases} \frac{A-B}{B} \lambda & \text{in the oblate case } (A > B): \text{ precession } \underline{\text{prograde}} \\ \frac{B-A}{B} \lambda & \text{in the prolate case } (B > A): \text{ precession } \underline{\text{retrograde}} \end{cases} \quad (68)$$

The situation is illustrated in FIGURE 14. For discussion of a geophysical instance of this kind of free rotor precession, and its relation to “Chandler wobble,” see Goldstein.<sup>21</sup>

Looking back to (59) on page 29, we see that axial symmetry of any type—be it oblate, spherical or prolate—immediately entails

$$\det \mathbb{J} = (I_1 - I_2)(I_1 - I_3)(I_2 - I_3) = (A - B)(A - C)(B - C) = 0$$

giving

$$\frac{d}{dt} \omega^2 = \frac{\det \mathbb{J}}{I_1 I_2 I_3} \omega_1 \omega_2 \omega_3 = 0 \quad : \quad \omega^2 \text{ is conserved} \quad (69)$$

For tops-in-general  $\boldsymbol{\omega}$  and  $\omega^2$  are *both* time-dependent (unless  $\boldsymbol{\omega}$  happens to be an eigenvector of  $\mathbb{J}$ , in which case both are constant: see again (54)), but for all symmetric tops  $\omega^2$  becomes constant, though  $\boldsymbol{\omega}$  typically continues to wander.

---

<sup>21</sup> H. Goldstein, *Classical Mechanics* (2<sup>nd</sup> edition 1980), page 212.



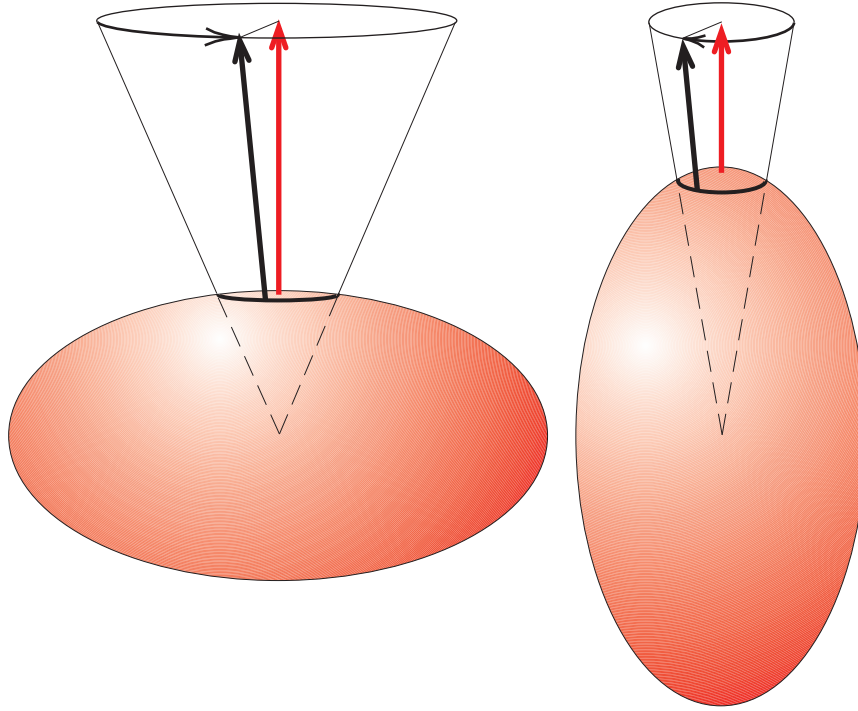


FIGURE 14:  $\boldsymbol{\omega}^0$  is seen in the body frame to precess prograde around the symmetry axis in the oblate case, retrograde in the prolate case. The component  $\boldsymbol{\omega}^0_{\parallel}$  of  $\boldsymbol{\omega}^0$  (shown here in red) that parallels the symmetry axis is stationary, while  $\boldsymbol{\omega}^0_{\perp}$  revolves with constant angular velocity  $\Omega$ . That the motion of  $\boldsymbol{\omega}^0$  is—for symmetric tops—length-preserving is seen in this light to be not at all surprising.

**12. Detailed account of the free gyration of a symmetric top.** Knowledge of  $\boldsymbol{\omega}^0(t)$ , or even of  $\boldsymbol{\omega}(t)$ , leaves one still an integration away from a description of the motion  $\mathbb{R}(t)$  of the physical top itself. To gain the latter kind of understanding we look to the motion of the Euler angles that serve to describe the relation of the body frame to the space frame.

We know that  $\boldsymbol{S}$  is conserved, and will (without real loss of generality) look to those motions with the property that  $\boldsymbol{S}$  is aligned with the 3-axis of the space frame:

$$\boldsymbol{S} = \begin{pmatrix} 0 \\ 0 \\ S \end{pmatrix} \quad (70)$$

It is also without loss of generality that we will identify the symmetry axis of the top with the 3<sup>0</sup>-axis of the body frame ... which is to say: we will stipulate

the symmetry of the top by setting  $I_1 = I_2$ . The Lagrangian then becomes (see again (47) on page 26)

$$\mathcal{L} = \frac{1}{2}B(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}A(\dot{\phi} \cos \theta + \dot{\psi})^2 \quad (71)$$

in which, it will be noticed, neither  $\phi$  nor  $\psi$  makes an appearance, so the conjugate momenta  $p_\phi$  and  $p_\psi$  are known already to be constant. Writing out the Lagrange equations, we have

$$\left. \begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} &= B\dot{\phi} \sin^2 \theta + A(\dot{\phi} \cos \theta + \dot{\psi}) \cos \theta = \text{constant} \\ \frac{\partial \mathcal{L}}{\partial \dot{\psi}} &= A(\dot{\phi} \cos \theta + \dot{\psi}) = \text{constant} \\ \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}} - \frac{\partial \mathcal{L}}{\partial \theta} &= B\ddot{\theta} - B\dot{\phi}^2 \sin \theta \cos \theta + A(\dot{\phi} \cos \theta + \dot{\psi})\dot{\phi} \sin \theta = 0 \end{aligned} \right\} \quad (72)$$

As a preliminary to discussion of the implications of (72) I digress to translate (70) into a statement relating Euler angles and their derivatives. We introduce (44) into

$$\mathbf{S}^0 = \begin{pmatrix} B & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & A \end{pmatrix} \boldsymbol{\omega}^0$$

and obtain

$$= \begin{pmatrix} B(\dot{\theta} \sin \phi \sin \psi + \dot{\theta} \cos \psi) \\ B(\dot{\theta} \sin \phi \cos \psi - \dot{\theta} \sin \psi) \\ A(\dot{\phi} \cos \theta + \dot{\psi}) \end{pmatrix}$$

Recalling now that  $\mathbf{S} = \mathbb{R} \mathbf{S}^0$  and taking  $\mathbb{R}$  from (41), we compute

$$\begin{aligned} \mathbf{S} &= \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \\ &= \begin{pmatrix} B(-\dot{\phi} \sin \theta \cos \theta \sin \phi + \dot{\theta} \cos \phi) + A(\dot{\phi} \cos \theta + \dot{\psi}) \sin \theta \sin \phi \\ B(+\dot{\phi} \sin \theta \cos \theta \cos \phi + \dot{\theta} \sin \phi) - A(\dot{\phi} \cos \theta + \dot{\psi}) \sin \theta \cos \phi \\ B\dot{\phi} \sin^2 \theta + A(\dot{\phi} \cos \theta + \dot{\psi}) \cos \theta \end{pmatrix} \end{aligned} \quad (73)$$

Evidently

$$S_1 \cos \phi + S_2 \sin \phi = B\dot{\theta}$$

But at (70) we set  $S_1 = S_2 = 0$ , so we have  $\dot{\theta} = 0$  giving

$$\theta(t) = \theta_0 \quad : \quad \text{constant} \quad (74)$$

It proves handy to note also that (taking  $\mathbb{R}^{-1}$  from (40.2))

$$\mathbf{S}^0 = \mathbb{R}^{-1} \begin{pmatrix} 0 \\ 0 \\ S \end{pmatrix} = \begin{pmatrix} S \sin \theta_0 \sin \psi \\ S \sin \theta_0 \cos \psi \\ S \cos \theta_0 \end{pmatrix}$$

From the last of the equations (73) we see that the first of the Lagrange equations (72) can be formulated

$$B\dot{\phi} \sin^2 \theta + A(\dot{\phi} \cos \theta + \dot{\psi}) \cos \theta = S$$

and when multiplied by  $\cos \theta$ , becomes

$$[B\dot{\phi} \cos \theta - A(\dot{\phi} \cos \theta + \dot{\psi})] \sin^2 \theta + A(\dot{\phi} \cos \theta + \dot{\psi}) = S \cos \theta$$

The third Lagrange equation, by  $\ddot{\theta} = 0$ , has become

$$[B\dot{\phi} \cos \theta - A(\dot{\phi} \cos \theta + \dot{\psi})] \dot{\phi} \sin \theta = 0$$

which when brought to the preceding equation tells us that the second Lagrange equation can be written

$$A(\dot{\phi} \cos \theta + \dot{\psi}) = S \cos \theta \quad (75)$$

The equation

$$[B\dot{\phi} \cos \theta - A(\dot{\phi} \cos \theta + \dot{\psi})] = 0$$

has therefore become  $[B\dot{\phi} - S] \cos \theta = 0$ , which supplies

$$\dot{\phi} = \frac{S}{B} \quad : \quad \text{constant} \quad (76)$$

Returning with this information to (75) we obtain

$$\dot{\psi} = \frac{B-A}{AB} S \cos \theta_0 \quad : \quad \text{constant} \quad (77)$$

The functions  $\phi(t)$  and  $\psi(t)$  are therefore *linear* in  $t$ , and the *exact solutions* of the Lagrange equations (72) can be presented

$$\left. \begin{aligned} \theta(t) &= \theta_0 \\ \phi(t) &= \Omega_\phi t + \phi_0 \quad \text{with} \quad \Omega_\phi \equiv \frac{S}{B} \\ \psi(t) &= \Omega_\psi t + \psi_0 \quad \text{with} \quad \Omega_\psi \equiv \frac{B-A}{AB} S \cos \theta_0 \end{aligned} \right\} \quad (78)$$

The fixed “angle of tilt”  $\theta_0$ —whence also  $\Omega_\phi$  and  $\Omega_\psi$ —is determined by joint specification of  $S$  and  $\mathcal{T}$  (together with  $A$  and  $B$ , which describe the

effective figure of the symmetric top), as I now show: Returning to (71) with descriptions of  $B\dot{\phi}$  and  $A(\dot{\phi}\cos\theta + \dot{\psi})$  that were developed just above, we obtain

$$\mathcal{T} = \mathcal{L} = \frac{1}{2} \left\{ \frac{1}{B} \sin^2 \theta_0 + \frac{1}{A} \cos^2 \theta_0 \right\} S^2 \quad (79.1)$$

$$\begin{aligned} &= \frac{1}{2A} S^2 - \frac{1}{2} \left( \frac{1}{A} - \frac{1}{B} \right) S^2 \sin^2 \theta_0 \\ &= \frac{1}{2B} S^2 + \frac{1}{2} \left( \frac{1}{A} - \frac{1}{B} \right) S^2 \cos^2 \theta_0 \end{aligned} \quad (79.2)$$

giving

$$\tan \theta_0 = \sqrt{\left( \frac{1}{A} - \frac{2\mathcal{T}}{S^2} \right) / \left( \frac{2\mathcal{T}}{S^2} - \frac{1}{B} \right)} \quad (80)$$

It is gratifying to note in this connection that (61) supplies

$$\begin{aligned} \frac{1}{B} &\leq \frac{2\mathcal{T}}{S^2} \leq \frac{1}{A} & : & \text{oblate case} \\ \frac{1}{A} &\leq \frac{2\mathcal{T}}{S^2} \leq \frac{1}{B} & : & \text{prolate case} \end{aligned}$$

so the parenthetic expressions under the radical have in all cases the same sign: we are at (80) never asked to take the square root of a negative number. Making use now of  $\cos^2 = 1/\sqrt{1 + \tan^2}$  we obtain finally

$$\begin{aligned} \Omega_\psi &= \frac{B-A}{AB} S \sqrt{\left( \frac{2\mathcal{T}}{S^2} - \frac{1}{B} \right) / \left( \frac{1}{A} - \frac{1}{B} \right)} \\ &= S \sqrt{\left( \frac{1}{A} - \frac{1}{B} \right) \left( \frac{2\mathcal{T}}{S^2} - \frac{1}{B} \right)} \end{aligned} \quad (81)$$

where again—for reasons just stated—the expression under the radical is in all cases non-negative.

I enter now upon a series of elementary remarks that culminate in a celebrated *geometrical interpretation* of the  $\mathbb{R}(t)$  implicit in (78):

- From  $\mathcal{T} = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{S} = \frac{1}{2} \omega S \cos \alpha$  and the established facts that for a symmetric free top not only  $\mathcal{T}$  and  $S$  but also  $\omega$  are constants of the motion, we see that for such a top the angle  $\mathbf{S} \angle \boldsymbol{\omega}$  (we have named it  $\alpha$ ) is *invariant*. We have

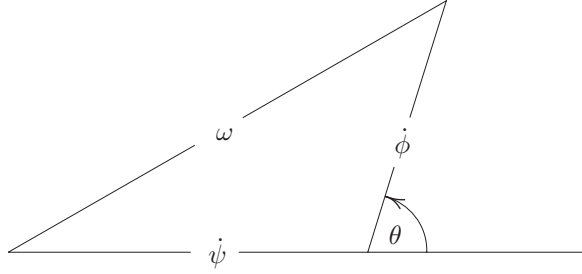
$$\cos \alpha = \frac{2\mathcal{T}/S}{\omega} \quad (82)$$

which in the notation of FIGURE 11 becomes simply  $\cos \alpha = d/\omega$

- Bringing  $\dot{\theta} = 0$  to the description (42) of  $\boldsymbol{\omega}$ , we find

$$\omega^2 = \dot{\psi}^2 + 2\dot{\psi}\dot{\phi}\cos\theta + \dot{\phi}^2 \quad (83)$$

which invites the diagrammatic representation shown in the following figure:



It follows in any event from (78) that

$$\begin{aligned}
 \omega^2 &= \Omega_{\dot{\psi}}^2 + 2\Omega_{\dot{\psi}}\Omega_{\dot{\phi}}\cos\theta_0 + \Omega_{\dot{\phi}}^2 \\
 &= \left\{ \left[ \left( \frac{1}{A} - \frac{1}{B} \right)^2 + 2\left( \frac{1}{A} - \frac{1}{B} \right) \frac{1}{B} \right] \cos^2\theta_0 + \left( \frac{1}{B} \right)^2 \right\} S^2 \\
 &= \left\{ \frac{1}{A^2} \cos^2\theta_0 + \frac{1}{B^2} \sin^2\theta_0 \right\} S^2
 \end{aligned} \tag{84}$$

- Returning with this information and (79) to (82) we have

$$\cos\alpha = \frac{\frac{1}{A} \cos^2\theta_0 + \frac{1}{B} \sin^2\theta_0}{\sqrt{\frac{1}{A^2} \cos^2\theta_0 + \frac{1}{B^2} \sin^2\theta_0}} \tag{85}$$

Evidently  $\alpha \rightarrow 0$  as  $A \rightarrow B$

- The angle  $\omega\angle$ (symmetry axis)—call it  $\beta$ —can be obtained from

$$\begin{aligned}
 \cos\beta &= \frac{\omega_3^0}{\omega} \\
 &= \frac{\dot{\phi} \cos\theta_0 + \dot{\psi}}{\omega} && \text{by (44)} \\
 &= \frac{\frac{S}{A} \cos\theta_0}{S \sqrt{\frac{1}{A^2} \cos^2\theta_0 + \frac{1}{B^2} \sin^2\theta_0}} && \text{by (75) and (84)} \\
 &= \frac{1}{\sqrt{1 + \left( \frac{A}{B} \right)^2 \tan^2\theta_0}}
 \end{aligned} \tag{86}$$

Drawing again upon  $\cos = 1/\sqrt{1 + \tan^2}$ , we have

$$\tan\beta = \pm \frac{A}{B} \tan\theta_0 \tag{87}$$

We have learned that  $\beta$  is (like  $\alpha$ ) a *dynamical invariant*. Evidently  $\beta \rightarrow \theta_0$  as  $A \uparrow B$ .

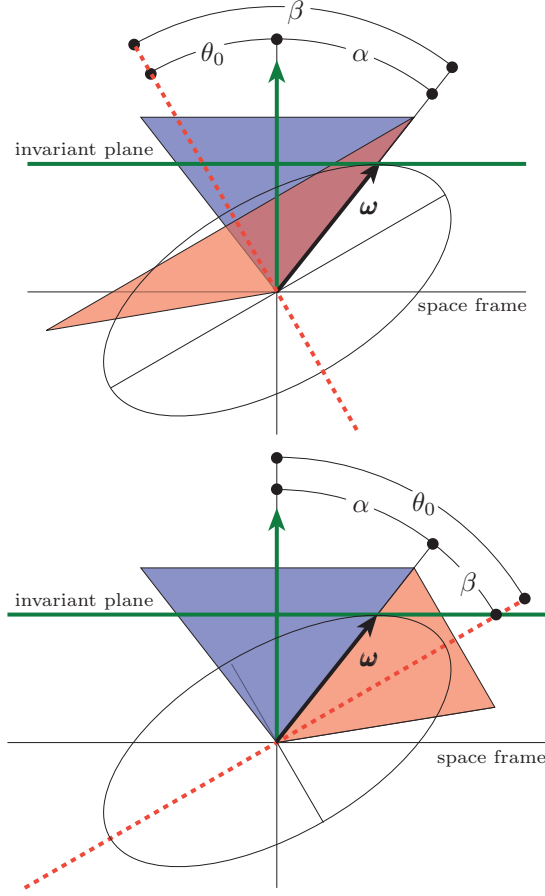


FIGURE 15: Two-dimensional sections of figures drawn in  $\omega$ -space, where the colored triangles become cones. Conserved  $\mathbf{S}$  sets the orientation of the invariant plane,  $\mathcal{T}$  sets its distance above the origin. The dashed line identifies the symmetry axis, so the top figure refers to the OBLATE case, the lower figure to the PROLATE case. The blue “space cone” is dynamically invariant, the red “body cone”—interior to the space cone in the oblate case, exterior in the prolate case—is fixed with respect to the top, and its motion (“rolling-without-slipping,” as described in the text) provides a representation of the motion of the physical top.

- From the figure, which was designed to make plain the *meanings* of the invariant angles  $\theta_0$ ,  $\alpha$  and  $\beta$ , we read

$$\theta_0 = \begin{cases} \beta - \alpha & : \text{ OBLATE cases} \\ \beta + \alpha & : \text{ PROLATE cases} \end{cases} \quad (88)$$

In either case we have

$$\begin{aligned}\cos \alpha &= \cos(\theta_0 - \beta) = \cos \beta \cos \theta_0 (1 + \tan \beta \tan \theta_0) \\ &= \cos \beta \cos \theta_0 (1 + \frac{A}{B} \tan^2 \theta_0) \quad \text{by (87)}\end{aligned}$$

where the minus sign in (87) has been dismissed as an artifact. It is gratifying to notice that if we draw upon (86) we obtain

$$= \cos \theta_0 \frac{(1 + \frac{A}{B} \tan^2 \theta_0)}{\sqrt{(1 + (\frac{A}{B})^2 \tan^2 \theta_0)}}$$

from which, after simplifications, we are led back again to precisely (85).

- At (79.2) we had an equation which by (76) and (77) can be written

$$2\mathcal{T}/S = \dot{\phi} + \dot{\psi} \cos \theta_0$$

so the equation (82) at which we introduced  $\alpha$  becomes

$$\cos \alpha = \frac{\dot{\phi} + \dot{\psi} \cos \theta_0}{\omega} \quad (89.1)$$

This equation is structurally reminiscent of an equation

$$\cos \beta = \frac{\dot{\psi} + \dot{\phi} \cos \theta_0}{\omega} \quad (89.2)$$

encountered in the derivation of (86). Looking to the square of (89.1) and drawing upon the description (83) of  $\omega^2$ , we have

$$\begin{aligned}\cos^2 \alpha &= \frac{\dot{\phi}^2 + 2\dot{\phi}\dot{\psi} \cos \theta_0 + \dot{\psi}^2 [1 - \sin^2 \theta_0]}{\dot{\psi}^2 + 2\dot{\psi}\dot{\phi} \cos \theta_0 + \dot{\phi}^2} \\ &= 1 - \frac{\dot{\psi}^2 \sin^2 \theta_0}{\omega^2} \implies \dot{\psi}^2 = \omega^2 \frac{\sin^2 \alpha}{\sin^2 \theta_0}\end{aligned} \quad (90.1)$$

while by a similar argument

$$\dot{\phi}^2 = \omega^2 \frac{\sin^2 \beta}{\sin^2 \theta_0} \quad (90.2)$$

- Evidently

$$\dot{\phi} \sin \alpha = \dot{\psi} \sin \beta \quad i.e., \quad \Omega_\phi \sin \alpha = \Omega_\psi \sin \beta \quad (91)$$

which provides the basis for the claim—developed in the following figures—that *the body cone (with vertex angle  $\beta$ ) rolls without slipping on the stationary space cone (vertex angle  $\alpha$ )*.

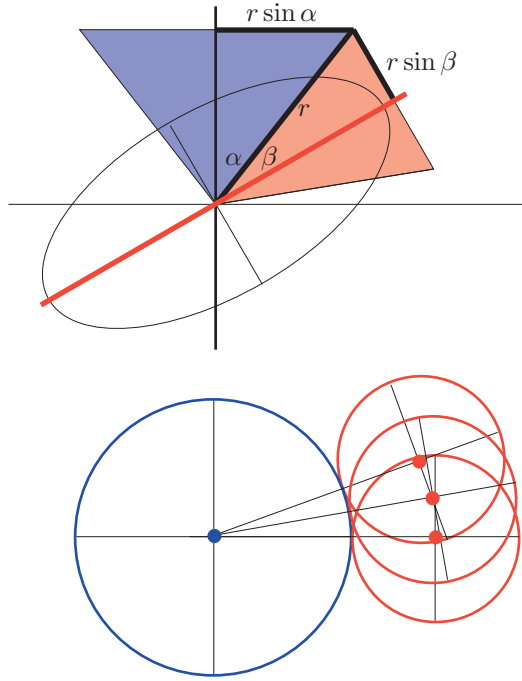


FIGURE 16: Attempts to represent the **body cone** rolling-without-slipping on the **space cone** in the PROLATE case. Rotation of the space cone through angle  $d\phi$  about the verticle **S** axis will mesh with rotation of the body cone through angle  $d\psi$  only if

$$d\phi \cdot r \sin \alpha = d\psi \cdot r \sin \beta$$

But that is precisely the upshot of (91). The lower figure provides another representation of the same principle, and makes clear the fact that  $\odot$  rotation of the body cone causes its center to advance around the space cone in that same  $\odot$  sense. The red axis in the upper figure represents the **symmetry axis** of the top.



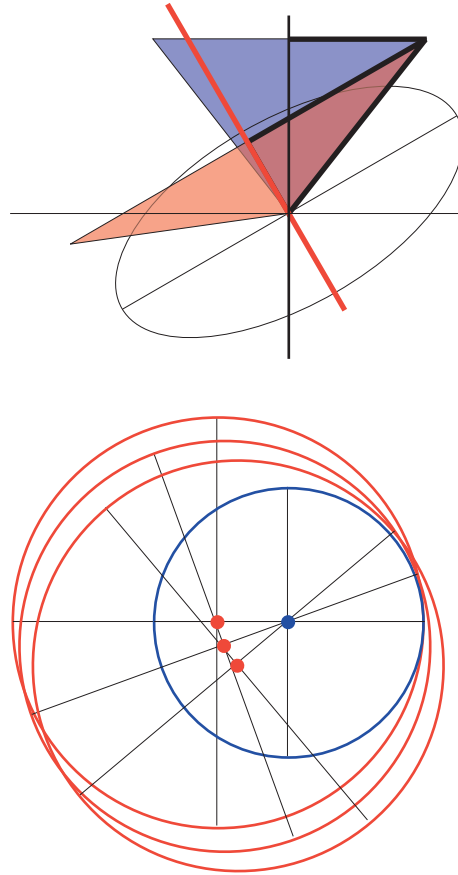


FIGURE 17: *Preceding constructions, here adapted to the OBLATE case. Note that, while the body cone lies exterior to the space cone the prolate case, it envelops the space cone in the oblate case. Note also that—here as before— $\odot$  rotation of the body cone produces a same-sense  $\odot$ -advance of its center.*

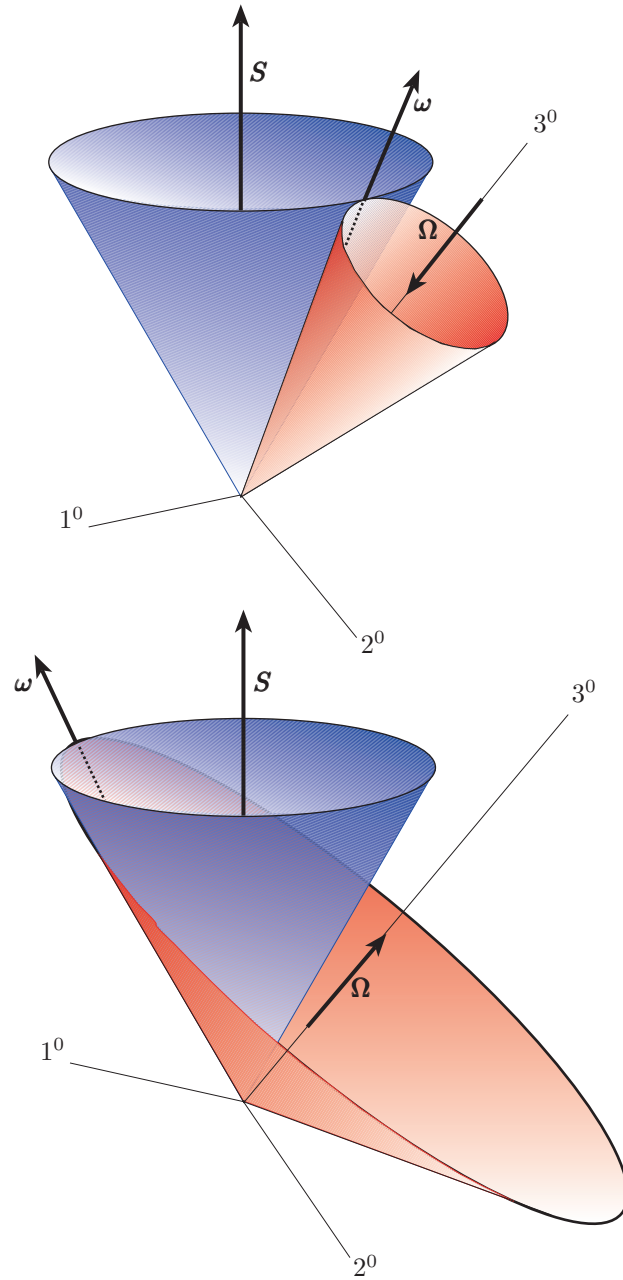
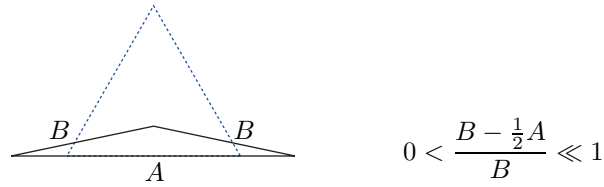


FIGURE 18: *Alternative representations of the body cone rolling without slipping on the space cone (prolate case above, oblate case below). The rolling body cone controls the motion of the symmetry axis of the body—the  $3^0$ -axis. The uniform rotation (about that axis) is controlled by the  $\Omega$ -vector that was introduced on page 40: it is retrograde in the prolate case, prograde in the oblate case.*

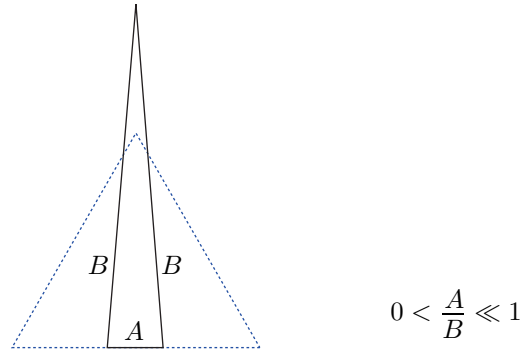
The general results developed above can (in leading approximation) be simplified in limiting special cases. One can, for example, readily imagine having geophysical/astrophysical interest in *slightly oblate* symmetrical bodies:



Flipped coins inspire interest in the gyroynamics of *highly oblate* tops



while tumbling needles are, in effect, *highly prolate* tops



I invite the reader to construct variants of figures 16–18 appropriate to those cases.

**13. Instability of spin about the intermediate axis.**<sup>22</sup> Let the Euler equations (35) of a *free asymmetric* top be written

$$\begin{bmatrix} I_1 \dot{\omega}_1 - (I_2 - I_3) \omega_2 \omega_3 = 0 \\ I_2 \dot{\omega}_2 - (I_3 - I_1) \omega_3 \omega_1 = 0 \\ I_1 \dot{\omega}_3 - (I_1 - I_2) \omega_1 \omega_2 = 0 \end{bmatrix}^0$$

<sup>22</sup> My primary source here has been the discussion presented by S. T. Thornton & J. B. Marion in §11.12 of their *Classical Dynamics of Particles and Systems*, (5<sup>th</sup> edition 2004).

Assume  $I_1 > I_2 > I_3$  and agree for the purposes of this discussion to omit the  $^0$ s. Write

$$\boldsymbol{\omega} = \begin{pmatrix} \omega_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \quad : \quad \epsilon_2 \text{ and } \epsilon_3 \text{ infinitesimal}$$

to express our presumption that  $\boldsymbol{\omega}$  lies in the immediate neighborhood of the principal axis of greatest moment. In leading order (*i.e.*, after abandoning terms of 2<sup>nd</sup> order) we have

$$\begin{aligned} I_1 \dot{\omega}_1 &= 0 \quad \Rightarrow \quad \omega_1 \text{ is constant} \\ I_2 \dot{\epsilon}_2 - (I_3 - I_1) \omega_1 \epsilon_3 &= 0 \\ I_1 \dot{\epsilon}_3 - (I_1 - I_2) \omega_1 \epsilon_2 &= 0 \end{aligned}$$

The last pair of equations can be “separated by differentiation.” One is led to the conclusion that  $\epsilon_1$  and  $\epsilon_2$  are *both* solutions of an equation

$$\ddot{\chi} = -\omega_1^2 \frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3} \chi$$

of which the general solution is

$$\chi(t) = P e^{+i\Omega_1 t} + Q e^{-i\Omega_1 t} \quad \text{with} \quad \Omega_1 \equiv \omega_1 \sqrt{\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3}}$$

Had we proceeded from

$$\boldsymbol{\omega} = \begin{pmatrix} \epsilon_1 \\ \omega_2 \\ \epsilon_3 \end{pmatrix} \quad : \quad \epsilon_1 \text{ and } \epsilon_3 \text{ infinitesimal}$$

we would have been led by a similar argument to the conclusion that  $\epsilon_1$  and  $\epsilon_3$  both move like

$$\chi(t) = P e^{+i\Omega_2 t} + Q e^{-i\Omega_2 t} \quad \text{with} \quad \Omega_2 \equiv \omega_2 \sqrt{\frac{(I_2 - I_3)(I_2 - I_1)}{I_3 I_1}}$$

while

$$\boldsymbol{\omega} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \omega_3 \end{pmatrix} \quad : \quad \epsilon_1 \text{ and } \epsilon_2 \text{ infinitesimal}$$

leads to

$$\chi(t) = P e^{+i\Omega_3 t} + Q e^{-i\Omega_3 t} \quad \text{with} \quad \Omega_3 \equiv \omega_3 \sqrt{\frac{(I_3 - I_1)(I_3 - I_2)}{I_1 I_2}}$$

The point to notice is that

$$I_1 > I_2 > I_3 \quad \Rightarrow \quad \Omega_1 \text{ and } \Omega_3 \text{ are real, but } \Omega_2 \text{ is imaginary}$$

The implication is that

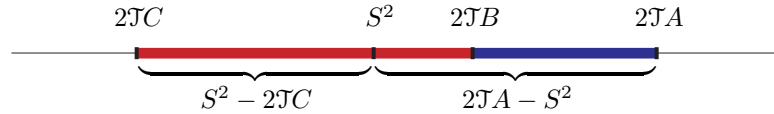
- $\omega^0$ , if initially nearly parallel to either the principal axis of greatest moment or the principal axis of least moment, moves like a 2-dimensional oscillator, tracing and retracing an ellipse: spin about either of those axes is stable;
- $\omega^0$ , if initially nearly parallel to the principal axis of intermediate moment, moves away from that neighborhood with exponential haste: *spin about the intermediate axis is unstable*.

This analytical result conforms very nicely to the lesson latent in figures 10b, 10c & 10d (though those relate to how  $\mathbf{S}^0$  wanders in  $\mathbf{S}$ -space, while we have been working here in  $\omega^0$ -space).

If we write  $\tau_i$  to denote the time it takes  $\omega^0$  to complete a circuit in the immediate neighborhood of the  $i^{\text{th}}$  principal axis, then we have

$$\left. \begin{aligned} \tau_1 &= 2\pi/\Omega_1 \\ \tau_2 &= \text{undefined} \\ \tau_3 &= 2\pi/\Omega_3 \end{aligned} \right\} \quad (92)$$

where  $\tau_2$  is “undefined” because  $\omega^0$  *does not remain confined* to neighborhoods that contain the intermediate axis: it cyclically departs from the neighborhood, visits its antipode, returns. Such excursions are actually remote tours around either the axis of greatest moment or the axis of least moment. It becomes of interest, therefore, to discover how to describe the the periods of such “remote excursions.” The problem is addressed in §150a of Routh,<sup>16</sup> who builds upon Kirchhoff’s account of Jacobi’s discovery—anticipated by Euler—that the **general solution of the asymmetric top problem** can be developed in terms of elliptic functions. Here I must be content merely to discuss some properties and implications of the results reported by Routh. Writing  $A > B > C$  in place of  $I_1 > I_2 > I_3$ , we will consider  $\mathcal{T}$  to be given/fixed, and  $S^2$  to range on what we discovered at (61) to be the physically allowed interval  $2\mathcal{T}C \leq S^2 \leq 2\mathcal{T}A$ :



Circulation in the immediate neighborhood of the axis of greatest moment requires that  $S^2$  lie very near the right end of the blue region in the preceding figure. At more remote blue points the circulation about that axis the period is reported by Routh to be given by

$$\tau_1 = 4\sqrt{\frac{ABC}{(A-B)(S^2-2\mathcal{T}C)}} \int_0^{\frac{1}{2}\pi} \frac{1}{\sqrt{1-k^2 \sin^2 \phi}} d\phi \quad (93)$$

with

$$k^2 = \frac{(B-C)(2\mathcal{T}A-S^2)}{(A-B)(S^2-2\mathcal{T}C)}$$

*Mathematica* recognizes the  $\int$  integral to be just the *complete elliptic integral* `EllipticK[k^2]`. If we construe  $\tau_1$  to be a function of  $S^2$  and, in order to learn

the value assumed by  $\tau_1(S^2)$  as  $S^2 \uparrow 2\mathcal{T}A$ , ask *Mathematica* to develop  $\tau_1(2\mathcal{T}A - x^2)$  as a power series in  $x$ , we obtain

$$\tau_1(2\mathcal{T}A - \sigma^2) = 2\pi \sqrt{\frac{ABC}{(A-B)(A-C)2\mathcal{T}}} + (\text{complicated term})\sigma^2 + \dots$$

which (since in leading order  $2\mathcal{T} = A\omega_1^2$ ) can be written

$$= 2\pi/\Omega_1 + \dots$$

with

$$\Omega_1 \equiv \omega_1 \sqrt{\frac{(A-B)(A-C)}{BC}}$$

—in precise agreement with the result obtained by simpler means on page 52.

Formulae appropriate to cases in which  $S^2$  falls on the red interval in the preceding figure (cases, that is to say, in which  $2\mathcal{T}C \leq S^2 < 2\mathcal{T}B$  and  $\omega^0$  circulates about the principal axis of least moment) can be obtained from the preceding formulae by  $A \rightleftharpoons C$  interchange.

Return now to (93) and set  $S^2 = 2\mathcal{T}B$  to obtain

$$\tau_1 \rightarrow \tau_2 = 4 \sqrt{\frac{ABC}{(A-B)(B-C)2\mathcal{T}}} \int_0^{\frac{1}{2}\pi} \frac{1}{\sqrt{1 - p^2 \sin^2 \phi}} d\phi \quad (94)$$

with

$$p^2 = \frac{(B-C)(A-B)}{(A-B)(B-C)} = 1$$

This result is invariant under  $A \rightleftharpoons C$  (we have therefore  $\tau_1 \rightarrow \tau_2 \leftarrow \tau_3$ , which is gratifying), and—since the  $\int$  diverges—leads to the conclusion that  $\tau_2 = \infty$ : a top set spinning about the intermediate axis does not wander. But that situation is, as we have recently established, unstable: if the alignment is not absolutely precise the top *does* wander, and we have interest in computing the period in such more realistic cases.

To signal our intent to approach  $S_{\text{critical}}^2 = 2\mathcal{T}B$  from above, we return to (93)—which we agree now to notate

$$\tau_1(S^2) = F_1(S^2) \cdot \text{EllipticK}[p_1^2(S^2)]$$

—and set  $S^2 = 2\mathcal{T}B + \sigma^2$ . *Mathematica* supplies

$$\begin{aligned} F_1(2\mathcal{T}B + \sigma^2) &= 4 \sqrt{\frac{ABC}{2\mathcal{T}(A-B)(B-C)}} \left\{ 1 - \frac{1}{\mathcal{T}(B-C)} \sigma^2 + \frac{3}{8} \frac{1}{[\mathcal{T}^2(B-C)]^2} \sigma^4 - \dots \right\} \\ p_1^2(2\mathcal{T}B + \sigma^2) &= 1 - \frac{A-C}{(A-B)[2\mathcal{T}(B-C)]} \sigma^2 + \frac{A-C}{(A-B)[2\mathcal{T}(B-C)]^2} \sigma^4 - \dots \\ &\equiv 1 - q^2(\sigma^2) \end{aligned}$$

Mathematical handbooks<sup>23</sup> supply moreover the information that

$$\text{EllipticK}[p^2] \equiv \text{EllipticK}[1 - q^2]$$

can, for small values of  $q^2$ , be developed

$$= Q + \frac{1}{4}(Q - 1)q^2 + \frac{9}{64}(Q - \frac{7}{6})q^4 + \frac{25}{256}(Q - \frac{37}{30})q^6 + \dots$$

where

$$Q \equiv \log(4/q) \quad : \quad \text{blows up logarithmically as } q \downarrow 0$$

The results now in hand could be used to compute  $\tau_1(S_{\text{critical}}^2 + \sigma^2)$ , and by  $A \rightleftharpoons C$  one could without labor obtain a description of  $\tau_3(S_{\text{critical}}^2 - \sigma^2)$ . For more detailed discussion see GYRODYNAMICS (1976/77), pages 139–144 or the previously cited paper by W. G. Harter & C. C. Kim.<sup>17</sup> But to obtain a good qualitative understanding of the situation it is, I think, most instructive to proceed not analytically but graphically: letting  $\tau_1(S^2; A, B, C, \mathcal{T})$  denote the expression that appears on the right side of (93), we define

$$\tau(S^2; A, B, C, \mathcal{T}) \equiv \begin{cases} \tau_1(S^2; A, B, C, \mathcal{T}) & : \quad 2\mathcal{T}B < S^2 \leq 2\mathcal{T}A \\ \tau_1(S^2; C, B, A, \mathcal{T}) & : \quad 2\mathcal{T}C \leq S^2 < 2\mathcal{T}B \end{cases}$$

Suppose, for example, we set  $\mathcal{T} = 1$  and assign to the principal moments the values  $A = 4$ ,  $B = \frac{8}{3}$  and  $C = 2$  that were used to construct the figures on pages 32–34: then  $2\mathcal{T}A = 8$ ,  $2\mathcal{T}B = \frac{16}{3} = 5.333$ ,  $2\mathcal{T}C = 4$  and *Mathematica* constructs the graph of  $\tau(S^2; 4, \frac{16}{3}, 2, 1)$  presented here as FIGURE 19.

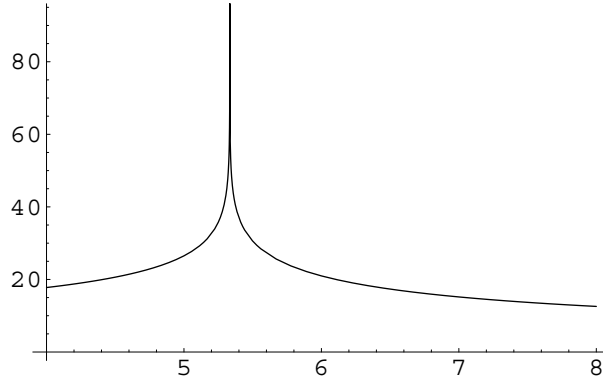


FIGURE 19: Graph of  $\tau(S^2)$  in a typical case.  $\omega^0$  circulates around the axis of least moment on the left side of the spike, the axis of greatest moment on the right. The spike is situated at  $S_{\text{critical}}^2$ .

<sup>23</sup> See E. Jahnke & F. Emde, *Tables of Functions* (1945), page 73; J. Spanier & K. B. Oldham, *An Atlas of Functions* (1987), page 612.

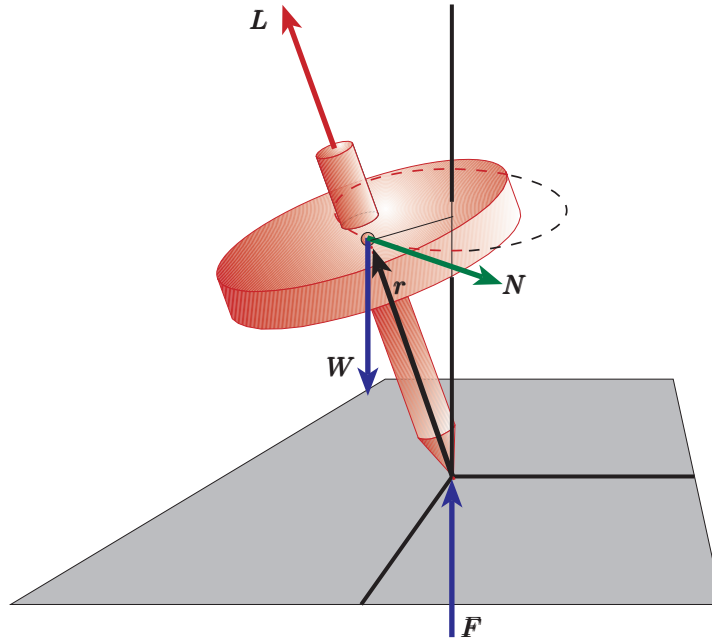


FIGURE 20: *Diagram of the sort standard to elementary discussions of the precession exhibited by spun toy tops. Not shown is the horizontal component of the support force. We will take  $\theta$  to be the “angle of tilt,” and will use  $\phi$  to describe “twirl about the vertical space-axis.”*

**14. Symmetric tops with a fixed point.** Astrophysical reality presents many examples of semi-rigid bodies rotating semi-freely in inertial space, but we are seldom inclined to call such objects “tops,” as has been my practice. The tops of playroom experience derive much of their fascination from the fact that—inevitably—they are spun in uniform gravitational fields, and are supported ... which is to say: they are *not* free.

Discussions of the precession of tops such as are found in introductory textbooks<sup>24</sup> standardly proceed from diagrams resembling FIGURE 20. Writing

$$\mathbf{F} = \mathbf{F}_{\text{vertical}} + \mathbf{F}_{\text{horizontal}}$$

one claims (though it is generally untrue!) that  $\mathbf{F}_{\text{vertical}} + \mathbf{W} = \mathbf{0}$ , and cleverly circumvents the awkward fact that  $\mathbf{F}_{\text{horizontal}}$  is generally unknown by taking the contact point to be the reference point with respect to which all torques and angular momenta will be defined: with respect to that point  $\mathbf{F}$  gives rise

<sup>24</sup> See, for example, Paul A. Tipler, *Physics for Scientists and Engineers* (3<sup>rd</sup> edition 1990), §8-8; Richard Wolfson & Jay M. Pasachoff, *Physics* (1990), pages 297–299.



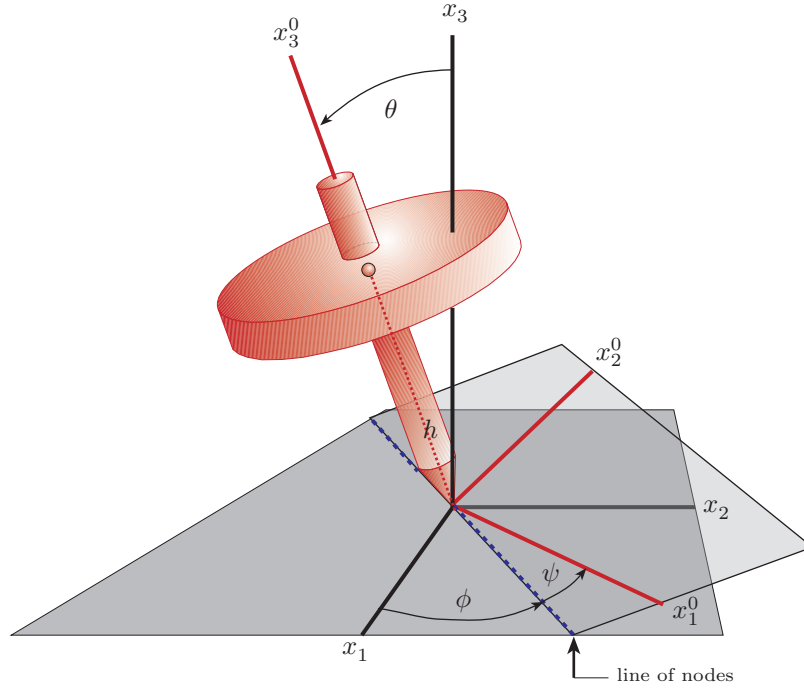


FIGURE 21: *Indication of the meanings of the variables employed in a more careful treatment of the problem.*

to no torque, and becomes effectively irrelevant. From  $\mathbf{N} = \mathbf{r} \times \mathbf{W}$  we obtain

$$\mathbf{N}_{\text{vertical}} = \mathbf{0} \quad \text{and} \quad \mathbf{N}_{\text{horizontal}} = mgr \sin \theta \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}$$

From  $\mathbf{N} = \dot{\mathbf{L}}$  we learn that  $\mathbf{L}_{\text{vertical}}$  is conserved, while from

$$\mathbf{L}_{\text{horizontal}} = L \sin \theta \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix}$$

we get

$$\dot{\mathbf{L}}_{\text{horizontal}} = L \dot{\phi} \sin \theta \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}$$

The implication appears to be that the top will *precess* with angular frequency

$$\dot{\phi} = \frac{mgr}{L}$$

The physical fact of the matter is, however, that tops exhibit much more complicated kinds of motion that this simple theory leads one to anticipate!

A full-blown dynamical theory of tops must account both for the motion *of* the center of mass and rotation *about* the center of mass, and might therefore appear (Chasle's theorem: page 3) to entail that we keep track of six variables.

We require, however, that the inertial coordinates of one point—the point of support—be fixed. That requirement imposes upon our six variables three holonomic constraints. Our system has, therefore, only three degrees of freedom, all of which refer to rotations about the support point. It becomes in this light natural to take as generalized coordinates the Euler angles that (FIGURE 21) relate

- an inertial Cartesian SPACE FRAME erected at the support point to
- a TRANSLATED COPY OF THE PRINCIPAL AXIS FRAME of the top.

This done, a slight modification of the argument that gave (71) gives

$$\mathcal{L} = \frac{1}{2}B(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}A(\dot{\phi} \cos \theta + \dot{\psi})^2 - mgh \cos \theta \quad (95)$$

where  $m$  refers to the total mass of the top, where the final term on the right is a potential energy term that was absent from the theory of free rigid rotators, and where it is to be understood that in the present instance<sup>25</sup>

$$A = (\text{center of mass value})$$

$$B = (\text{center of mass value}) + mh^2$$

In place of the equations of motion (72) we now have

$$\left. \begin{aligned} p_\phi &= B\dot{\phi} \sin^2 \theta + A(\dot{\phi} \cos \theta + \dot{\psi}) \cos \theta = \text{constant} \\ p_\psi &= A(\dot{\phi} \cos \theta + \dot{\psi}) = \text{constant} \\ \underbrace{B\ddot{\theta} - B\dot{\phi}^2 \sin \theta \cos \theta + A(\dot{\phi} \cos \theta + \dot{\psi})\dot{\phi} \sin \theta}_{\dot{p}_\theta} - mgh \sin \theta &= 0 \end{aligned} \right\} \quad (96)$$

where  $p_\phi$ ,  $p_\psi$  and  $p_\theta$  are (angular) momenta conjugate to the angles  $\phi$ ,  $\psi$  and  $\theta$ . From the second of the preceding equations we get

$$\dot{\psi} = \frac{p_\psi - A\dot{\phi} \cos \theta}{A}$$

which when brought to the first equation gives

$$\dot{\phi} = \frac{p_\phi - p_\psi \cos \theta}{B \sin^2 \theta} \quad (97.1)$$

whence

$$\dot{\psi} = \frac{p_\psi}{A} - \frac{p_\phi - p_\psi \cos \theta}{B \sin^2 \theta} \cos \theta \quad (97.2)$$

---

<sup>25</sup> This follows directly from (14) if one makes the replacements

$$\begin{aligned} r_1 &\mapsto r_1 \\ r_2 &\mapsto r_2 \\ r_3 &\mapsto r_3 + h \end{aligned}$$

and uses (3) to eliminate terms of the form  $\int \mathbf{r} h \rho(\boldsymbol{\rho}) dr_1 dr_2 dr_3$ .

Note that if  $\theta(t)$  were known then we could in principle use (97) to figure out  $\phi(t)$  and  $\psi(t)$ .

From the design (95) of  $\mathcal{L}$  it follows that total energy of the spinning top is conserved:

$$E = \frac{1}{2}B(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + \frac{1}{2}A(\dot{\phi} \cos \theta + \dot{\psi})^2 + mgh \cos \theta = \text{constant} \quad (98)$$

But  $\frac{1}{2}A(\dot{\phi} \cos \theta + \dot{\psi})^2 = \frac{1}{2}p_\psi^2/A$  was seen at (96) to be conserved all by itself, so we have conservation of

$$\mathcal{E} \equiv \frac{1}{2}B(\dot{\phi}^2 \sin^2 \theta + \dot{\theta}^2) + mgh \cos \theta$$

which upon elimination of  $\dot{\phi}$  becomes

$$= \frac{1}{2}B\dot{\theta}^2 + \underbrace{\frac{(p_\phi - p_\psi \cos \theta)^2}{2B \sin^2 \theta}}_{\text{“effective potential,” call it } \mathcal{V}(\theta)} + mgh \cos \theta \quad (99)$$

At this point it becomes natural to mimic methods borrowed from the mechanics of one-dimensional conservative systems, writing (for example)

$$\begin{aligned} \frac{d\theta}{dt} &= \sqrt{\frac{2}{B}[\mathcal{E} - \mathcal{V}(\theta)]} \\ \Downarrow \\ \text{transit time } \theta' \rightarrow \theta'' &= \int_{\theta'}^{\theta''} \frac{1}{\sqrt{\frac{2}{B}[\mathcal{E} - \mathcal{V}(\vartheta)]}} d\vartheta \end{aligned}$$

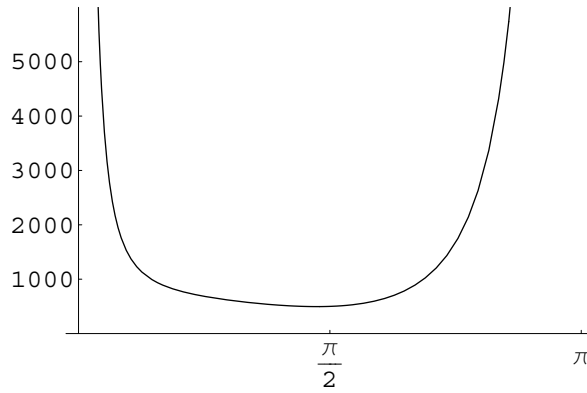


FIGURE 22: Graph of the effective potential  $\mathcal{V}(\theta)$  in the case  $p_\phi = 5$ ,  $p_\psi = 3$ ,  $2B = 0.05$  and  $mgh = 500$ . It is claimed not that the numbers are (or are not) physically reasonable, only that the figure is qualitatively typical.

For many purposes it is, however, very useful to notice that (99) can, by a change of variables  $\theta \mapsto u \equiv \cos \theta$ , be brought to the form

$$\begin{aligned}\dot{u}^2 &= (\alpha - \beta u)(1 - u^2) - (p - qu)^2 \\ &= f(u) \quad : \quad \text{cubic in } u\end{aligned}\tag{100}$$

and to proceed under the presumption that physically self-consistent values have been assigned to  $\alpha \equiv 2\mathcal{E}/B$ ,  $\beta \equiv 2mgh/B$ ,  $p \equiv p_\phi/B$  and  $q \equiv p_\psi/B$ . From

$$f(u) \sim \beta u^3 \quad \text{for } u \text{ large}$$

from

$$f(\pm 1) = -(p \mp q)^2 < 0$$

(we agree to exclude temporarily the exceptional cases  $p \mp q = 0$ ) and from the fact that for our results to admit of physical interpretation it must be the case that  $-\pi < \theta < \pi$  ( $-1 < u < +1$ ), we conclude that in physically realistic cases  $f(u)$  must be of the form graphed in the following figure:

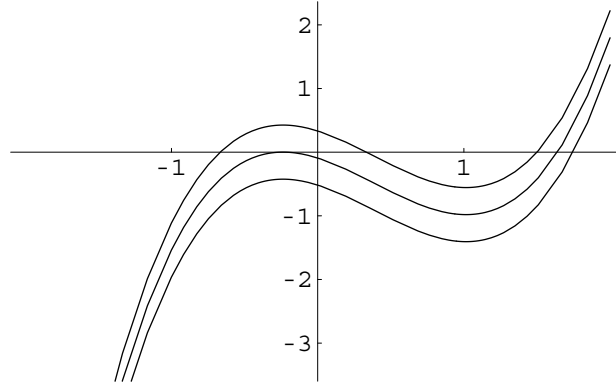


FIGURE 23: Graphical demonstration that the largest root of  $f(u)$ —call it  $u_3$ —must necessarily be unphysical:  $u_3 \equiv \cos \theta_3 > 1$  (unless, exceptionally,  $u_2 = u_3 = 1$ ). In all physically possible circumstances the other roots ( $u_1$  and  $u_2 \geq u_1$ ) must lie on the interval  $[-1, +1]$ . From (100), i.e., from  $f(u) = \dot{u}^2 \geq 0$ , we learn that only  $u$ -values on the interval  $u_1 \leq u \leq u_2$  refer to physical reality.

*Mathematica* is happy to provide explicit descriptions of  $u_1$ ,  $u_2$  and  $u_3$ , but they are, generally speaking, so complicated as to be worthless (except in concrete cases where they can be presented as numbers). *Mathematica* is happy also to supply

$$\begin{aligned}\text{transit time } u' \rightarrow u &= \int_{u'}^u \frac{1}{\sqrt{f(w)}} dw \\ &= \text{sum of incomplete elliptic functions} \\ &\quad \text{with complicated arcsine arguments} \\ &\equiv t(u; u')\end{aligned}$$

from which it would not be feasible to extract  $u(t; u')$ . We proceed therefore *qualitatively*:

- the top **spins** under the control of  $\psi(t)$ , the detailed motion of which is, by (97.2), under the control of  $u(t)$ :

$$\dot{\psi} = (B/A)q - \frac{p-qu}{1-u^2}u$$

- as it spins it **precesses** under the control of  $\phi(t)$ , the detailed motion of which is, by (97.1), again under the control of  $u(t)$ :

$$\dot{\phi} = \frac{p-qu}{1-u^2} \quad (101)$$

- as it spins and precesses it **nutates** under the control of  $\theta(t)$ , the detailed motion of which is, by

$$\theta = \arccos u$$

again under the control of  $u(t)$ , which *oscillates back and forth*—periodically but non-sinusoidally—between the turning points  $u_1$  and  $u_2$ .

Looking to (101) we see that  $\dot{\phi}$  vanishes at  $u = u' \equiv p/q$ . If  $u'$  lies *between*  $u_1$  and  $u_2$  then  $\dot{\phi}$  *reverses sign* as  $u$  proceeds  $u_1 \rightarrow u_2$  and the symmetry axis of the top traces a looping curve such as appears at the top of the following figure. If—exceptionally— $u' = u_2$  we get cusps (middle of the figure),<sup>26</sup> while if  $u' > u_2$  then  $\dot{\phi}$  *retains the same sign* as  $u$  proceeds  $u_1 \rightarrow u_2$  and we get the undulating curve shown at the bottom of the figure.

The elementary theory sketched on pages 56–57 provided no indication of the nutation exhibited by real tops. We are in position now to recognize that a top will display *nutation-free precession* if and only if the turning points  $u_1$  and  $u_2$  are *coincident*:  $u_1 = u_2 = u_0$ , where  $u_0$  marks the point at which the effective potential

$$V(u) \equiv \mathcal{V}(\arccos u) = \frac{(p-qu)^2}{1-u^2} + \beta u$$

assumes its minimal value. To discover the value of  $u_0$  we construct

$$\begin{aligned} \frac{dV(u)}{du} &= \frac{2u(p-qu)^2}{(1-u^2)^2} - \frac{2q(p-qu)}{1-u^2} + \beta \\ &= \frac{2u(p-qu)^2 - 2q(1-u^2)(p-qu) + \beta(1-u^2)^2}{(1-u^2)^2} \end{aligned}$$

and look for the root of the quartic numerator that lies on the physical interval  $[-1, +1]$ . We note in this connection that the numerator is quadratic in  $(p-qu)$ ,

---

<sup>26</sup> Though mathematically exceptional, such cusps are in fact observed if one spins up a top and then “drops” it with (initially)  $\dot{\phi} = \dot{\theta} = 0$ .

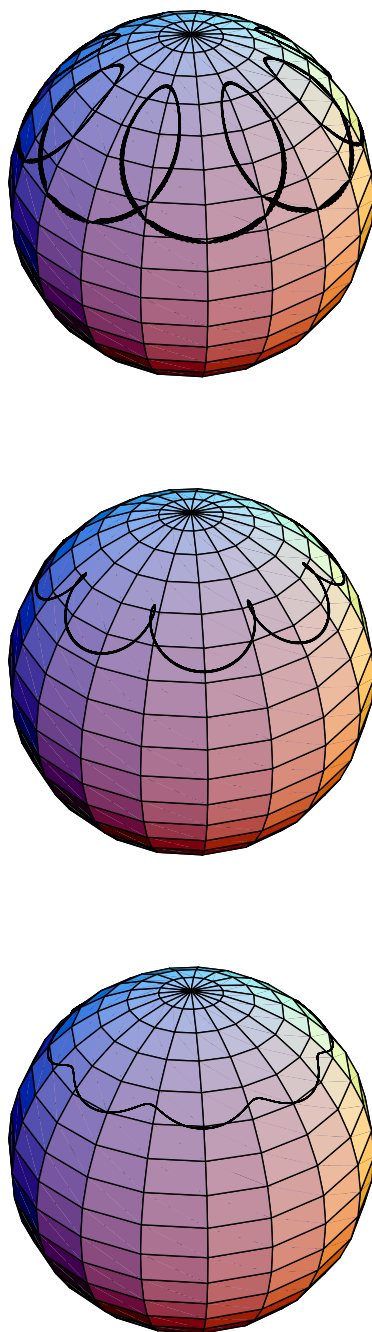


FIGURE 24: *Typical patterns traced by the symmetry axis of a nutating top as it precesses.*

and that solving the quadratic gives

$$p - qu_0 = \frac{(1 - u_0^2)}{2u_0} \left\{ q \pm \sqrt{q^2 - 2\beta u_0} \right\}$$

which by (101) becomes

$$\dot{\phi}_{0\pm} = \frac{1}{2u_0} \left\{ q \pm \sqrt{q^2 - 2\beta u_0} \right\}$$

We conclude that nutationless (or “steady”) precession can occur only if  $q^2 - 2\beta u_0 \geq 0$ , and that if  $q^2 - 2\beta u_0 > 0$  then such precession can be either *fast* or *slow*. It follows, moreover, that if  $q^2 - 2\beta u_0 \gg 0$  then

$$\dot{\phi}_{0\pm} = \frac{q}{2u_0} \cdot \begin{cases} 2 - \frac{1}{2}k - \frac{1}{8}k^2 - \dots \\ \frac{1}{2}k + \frac{1}{8}k^2 + \dots \end{cases} \quad \text{with} \quad k \equiv \frac{2\beta u_0}{q^2} = \frac{4Bmgh \cos \theta_0}{p_\psi^2} < 1$$

so in leading order

$$\begin{aligned} \dot{\phi}_{0+} &= \frac{q}{u_0} = \frac{p_\psi}{B \cos \theta_0} = \frac{A}{B \cos \theta_0} \omega_{\text{spin}} \\ \dot{\phi}_{0-} &= \frac{\beta}{2q} = \frac{mgh}{p_\psi} = \frac{mgh}{A \omega_{\text{spin}}} \end{aligned}$$

It is interesting that  $\dot{\phi}_{0-}$  is (in leading order)  $\theta_0$ -independent; *i.e.*, that steady precession can occur at any tilt. This may account for why it is that—according to Thornton & Marion<sup>22</sup> that “it is the slower of the two [steady] precessional velocities that is usually observed.” In any event (to rephrase in more physical terms a point established just above), steady precession of neither sort can occur unless

$$\omega_{\text{spin}} > \sqrt{\frac{4Bmgh \cos \theta_0}{A^2}}$$

Much more could be said about the physics of toy tops, a subject which has first charmed, then challenged, many of the greatest classical theorists. Here I quote the 25-year-old Maxwell, writing in 1857:<sup>27</sup>

*“To those who study the progress of exact science, the common spinning-top is a symbol of the labours and the perplexities of men who had successfully threaded the mazes of the planetary motions. The mathematicians of the last age, searching through nature for problems worthy of their analysis, found in this toy of their youth, ample occupation for their highest mathematical powers. . . We find Euler and D’Alembert devoting their talent and their patience to the establishment of the laws of the rotation of solid bodies. Lagrange has incorporated his own analysis of the problem with his general*

<sup>27</sup> See pages 246–262 in Volume I of W. D. Niven (editor), *The Scientific Papers of James Clerk Maxwell* (1927).

*treatment of mechanics, and since his time Poinsôt has brought the subject under the power of a more searching analysis than that of the calculus, in which ideas take the place of symbols, and intelligible propositions supersede equations.”* Maxwell continues with a reference to “...the top which I have the honour to spin before the Society...”

But for further particulars and finer details I must refer my reader to §5-7 in the 1<sup>st</sup> and 2<sup>nd</sup> editions (1950 and 1980) of Goldstein, and to additional references cited there.

**15. Nonparallelism of angular velocity and spin.** It was remarked in passing already on page 6 that while

- the linear momentum  $\mathbf{p}$  and linear velocity  $\mathbf{v}$  of a point particle stand in the relation  $\mathbf{p} = m\mathbf{v}$ , where  $m$  is a scalar,
- the intrinsic angular momentum (or spin)  $\mathbf{S}$  and intrinsic angular velocity  $\boldsymbol{\omega}$  of a rigid body stand in the relation  $\mathbf{S} = \mathbb{I}\boldsymbol{\omega}$ , where  $\mathbb{I}$  is a symmetric matrix.

So while  $\mathbf{p}$  and  $\mathbf{v}$  are invariably parallel,  $\mathbf{S}$  and  $\boldsymbol{\omega}$  are typically *not* parallel but stand in an ever-shifting angular relationship, even in the total absence of impressed torques. We have learned to attribute largely to this circumstance the fact that the motion of free rigid bodies is so much more intricate than the motion of free point particles. I propose to address this question: *How great can the angle  $\mathbf{S} \angle \boldsymbol{\omega}$  be?*

We look first, by way of preparation, to a 2-dimensional model of the 3-dimensional issue. Let

$$\boldsymbol{\omega} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Let  $\mathbb{I}$  range over the set  $\mathfrak{I}$  of all real symmetric  $2 \times 2$  matrices with prescribed eigenvalues  $A$  and  $B$ ; *i.e.*, let  $\mathbb{I}$  be rotationally equivalent to the diagonal matrix

$$\mathbb{I}_0 = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

and look to the set of vectors  $\mathbf{S} = \mathbb{I}\boldsymbol{\omega}$  that is generated as  $\mathbb{I}$  ranges over  $\mathfrak{I}$ . Typical elements of  $\mathfrak{I}$  can be described

$$\begin{aligned} \mathbb{I}(\theta) &= \mathbb{R}^\top(\theta) \mathbb{I}_0 \mathbb{R}(\theta) \quad \text{with} \quad \mathbb{R}(\theta) \equiv \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} A \cos^2 \theta + B \sin^2 \theta & (A - B) \cos \theta \sin \theta \\ (A - B) \cos \theta \sin \theta & B \cos^2 \theta + A \sin^2 \theta \end{pmatrix} \end{aligned}$$

Multiplication into  $\boldsymbol{\omega}$  gives

$$\mathbf{S}(\theta) = \begin{pmatrix} A \cos^2 \theta + B \sin^2 \theta \\ (A - B) \cos \theta \sin \theta \end{pmatrix}$$



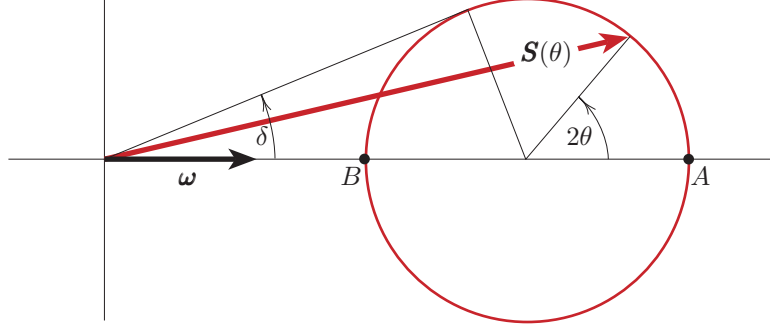


FIGURE 25: The vector  $\mathbf{S}(\theta) = \mathbb{I}(\theta)\boldsymbol{\omega}$  ranges—twice—around the red circle as  $\theta$  ranges on  $[0, 2\pi]$ . The circle intersects the  $\boldsymbol{\omega}$ -axis at the eigenvalues of  $\mathbb{I}(\theta)$ , so is centered at  $\frac{1}{2}(A+B)$  and has radius  $\frac{1}{2}(A-B)$ . The double-valuedness of the construction reflects the fact that the map  $\mathbb{I}_0 \rightarrow \mathbb{R}^T \mathbb{I}_0 \mathbb{R}$  is insensitive  $\mathbb{R} \mapsto -\mathbb{R}$ .

Writing

$$\begin{aligned} S_1 &= \frac{1}{2}A(\cos^2 \theta + 1 - \sin^2 \theta) + \frac{1}{2}B(\sin^2 \theta + 1 - \cos^2 \theta) \\ &= \frac{1}{2}(A+B) + \frac{1}{2}(A-B)\cos 2\theta \\ S_2 &= \frac{1}{2}(A-B)\sin 2\theta \end{aligned}$$

we have

$$\left(S_1 - \frac{A+B}{2}\right)^2 + (S_2)^2 = \left(\frac{A-B}{2}\right)^2$$

The implication is that the vectors  $\mathbf{S}(\theta)$  all lie on the circle shown in the preceding figure. On two occasions  $\mathbf{S}(\theta)$  and  $\boldsymbol{\omega}$  are parallel:

$$\mathbf{S}(0) = A\boldsymbol{\omega} \quad \text{and} \quad \mathbf{S}(\pi) = B\boldsymbol{\omega}$$

The greatest angular deviation is

$$\delta = \arcsin \left\{ \frac{A-B}{A+B} \right\} \tag{102}$$

and occurs at the solution  $\theta_{\max}$  of

$$2\theta = \frac{1}{2}\pi + \delta$$

Turning now to the 3-dimensional case, we set

$$\boldsymbol{\omega} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbb{I}_0 = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}$$

and—assuming our rotation matrix  $\mathbb{R}$  to have been presented in the Euler representation  $\mathbb{R}(\phi, \theta, \psi)$  that was spelled out at (42), and entrusting all calculation to *Mathematica*—obtain a description of

$$\mathbb{I}(\phi, \theta, \psi) = \mathbb{R}^\top(\phi, \theta, \psi) \mathbb{I}_0 \mathbb{R}(\phi, \theta, \psi)$$

that when multiplied into  $\boldsymbol{\omega}$  gives

$$\begin{aligned} \mathbf{S}(\phi, \theta, \psi) &= \mathbb{R}^\top(\phi, \theta, \psi) \begin{pmatrix} A \sin \theta \sin \psi \\ B \sin \theta \cos \psi \\ C \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Sigma_1(\theta, \psi) \\ \Sigma_2(\theta, \psi) \\ \Sigma_3(\theta, \psi) \end{pmatrix} \\ &\equiv \mathbb{R}(\phi) \boldsymbol{\Sigma}(\psi, \theta) \end{aligned}$$

with

$$\begin{aligned} \Sigma_1 &\equiv \frac{1}{2}(A - B) \sin 2\psi \sin \theta \\ \Sigma_2 &\equiv \frac{1}{2}[A \sin^2 \psi + B \cos^2 \psi - C] \sin 2\theta \\ \Sigma_3 &\equiv (A \sin^2 \psi + B \cos^2 \psi) \sin^2 \theta + C \cos^2 \theta \end{aligned}$$

Our assignment is to describe the  $\boldsymbol{\Sigma}(\psi, \theta)$ -vector, which  $\mathbb{R}(\phi)$  serves simply to twirl about the 3-axis (the  $\boldsymbol{\omega}$ -axis), with these consequences:

$$\begin{aligned} S_1^2 + S_2^2 &= \Sigma_1^2 + \Sigma_2^2 \\ S_3 &= \Sigma_3 \end{aligned}$$

We begin with the observations that  $\Sigma_3$  is manifestly non-negative, and can be written

$$\Sigma_3 = [A \sin^2 \psi + B \cos^2 \psi - C] \sin^2 \theta + C$$

We have had occasion to note the identity

$$A \sin^2 \psi + B \cos^2 \psi = \frac{1}{2}(A + B) - \frac{1}{2}(A - B) \cos 2\psi$$

so with the introduction of  $a \equiv A - C$  and  $b \equiv B - C$  we have

$$A \sin^2 \psi + B \cos^2 \psi - C = \frac{1}{2}(a + b) - \frac{1}{2}(a - b) \cos 2\psi$$

giving

$$\begin{aligned} \Sigma_1 &= \left[ \frac{1}{2}(a - b) \sin 2\psi \right] \sin \theta \\ \Sigma_2 &= \left[ \frac{1}{2}(a + b) - \frac{1}{2}(a - b) \cos 2\psi \right] \sin \theta \cos \theta \\ \Sigma_3 &= \Sigma_3 = \left[ \frac{1}{2}(a + b) - \frac{1}{2}(a - b) \cos 2\psi \right] \sin \theta \sin \theta + C \end{aligned}$$



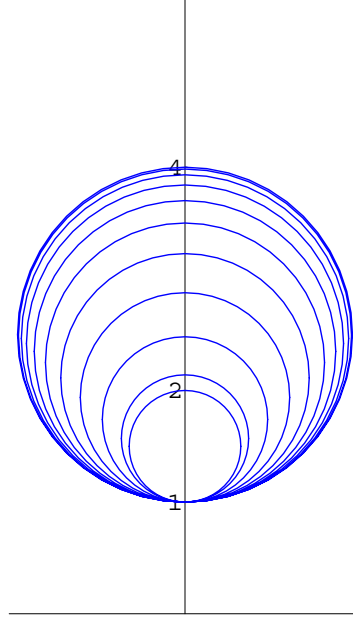


FIGURE 26: Cross section of a family of nested spheres obtained from (104), with  $A = 4$ ,  $B = 3$ ,  $C = 1$ . The angular parameter  $\psi$  has been stepped through the values  $n\frac{\pi}{20}$  ( $n = 0, 1, 2, \dots, 10$ ) and is constant on each sphere.

and where it is evident that

$$\ell_{\min} = \ell\left(\frac{1}{2}\pi\right) = \frac{1}{2}(A - B) \quad \text{and} \quad \ell_{\max} = \ell(0) = \frac{1}{2}(A + B) - C \\ = \frac{1}{2}(A - B) + (B - C)$$

From (105) we are led to FIGURE 27, which again tells only half—the *other* half—of the story. It is only by conflating those figures—by taking (104) and (105) in combination—that we obtain a description of the set of points to which parameters  $\phi$ ,  $\theta$  and  $\psi$  can be simultaneously assigned, a description of the curious region to which  $\mathbf{S}(\phi, \theta, \psi)$  is necessarily confined.

The  $\omega$ -axis punctures that “crescent of revolution” at only three points, and those mark the eigenvalues of  $\mathbb{I}$ .

It is evident from FIGURE 28 that

$$\delta = \arcsin \left\{ \frac{\frac{1}{2}[(\text{greatest eigenvalue}) - (\text{least eigenvalue})]}{\frac{1}{2}[(\text{greatest eigenvalue}) + (\text{least eigenvalue})]} \right\} \quad (106)$$

which serves very nicely as a generalization of (102). I invite my reader to consider the limiting cases  $B \uparrow A$ ,  $B \downarrow C$  and  $A = B = C$ .

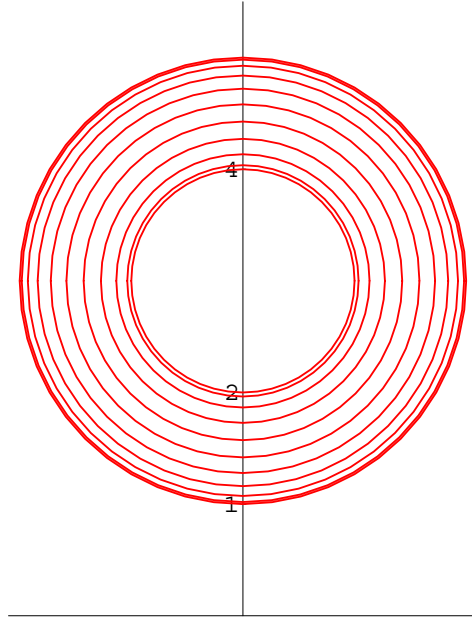


FIGURE 27: Cross section of a family of nested spheres obtained from (105), with  $A = 4$ ,  $B = 3$ ,  $C = 1$ . The angular parameter  $\theta$  has been stepped through the values  $n\frac{\pi}{20}$  ( $n = 0, 1, 2, \dots, 10$ ) and is constant on each sphere.

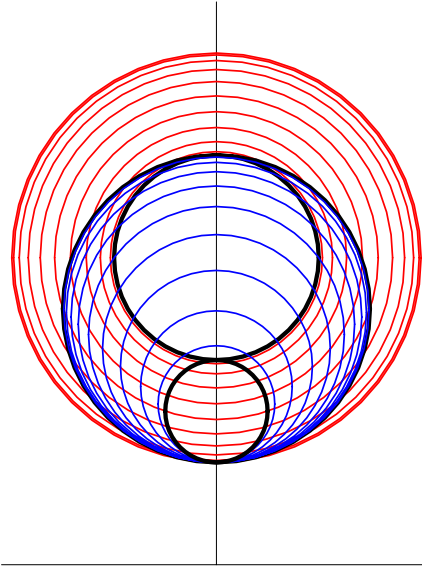


FIGURE 28: Superimposed figures.  $\mathcal{S}(\phi, \theta, \psi)$  lies necessarily in the crosshatched region, which has the form of a sphere with two interior spherical exclusions

I originally developed this material to resolve some procedural problems that arose in connection with some experiments I was performing (1977) with air-supported gyros of various designs. Only later did I realize that the mathematics has nothing specific to do with the relationship of spin to angular velocity: it pertains simultaneously to *all* statements of a form

$$\mathbf{y} = \mathbb{M}\mathbf{x} \quad : \quad \mathbb{M} \text{ a real } 3 \times 3 \text{ symmetric matrix}$$

that is encountered very commonly in physics, especially in linearized phenomenological theories of all sorts.<sup>28</sup> It is therefore not surprising that the essentials—at least the 2-dimensional essentials—of the material have been reinvented many times by many people. The first occurrence of my FIGURE 25 appears to have been in a publication of Christian Otto Mohr (1882), who had himself built upon a suggestion of Karl Culmann (1866). Culmann and Mohr were concerned not with the dynamics of tops but with stress analysis and the fracture of brittle materials.<sup>29</sup> Some variants and generalizations of “Mohr’s construction” are discussed in my “Non-standard applications of Mohr’s construction” (1998).

**16. Theory of celts.** Footballs, hardboiled eggs, tippy tops ... all behave in counterintuitive ways when spun, and each has generated a literature.<sup>30</sup> Here I propose to discuss only one of those curiosities. The story begins in the British Museum, where one day in the 1890s the physicist G. T. Walker had reason to examine that museum’s collection of “celts”—smooth axhead-like stones found in abundance at paleolithic sites all over Europe and the British Isles—and chanced to notice that many of them, while they spun easily in one direction, first wobbled and then reversed course when spun in the opposite direction.

---

<sup>28</sup> Such an equation relates stress to strain in elastic media, polarization to electric field strength in dielectric media, *etc.*: the list could be very greatly extended.

<sup>29</sup> Culmann (1821–1881) was a German professor of civil engineering who is remembered today mainly for his contributions—some of which had been anticipated by Maxwell—to “graphical statics.” Mohr (1835–1918) taught civil engineering first in Stuttgart and then (from 1873 until his retirement in 1900) in Dresden. He was said by his student A. Föppl (who himself figures importantly in the history of electrodynamics, and whose texts influenced the development of the young Einstein) to have been an outstanding teacher: a tall, proud and taciturn man who spoke and wrote with simplicity, clarity and conciseness. “Mohr’s stress circle”—the contribution for which he is today remembered—provided the basis for his theory of stress failure; for an account of something called the “Coulomb-Mohr fracture criterion” see (for example) C. C. Mei, *Mathematical Analysis in Engineering* (1995), p. 150. For an electrodynamical application of Mohr’s idea—having nothing at all to do either with tops or with fracture—see my CLASSICAL ELECTRODYNAMICS (1980), p. 127.

<sup>30</sup> A fairly extensive bibliography—which is, however, by no means complete—begins at page 146 in GYRODYNAMICS (1976/77).

Walker soon demonstrated<sup>31</sup> that this odd behavior could be attributed to the circumstance that the geometrical axes defined by the (approximately) ellipsoidal base of the object and its principal axes are misaligned. Walker's lectures and demonstrations were witnessed by Arnold Sommerfeld (Trinity College, Cambridge, 1899), who later recalled the powerful impression they made upon him, and who gave brief attention to the subject on pages 149–150 of his *Lectures on Theoretical Physics: Volume 1. Mechanics* (1952). I myself learned of the “celt phenomenon” from §2.72 of J. Walker's *Flying Circus of Physics*: I wrote up a modernized version of G. T. Walker's original theory, and—taking as my model the jade celt that Frank Oppenheimer one day pulled from his desk drawer and showed me—fashionedd from Brazilian rosewood what has become known locally as “Wheeler's banana top.” My work came to the attention of J. Walker, and is mentioned in the “Amateur Scientist” column (of which Walker was then the editor) in the October 1979 issue of *Scientific American*. It was J. Walker who on that occasion attached the name “rattleback” to these objects, and it is under that head that one should approach Google for recent references.

The *Scientific American* article generated a flood of correspondence—much of it goofy, some of it not—that continues to this day, 25 years later. It was one of my correspondents who directed my attention to a then-recent article by Sir Herman Bondi.<sup>32</sup> Bondi's objective was to write an improved and more complete revision of G. T. Walker's paper. Close study of Bondi's paper leaves me unconvinced, however, that he achieved his objective. Here I have taken as my source the account of Walker's paper that appears in Chapter 17, §§1–3 of A. Gray's *A Treatise on Gyrostatics and Rotational Motion* (1918).

#### DESIGN CONSIDERATIONS

The convex surface of our resting celt, in the immediate vicinity of its support point, can in leading approximation be considered to be ellipsoidal. To deal most simply with that fact we will consider the celt to be ellipsoidal not just locally but globally; *i.e.*, to have a surface of which

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 \quad (107)$$

provides an implicit description. We will assume  $a > b > c$ : the point of static support resides then (see FIGURE 29) at  $(0, 0, -c)$ . Stability requires that

$$\text{least radius of curvature at support point} > c$$

We are led to look at the bottom of the  $x = 0$  cross-section of the celt, where

$$z(y) = -c\sqrt{1 - (y/b)^2}$$

<sup>31</sup> Quarterly Journal of Pure & Applied Mathematics **28** (1896), pages 175–184.

<sup>32</sup> “The rigid body dynamics of unidirectional spin,” Proceedings of the Royal Society (London) **405A**, 265 (1986).

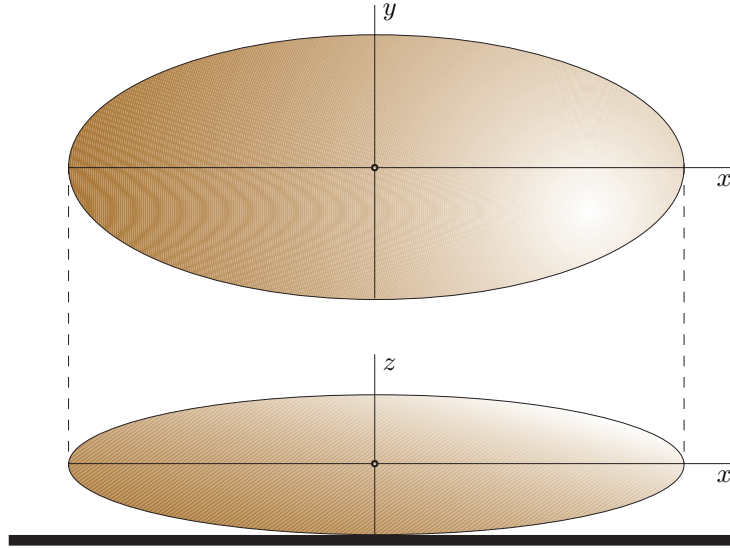


FIGURE 29: *Assumed shape of the celt. The coordinate system—with origin at the geometric center of the ellipsoid, and aligned in coincidence with its symmetry axes—will eventually (after the top has been properly “loaded”) be abandoned in favor of a coordinate system that diagonalizes the moment of inertia matrix.*

The curvature of such a plain curve can be described<sup>33</sup>

$$K \equiv \frac{d}{ds}(\text{slope}) = \frac{1}{\sqrt{1 + (z')^2}} \frac{d}{dy} \arctan \frac{dz}{dy} = \frac{z''}{[1 + (z')^2]^{\frac{3}{2}}}$$

which in the instance at hand supplies

$$\begin{aligned} K(0, y) &= \text{complicated expression (ask Mathematica)} \\ &\downarrow \\ K(0, 0) &= c/b^2 \\ &= \frac{1}{\text{least radius of curvature}} \end{aligned}$$

so to achieve stability-at-rest we have only to require that  $b > c$ , and this we have in fact already done.

It is our intention to load the ellipsoid in such a way as to cause the horizontal principal axes to be slightly misaligned with respect to their geometrical counterparts (see FIGURE 30). We confront therefore this small mathematical problem: *How to describe an ellipsoid that has been thus slewed with respect to its principal axes?* The answer, as will emerge, lies already at

<sup>33</sup> Here  $s$  denotes arc length:  $ds = \sqrt{1 + (z')^2} dy$ .



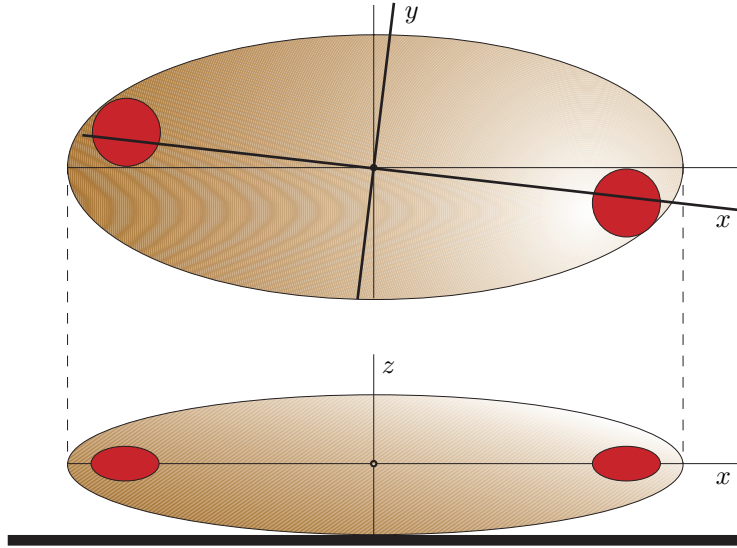


FIGURE 30: *The formerly homogeneous celt has now been “loaded” in such a way as to preserve the location of the center of mass, and to preserve also the  $z$ -axis as a principal axis, but to slew the other principal axes with respect to the associated geometrical axes.*

hand. Let (107) be notated

$$px^2 + qy^2 = k^2$$

with  $p \equiv 1/a^2$ ,  $q \equiv 1/b^2$ ,  $k^2 \equiv 1 - (z/c)^2$ .<sup>34</sup> Or again

$$\begin{pmatrix} x \\ y \end{pmatrix}^\top \mathbb{M}_0 \begin{pmatrix} x \\ y \end{pmatrix} = k^2 \quad \text{with} \quad \mathbb{M}_0 \equiv \begin{pmatrix} p & 0 \\ 0 & q \end{pmatrix}$$

To rotate that ellipse through an angle  $\psi$  we have—as was established already on page 64—only to make the replacement

$$\mathbb{M}_0 \longmapsto \mathbb{M}(\psi) \equiv \begin{pmatrix} p \cos^2 \psi + q \sin^2 \psi & (p - q) \cos \psi \sin \psi \\ (p - q) \cos \psi \sin \psi & q \cos^2 \psi + p \sin^2 \psi \end{pmatrix} \quad (108)$$

—the effect of which is illustrated in FIGURE 31.<sup>35</sup> Assuming  $\psi$  to have been prescribed/fixed, we will write

$$= \begin{pmatrix} P & R \\ R & Q \end{pmatrix}$$

<sup>34</sup> Note that  $a > b$  entails  $p < q$ .

<sup>35</sup> *Mathematica* confirms that the eigenvalues of  $\mathbb{M}(\psi)$  are  $\{p, q\}$  and that  $\det \mathbb{M}(\psi) = pq$  for all values of  $\psi$ .

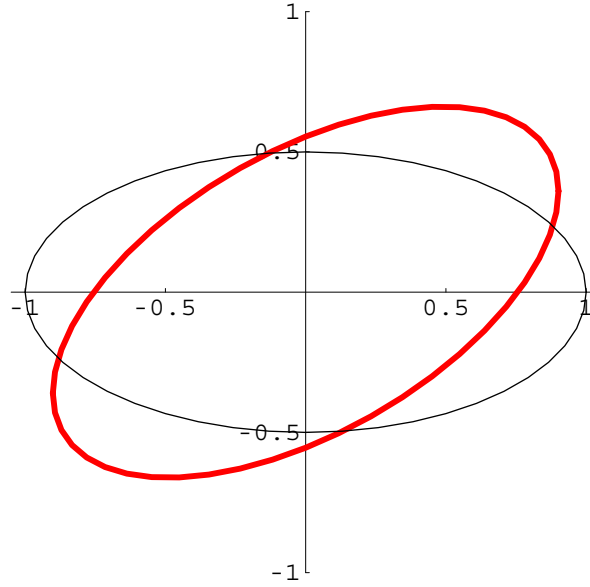


FIGURE 31: Graph of the ellipse  $x^2 + 4y^2 = 1$  and (in **red**) the result of rotation through angle  $\psi = \pi/6$ , constructed with the aid of (108).

in which notation (107) becomes

$$Px^2 + 2Rxy + Qy^2 + \left(\frac{z}{c}\right)^2 = 1$$

To say the same thing another way, we have

$$z = c \sqrt{1 - (Px^2 + 2Rxy + Qy^2)}$$

which in the near neighborhood of the point of static support (*i.e.*, for small  $x$  and  $y$ ) becomes

$$= -c \left\{ 1 - \frac{1}{2}(Px^2 + 2Rxy + Qy^2) - \dots \right\} \quad (109)$$

This equation will serve to describe—relative to the principal axes—all relevant aspects of the SHAPE of the celt. It is the presence of the  $R$ -term, which is under the control of the angular parameter  $\psi$ , that accounts for the *chirality* of celts.

Importance will attach in the dynamical theory to the *unit normal vector at the point to contact* (see FIGURE 32). To obtain a description of that vector,

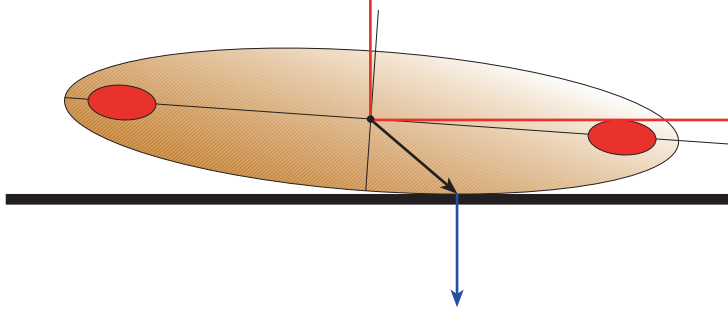


FIGURE 32: *Celt in a typical non-equilibrium position. Shown in blue is the unit vector that is normal to the celtic surface (and normal also to the support surface) at the point of contact. Shown in red is the space frame, the origin of which rides irrotationally with the center of mass.*

introduce the “scalar field”

$$\varphi(x, y, z) \equiv \frac{1}{2}c(Px^2 + 2Rxy + Qy^2) - z$$

and construe (109) to describe a “surface of constant  $\varphi$ ”:

$$\varphi(x, y, z) = c$$

We then have

$$\nabla\varphi = \begin{pmatrix} c(Px + Ry) \\ c(Rx + Qy) \\ -1 \end{pmatrix}$$

giving

$$\begin{aligned} \mathbf{n} &= \frac{\nabla\varphi}{|\nabla\varphi|} \\ &= \frac{1}{\sqrt{1 + c^2(Px + Ry)^2 + c^2(Rx + Qy)^2}} \begin{pmatrix} c(Px + Ry) \\ c(Rx + Qy) \\ -1 \end{pmatrix} \end{aligned}$$

This equation describes—*relative to the body frame*—the unit normal at the point of contact, but only if  $x$  and  $y$  are assigned the values that *describe* the instantaneous point of contact. In the near neighborhood of the origin (point of resting contact) we have Hoyle’s

$$\mathbf{n} = \begin{pmatrix} c(Px + Ry) \\ c(Rx + Qy) \\ -1 + \frac{1}{2}[c^2(Px + Ry)^2 + c^2(Rx + Qy)^2] \end{pmatrix} + \dots$$

where the abandoned terms are of higher than second order in  $x$  and  $y$ .

If, however, we elect with Walker/Gray to *work in first order*<sup>36</sup> we have

$$\mathbf{n}^0 = \begin{pmatrix} c(Px + Ry) \\ c(Rx + Qy) \\ -1 \end{pmatrix} + \cdots \quad (110)$$

In that same leading approximation the vector that extends from the center of mass to the instantaneous contact point becomes

$$\mathbf{r}^0 = \begin{pmatrix} x \\ y \\ -c \end{pmatrix} \quad (111)$$

from which it follows that  $\mathbf{n}^0$  can be described

$$\mathbf{n}^0 = \begin{pmatrix} cP & cR & 0 \\ cR & cQ & 0 \\ 0 & 0 & c^{-1} \end{pmatrix} \mathbf{r}^0 \quad (112.1)$$

In the preceding equations I have installed <sup>0</sup>s to emphasize that they refer to the *body frame*.

EQUATIONS OF MOTION
---------------------

Looking first to the motion of the center of mass, Newton's 2<sup>nd</sup> law supplies

$$m\dot{\mathbf{v}} = mg\mathbf{n} + \mathbf{f}$$

where  $\mathbf{f}$  refers to the net force (support and friction) exerted on the celt by the table. Here all vectors are relative to the inertial “table frame,” of which the “space frame” (origin riding on the center of mass) is a non-inertial translated copy. To express inertial vectors in terms of body-frame vectors we write  $\mathbf{v} = \mathbb{R}\mathbf{v}^0$ ,  $\mathbf{f} = \mathbb{R}\mathbf{f}^0$ ,  $\mathbf{n} = \mathbb{R}\mathbf{n}^0$ , *etc.* The time-dependence of  $\mathbb{R}$  entails

$$\begin{aligned} \dot{\mathbf{v}} &= \mathbb{R}\dot{\mathbf{v}}^0 + \dot{\mathbb{R}}\mathbb{R}^{-1}\mathbf{v} \\ &= \mathbb{R}\dot{\mathbf{v}}^0 + \boldsymbol{\omega} \times \mathbf{v} \\ &= \mathbb{R}(\dot{\mathbf{v}}^0 + \boldsymbol{\omega}^0 \times \mathbf{v}^0) \end{aligned}$$

so we have

$$m(\dot{\mathbf{v}}^0 + \boldsymbol{\omega}^0 \times \mathbf{v}^0) = mg\mathbf{n}^0 + \mathbf{f}^0$$

The motion of  $\boldsymbol{\omega}^0$  is described by Euler's equation

$$\mathbb{I}^0\dot{\boldsymbol{\omega}}^0 + \boldsymbol{\omega}^0 \times \mathbb{I}^0\boldsymbol{\omega}^0 = \mathbf{r}^0 \times \mathbf{f}^0$$

---

<sup>36</sup> It is principally because Hoyle works in 2<sup>nd</sup> order that his equations are relatively so complicated, and the significance of his results so hard to grasp intuitively.

which, if we use Newton's law to eliminate reference to the presently unknown force  $\mathbf{f}^0$ , becomes

$$\mathbb{I}^0 \dot{\boldsymbol{\omega}} + \boldsymbol{\omega}^0 \times \mathbb{I}^0 \boldsymbol{\omega}^0 - m \mathbf{r}^0 \times \{\dot{\mathbf{v}}^0 + \boldsymbol{\omega}^0 \times \mathbf{v}^0 - g \mathbf{n}^0\} = \mathbf{0} \quad (112.2)$$

Here

$$\mathbb{I}^0 = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix} \quad (112.3)$$

Additionally we have  $\dot{\mathbf{n}} = \mathbf{0}$  because the table is (by assumption) flat, giving

$$\dot{\mathbf{n}}^0 + \boldsymbol{\omega}^0 \times \mathbf{n}^0 = \mathbf{0} \quad (112.4)$$

And finally, because we assume the celt rolls without slipping, we have

$$\mathbf{v}^0 + \boldsymbol{\omega}^0 \times \mathbf{r}^0 = \mathbf{0} \quad (112.5)$$

Equations (112) provide the physical basis of Walker's theory (also of Hoyle's). In working out the consequences of these equations let us now agree, as a matter of typographic convenience, to drop the  $^0$ s.

Cross (112.4) into the unit vector  $\mathbf{n}$  and use  $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{c} \cdot \mathbf{a})\mathbf{b} - (\mathbf{c} \cdot \mathbf{b})\mathbf{a}$  to obtain  $\dot{\mathbf{n}} \times \mathbf{n} + (\mathbf{n} \cdot \boldsymbol{\omega})\mathbf{n} - (\mathbf{n} \cdot \mathbf{n})\boldsymbol{\omega} = \mathbf{0}$  or

$$\boldsymbol{\omega} = \dot{\mathbf{n}} \times \mathbf{n} + \omega \mathbf{n} \quad \text{with} \quad \omega \equiv (\mathbf{n} \cdot \boldsymbol{\omega}) \\ = \text{magnitude of } \boldsymbol{\omega}_{\parallel} \quad (113.1)$$

Returning with this result to (112.5) we have

$$\mathbf{v} = \mathbf{r} \times (\dot{\mathbf{n}} \times \mathbf{n} + \omega \mathbf{n}) \quad (113.2)$$

Equations (113)—taken together with their time-partials<sup>37</sup>

$$\dot{\boldsymbol{\omega}} = \ddot{\mathbf{n}} \times \mathbf{n} + \omega \dot{\mathbf{n}} \\ \dot{\mathbf{v}} = \dot{\mathbf{r}} \times (\dot{\mathbf{n}} \times \mathbf{n} + \omega \mathbf{n}) + \mathbf{r} \times (\ddot{\mathbf{n}} \times \mathbf{n} + \dot{\omega} \mathbf{n})$$

—can be used to turn (112.2) into an equation involving only  $\mathbf{n}$ ,  $\mathbf{r}$  and their derivatives. I postpone that substitutional exercise.

The unit normal  $\mathbf{n}$ , though fixed with respect to the table frame, moves relative to the body frame (in which we are now working), and Walker takes that apparent motion to be the indicator of what the celt is doing. Hoyle, on the other hand, elects to watch the motion of  $\mathbf{r}$  (*i.e.*, of  $x$  and  $y$ , which refer

---

<sup>37</sup> Here we accept Walker's intuition-based assertion that  $\omega$  will be constant in leading order, that its temporal variation will be a higher-order effect. It was in an effort to avoid such *ad hoc* assertions that Hoyle worked in second order.

to the instantaneous location of the contact point). Methodologically it is six one way, half a dozen the other: I will follow Walker's lead, which calls for elimination of all  $\mathbf{r}$  and  $\dot{\mathbf{r}}$  terms. To that end we return to (112.1), which supplies

$$\begin{aligned}\mathbf{r} &= \frac{1}{c^2(PQ - R^2)} \begin{pmatrix} cQ & -cR & 0 \\ -cR & cP & 0 \\ 0 & 0 & c^3(PQ - R^2) \end{pmatrix} \mathbf{n} \\ &\equiv \begin{pmatrix} \alpha & \rho & 0 \\ \rho & \beta & 0 \\ 0 & 0 & c \end{pmatrix} \mathbf{n} \\ &\equiv \mathbb{J} \mathbf{n}\end{aligned}\tag{113.3}$$

$$\dot{\mathbf{r}} = \mathbb{J} \dot{\mathbf{n}}$$

With Walker, we proceed now in the assumption that in first order<sup>38</sup>

$$\mathbf{n} = \begin{pmatrix} \epsilon n_1 \\ \epsilon n_2 \\ -1 \end{pmatrix} \quad \text{whence} \quad \dot{\mathbf{n}} = \begin{pmatrix} \epsilon \dot{n}_1 \\ \epsilon \dot{n}_2 \\ 0 \end{pmatrix}, \quad \ddot{\mathbf{n}} = \begin{pmatrix} \epsilon \ddot{n}_1 \\ \epsilon \ddot{n}_2 \\ 0 \end{pmatrix}$$

It then follows by (113.1) that

$$\boldsymbol{\omega} = \begin{pmatrix} \epsilon(-\dot{n}_2 + \omega n_1) \\ \epsilon(\dot{n}_1 + \omega n_2) \\ -\omega \end{pmatrix}, \quad \dot{\boldsymbol{\omega}} = \begin{pmatrix} \epsilon(-\ddot{n}_2 + \omega \dot{n}_1) \\ \epsilon(\ddot{n}_1 + \omega \dot{n}_2) \\ 0 \end{pmatrix}$$

and from (113.3) that

$$\mathbf{r} = \begin{pmatrix} \epsilon(\alpha n_1 + \rho n_2) \\ \epsilon(\rho n_1 + \beta n_2) \\ -c \end{pmatrix}, \quad \dot{\mathbf{r}} = \begin{pmatrix} \epsilon(\alpha \dot{n}_1 + \rho \dot{n}_2) \\ \epsilon(\rho \dot{n}_1 + \beta \dot{n}_2) \\ 0 \end{pmatrix}$$

Bringing this information to (113.2) we find

$$\begin{aligned}\mathbf{v} &= \begin{pmatrix} \epsilon(c \dot{n}_1 + \omega[c n_2 - \beta n_2 - \rho n_1]) \\ \epsilon(c \dot{n}_2 - \omega[c n_1 - \alpha n_1 - \rho n_2]) \\ 0 \end{pmatrix} \\ \dot{\mathbf{v}} &= \begin{pmatrix} \epsilon(c \ddot{n}_1 + \omega[c \dot{n}_2 - \beta \dot{n}_2 - \rho \dot{n}_1]) \\ \epsilon(c \ddot{n}_2 - \omega[c \dot{n}_1 - \alpha \dot{n}_1 - \rho \dot{n}_2]) \\ 0 \end{pmatrix}\end{aligned}$$

---

<sup>38</sup> The  $\epsilon$  factors have been introduced to identify terms that we imagine to be “small,” and to provide *Mathematica* with means to identify and discard second order terms as they arise: at the end of the day we will set  $\epsilon = 1$ .

Returning with this information to (112.2) we are led—after abandoning terms of  $O[\epsilon^2]$ —to three equations, the third of which reads  $0 = 0$  and the first pair of which (after we reverse their order and change a sign) read

$$\begin{aligned} 0 = & +\ddot{n}_1[B + mc^2] \\ & +\dot{n}_1[-mc\rho\omega] \\ & +\dot{n}_2[(A + B - C)\omega + 2mc^2\omega - mc\beta\omega] \\ & +n_1[(-A + C - mc^2 + mc\alpha)\omega^2 - mg(c - \alpha)] \\ & +n_2[m\rho(g + c\omega^2)] \end{aligned}$$

$$\begin{aligned} 0 = & +\ddot{n}_2[A + mc^2] \\ & +\dot{n}_1[-(A + B - C)\omega - 2mc^2\omega + mc\alpha\omega] \\ & +\dot{n}_2[+mc\rho\omega] \\ & +n_1[m\rho(g + c\omega^2)] \\ & +n_2[(-B + C - mc^2 + mc\beta)\omega^2 - mg(c - \beta)] \end{aligned}$$

We have here a pair of coupled *linear* equations that can be written

$$\mathbb{M}\ddot{\mathbf{n}} + (\mathbb{S} + \mathbb{A})\dot{\mathbf{n}} + \mathbb{K}\mathbf{n} = \mathbb{O} \quad (114)$$

where

$$\left. \begin{aligned} \mathbb{M} &= \begin{pmatrix} B+mc^2 & 0 \\ 0 & A+mc^2 \end{pmatrix} \\ \mathbb{S} &= \begin{pmatrix} -mc\rho\omega & \frac{1}{2}mc(\alpha-\beta)\omega \\ \frac{1}{2}mc(\alpha-\beta)\omega & +mc\rho\omega \end{pmatrix} \\ \mathbb{A} &= \begin{pmatrix} 0 & +[A+B-C+2mc^2-\frac{1}{2}mc(\alpha+\beta)]\omega \\ -[A+B-C+2mc^2-\frac{1}{2}mc(\alpha+\beta)]\omega & 0 \end{pmatrix} \\ \mathbb{K} &= \begin{pmatrix} (-A+C-mc^2+mc\alpha)\omega^2-mg(c-\alpha) & m\rho(g+c\omega^2) \\ m\rho(g+c\omega^2) & (-B+C-mc^2+mc\beta)\omega^2-mg(c-\beta) \end{pmatrix} \end{aligned} \right\} \quad (115)$$

and where  $\mathbf{n}$  is understood now to be the 2-vector

$$\mathbf{n} \equiv \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}$$

The matrices (115) are assembled from

- physical parameters  $m$  and  $g$ ;
- principal moments  $A$ ,  $B$  and  $C$ ;
- parameters  $a$ ,  $b$  and  $c$  that set the shape of the celt (particularly of its foot); and
- an angle  $\theta$  that describes the misalignment of the symmetry and principal axis systems.

The later four parameters are hidden in the designs of  $\alpha$ ,  $\beta$  and  $\rho$ . In the absence of misalignment we have  $\rho = 0$ . The matrices  $\mathbb{M}$ ,  $\mathbb{S}$  and  $\mathbb{K}$  are symmetric;  $\mathbb{A}$  is antisymmetric.

Notice that reversing the sign of what Hoyle calls the “spin” ( $\omega \mapsto -\omega$ ) sends

$$\begin{aligned} \mathbb{M}\ddot{\mathbf{n}} + (\mathbb{S} + \mathbb{A})\dot{\mathbf{n}} + \mathbb{K}\mathbf{n} &= \mathbb{O} \\ \downarrow \\ \mathbb{M}\ddot{\mathbf{n}} - (\mathbb{S} + \mathbb{A})\dot{\mathbf{n}} + \mathbb{K}\mathbf{n} &= \mathbb{O} \end{aligned}$$

Which is to say: Celts spun  $\odot$  or  $\ominus$  are described by *distinct equations of motion*, already in first-order theory. The chirality of celt dynamics is thus made immediately apparent.

The Walker/Gray argument becomes at this point a cleverly executed exercise in stability theory. To expose the elegance of their idea without the distraction of notational clutter, let us write

$$\left. \begin{aligned} \mathbb{M} &= \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \\ \mathbb{S} &= \begin{pmatrix} -\rho\sigma & s \\ s & \rho\sigma \end{pmatrix} \\ \mathbb{A} &= \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \\ \mathbb{K} &= \begin{pmatrix} k_1 & \rho\kappa \\ \rho\kappa & k_2 \end{pmatrix} \end{aligned} \right\} \quad (116)$$

to abbreviate the structural essentials of (115). We look for solutions of the form

$$\mathbf{n}(t) = \boldsymbol{\nu} e^{i\Omega t} \quad (117)$$

From (114) we obtain

$$[-\Omega^2 \mathbb{M} + i\Omega(\mathbb{S} + \mathbb{A}) + \mathbb{K}] \boldsymbol{\nu} = \mathbf{0} \quad (118)$$

which entails

$$\begin{aligned} 0 &= \det[-\Omega^2 \mathbb{M} + i\Omega(\mathbb{S} + \mathbb{A}) + \mathbb{K}] \\ &= (k_1 k_2 - \rho^2 \kappa^2) + i\rho[\sigma(k_1 - k_2) - 2s\kappa]\Omega \\ &\quad - (m_1 k_2 + m_2 k_1 + a^2 - s^2 - \rho^2 \sigma^2)\Omega^2 \\ &\quad - i\rho\sigma(m_1 - m_2)\Omega^3 \\ &\quad + m_1 m_2 \Omega^4 \quad \text{according to } \textit{Mathematica} \\ &\equiv (K - \rho^2 \kappa^2) - (\mu - \rho^2 s^2)\Omega^2 + M\Omega^2 \\ &\quad + \rho\{i(\sigma \cdot \Delta k - 2s\kappa)\Omega - i\sigma \cdot \Delta m \Omega^3\} \\ &\equiv f(\Omega) + \rho \cdot g(\Omega) + \rho^2 \cdot h(\Omega) \end{aligned} \quad (119)$$



Quartics are awkward. Note, however, that at  $\rho = 0$ ; *i.e.*, in the *absence* of misalignment, the preceding equation becomes quadratic in  $\Omega^2$ : it becomes

$$f(\Omega) \equiv K - \mu\Omega^2 + M\Omega^4 = 0$$

and supplies

$$\Omega_0^2 = \frac{\mu \pm \sqrt{\mu^2 - 4MK}}{2M}$$

We conclude—since  $M = (A + mc^2)(B + mc^2) > 0$  in all cases—that  $\Omega_0$  will be real only in those parts of parameter space where

$$\mu^2 - 4MK \geq 0 \quad \text{and} \quad \mu > 0 \quad (120)$$

When those conditions are satisfied we have

$$\mathbf{n}(t) = \boldsymbol{\nu}_{\text{fast}} \cos(\Omega_{\text{fast}} t) + \boldsymbol{\nu}_{\text{slow}} \cos(\Omega_{\text{slow}} t + \delta) \quad (121)$$

where

$$\begin{aligned} \Omega_{\text{fast}} &\equiv \left[ \frac{\mu + \sqrt{\mu^2 - 4MK}}{2M} \right]^{\frac{1}{2}} \\ \Omega_{\text{slow}} &\equiv \left[ \frac{\mu - \sqrt{\mu^2 - 4MK}}{2M} \right]^{\frac{1}{2}} \end{aligned}$$

and where  $\boldsymbol{\nu}_{\text{fast}}$ ,  $\boldsymbol{\nu}_{\text{slow}}$  are associated solutions of the homogeneous system (118), our present assumption being that  $\rho$  has been set to zero.<sup>39</sup> Equation (121) shows the small-amplitude motion of  $\mathbf{n}$  to trace a (generally aperiodic) “skew-Lissajous figure”:

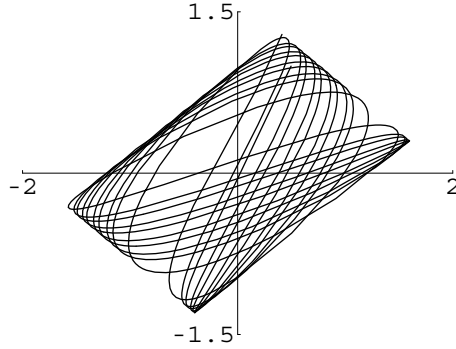


FIGURE 33: “Skew-Lissajous figure” generated by (121) in the case

$$\Omega_{\text{fast}} = 1.00, \quad \Omega_{\text{slow}} = 0.68, \quad \boldsymbol{\nu}_{\text{fast}} = \begin{pmatrix} 1.0 \\ 0.8 \end{pmatrix}, \quad \boldsymbol{\nu}_{\text{slow}} = \begin{pmatrix} -0.6 \\ 0.5 \end{pmatrix}$$

<sup>39</sup> It is tempting but would be incorrect to call the  $\Omega$ ’s “eigenvalues”—though they are, like eigenvalues, roots of a polynomial—and it would for that same reason be incorrect to call the  $\boldsymbol{\nu}$ ’s—which are in general *not* orthogonal—“eigenvectors.”

But setting  $\rho = 0$  destroys the phenomenon of interest, for on looking back again to (115) we see that the parameters  $\{\mu, M, K\}$  that enter into the construction of  $\Omega_0$  are all *invariant* under  $\omega \mapsto -\omega$ .

At this point it is Walker/Gray's resourceful idea to assume  $\rho$  to be non-zero *but small*, and to pass from quadratic to quintic by the methods of perturbation theory.

**MATHEMATICAL DIGRESSION: A toy perturbation theory.** Let  $x_0$  be a zero of  $f(x)$ . What values should be assigned to  $\{x_1, x_2, \dots\}$  to make  $x_0 + \epsilon x_1 + \epsilon^2 x_2 + \dots$  a zero of the perturbed function

$$F_\epsilon(x) \equiv f(x) + \epsilon g(x) + \epsilon^2 h(x) + \dots$$

Asking *Mathematica* to expand  $F_\epsilon(x_0 + \epsilon x_1 + \epsilon^2 x_2 + \dots)$ , we are led to

$$\begin{aligned} x_1 f'(x_0) + g(x_0) &= 0 \\ x_2 f'(x_0) + x_1 g'(x_0) + \frac{1}{2} x_1^2 f''(x_0) + h(x_0) &= 0 \\ &\vdots \end{aligned}$$

which can be solved recursively. In leading order we have

$$x_1 = -\frac{g(x_0)}{f'(x_0)} \quad (122)$$

Taking our definitions of  $f(\Omega)$  and  $g(\Omega)$  from (119)

$$\begin{aligned} f(\Omega) &= K - \mu\Omega^2 + M\Omega^4 \\ g(\Omega) &= i(\sigma \cdot \Delta k - 2s\kappa)\Omega - i\sigma \cdot \Delta m \Omega^3 \end{aligned}$$

we have first-order interest in the roots  $\Omega = \Omega_0 + \rho \cdot \Omega_1$  of  $f(\Omega) + \rho \cdot g(\Omega)$ , which according to (122) are given by

$$\begin{aligned} &\pm \Omega_{\text{fast}} + \rho \cdot i\Gamma_{\text{fast}} \\ \Gamma_{\text{fast}} &= + \frac{\sigma[\Delta m(\mu + \sqrt{\mu^2 - 4KM}) - 2M\Delta k] + s[4\kappa M]}{4M\sqrt{\mu^2 - 4KM}} \end{aligned}$$

and

$$\begin{aligned} &\pm \Omega_{\text{slow}} + \rho \cdot i\Gamma_{\text{slow}} \\ \Gamma_{\text{slow}} &= - \frac{\sigma[\Delta m(\mu - \sqrt{\mu^2 - 4KM}) - 2M\Delta k] + s[4\kappa M]}{4M\sqrt{\mu^2 - 4KM}} \end{aligned}$$

where

$$\begin{aligned} \Delta m &\equiv m_1 - m_2 \\ &= B - A \end{aligned}$$

and

$$\begin{aligned} \Delta k &= k_1 - k_2 \\ &= B - A \end{aligned}$$

are, by our assumptions, positive.



The stable solutions of (123) die exponentially, which is to say: they lose energy—an effect we normally attribute to dissipation. Odd in the present instance, since *no dissipation mechanism was built into the model*. The resolution of this little paradox must lie in the circumstance that energy leaks into higher-order aspects of the motion which the first-order theory is powerless to take into account...just as the first-order theory is powerless to temper the seeming “explosions” in the unstable case.<sup>41</sup>

I draw attention finally to the fact that it is the  $\mathbb{S}$ -term in (115) that accounts for celtic chirality. One has

$$\lim_{\rho \downarrow 0} \mathbb{S} = \begin{pmatrix} 0 & s \\ s & 0 \end{pmatrix} : \text{ resembles } \mathbb{A} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$$

where both  $s$  and  $a$  depend linearly upon  $\omega$ . But in the 0<sup>th</sup>-order theory one encounters only  $s^2$  and  $a^2$ , both of which are insensitive to the sign of  $\omega$ . In first order one encounters  $s$  (also  $\sigma$ ) and  $a^2$ :  $\mathbb{A}$  is again blameless.

**17. Deformable bodies.** Reasonably good approximations to rigid bodies can be found in toy rooms, on ball fields, on battle fields (among the small hunks of matter hurled back and forth), in the astroid belt. But the earth is “rigid” only in zeroth approximation: the circulation of atmosphere, oceans—even, on a longer time scale, of continents—has an observable effect on its rotational dynamics. Similar effects afflict almost all astrophysical bodies, and pertain with especial importance to systems of *interacting* deformable bodies: it is because the moon is deformable that it has come to present always the same face to earth. Even an isolated body, made of material of finite strength, will deform (expand at the equator) *in response* to its own rotation. Acrobats and divers are deformation virtuosi: they tour many points in “shape space” with rapid precision, and can never be accused of suffering from rigor mortis. We evidently stand in need of a rotational dynamics of deformable bodies. But confront at the outset several perplexing questions:

Once we abandon the concept of “rigidity,” what is left? Cannot *every* isolated many-body system be considered to be a “deformable body”? How are we to give physical meaning to our intuitive sense that some bodies are “almost rigid”? Can it be of dynamical relevance that the deformable bodies that come most naturally to mind possess (in leading approximation) a well-defined and shape-independent fixed volume? How many degrees of freedom has a deformable body? The number appears to be indefinite: a swarm of  $N$  bees has  $3N$  degrees of freedom, but the system comprised of a bug walking around on a rigid sphere has only eight (six for the sphere, two for the bug). The great simplification brought to rigid body dynamics by Chasle’s theorem is clearly no

---

<sup>41</sup> Though Hoyle works in second order, he does not escape the paradox: escape would appear to require that one work not in truncated order, but exactly.

longer operative. Finally, how—if at all—is one to attach a “body frame” to a deformable body?<sup>42</sup>

We begin by looking to one respect in which the rotational physics of deformable bodies differs profoundly from that of rigid bodies. Consider the mechanism shown in FIGURES 35. No external forces/torques are impressed upon the device, but an internal energy supply enables it to

- flex its elbows
- twist at the waist

We have

$$\varphi_1 - \varphi_2 = \alpha \quad \text{whence} \quad \dot{\varphi}_1 - \dot{\varphi}_2 = \dot{\alpha} \quad (124.1)$$

while by angular momentum conservation<sup>43</sup>

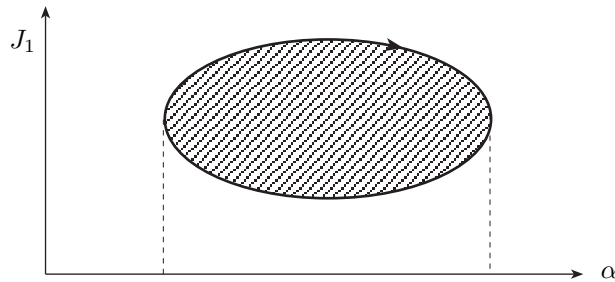
$$I_1 \dot{\varphi}_1 + I_2 \dot{\varphi}_2 = 0 \quad (124.2)$$

It follows that

$$\left. \begin{aligned} \dot{\varphi}_1 &= + \frac{I_2}{I_1 + I_2} \dot{\alpha} \\ \dot{\varphi}_2 &= - \frac{I_1}{I_1 + I_2} \dot{\alpha} \end{aligned} \right\} \iff \left\{ \begin{aligned} d\varphi_1 &= + \frac{I_2}{I_1 + I_2} d\alpha \\ d\varphi_2 &= - \frac{I_1}{I_1 + I_2} d\alpha \end{aligned} \right. \quad (125)$$

from which (124) are readily recovered as corollaries. Suppose now that the device has been programmed so as to cause  $\alpha$  to increase/decrease periodically, and to synchronously flex its elbows in such a way as to make

$$J_1 \equiv I_2/(I_1 + I_2) \begin{cases} \text{large when } \dot{\alpha} > 0 : \text{“inhaling”} \\ \text{small when } \dot{\alpha} < 0 : \text{“exhaling”} \end{cases}$$



<sup>42</sup> This last problem has been addressed in a profound way by A. Shapere & F. Wilczek in “Gauge kinematics of deformable bodies,” AJP **57**, 514 (1989). The paper appears also as §8.3 in *Geometric Phases in Physics* (1989), which they edited, and provides the basis of the discussion which begins on page 89 below.

<sup>43</sup> Actually, *spin* conservation. Without essential loss of generality we will assume that initially—and therefore for all time, in the continued absence of impressed torques— $\mathbf{S} = \mathbf{0}$ .

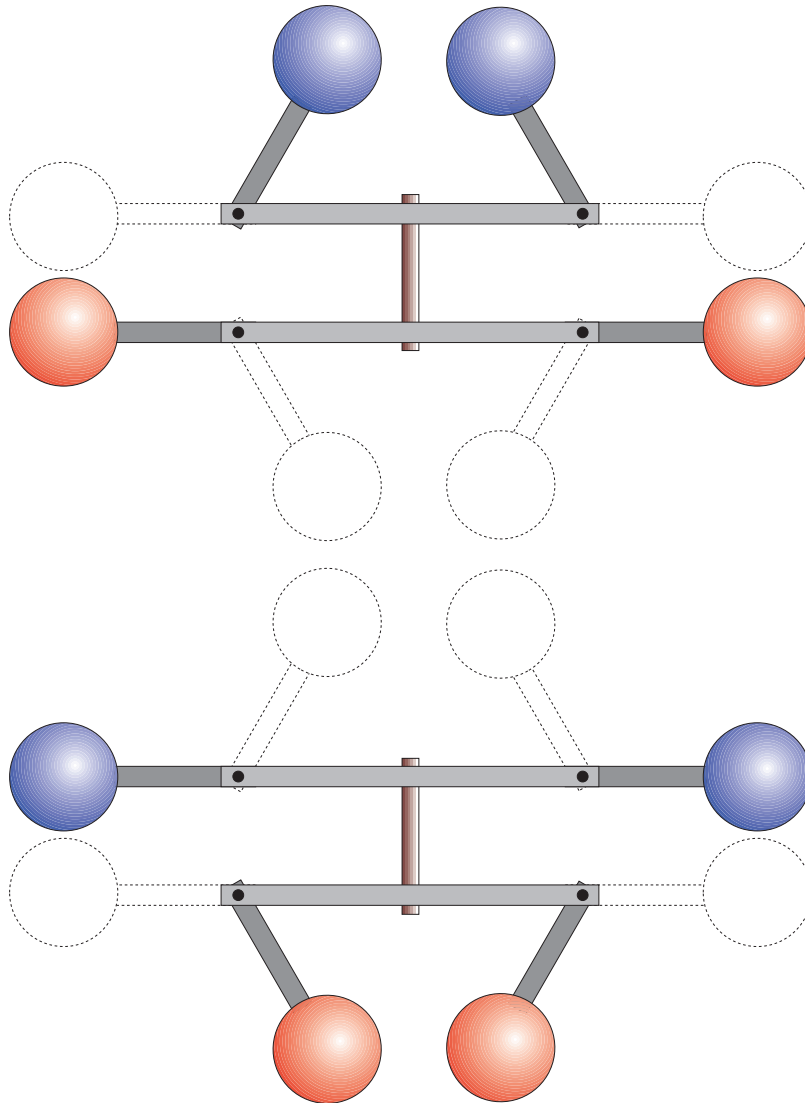


FIGURE 35A: Side views of the automaton discussed in the text, showing (above) the elbows as they might be flexed when the device is “inhaling” ( $\dot{\alpha} > 0$ ) and (below) as they might be flexed when the device is “exhaling” ( $\dot{\alpha} < 0$ ).

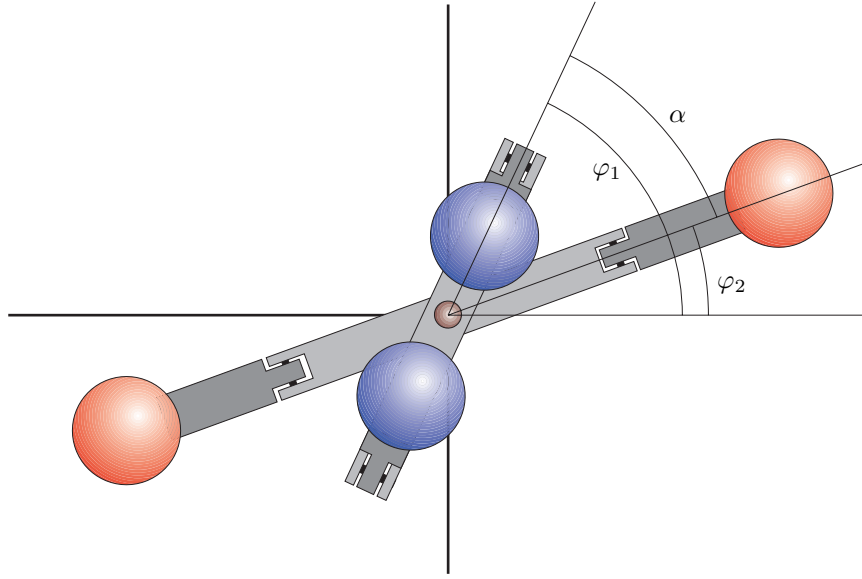


FIGURE 35B: View from above of the automaton when configured as shown at the top of the preceding figure. All angles are to be assigned the same counterclockwise sense. By repeatedly performing its carefully choreographed routine the device manages to rotate  $\odot$  without standing ever in violation of the condition  $\mathbf{S} = \mathbf{0}$ .

It becomes immediately evident from the diagram at the bottom of page 85 that the device achieves a net  $\varphi_1$ -advance per  $\alpha$ -cycle (and an identical  $\varphi_2$ -advance) that can be described

$$\Delta\varphi_1 = \oint J_1(\alpha) d\alpha \quad (126)$$

It is clear also that  $\Delta\varphi_1$  is independent of all *temporal* aspects of the cycle: in that respect the striking phenomenon here at issue is (as Shapere & Wilczek have emphasized) not so much “physical” as “geometrical.”

Suppose, for example, that the elbows of our automaton flex in such a way as to achieve

$$\begin{aligned} I_1(t) &= a - b \cos \omega t \\ I_2(t) &= a + b \cos \omega t \end{aligned} \quad (127.1)$$

and that the breathing of  $\alpha$  can be described (see FIGURE 36)

$$\alpha(t) = \frac{1}{2} \alpha_{\max} (1 + \sin \omega t) \quad (127.2)$$

We then—by (125)—have

$$\begin{aligned} \varphi(t) &= \int_0^t \frac{a + b \cos \omega t}{4a} \alpha_{\max} \omega \cos \omega t dt \\ &= \frac{\alpha_{\max}}{4a} \left[ \frac{1}{2} b \omega t + a \sin \omega t + b \sin 2\omega t \right] \\ &= \frac{b}{8a} \alpha_{\max} \cdot \omega t + \text{oscillatory term} \end{aligned}$$

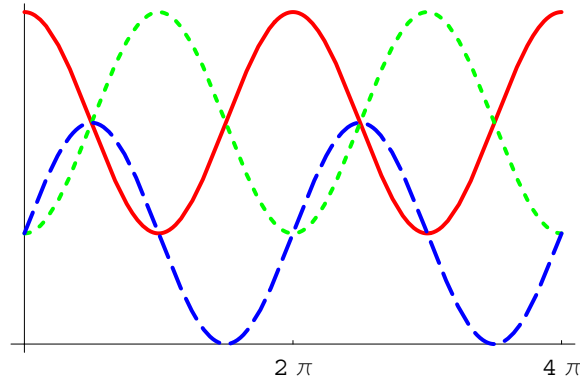


FIGURE 36: The **green** curve (short dashes) describes the motion of  $I_1$ , the solid **red** curve the motion of  $I_2$ , the **blue** curve (long dashes) the motion of  $\alpha$ . Notice that  $I_2$  is dominant when  $\alpha$  is increasing,  $I_1$  is dominant when  $\alpha$  is decreasing. In constructing the figure I have set  $a = 2$ ,  $b = \omega = \alpha_{\max} = 1$ .

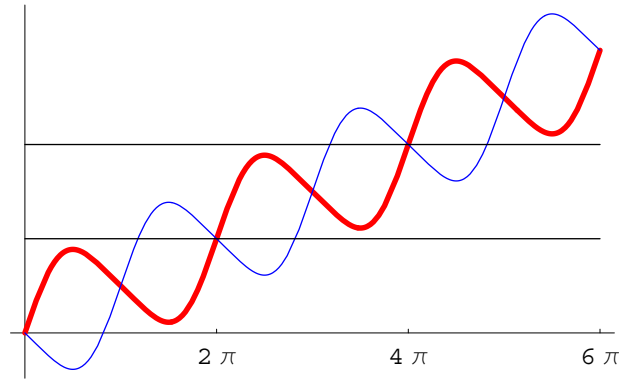


FIGURE 37: Graph of the resulting motion of  $\varphi_1$ . The horizontal lines indicate the  $\varphi_1$ -advance/period, which in the present instance is  $\Delta\varphi_1 = \pi/8$ . The finer blue curve traces the advance of  $\varphi_2$ , which for obvious reasons must achieve the same advance per period.

which yields a  $\varphi_1$ -advance per period (equal necessarily to the  $\varphi_2$ -advance per period) given by

$$\Delta\varphi = \frac{\pi}{4}(b/a)\alpha_{\max}$$

Notice that  $t$  enters into the preceding integral only *via* the dimensionless product  $\omega t$ : it is for this reason that  $\Delta\varphi$  is independent of the time  $\tau = 2\pi/\omega$  that it takes for the device to complete a stroke, a deformation cycle.



Though striking, the phenomenon of **rotation without angular momentum** accounts for how dropped cats manage to land on their feet,<sup>44</sup> and is in evidence whenever a diver departs the diving platform with zero angular momentum and yet manages to perform complicated somersaults.<sup>45</sup> I myself first became interested in the phenomenon when, in 1956, I attended a lecture by Thomas Gold in which he advanced the thesis that in view of the plasticity of the earth one should expect to find paleomagnetic evidence of large-scale polar wander (mantle shifting with respect to the spin axis) over geologic time.<sup>46,47</sup>

I propose now to discuss the operation of our automaton from the more readily generalizable point of view advocated by Shapere & Wilczek.<sup>42</sup> The device has five adjustable features:  $\alpha$  and the four elbow angles  $\beta_1, \beta_2, \beta_3$  and  $\beta_4$ . But if we insist that deformations preserve (relative to the device itself) the location of the center of mass then three of the elbow angles become slaves of the fourth (see FIGURE 38). To describe the shape of the device it would suffice then to specify the values of  $\alpha$  and  $\beta$ : though the “shape space”  $\mathcal{S}$  of deformable bodies can, in general, be expected to be  $\infty$ -dimensional, it is in the present instance only 2-dimensional. We would write  $\{\alpha(t), \beta(t)\}$ —which is to say: we would inscribe a  $t$ -parameterized curve  $\mathcal{C}$  on shape space—to describe a temporal sequence of deformations. If the deformations are cyclic (as deformations with locomotive intent tend to be) then  $\mathcal{C}$  would have the form of a closed loop. Note, however, that cyclicity does not, of itself, imply temporal periodicity.

Imagine now that onto each shaped object we have—for future reference—stamped a Cartesian frame, with origin at the center of mass. *How* this is accomplished is a matter of fundamental indifference (see FIGURE 39), though some frame-assignment procedures (adoption of the principal axis frame?) may prove more useful—or at least feel more natural—in specific contexts. Clearly, a rotation-matrix-valued function  $\mathbb{Q}(\alpha, \beta)$  would serve to relate any such frame assignment to any alternative assignment.

Consider now a blob—our carefully crafted device has, for the purposes

---

<sup>44</sup> T. R. Kane & M. P. Scher, “A dynamical explanation of the falling cat phenomenon,” *J. Solids Struct.* **5**, 663 (1969).

<sup>45</sup> C. Frohlich, “The physics of somersaulting & twisting,” *Scientific American* **263**, 155 (March 1980); “Do springboard divers violate angular momentum conservation?” *AJP* **47**, 583 (1979).

<sup>46</sup> T. Gold, “Instability of the earth’s axis of rotation,” *Nature* **175**, 526 (1955).

<sup>47</sup> It is important to notice that, while **one can, by contortion, rotate about one’s center of mass**/change the way one faces in inertial space/“translate in an angular sense” without external assistance, **one cannot, by any amount of contortion, translate one’s center of mass**. And it is in this light interesting that, according to Jack Wisdom (“Swimming in spacetime: motion by cyclic changes in body shape,” *Science* **299**, 1865 (March 2003)), one *can* do so in *curved* spacetime: the effect is relativistic (disappears in the limit  $c \uparrow \infty$ ), and such swimming becomes impossible in flat spacetime.

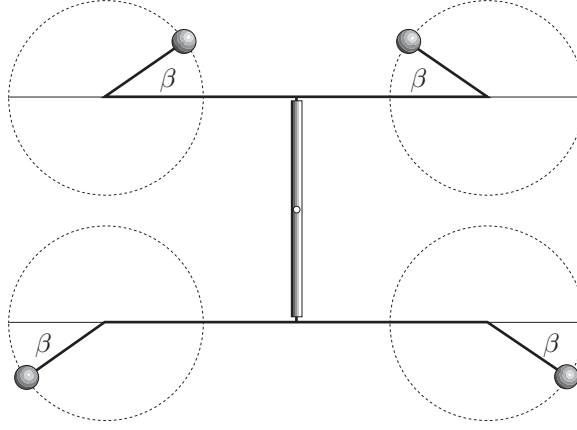


FIGURE 38: *If the automaton flexes its elbows in such a way as to preserve the location of its center of mass, then all four elbows must be under the control of a single parameter,  $\beta$ .*

of this discussion, become a generic “blob”—which moves through some  $t$ -parameterized continuous sequence of shapes:

$$\{\alpha, \beta\}_{\text{initial}} \xrightarrow{\{\alpha(t), \beta(t)\}} \{\alpha, \beta\}_{\text{final}}$$

Identifying the frame attached to  $\{\alpha, \beta\}_{\text{initial}}$  with a fixed reference frame in physical space (the **space frame**), we introduce  $\mathbb{R}(t)$  to describe the relationship of that frame to the frame carried by the deformed blob after it has been transported along  $\mathcal{C}$  to the shape  $\{\alpha(t), \beta(t)\}$  (see FIGURE 40). The orientation of the deformed blob is determined by *physical* principle (conservation of angular momentum), but the way it wears its frame is arbitrary. Our problem is to find some way to distinguish what’s physical from what’s merely conventional.

From  $\mathbb{R}(t)^\top \mathbb{R}(t) = \mathbb{I}$  it follows familiarly that it is always possible to write

$$\dot{\mathbb{R}} = \mathbb{A} \mathbb{R} \quad \text{with} \quad \mathbb{A}^\top = -\mathbb{A}$$

or—which is to say the same thing another way—

$$\mathbb{R}(t) = \mathbb{R}_0 + \int_0^t \mathbb{A}(t') \mathbb{R}(t') dt'$$

Now, the solution of  $\dot{r}(t) = a(t)r(t)$  (*i.e.*, of  $r(t) = r_0 + \int_0^t a(t')r(t') dt'$ ) is easily seen to be

$$r(t) = \exp \left\{ \int_0^t a(t') dt' \right\} \cdot r_0 \quad (128)$$

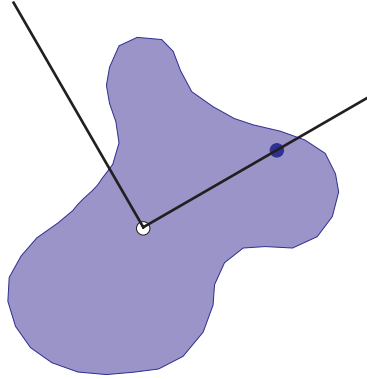


FIGURE 39: To assign frames to the various shapes of a 2-blob we have here used center of mass  $\circ$  and a pimple  $\bullet$  as our guide—a procedure that does not work for 3-blobs. We might alternatively have adopted (say) the principal axis frame, though such a procedure would become ambiguous when the principal moments are identical.

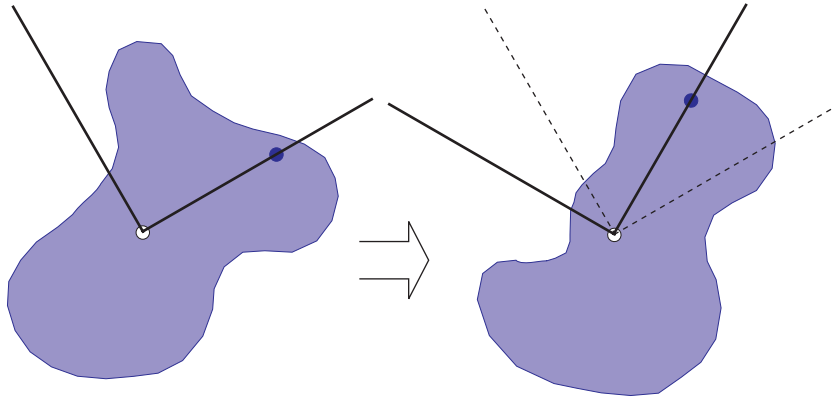


FIGURE 40: A rotation matrix  $\mathbb{R}(t)$  relates the frame of a deformed blob to the frame of the original blob.

in which connection we note that a temporal rescaling  $t \rightarrow \tau = \tau(t)$  sends

$$\frac{d}{dt}r(t) = a(t)r(t) \implies \frac{d}{d\tau}R(\tau) = A(\tau)R(\tau)$$

with  $R(\tau) \equiv r(t)$  and

$$A(\tau) = a(t)/\dot{\tau}(t)$$

and this, by  $d\tau = \dot{\tau}(t)dt$ , means that the solution of  $\dot{r} = ar$  is scale-invariant:

$$\int_0^\tau A(\tau) d\tau = \int_0^t a(t) dt$$

A similar remark pertains to the iterative solution<sup>48</sup>

$$\begin{aligned}\mathbb{R}(t) &= \left\{ \mathbb{I} + \int_0^t \mathbb{A}(t') dt' + \int_0^t \int_0^{t'} \mathbb{A}(t') \mathbb{A}(t'') dt' dt'' + \dots \right\} \cdot \mathbb{R}_0 \\ &= \mathcal{P} \exp \left\{ \int_0^t \mathbb{A}(t) dt \right\} \cdot \mathbb{R}_0\end{aligned}\quad (129)$$

of  $\hat{\mathbb{R}} = \mathbb{A}\mathbb{R}$ . Shapere & Wilczek interpret this to mean that the  $\mathbb{R}(t)$  that results from transport along a “curve in shape space” depends not at all upon temporal specifics of the process, but only upon the geometry of the curve  $\mathcal{C}$ . Equation (129) makes clear also that specification of  $\mathbb{A}(t)$  is sufficient in principle to determine  $\mathbb{R}(t)$ .

Our further progress will be facilitated by some notational adjustment. Let us write  $\alpha^1 = \alpha$ ,  $\alpha^2 = \beta$ , and let us recognize that  $\mathbb{R}(t)$  means  $\mathbb{R}(\alpha^1(t), \alpha^2(t))$  which we will abbreviate  $\mathbb{R}(\alpha(t))$ . Then

$$d\mathbb{R} = \mathbb{A}\mathbb{R} dt \quad \text{becomes} \quad d\mathbb{R} = \mathbb{R}_{,i} \dot{\alpha}^i dt = \mathbb{R}_{,i} d\alpha^i \quad (130)$$

where  $\mathbb{R}_{,i} \equiv \partial_i \mathbb{R} \equiv \partial \mathbb{R}(\alpha) / \partial \alpha^i$  and where the Einstein summation convention is understood to be in force. The differential  $d\alpha^i$  is tangent to  $\mathcal{C}$  at the point  $\alpha$ .

Immediately we confront a major problem: if we exercise our **local frame reassignment option**

$$\mathbb{R}(\alpha) \mapsto \hat{\mathbb{R}}(\alpha) = \mathbb{Q}(\alpha) \mathbb{R}(\alpha)$$

then

$$\hat{\mathbb{R}}_{,i} = \mathbb{Q} \mathbb{R}_{,i} + \mathbb{Q}_{,i} \mathbb{R}$$

and the added term destroys the “reassignment covariance” of (130). To remedy this defect we resort to a standard device: we introduce “compensating terms” or “gauge fields”  $\mathbb{A}_i$ —one for each degree of freedom in shape space—writing

$$\mathbb{R}_{,i} \equiv \hat{\mathbb{R}}_{,i} - \mathbb{A}_i \mathbb{R} \quad (131)$$

We then have

$$\begin{aligned}\mathbb{Q} \mathbb{R}_{,i} &= \mathbb{Q} \hat{\mathbb{R}}_{,i} - \mathbb{Q} \mathbb{A}_i \mathbb{R} \\ &= (\hat{\mathbb{R}}_{,i} - \mathbb{Q}_{,i} \mathbb{R}) - \mathbb{Q} \mathbb{A}_i \mathbb{R} \\ &= (\hat{\mathbb{R}}_{,i} - \mathbb{Q}_{,i} \mathbb{Q}^{-1} \hat{\mathbb{R}}) - \mathbb{Q} \mathbb{A}_i \mathbb{Q}^{-1} \hat{\mathbb{R}}\end{aligned}$$

and insist upon

$$= \hat{\mathbb{R}}_{,i} - \hat{\mathbb{A}}_i \hat{\mathbb{R}}$$

---

<sup>48</sup> This is obtained by iteration of  $\mathbb{R}(t) = \mathbb{R}_0 + \int_0^t \mathbb{A}(t') \mathbb{R}(t') dt'$ , and gives back (128) when all  $\mathbb{A}(t)$ -matrices commute with one another. Here  $\mathcal{P}$  is the “chronological ordering” operator, the characteristic action of which becomes evident from

$$\mathcal{P}[\mathbb{A}(t_1) \mathbb{A}(t_2)] = \begin{cases} \mathbb{A}(t_1) \mathbb{A}(t_2) & : t_1 \geq t_2 \\ \mathbb{A}(t_2) \mathbb{A}(t_1) & : t_2 \geq t_1 \end{cases}$$

which entails setting

$$\hat{\mathbb{A}}_i = \mathbb{Q}\mathbb{A}_i\mathbb{Q}^{-1} + \mathbb{Q}_{,i}\mathbb{Q}^{-1} \quad (132)$$

In short: if the “covariant derivative”  $\mathbb{R}_{;i}$  is understood to be defined by (131), and if the gauge fields are understood to transform  $\mathbb{A}_i \mapsto \hat{\mathbb{A}}_i$  by the rule (132), then

$$\mathbb{R} \mapsto \hat{\mathbb{R}} = \mathbb{Q}\mathbb{R} \quad \text{induces} \quad \mathbb{R}_{;i} \mapsto \hat{\mathbb{R}}_{;i} = \mathbb{Q}\mathbb{R}_{;i} \quad (133)$$

Notice also that, since  $\mathbb{Q}$  is a rotation matrix ( $\mathbb{Q}^{-1} = \mathbb{Q}^\top$ ), it is an implication of (132) that *if the matrices  $\mathbb{A}_i$  are antisymmetric, then so are the matrices  $\hat{\mathbb{A}}_i$* .

Associated with every infinitesimal displacement in shape space  $\alpha \rightarrow d\alpha$  are not one but *two distinct types of infinitesimal rotation*. On the one hand we have

$$\mathbb{R}(\alpha) \longrightarrow \mathbb{R}(\alpha + d\alpha) = \mathbb{R}(\alpha) + \mathbb{R}_{;i}(\alpha)d\alpha^i \quad (134.1)$$

that refers straightforwardly to the gradient structure of the frame-field that we have (arbitrarily) deposited on shape space. On the other hand, we have<sup>49</sup>

$$\begin{aligned} \mathbb{R}(\alpha) \longrightarrow \mathbb{R}(\alpha + d\alpha) &= \mathbb{R}(\alpha + d\alpha) - \mathbb{R}_{;i}(\alpha)d\alpha^i \\ &= \mathbb{R}(\alpha) + \mathbb{A}_i(\alpha)\mathbb{R}(\alpha)d\alpha^i \end{aligned} \quad (134.2)$$

the precise meaning of which depends upon the structure assigned to the gauge fields  $\mathbb{A}_i(\alpha)$ , which are constrained only by (132): it is here—by contrived specification of  $\mathbb{A}_i(\alpha)$ —that we will have an opportunity to slip some *physics* into this formal scheme. The matrices  $\mathbb{R}(\alpha + d\alpha)$  and  $\mathbb{R}(\alpha + d\alpha)$  will, in general, be distinct. If, however, they are identical then we say that  $\mathbb{R}(\alpha) \longrightarrow \mathbb{R}(\alpha + d\alpha)$  has proceeded by **parallel transport**, and can write

$$\mathbb{R}(\alpha + d\alpha) = \mathbb{R}(\alpha) + \mathbb{A}_i(\alpha)\mathbb{R}(\alpha)d\alpha^i$$

If the parallel transport is along a curve  $\alpha(t)$  in shape space—here  $t$  might but *need not* signify time—then we have  $\mathbb{R}(t + dt) = \mathbb{R}(t) + \mathbb{A}_i(t)\mathbb{R}(t)\dot{\alpha}^i dt$  or

$$\dot{\mathbb{R}} = \mathbb{A}\mathbb{R} \quad \text{with} \quad \mathbb{A} \equiv \mathbb{A}_i\dot{\alpha}^i \quad (135)$$

If we assume without real loss of generality that  $\mathbb{R}(0) = \mathbb{I}$  then, by (129), we have

$$\mathbb{R}(t) = \left\{ \mathbb{I} + \int_0^t \mathbb{A}(t')dt' + \frac{1}{2} \int_0^t \int_0^t \mathcal{P}[\mathbb{A}(t')\mathbb{A}(t'')] dt'dt'' + \dots \right\} \quad (136)$$

---

<sup>49</sup> Said another way, we have

$$\begin{aligned} \mathbb{R}(\alpha + d\alpha) - \mathbb{R}(\alpha) &= \mathbb{R}_{;i}(\alpha)d\alpha^i \\ \mathbb{R}(\alpha + d\alpha) - \mathbb{R}(\alpha + d\alpha) &= \mathbb{R}_{;i}(\alpha)d\alpha^i \end{aligned}$$

It is from the circumstance that both matrices on the left side of the second equation attach to the *same point* in shape space that  $\mathbb{R}_{;i}(\alpha)$  acquires its superior transformation properties.

It is important to appreciate that  $\mathbb{R}(t)$  is a path-dependent object: it is the result of parallel-transporting  $\mathbb{I}$  from  $\alpha(0)$  to  $\alpha(t)$  along path  $\mathcal{C}$ . Transport along a different path linking the same endpoints can in general be expected to yield a different result. To get a handle on the situation we ask: *What is the difference  $\delta\mathbb{R} = \mathbb{R}_2 - \mathbb{R}_1$  that results when  $\mathcal{C}_2$  and  $\mathcal{C}_1$  differ only infinitesimally?*

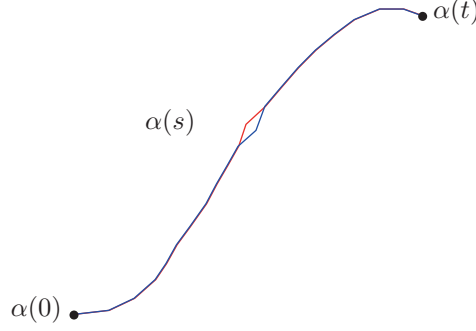


FIGURE 41: *Two curves inscribed on shape space that are coincident except in the localized neighborhood of a pimple, where they differ only infinitesimally.*

Assuming the curves to differ only in the neighborhood of  $\alpha(s)$ , we have

$$\delta_s \mathbb{R} = \mathcal{P} \exp \left\{ \int_s^t \mathbb{A}(\tau) d\tau \right\} \cdot \delta \mathbb{A}(s) \cdot \mathcal{P} \exp \left\{ \int_0^s \mathbb{A}(\tau) d\tau \right\}$$

and to distribute such pimples along the length of the curve we have only to write

$$\delta \mathbb{R} = \int_0^t \left[ \mathcal{P} \exp \left\{ \int_s^t \mathbb{A}(\tau) d\tau \right\} \cdot \delta \mathbb{A}(s) \cdot \mathcal{P} \exp \left\{ \int_0^s \mathbb{A}(\tau) d\tau \right\} \right] ds$$

But

$$\begin{aligned} \delta \mathbb{A}(s) &= \mathbb{A}_j(s) \cdot \delta \dot{\alpha}^j(s) + \dot{\alpha}^i(s) \cdot \delta \mathbb{A}_i(s) \\ &= \mathbb{A}_j \cdot \frac{d}{ds} \delta \alpha^j(s) + \dot{\alpha}^i(s) \cdot \frac{\partial \mathbb{A}_i}{\partial \alpha^j} \delta \alpha^j(s) \end{aligned}$$

so

$$\begin{aligned} \delta \mathbb{R} &= \int_0^t \left[ \mathcal{P} \exp \left\{ \int_s^t \text{etc.} \right\} \cdot \mathbb{A}_j \frac{d}{ds} \delta \alpha^j \cdot \mathcal{P} \exp \left\{ \int_0^s \text{etc.} \right\} \right] ds \\ &\quad + \int_0^t \left[ \mathcal{P} \exp \left\{ \int_s^t \text{etc.} \right\} \cdot \frac{\partial \mathbb{A}_i}{\partial \alpha^j} \dot{\alpha}^i \cdot \mathcal{P} \exp \left\{ \int_0^s \text{etc.} \right\} \right] \delta \alpha^j(s) ds \end{aligned}$$

where it is understood that the factors between dots are to be evaluated at  $s$ . The first term we integrate by parts to obtain (after noting that by assumption  $\delta \alpha^i(s)$  vanishes at the endpoints:  $\delta \alpha^i(0) = \delta \alpha^i(t) = 0$ )

$$- \int_0^t \left[ \mathcal{P} \exp \left\{ \int_s^t \text{etc.} \right\} \left( -\mathbb{A}_i \dot{\alpha}^i \mathbb{A}_j + \frac{\partial \mathbb{A}_j}{\partial \alpha^i} \dot{\alpha}^i + \mathbb{A}_j \mathbb{A}_i \dot{\alpha}^i \right) \mathcal{P} \exp \left\{ \int_0^s \text{etc.} \right\} \right] \delta \alpha^j(s) ds$$

giving finally

$$\delta \mathbb{R} = \int_0^t \delta \alpha^j(s) \left[ \mathcal{P} \exp \left\{ \int_s^t etc. \right\} \cdot \mathbb{F}_{ji} \cdot \mathcal{P} \exp \left\{ \int_0^s etc. \right\} \right] \dot{\alpha}^i(s) ds \quad (137)$$

where  $\mathbb{F}_{ij}$  refers to the antisymmetric array of  $3 \times 3$  antisymmetric matrices defined

$$\begin{aligned} \mathbb{F}_{ij} &\equiv \partial_i \mathbb{A}_j - \partial_j \mathbb{A}_i - [\mathbb{A}_i, \mathbb{A}_j] \\ &= -\mathbb{F}_{ji} \end{aligned} \quad (138)$$

We have come here—by an argument adapted from a paper published by Peter G. Bergman<sup>50</sup>—upon a particular manifestation of object known to differential geometers and general relativists as the “Riemann-Christoffel curvature tensor” and to field theorists as the “gauge field tensor.”<sup>51</sup> Of the many remarkable properties with which  $\mathbb{F}_{ij}$  is endowed, I will mention only one: working from (132) we compute

$$[\hat{\mathbb{A}}_i, \hat{\mathbb{A}}_j] = \mathbb{Q} [\mathbb{A}_i, \mathbb{A}_j] \mathbb{Q}^{-1} + \boxed{\text{stuff}}$$

and<sup>52</sup>

$$\partial_i \hat{A}_j - \partial_j \hat{A}_i = \mathbb{Q} (\partial_i A_j - \partial_j A_i) \mathbb{Q}^{-1} + \boxed{\text{same stuff!}}$$

—the implication being that  $\mathbb{F}_{ij}$  responds tensorially to gauge transformations:

$$\hat{\mathbb{F}}_{ij} = \mathbb{Q} \mathbb{F}_{ij} \mathbb{Q}^{-1} \quad (139)$$

---

<sup>50</sup> “On Einstein’s  $\lambda$  transformations,” Phys. Rev. **103**, 780 (1956). For my immediate source, see pages 134–137 in Chapter 2 of CLASSICAL DYNAMICS (1964/65).

<sup>51</sup> Note in this connection that if the matrices  $\mathbb{A}_i$  and  $\mathbb{Q}$  were number-valued instead of matrix-valued, then (132) would read

$$\hat{A}_i = A_i + \partial_i Q$$

and (138) would become

$$F_{ij} = \partial_i A_j - \partial_j A_i = \hat{F}_{ij}$$

These are equations that we recognize to be fundamental to electrodynamics. “Non-abelian gauge field theory” is a generalization of electrodynamics in which importance is assigned to the *non-commutativity* the gauge fields  $\mathbb{A}_i$ .

<sup>52</sup> Here one has need of  $(\mathbb{Q}^{-1})_{,i} = -\mathbb{Q}^{-1} \mathbb{Q}_{,i} \mathbb{Q}^{-1}$ , which follows directly from  $(\mathbb{Q}^{-1} \mathbb{Q})_{,i} = \mathbb{O}$  and is the non-commutative analog of  $\frac{d}{dt} q^{-1} = -q^{-2} \dot{q}$ .

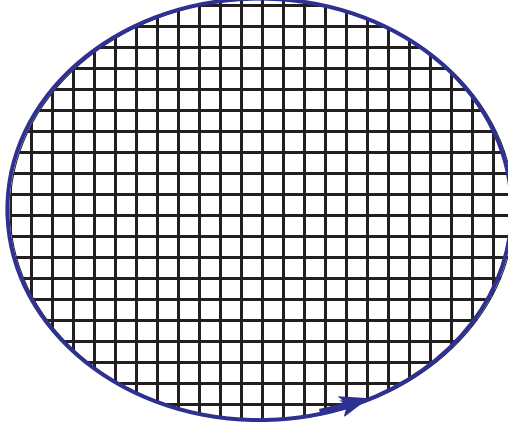


FIGURE 42: A closed curve  $\mathcal{C} = \partial\mathcal{R}$  bounds a region  $\mathcal{R}$  in shape space, which has been resolved into differential patches. Parallel transport around  $\mathcal{C}$  can be achieved by superimposing the results of transport around each of the patches.

At (137) we managed to establish in effect that the result of parallel transport around a differential patch (or “pimple”) at shape  $\alpha$  can be described

$$\delta\mathbb{R} = \mathbb{F}_{ij}(\alpha) d\alpha^i \wedge d\alpha^j$$

We conclude (see the preceding figure) that transport of  $\mathbb{I}$  around a finite closed curve  $\mathcal{C}$ —a cycle of shapes—has a rotational consequence that can be described

$$\mathbb{R}_{\mathcal{C}} \equiv \mathcal{P} \exp \left\{ \oint_{\mathcal{C}} \mathbb{A}(\tau) d\tau \right\} \mathbb{I} = \iint_{\mathcal{R}} \mathbb{F}_{ij}(\alpha) d\alpha^i \wedge d\alpha^j$$

and that (since the initial and final reference frames are *identical*)  $\mathbb{R}_{\mathcal{C}}$  is gauge invariant.<sup>53</sup>

We have now to pour some physics into the mathematical vessel that we have been at such pains to construct. Erect an inertial frame at (let us say) the center of mass of a (let us say) spinless system of particles:

$$\mathbf{S} = \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i = \mathbf{0}$$

<sup>53</sup> Recall from electrodynamics that

$$\oint_{\partial\mathcal{R}} \mathbf{A} \cdot d\mathbf{s} = \iint_{\mathcal{R}} (A_{i,j} - A_{j,i}) dx^i \wedge dx^j$$

is invariant under  $\mathbf{A} \mapsto \mathbf{A} + \nabla Q$ .



Imagine now that we have—in some continuous but otherwise arbitrary way—associated a frame (origin coincident with that of the space frame) with every possible configuration of the system. At time  $t$  the system finds itself in some specific configuration, to which we have associated a frame. We agree to write  $\mathbf{r}_i$  to describe the position of  $m_i$  relative to that momentary frame, and

$$\mathbf{r}_i = \mathbb{R}(t) \mathbf{r}_i$$

to describe the relationship of that frame to the space frame. In (non-inertial) red variables the statement  $\mathbf{S} = \mathbf{0}$  becomes

$$\begin{aligned} \sum_i m_i [\mathbb{R} \mathbf{r}_i] \times [\mathbb{R} \dot{\mathbf{r}}_i + \dot{\mathbb{R}} \mathbf{r}_i] &= \mathbf{0} \\ &= \mathbb{R} \left\{ \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i + \sum_i m_i \mathbf{r}_i \times \mathbb{R}^{-1} \dot{\mathbb{R}} \mathbf{r}_i \right\} \end{aligned}$$

Because  $\mathbb{R}$  is a rotation matrix we have  $\dot{\mathbb{R}} = \mathbb{A} \mathbb{R} = \mathbb{R} \mathbb{B}$  where  $\mathbb{A}$  and  $\mathbb{B}$  are both antisymmetric, but generally distinct. Writing  $\mathbb{B} = \boldsymbol{\Omega} \times$ , we have (see again page 5)

$$\sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i + \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\Omega} \times \mathbf{r}_i) = \mathbf{S} + \mathbb{I} \boldsymbol{\Omega} = \mathbf{0}$$

giving

$$\boldsymbol{\Omega} = -\mathbb{I}^{-1} \mathbf{S}$$

whence<sup>54</sup>

$$\mathbb{B} = \|B_{ij}\| \quad \text{with} \quad B_{ij} = \epsilon_{ijk} (\mathbb{I}^{-1} \mathbf{S})_k$$

At this point we

- can (but are under no obligation to) identify the particles  $m_i$  with the component parts of our deformable blob;
- can (but are under no obligation to) take our configuration-associated frames to be principal axis frames.

Whatever our position with respect to the exercise of those options, we

- extract from the physics of the system a description of (compare (135))  $\mathbb{B} = \mathbb{B}_i \dot{\alpha}^i$  whence of  $\mathbb{B}_i$  (which will be defined not everywhere in shape space, but only where it is needed: on the curve pursued by the system), with the aid of which we play the parallel transport game. Should the system ever *revisit a point in shape space* we will be able to announce whether it has experienced a net rotation as a result of its dynamical zero-spin adventures (contortions).

---

<sup>54</sup> It becomes clear only at this point why Shapere & Wilczek look to the iterative solution of  $\dot{\mathbb{R}} = \mathbb{R} \mathbb{B}$  rather than (which is more common) of  $\dot{\mathbb{R}} = \mathbb{A} \mathbb{R}$ . The resulting formalism is literally the *transpose* of that described on pages 90–96. Transposition entails *reversal of the chronological ordering*.

It is an ambitious program, which I can expect to be computationally feasible only in the simplest cases.<sup>55</sup> It would be interesting to see whether it can be brought to bear on the motion of our simple automaton (FIGURES 35) which, as we have seen, admits of detailed analysis by elementary means. But that is an exercise I must leave for another day.

**18. Transformational aspects of rigid body mechanics.** At (31) we obtained rigid body equations of motion that read

$$\begin{aligned} \mathbf{N}_{\text{intrinsic}} &= \dot{\mathbf{S}} \\ &= \mathbb{I}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbb{I}\boldsymbol{\omega} \end{aligned}$$

when referred to the **space frame** (generally non-inertial translated copy of the inertial **lab frame**), on page 15 we drew attention to the fact that those equations read

$$\mathbf{N}^0 = \mathbb{I}^0 \dot{\boldsymbol{\omega}}^0 + \boldsymbol{\omega}^0 \times \mathbb{I}^0 \boldsymbol{\omega}^0$$

—which is to say: *they preserve their structure*—when referred to the wobbly **body frame**. I want now to discuss how this remarkable fact comes about.

Let a wobbly **red frame** which shares the origin of—but be in a state of arbitrary rotation with respect to—the **space frame**, and write

$$\mathbf{r} = \mathbb{W} \mathbf{r} \quad : \quad \mathbb{W} \text{ is an arbitrarily } t\text{-dependent rotation matrix} \quad (140)$$

to describe the relationship between the red and black coordinates of any given point. Immediately

$$\begin{aligned} \dot{\mathbf{r}} &= \mathbb{W} \dot{\mathbf{r}} + \dot{\mathbb{W}} \mathbf{r} \\ &= \mathbb{W} \dot{\mathbf{r}} + \mathbb{U} \mathbf{r} \quad \text{with} \quad \mathbb{U} \equiv \dot{\mathbb{W}} \mathbb{W}^{-1} = -\mathbb{U}^\top \end{aligned}$$

which can be written

$$\begin{aligned} \mathbb{W} \dot{\mathbf{r}} &= \dot{\mathbf{r}} + \mathbb{U} \mathbf{r} \quad \text{with} \quad \mathbb{U} \equiv -\mathbb{U} \\ &= \left( \frac{d}{dt} + \mathbb{U} \right) \mathbf{r} \end{aligned} \quad (141.1)$$

Extensions of the same basic line of argument give

$$\begin{aligned} \mathbb{W} \ddot{\mathbf{r}} &= \left( \frac{d}{dt} + \mathbb{U} \right)^2 \mathbf{r} \\ &= \ddot{\mathbf{r}} + 2\mathbb{U} \dot{\mathbf{r}} + (\dot{\mathbb{U}} + \mathbb{U}^2) \mathbf{r} \\ &\vdots \\ \mathbb{W} \mathbf{r}^{(n)} &= \left( \frac{d}{dt} + \mathbb{U} \right)^n \mathbf{r} \end{aligned} \quad (141.2)$$

---

<sup>55</sup> Some specific examples are discussed by Shapere & Wilczek, both in the paper cited previously<sup>42</sup> and in a companion paper of slightly earlier date: “Geometry of self-propulsion at low Reynolds number,” J. Fluid. Mech. **198**, 557 (1989). This paper also is reprinted (as §8.4) in *Geometric Phases in Physics*.

If the space frame were inertial (which generally it is not) then to describe the dynamics of a *single* particle we would write  $\mathbf{F} = m\ddot{\mathbf{r}}$ , which when referred to the wobbly frame becomes

$$\mathbf{F} = m\{\ddot{\mathbf{r}} + 2\mathbf{U}\dot{\mathbf{r}} + (\dot{\mathbf{U}} + \mathbf{U}^2)\mathbf{r}\}$$

with  $\mathbf{F} \equiv \mathbb{W}\mathbf{F}$  and  $m \equiv m$ . The preceding equation is often written

$$\mathbf{F} + \mathbf{F}^{\text{Coriolis}} + \mathbf{F}^{\text{centrifugal}} = m\ddot{\mathbf{r}} \quad (142)$$

with

$$\begin{aligned} \mathbf{F}^{\text{Coriolis}} &\equiv -2m\mathbf{U}\dot{\mathbf{r}} \\ &= -2m\boldsymbol{\Omega} \times \dot{\mathbf{r}} \end{aligned} \quad (143.1)$$

$$\begin{aligned} \mathbf{F}^{\text{centrifugal}} &\equiv -m(\dot{\mathbf{U}} + \mathbf{U}^2)\mathbf{r} \\ &= -m\dot{\boldsymbol{\Omega}} \times \mathbf{r} - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) \end{aligned} \quad (143.2)$$

Look now to the intrinsic angular momentum (“spin”) of a loose *system* of particles. Hitting  $\mathbf{S} = \sum_i m_i \mathbf{x}_i \times \dot{\mathbf{x}}_i$  with  $\mathbb{W}$  gives

$$\begin{aligned} \mathbb{W}\mathbf{S} &= \sum_i m_i \mathbb{W}(\mathbf{r}_i \times \mathbb{W}^{-1}\mathbb{W}\dot{\mathbf{r}}_i) \\ &= \sum_i m_i \mathbf{r}_i \times \mathbb{W}\dot{\mathbf{r}}_i \quad \text{by the LEMMA of page 15} \\ &= \sum_i m_i \mathbf{r}_i \times (\dot{\mathbf{r}}_i + \mathbf{U}\mathbf{r}_i) \\ &= \sum_i m_i \mathbf{r}_i \times (\dot{\mathbf{r}}_i + \boldsymbol{\Omega} \times \mathbf{r}_i) \\ &= \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i - \sum_i m_i (\mathbf{r}_i \times)^2 \boldsymbol{\Omega} \\ &\equiv \mathbf{S} + \mathbb{I}\boldsymbol{\Omega} \end{aligned} \quad (144)$$

which is not the result ( $\mathbb{W}\mathbf{S} = \mathbf{S}$ ) that one might have anticipated. We note also in this connection that, by appeal once again to the LEMMA,

$$\mathbb{W}\mathbb{I}\mathbb{W}^{-1} = - \sum_i m_i \mathbb{W}(\mathbf{r}_i \times) \mathbb{W}^{-1} \mathbb{W}(\mathbf{r}_i \times) \mathbb{W}^{-1} = - \sum_i m_i (\mathbf{r}_i \times)^2 \equiv \mathbb{I} \quad (145)$$

Look now to the *motion* of  $\mathbf{S}$ . By differentiation of (144) we have

$$\begin{aligned} \mathbf{N} &\equiv \mathbb{W}\dot{\mathbf{S}} \\ &= \mathbb{W}\dot{\mathbf{S}} \\ &= \dot{\mathbf{S}} + \dot{\mathbb{I}}\boldsymbol{\Omega} + \mathbb{I}\dot{\boldsymbol{\Omega}} - \dot{\mathbb{W}}\mathbf{S} \\ &= \dot{\mathbf{S}} + \dot{\mathbb{I}}\boldsymbol{\Omega} + \mathbb{I}\dot{\boldsymbol{\Omega}} - \dot{\mathbb{W}}\mathbb{W}^{-1} \cdot \mathbb{W}\mathbf{S} \\ &= \dot{\mathbf{S}} + \dot{\mathbb{I}}\boldsymbol{\Omega} + \mathbb{I}\dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times (\mathbf{S} + \mathbb{I}\boldsymbol{\Omega}) \\ &= (\dot{\mathbf{S}} + \boldsymbol{\Omega} \times \mathbf{S}) + ([\dot{\mathbb{I}}\boldsymbol{\Omega} + \mathbb{I}\dot{\boldsymbol{\Omega}}] + \boldsymbol{\Omega} \times \mathbb{I}\boldsymbol{\Omega}) \end{aligned}$$

which—by seemingly trivial rearrangement—becomes

$$\mathbf{N} - (\boldsymbol{\Omega} \times \mathbf{S} + \mathbb{I} \dot{\boldsymbol{\Omega}}) - (\mathbb{I} \dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times \mathbb{I} \boldsymbol{\Omega}) = \dot{\mathbf{S}} \quad (146)$$

I turn now to manipulations intended to clarify the meaning of (146). We have

$$\begin{aligned} (\boldsymbol{\Omega} \times \mathbf{S} + \mathbb{I} \dot{\boldsymbol{\Omega}}) &= \boldsymbol{\Omega} \times \sum_i m_i (\mathbf{r}_i \times \dot{\mathbf{r}}_i) \\ &\quad - \left\{ \sum_i m_i (\dot{\mathbf{r}}_i \times) (\mathbf{r}_i \times) + \sum_i m_i (\mathbf{r}_i \times) (\dot{\mathbf{r}}_i \times) \right\} \boldsymbol{\Omega} \\ &= \sum_i m_i \left\{ \boldsymbol{\Omega} \times (\mathbf{r}_i \times \dot{\mathbf{r}}_i) + \dot{\mathbf{r}}_i \times (\boldsymbol{\Omega} \times \mathbf{r}_i) \right\} \\ &\quad + \sum_i m_i \mathbf{r}_i \times (\boldsymbol{\Omega} \times \dot{\mathbf{r}}_i) \end{aligned}$$

which by  $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) = \mathbf{0}$  becomes

$$= \sum_i \mathbf{r}_i \times \left\{ 2m_i (\boldsymbol{\Omega} \times \dot{\mathbf{r}}_i) \right\} \quad (147.1)$$

A similar argument (I omit the details) supplies

$$(\mathbb{I} \dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times \mathbb{I} \boldsymbol{\Omega}) = \sum_i \mathbf{r}_i \times \left\{ m_i \dot{\boldsymbol{\Omega}} \times \mathbf{r}_i + m_i \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}_i) \right\} \quad (147.2)$$

But the expressions in braces are familiar already from (143). The striking implication is that (146) can be written

$$\mathbf{N} + \mathbf{N}^{\text{Coriolis}} + \mathbf{N}^{\text{centrifugal}} = \dot{\mathbf{S}} \quad (148)$$

where

$$\mathbf{N}^{\text{Coriolis}} \equiv \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{Coriolis}} \quad (149.1)$$

$$\mathbf{N}^{\text{centrifugal}} \equiv \sum_i \mathbf{r}_i \times \mathbf{F}_i^{\text{centrifugal}} \quad (149.2)$$

serve to define what might be called—though I have never encountered such terminology in the literature—the “net Coriolis and centrifugal torques.” These we recognize to be “fictitious torques,” artifacts of the non-inertiality of the wobbly red frame.

If, as a special circumstance, our many-particle system is *rigidly* assembled then it becomes natural—not mandatory, but natural—to

identify the wobbly **red frame** with the **body frame**,

with respect to which all particles are at rest:  $\dot{\mathbf{r}}_i = \mathbf{0}$  (all  $i$ ). From (147.1) and  $\mathbf{S} \equiv \sum m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i$  it then follows that  $(\boldsymbol{\Omega} \times \mathbf{S} + \mathbb{I} \dot{\boldsymbol{\Omega}})$  and  $\dot{\mathbf{S}}$  both vanish: (146) therefore becomes

$$\mathbf{N} - (\mathbb{I} \dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times \mathbb{I} \boldsymbol{\Omega}) = \mathbf{0} \quad (150.1)$$

which are usually written

$$\mathbf{N} = \mathbb{I} \dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times \mathbb{I} \boldsymbol{\Omega} \quad (150.2)$$

and called the **Euler equations**. Equations (150.1) serve at once to illuminate and to deepen the mystery that motivated this discussion, for they establish that Euler's equations might most properly be expressed

$$\mathbf{N} + \mathbf{N}^{\text{centrifugal}} = \mathbf{0} \quad \text{in the non-inertial body frame of a rigid system}$$

Notice that if the system were deformable (which is to say: *not* rigid) then it would be *impossible* to select a frame with respect to which all  $\dot{\mathbf{r}}_i$  terms vanish: one would be forced to work with some instance of (148). The implication is that it should be possible to get from (148) to the **Liouville equations**,<sup>56</sup> which are used by astrophysicists to study the rotational dynamics of stars, planets and asteroids. I must admit, however, that I do not at present know how to do so.

---

<sup>56</sup> See Problem 10 in Chapter 5 in H. Goldstein's *Classical Mechanics* (2<sup>nd</sup> or 3<sup>rd</sup> editions).