SHORT COURSE IN CLASSICAL MECHANICS

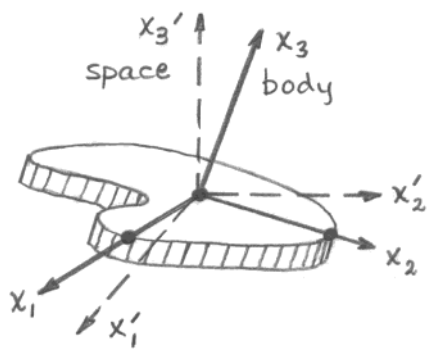
M. Strovink
University of California, Berkeley

January 2, 2001

1. Vectors and Transformations.

1.1. Body and space axes.

In ordinary 3-dimensional space, we require six independent quantities to specify the configuration of a rigid body. (Take $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ to be vectors from the origin to each of three reference points in the body. If the body is rigid, $|\mathbf{r}_1 - \mathbf{r}_2|$, $|\mathbf{r}_2 - \mathbf{r}_3|$, and $|\mathbf{r}_3 - \mathbf{r}_1|$ are fixed, so the number of independent quantities is only six.) Three of these six may be identified with a vector $\mathbf{R}$ from the origin to some basic reference point, e.g. the center of mass. The remaining three quantities are *orientation variables*.



To study these orientation variables, we use a set of unprimed "body axes" (x_1, x_2, x_3) attached to the rigid body, and a set of primed "space axes" (x_1', x_2', x_3') using the same origin as the body axes, but having their directions fixed to be the same as those of a particular set of external reference axes. The external frame is usually taken to be inertial, i.e. not accelerating. A point on the rigid body can be represented in either coordinate system.

Going from one representation to another requires a *linear transformation*:

$$\begin{aligned} x_1' &= \lambda_{11}x_1 + \lambda_{12}x_2 + \lambda_{13}x_3 \\ x_2' &= \lambda_{21}x_1 + \lambda_{22}x_2 + \lambda_{23}x_3 \\ x_3' &= \lambda_{31}x_1 + \lambda_{32}x_2 + \lambda_{33}x_3 \end{aligned} \tag{1.1}$$

(We require the transformation to be linear so that it does not depend on the dimensions of x and x'.) Other notation for (1.1) is:

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_j$$

$$x_i' = \lambda_{ij} x_j$$

where in the last expression summation from 1 to 3 over the *repeated index* j is *assumed by convention*. In matrix notation, we could also write

$$\tilde{x}' = \Lambda \tilde{x}$$

where

$$\tilde{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \tilde{x}' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$$

are column vectors, and

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

is a 3×3 matrix. Of course, the rules of matrix multiplication are followed.

To test your understanding of matrix multiplication and the convention that repeated indices are summed, consider the product

$$C = AB$$

where A, B, and C are 3×3 matrices. Then the ij element of C is given by

$$C_{ij} = A_{ik} B_{kj}.$$

1.2. Properties of the transformation matrix.

The nine matrix elements $\lambda_{11} \ldots \lambda_{33}$ depend on only three (as yet unspecified) orientation

variables. Therefore, there must be six equations that relate the matrix elements to each other (more on that below). The physical significance of the λ_{ij} is revealed by transforming the *unit vectors*

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \hat{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

into the primed system, as usual by multiplying them by Λ:

$$\Lambda \hat{e}_1 = \begin{pmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{pmatrix}.$$

Therefore, λ_{11} is the projection of $\hat{e}_1$ on the x'_1 axis, λ_{21} is the projection of $\hat{e}_1$ on the x'_2 axis, etc.

Using the well-known property of the dot product

$$\cos\theta_{ab} = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|},$$

we obtain

$$\begin{aligned} \lambda_{11} &= \hat{e}'_1 \cdot \hat{e}_1 = \cos\theta_{1'1} \\ \lambda_{21} &= \hat{e}'_2 \cdot \hat{e}_1 = \cos\theta_{2'1} \\ \lambda_{31} &= \hat{e}'_3 \cdot \hat{e}_1 = \cos\theta_{3'1}, \end{aligned}$$

etc. That is, the λ_{ij} are the *direction cosines* relating axis i' to axis j.

We return to the six equations relating the λ_{ij} to each other. These may be obtained from trigonometry, or more instructively by considering the need to *preserve the dot product* under transformation. In matrix notation,

$$\mathbf{a}\cdot\mathbf{b} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_i b_i.$$

It is useful to express the dot product in terms of the *transpose* of a matrix. We denote the transpose of A by A^t, following the definition

$$A^t_{ij} \equiv A_{ji},$$

signifying the interchange of rows and columns. Then the dot product becomes

$$\mathbf{a}\cdot\mathbf{b} = \tilde{a}^t\tilde{b},$$

where, as usual, $\tilde{a}$ and $\tilde{b}$ are the column vectors.

Now we require that $\tilde{a}^t\tilde{b}$ remain invariant in the primed system, related to the unprimed system by the transformation Λ:

$$\tilde{a}' = \Lambda\tilde{a}; \;\; \tilde{b}' = \Lambda\tilde{b}.$$

We want

$$\tilde{a}'^t\tilde{b}' = \tilde{a}^t\tilde{b}.$$

To proceed further, we need to know how to take the transpose of a product of matrices. In general, for any two matrices A and B,

$$[(AB)^t]_{ij} = [AB]_{ji} = A_{jk}B_{ki} = (A^t)_{kj}(B^t)_{ik}$$

$$= (B^t)_{ik}(A^t)_{kj} = (B^tA^t)_{ij}.$$

That is, the transpose of a product of matrices is the product of the transposed matrices multiplied in the opposite order:

$$(AB)^t = B^tA^t.$$

Returning to the invariance of the dot product, we have

$$\tilde{a}'^t\tilde{b}' = (\Lambda\tilde{a})^t\Lambda\tilde{b} = \tilde{a}^t\Lambda^t\Lambda\tilde{b}.$$

This can be equal to $\tilde{a}^t\tilde{b}$ for *all possible* $\tilde{a}$ and $\tilde{b}$ if and only if the product $\Lambda^t\Lambda$ reduces to the unit matrix:

$$\Lambda^t\Lambda = \mathrm{I}.$$

If this condition is satisfied, Λ is said to be *orthogonal*. Expressed in component form, the orthogonality requirement is

$$\lambda^t_{ik}\lambda_{kj} = \lambda_{ki}\lambda_{kj} = \delta_{ij},$$

where the *Kronecker delta* δ_{ij} is 1 if $i = j$, 0 otherwise.

Since i and j each can range through three values, this last equation is really nine equations. They are:

$$\begin{aligned}
\lambda_{11}\lambda_{11} + \lambda_{21}\lambda_{21} + \lambda_{31}\lambda_{31} &= 1 \\
\lambda_{12}\lambda_{12} + \lambda_{22}\lambda_{22} + \lambda_{32}\lambda_{32} &= 1 \\
\lambda_{13}\lambda_{13} + \lambda_{23}\lambda_{23} + \lambda_{33}\lambda_{33} &= 1 \\
\lambda_{11}\lambda_{12} + \lambda_{21}\lambda_{22} + \lambda_{31}\lambda_{32} &= 0 \\
\lambda_{12}\lambda_{11} + \lambda_{22}\lambda_{21} + \lambda_{32}\lambda_{31} &= 0 \\
\lambda_{11}\lambda_{13} + \lambda_{21}\lambda_{23} + \lambda_{31}\lambda_{33} &= 0 \\
\lambda_{13}\lambda_{11} + \lambda_{23}\lambda_{21} + \lambda_{33}\lambda_{31} &= 0 \\
\lambda_{12}\lambda_{13} + \lambda_{22}\lambda_{23} + \lambda_{32}\lambda_{33} &= 0 \\
\lambda_{13}\lambda_{12} + \lambda_{23}\lambda_{22} + \lambda_{33}\lambda_{32} &= 0
\end{aligned}$$

As we expected, these nine equations are really only six, because the last three pairs are identical.

The following complex generalizations are useful in quantum mechanics and elsewhere:

$$\begin{aligned}
\text{transpose } (A^t)_{ij} \equiv A_{ji} &\to \text{adjoint } (A^\dagger)_{ij} \equiv A^*_{ji}; \\
\text{orthogonality } A^t A = \mathrm{I} &\to \text{unitarity } A^\dagger A = \mathrm{I}; \\
\text{symmetric } A^t = A &\to \text{Hermitian } A^\dagger = A.
\end{aligned}$$

1.3. Parity inversion.

Now we consider the determinant of Λ. In order to express the determinant in component form, we need the *Levi-Civita density* ϵ_{ijk}. This object is most straightforwardly used in the cross product:

$$\mathbf{a} \times \mathbf{b} \equiv \epsilon_{ijk}\hat{e}_i a_j b_k. \tag{1.2}$$

Given the rules for cross products, it must be true that

$$\begin{aligned}
\epsilon_{ijk} =& 0 \text{ unless } i \neq j \neq k \\
=& 1 \text{ for } i,j,k = \text{cyclic permutation of } 1,2,3 \\
=& -1 \text{ for } i,j,k = \text{cyclic permutation of } 3,2,1.
\end{aligned}$$

Using the Levi-Civita density, we can write the determinant of a 3×3 matrix as

$$\det A \equiv |A| \equiv \frac{1}{3!}\epsilon_{ijk}A_{il}A_{jm}A_{kn}\epsilon_{lmn}. \tag{1.3}$$

As i, j, k, l, m, n each run from 1 to 3, there are a total of $3^6 = 729$ terms. However, because the Levi-Civita density is usually zero, only 36 terms survive. Each is the product of three matrix elements that come from rows and columns which must be different. Each of these products, for example $A_{11}A_{22}A_{33}$, comes in six permutations, e.g. 123, 231, 312, 321, 132, 213. That's the reason for the 1/3! factor – only six of the 36 terms are not duplicates. Thus we confirm that this expression for the determinant gives the same result as does the standard "diagonal lines" mnemonic.

Writing the determinant using the Levi-Civita density enables us to conclude immediately that, for any 3×3 matrix,

$$\det A = \det A^t. \tag{1.4}$$

(In (1.3), changing A to A^t just interchanges i, j, k with l, m, n, yielding the same result.) It is also true that

$$\det AB = \det A \det B. \tag{1.5}$$

Equations (1.4) and (1.5) are true for (square) matrices of any size.

Returning to the rotation matrix Λ,

$$\begin{aligned}
1 = \det \mathrm{I} = \det \Lambda^t \Lambda &= (\det \Lambda)^2 \\
\pm 1 = \det \Lambda.
\end{aligned}$$

An *infinitesimal rotation*,

$$\Lambda = \mathrm{I} + \mathcal{E}; \;\; \det \mathcal{E} \ll 1$$

must be of the class

$$\det \Lambda = +1$$

since $\det \mathrm{I} = +1$. Any *finite* rotation can be built up from infinitesimal rotations and so must also have $\det \Lambda = +1$. If instead we have a transformation with $\det \Lambda = -1$, it must be related through a set of infinitesimal rotations not to I but to the *parity inversion* matrix

$$P \equiv \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \;\; \det P = -1.$$

1.4. Infinitesimal rotations.

Having found that the rotation matrix Λ is specified by three independent quantities, we arrive at a philosophical problem: can Λ be represented by a (3-component) vector, instead of a matrix, or tensor? For example, its direction could be the rotation axis, and its magnitude could be the angle of rotation. The answer turns out to be yes, but only if the rotation is infinitesimal.

To pursue this question further, we are reminded that, formally, a vector is any 3-component object whose components transform under rotations according to Λ. Vector addition is *commutative* and *associative*; multiplication by a *scalar* (invariant to rotations) is *commutative* and *distributive*. The dot product is *commutative*; and, as we see in (1.2), the cross product is *anticommutative*,

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a},$$

due to the change of sign of ϵ_{ijk} when any two indices are switched.

If the rotation could be represented by a vector, two successive rotations $\vec{\lambda}_1$ and $\vec{\lambda}_2$ would be described by another vector which is their sum $\vec{\lambda}_1 + \vec{\lambda}_2$. However, since vector addition is commutative, $\vec{\lambda}_1 + \vec{\lambda}_2 = \vec{\lambda}_2 + \vec{\lambda}_1$, the order of the rotations wouldn't matter. But physically it does! (Try rotating a book by 90° first through a vertical axis, then through a horizontal axis; and then vice versa.) On the other hand, if the rotation were represented by a tensor, as in general it is, we would ask that $\Lambda_1\Lambda_2 \neq \Lambda_2\Lambda_1$, as is generally the case, so that the order of rotations *would* matter.

Infinitesmal rotations *can* be described by a vector. Suppose

$$\begin{aligned} \Lambda_1 &= \mathrm{I} + \mathcal{E}_1, \quad \det \mathcal{E}_1 \ll 1 \\ \Lambda_2 &= \mathrm{I} + \mathcal{E}_2, \quad \det \mathcal{E}_2 \ll 1. \end{aligned}$$

Then the order of rotations doesn't matter:

$$\begin{aligned} \Lambda_1\Lambda_2 &= (\mathrm{I} + \mathcal{E}_1)(\mathrm{I} + \mathcal{E}_2) \\ &= \mathrm{I} + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_1\mathcal{E}_2 \\ &\approx \mathrm{I} + \mathcal{E}_1 + \mathcal{E}_2 \quad \text{(last term negligible)} \\ &= \Lambda_2\Lambda_1 \end{aligned}$$

Requiring Λ to be orthogonal demands that $\mathcal{E}$ be antisymmetric:

$$\begin{aligned} \mathrm{I} &= \Lambda^t\Lambda \\ &= (\mathrm{I} + \mathcal{E}^t)(\mathrm{I} + \mathcal{E}) \\ &= \mathrm{I} + \mathcal{E}^t + \mathcal{E} + \mathcal{E}^t\mathcal{E} \\ &\approx \mathrm{I} + \mathcal{E}^t + \mathcal{E} \\ \mathcal{E}^t &= -\mathcal{E}. \end{aligned}$$

An antisymmetric matrix has diagonal elements equal to zero, with only three independent off-diagonal elements. We choose the following general form for the infinitesimal rotation:

$$\mathcal{E} \equiv \begin{pmatrix} 0 & -d\Omega_3 & +d\Omega_2 \\ +d\Omega_3 & 0 & -d\Omega_1 \\ -d\Omega_2 & +d\Omega_1 & 0 \end{pmatrix},$$

where the d's emphasize smallness.

With $\mathcal{E}$ in this form, application to the vector $\mathbf{r}$ of the infinitesimal rotation produces an elegant formula:

$$\begin{aligned} \tilde{r}' &= \Lambda\tilde{r} \\ &= (\mathrm{I} + \mathcal{E})\tilde{r} \\ &= \tilde{r} + \mathcal{E}\tilde{r} \\ \tilde{r}' - \tilde{r} &= \mathcal{E}\tilde{r} \\ &= \begin{pmatrix} d\Omega_2 x_3 - d\Omega_3 x_2 \\ d\Omega_3 x_1 - d\Omega_1 x_3 \\ d\Omega_1 x_2 - d\Omega_2 x_1 \end{pmatrix} \end{aligned}$$

Defining $\tilde{r}' - \tilde{r} \equiv d\tilde{r}'$, this becomes

$$d\mathbf{r}' = d\vec{\Omega} \times \mathbf{r}. \tag{1.6}$$

In (1.6) $d\vec{\Omega}$ is a "vector" with components defined by the matrix elements of $\mathcal{E}$. Its direction is the axis about which the body (unprimed) axes have been rotated relative to the space axes. Equation (1.6) describes the difference $d\mathbf{r}'$ between the description $\mathbf{r}'$ of a space point as seen in the space axes and the description $\mathbf{r}$ of the same space point as seen in the (rotated) body axes. If the space point is *at rest* in the body

system, all its motion as seen in the space system will be due to the body rotation and will be "tangential", or perpendicular to $\mathbf{r}$:

$$\begin{aligned} \mathbf{v}'_{\text{tang}} \equiv \frac{d\mathbf{r}'}{dt} &= \frac{d\vec{\Omega}}{dt} \times \mathbf{r} \\ &\equiv \vec{\omega} \times \mathbf{r}, \end{aligned} \tag{1.7}$$

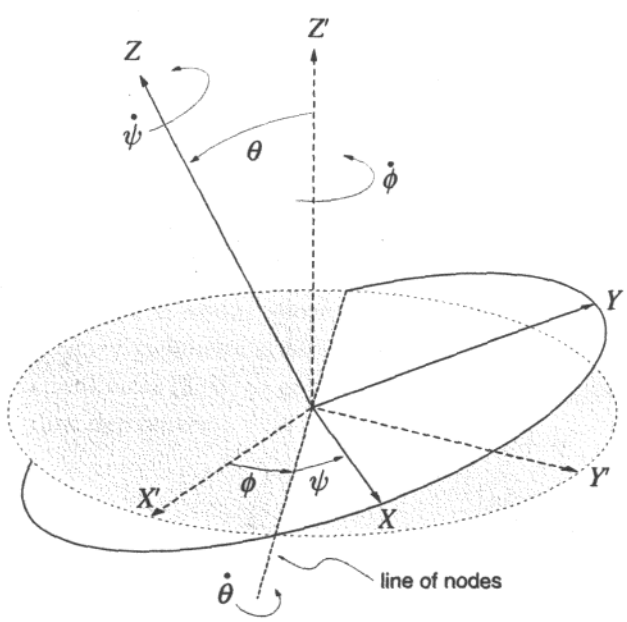
where $\vec{\omega}$ is the *angular velocity*. More generally, the space point will be moving with respect to the body system. Then its velocity in the space system will be the sum of the velocity in the body system and the velocity due to rotation:

$$\begin{aligned} \mathbf{v}' &= \mathbf{v} + \vec{\omega} \times \mathbf{r} \\ \frac{d\mathbf{P}'}{dt} &= \frac{d\mathbf{P}}{dt} + \vec{\omega} \times \mathbf{P} \end{aligned} \tag{1.8}$$

In the last equation, $\mathbf{v}'$ has been replaced by $d\mathbf{P}'/dt$ to emphasize that this transformation rule is valid for any vector $\mathbf{P}'$, not just the position vector.

We can prove that $\vec{\omega}$ is a vector only by examining in detail its transformation properties under coordinate rotation. It turns out that $\vec{\omega}$ *is* a type of vector ("axial vector" or "pseudovector") which is like an ordinary vector ("polar vector") exept that it does *not* change sign under parity inversion. Another familiar example of an axial vector is the magnetic field $\mathbf{B}$.

1.5. Euler rotation.



As an example of a 3-dimensional rotation, we introduce the *Eulerian angles*. The Euler rotation is *passive* (axes are rotated, not objects) and, by convention, it transforms from the space axes to the body axes. This is the *inverse* of the transformation we have been considering, so we denote the Euler rotation matrix by Λ^t rather than Λ:

$$\tilde{x} = \Lambda^t \tilde{x}'.$$

The Euler rotation is useful because it simplifies the analysis of certain problems such as tops, and because it is an established convention.

The Euler rotation consists of three steps. The first transformation, between the ($'$) and ($''$) frames, is a counterclockwise ("CCW") rotation about the $3'$ axis by the first Euler angle ϕ. The second transformation, between the ($''$) and ($'''$) frames, is a CCW rotation about the $1''$ axis (the "line of nodes") by the second Euler angle θ. The final step, a transformation between the ($'''$) and unprimed (body) frames, is a CCW rotation about the $3'''$ axis by the third Euler angle ψ. Note that, if θ were zero, the ϕ and ψ rotations would occur about the same axis and could be combined into a single rotation.

Expressed in rotation matrices,

$$\begin{aligned} \tilde{x}'' &= \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tilde{x}' \\ &\equiv \Lambda^t_\phi \tilde{x}' \\ \tilde{x}''' &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tilde{x}'' \\ &\equiv \Lambda^t_\theta \tilde{x}'' \\ \tilde{x} &= \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tilde{x}''' \\ &\equiv \Lambda^t_\psi \tilde{x}''' \end{aligned} \tag{1.9}$$

Within each of these three matrices, the nontrivial 2×2 submatrices are identical to the usual two-dimensional rotation matrices that are used for more elementary transformations.

To obtain the full rotation matrix, one applies the individual rotations in order:

$$\tilde{x} = \Lambda^t_\psi \Lambda^t_\theta \Lambda^t_\phi \tilde{x}' = \Lambda^t \tilde{x}'.$$

Note that the first rotation, Λ^t_ϕ, is the *right-hand* factor in the product of matrices. The elements of the full rotation matrix, obtained by carrying out the matrix multiplication, are:

$$\Lambda^t = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix} \tag{1.10}$$

2. Selective Review of Newtonian Mechanics.

2.1. Definitions.

Newton's laws are the consequence of the definition of force

$$\mathbf{F} \equiv \frac{d\mathbf{p}}{dt} \equiv \frac{d}{dt} m\mathbf{v} \tag{2.1}$$

in which the inertial mass m is the "resistance to change in velocity" and is proportional to the gravitational coupling constant, or gravitational mass. By convention, this constant of proportionality is unity. The fact that the two types of mass are proportional to each other is a consequence of general relativity and has been tested by balancing materials of high and low atomic number Z (in high Z materials, a larger fraction of the mass is due to relativistic effects that depend on the inertial mass). This proportionality has been verified to better than one part in 10^{14}. Equation (2.1) can be integrated to solve routine problems involving ballistics, rockets, etc.

The following analogies between linear and angular motion are important:

Quantity	**Linear motion**	**Angular motion**
Coordinate	$\mathbf{r}$	θ
Derivative of coordinate	$\mathbf{v} \equiv d\mathbf{r}/dt$	$\vec{\omega} \equiv \hat{e}d\theta/dt$
		($\hat{e}$ along axis of CCW rotation)
Momentum	$\mathbf{p} \equiv m\mathbf{v}$	$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$
Derivative of momentum	$\mathbf{F} \equiv d\mathbf{p}/dt$	$\mathbf{N} \equiv d\mathbf{L}/dt = \mathbf{r} \times \mathbf{F}$

Missing in the above table is the relationship between $\mathbf{L}$ and $\vec{\omega}$. For a sufficiently simple rigid body (details later),

$$\mathbf{L} = I\vec{\omega}$$

where the moment of inertia I is a scalar. More generally, $\mathbf{L}$ is not parallel to $\vec{\omega}$, so I must be a tensor, the "inertia tensor", represented by a symmetric 3×3 matrix.

2.2. Relations between $\mathbf{r}$, $\mathbf{v}$, and $\vec{\omega}$ for a point particle.

When the body is a *point*, no orientation is defined, and only three coordinates are needed to specify its position. We do not need body axes, so, for the time being, we drop the primes from the space axes.

Without the primes, Eq. (1.8) becomes

$$\mathbf{v}_{\text{tang}} = \vec{\omega} \times \mathbf{r}.$$

Taking the cross product of $\mathbf{r}$ with it,

$$\mathbf{r} \times \mathbf{v}_{\text{tang}} = \mathbf{r} \times \mathbf{v} = \mathbf{r} \times (\vec{\omega} \times \mathbf{r}).$$

Using the "*bac cab*" rule,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}),$$

$$\mathbf{r} \times \mathbf{v} = \vec{\omega}(\mathbf{r} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{r} \cdot \vec{\omega}).$$

The last term is that part of $r^2\vec{\omega}$ which is parallel to $\mathbf{r}$. After this parallel part is subtracted, what remains is the perpendicular part:

$$\begin{aligned} \mathbf{r} \times \mathbf{v} &= r^2 \vec{\omega}_{\perp \text{ to } \mathbf{r}} \\ \mathbf{L} &= mr^2 \vec{\omega}_{\perp \text{ to } \mathbf{r}}. \end{aligned}$$

Even for a point particle, $\mathbf{L}$ is equal to $mr^2\vec{\omega}$ only if $\vec{\omega}$ is $\perp$ to $\mathbf{r}$.

2.3. Work and energy.

The *work* done on a particle on a path between points 1 and 2 is

$$W_{12} \equiv \int_1^2 \mathbf{F} \cdot d\mathbf{r} \equiv T_2 - T_1,$$

where T is the *kinetic energy* $\frac{1}{2}mv^2$ and the integral is taken over the path. The vector force field $\mathbf{F}$ is called *conservative* if it can obtained from the gradient of a scalar field. The scalar field is conventionally written as $-U$ where U is the *potential energy*. If the force is conservative,

$$\begin{aligned} \mathbf{F} &= -\nabla U \\ W_{12} &= \int_1^2 -\nabla U \cdot d\mathbf{r} \\ &= -(U_2 - U_1) \\ T_2 - T_1 &= U_1 - U_2 \\ T_2 + U_2 &= T_1 + U_1, \end{aligned} \tag{2.2}$$

and total energy $T + U$ is conserved.

In the above situation in which $\mathbf{F}$ is conservative, W_{12} is path-independent. More generally, what conditions does path-independence of W_{12} place on $\mathbf{F}$? We consider any path from 1 to 2, combined with any other path from 2 back to 1. Since W_{12} is path-independent, and $W_{21} = -W_{12}$, it follows directly that the circuital integral of $\mathbf{F}$ around the combined path must vanish:

$$\oint \mathbf{F} \cdot d\mathbf{r} = W_{12} + W_{21} = 0.$$

What conditions on $\mathbf{F}$ are imposed?

First, $\mathbf{F}$ must be velocity-independent. Otherwise, we could negotiate the path 12 rapidly, and 21 slowly, spoiling the cancellation. Second, $\mathbf{F}$ must not depend explicitly on the time (as opposed to implicitly, for example as a result of particle motion). Otherwise, we could negotiate paths 12 and 21 at different times, again spoiling the cancellation. Finally, assuming that $\mathbf{F}$ is both velocity- and time-independent, it must be free of *circulation* in order that the circuital integral vanish. Qualitatively, a vector field with circulation loops back on itself. To get a more quantitative condition, we need Stokes' theorem:

$$\oint \mathbf{F} \cdot d\mathbf{r} = \iint (\nabla \times \mathbf{F}) \cdot d\mathbf{A},$$

where the surface integral applies to any surface bounded by the circuital path, and the element of area $d\mathbf{A}$ points outward according to the right-hand rule when applied to the loop. For the right-hand integral to be zero for any path and thus over any area, the integrand must vanish:

$$\nabla \times \mathbf{F} = 0.$$

The most general form of $\mathbf{F}$ with vanishing curl is:

$$\mathbf{F} = -\nabla U$$

because, formally,

$$\nabla \times (\nabla U) = 0.$$

Thus we are back to the original definition of a conservative force.

A qualitative evaluation of whether a vector field has circulation may not be completely reliable. Consider the magnetic field outside a long straight thin current-carrying wire oriented along $\hat{z}$. The field points in the θ direction and so obviously loops back along itself. However, since this field varies as $1/r$, where r is the perpendicular distance to the wire, it has no curl and therefore no circulation away from the singularity at $r = 0$. This can be seen from evaluating the curl in cylindrical coordinates, or more easily by noting that

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} = 0,$$

since the current density $\mathbf{j}$ vanishes outside the wire.

2.4. Conservation laws for multiparticle systems.

For a system of N particles i, $1 \le i \le N$, the total momentum is

$$\mathbf{P} = \sum_i m_i \mathbf{v}_i,$$

where sums are understood to run from 1 to N. The total force on the system is the time rate of change of $\mathbf{P}$:

$$\mathbf{F}_{\text{tot}} = \frac{d\mathbf{P}}{dt} = \sum_i \mathbf{F}_i^{\text{ext}} + \sum_{i,j} \mathbf{F}^j_{\text{on } i}$$

where on the right-hand side we sum firstly over external forces on the particles, and secondly over forces on one particle from another.

However, by Newton's third law (action = reaction),

$$\mathbf{F}^j_{\text{on } i} = -\mathbf{F}^i_{\text{on } j}.$$

The last term is a sum over cancelling pairs and therefore vanishes. Then

$$\mathbf{F}_{\text{ext}} \equiv \sum_i \mathbf{F}_i^{\text{ext}} = \frac{d}{dt}\sum_i m_i \mathbf{v}_i = \frac{d^2}{dt^2}\sum_i m_i \mathbf{r}_i.$$

Defining

$$M \equiv \sum_i m_i$$
$$\mathbf{R} \equiv \frac{1}{M}\sum_i m_i \mathbf{r}_i,$$

we obtain

$$\mathbf{F}_{\text{ext}} = M\frac{d^2\mathbf{R}}{dt^2}. \tag{2.3}$$

That is, the motion of the center of mass coordinate $\mathbf{R}$ of a system of particles depends only on the total external force $\mathbf{F}_{\text{ext}}$, as if the system were merely a single particle located at that coordinate. If the total external force is zero, the velocity of the center of mass is constant.

Using the center of mass coordinate makes further simplifications possible, in the form of *decomposition theorems*. Consider the total angular momentum

$$\mathbf{L} = \sum_i \mathbf{r}_i \times m_i \mathbf{v}_i.$$

Substitute

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i^*$$
$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt} = \frac{d\mathbf{R}}{dt} + \frac{d\mathbf{r}_i^*}{dt},$$

where $\mathbf{r}_i^*$ is the coordinate of particle i with respect to the center of mass coordinate $\mathbf{R}$:

$$\mathbf{L} = \sum_i (\mathbf{R} + \mathbf{r}_i^*) \times m_i (\frac{d\mathbf{R}}{dt} + \frac{d\mathbf{r}_i^*}{dt}).$$

Since

$$\sum m_i \mathbf{r}_i^* \equiv 0,$$

the two cross terms vanish, and

$$\mathbf{L} = \mathbf{R} \times \mathbf{P} + \sum_i \mathbf{r}_i^* \times \mathbf{p}_i^*. \tag{2.4}$$

That is, the total angular momentum is equal to the angular momentum *of* the center of mass plus the angular momentum *with respect to* the center of mass.

A similar proof yields the decomposition theorem for the kinetic energy:

$$T = \tfrac{1}{2}M(\frac{d\mathbf{R}}{dt})^2 + \sum_i \tfrac{1}{2}m_i(\frac{d\mathbf{r}_i^*}{dt})^2. \tag{2.5}$$

The total kinetic energy is equal to the kinetic energy of the center of mass plus the kinetic energy with respect to the center of mass.

2.5. Gravitational potential due to a spherically symmetric mass distribution.

The gravitational force on a point test particle of mass m_t due to another point mass m is, as usual,

$$\mathbf{F} = -\frac{Gmm_t}{r^2}\hat{e}_r,$$

where $\mathbf{r}$ is a vector from m (put at the origin) to m_t. To understand the force on m_t due to a

mass *distribution*, it is useful to consider Gauss's theorem for the surface integral of $\mathbf{F}$:

$$\oint\!\!\oint \mathbf{F}\cdot d\mathbf{A} = \iiint (\nabla\cdot\mathbf{F})dv \tag{2.6}$$

The radial part of the divergence in spherical coordinates is

$$\nabla\cdot\mathbf{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_r) + \ldots$$

so that $\nabla\cdot\mathbf{F}$ vanishes except at the origin, where it is singular. The spherical symmetry of the gravitational force makes the left-hand side of (2.6) easy to evaluate over a sphere of radius r:

$$\begin{aligned}\oint\!\!\oint \mathbf{F}\cdot d\mathbf{A} &= 4\pi r^2\frac{-Gmm_t}{r^2}\\ &= -4\pi Gmm_t = \iiint(\nabla\cdot\mathbf{F})dv\end{aligned}$$

The last equality states that $\nabla\cdot\mathbf{F}$, already shown to be infinite at the origin, has a finite volume integral. This means that it is proportional to a 3-dimensional δ function $\delta^3(\mathbf{r})$, whose volume integral is defined to be unity:

$$\nabla\cdot\mathbf{F} = -4\pi Gmm_t\delta^3(\mathbf{r}).$$

In the case of a mass *distribution* instead of a mass *point*, the quantity $m\delta^3(\mathbf{r})$, which has dimensions mass per unit volume, is replaced by $\rho(\mathbf{r})$, the *mass density*. Equation (2.6) is replaced by

$$\begin{aligned}\oint\!\!\oint \mathbf{F}\cdot d\mathbf{A} &= \iiint(-4\pi Gm_t\rho(\mathbf{r}))dv.\\ &= -4\pi Gm_tM,\end{aligned} \tag{2.7}$$

using the fact that the integral of ρ over the volume is just the mass M inside.

If the mass distribution $\rho(\mathbf{r})$ is spherically symmetric, we can use a spherical surface of radius r again to evaluate the left-hand side of Eq. (2.7). It becomes:

$$\begin{aligned}4\pi r^2F_r(r) &= -4\pi Gm_tM\\ \mathbf{F}_{\text{grav}} &= -\hat{e}_r\frac{GMm_t}{r^2}\end{aligned} \tag{2.8}$$

This proves (without any messy integrations) that the gravitational force due to a spherically symmetric mass distribution is the same as that from a point at the origin with the same mass, provided that the force is observed outside the mass distribution.

The *gravitational field* $\mathbf{g}$ is defined as the gravitational force $\mathbf{F}_{\text{grav}}$ divided by the test mass m_t. Thus it is analagous to the electrostatic field in that its value is independent of the constant (test mass or test charge) with which the test particle couples to the field. Again, if the source of the gravitational field is a spherically symmetric distribution of total mass M,

$$\mathbf{g} = -\hat{e}_r\frac{GM}{r^2}.$$

The potential U_{grav} from which this gravitational field is derived, defined by $\mathbf{g}\equiv-\nabla U_{\text{grav}}$, is

$$U_{\text{grav}} = -\frac{GM}{r},$$

adopting the convention that $U_{\text{grav}} = 0$ at $r=\infty$.

3. Oscillations.

3.1. Differential equation for linear oscillations.

Any potential energy $U(x')$ can be expanded about a minimum, for example at $x'=a$:

$$\begin{aligned}U(x') = U(a) + (x'-a)\frac{\partial U}{\partial x'}\Big|_a &+\\ + \tfrac{1}{2}(x'-a)^2\frac{\partial^2U}{\partial x'^2}\Big|_a &+\ldots\end{aligned}$$

Redefining $U(a)\equiv 0$, introducing $x\equiv x'-a$, and expressing the (single-particle) kinetic energy as $T=\frac{1}{2}m\dot{x}^2$, energy conservation demands

$$\begin{aligned}E = \text{constant} = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}kx^2,\ k\equiv\frac{\partial^2U}{\partial x^2}\Big|_0\\ \frac{dE}{dt} = 0 = \tfrac{1}{2}m(2\dot{x}\ddot{x}) + \tfrac{1}{2}k(2x\dot{x}),\end{aligned}$$

with the trivial solution $\dot{x} = 0$ and the nontrivial solution

$$m\ddot{x} + kx = 0.$$

Therefore the familiar "mass-spring" force equation is obtained for any potential that has a minimum, provided that the excursions are kept small.

Adding a viscous damping force and a driving force,

$$F_x^{\text{damp}} = -b\dot{x}$$
$$F_x^{\text{drive}} = F_0 \cos \omega t,$$

and defining $\gamma \equiv b/m$, $\omega_0^2 \equiv k/m$, one obtains the simple form

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = (F_0/m)\cos\omega t. \tag{3.1}$$

This is the "full" differential equation, in the sense that it has a driving term on the right-hand side (in the absence of a driving term, it would be called "homogeneous"). We seek a *particular* solution x_p of the full equation which is nontrivial. In order to match boundary conditions, later we will add a *general* solution x_h of the homogeneous equation. It will turn out that x_h vanishes at sufficiently large times; therefore x_p is the asymptotic solution as $t \to \infty$.

3.2. Particular solution.

To avoid unnecessary algebra in solving this equation, we will employ the *complex exponential method*. The first step is to substitute

$$x_p = \Re(\tilde{A}e^{i\omega' t})$$
$$F_0 \cos\omega t = \Re(F_0 e^{i\omega t}),$$

where ω' is some as-yet-undetermined trial frequency, and $\tilde{A}$ is complex to allow for differences in phase between the driving force and the response. We *hope* that it is possible to find a solution for which $\tilde{A}$ is time-independent, a hope that will be fulfilled in this case. With these substitutions, (3.1) is

$$\begin{aligned} &-\omega'^2\Re(\tilde{A}e^{i\omega' t}) + \gamma\omega'\Re(i\tilde{A}e^{i\omega' t}) + \\ &+ \omega_0^2\Re(\tilde{A}e^{i\omega' t}) = \Re((F_0/m)e^{i\omega t}). \end{aligned} \tag{3.2}$$

As the second step, to avoid unnecessary algebra, we *choose* to solve the complex equation of which (3.2) is the real part, rather than (3.2) itself:

$$\begin{aligned} &-\omega'^2\tilde{A}e^{i\omega' t} + \gamma\omega' i\tilde{A}e^{i\omega' t} + \\ &+ \omega_0^2\tilde{A}e^{i\omega' t} = (F_0/m)e^{i\omega t}. \end{aligned} \tag{3.3}$$

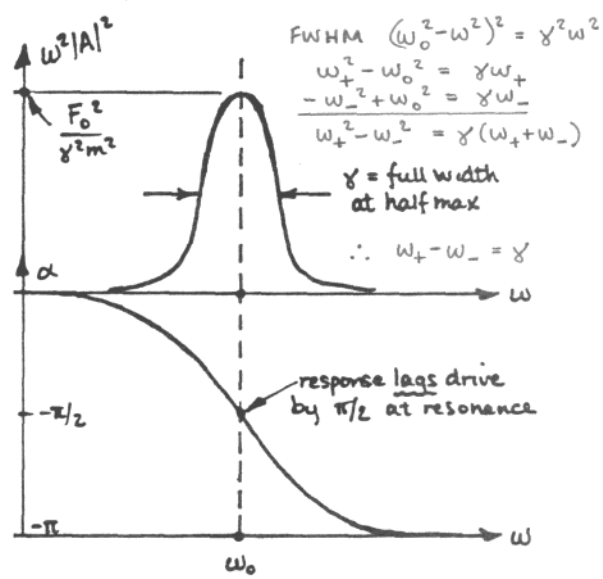
A nontrivial solution is impossible unless $\omega' = \omega$. Then the $e^{i\omega t}$ factors cancel, and

$$\tilde{A} = \frac{F_0/m}{(\omega_0^2 - \omega^2 + i\gamma\omega)}.$$

This is written in terms of a magnitude and a phase as

$$\begin{aligned} \tilde{A} &= \frac{F_0/m}{[(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2]^{1/2}} e^{i\alpha} \\ &\equiv |\tilde{A}|e^{i\alpha} \\ \alpha &= -\arctan\frac{\gamma\omega}{\omega_0^2 - \omega^2}. \end{aligned} \tag{3.4}$$

This result has the following properties: $\omega^2|\tilde{A}|^2$ (which is proportional to the oscillator's average rate of energy dissipation) reaches a peak of value $F_0^2/\gamma^2 m^2$ at the resonant frequency $\omega = \omega_0$. The half maximum of $\omega^2|\tilde{A}|^2$ occurs when $(\omega_2^2 - \omega_0^2) = +\gamma\omega_2$ or $(\omega_1^2 - \omega_0^2) = -\gamma\omega_1$. It follows easily that the full width at half maximum ("FWHM") of the resonant peak is $\omega_2 - \omega_1 = \gamma$.



As for the phase, assuming

$$Q \equiv \frac{\omega_0}{\gamma} \gg 1,$$

α starts just below zero at low frequency, falls through $-\pi/2$ at resonance, and approaches $-\pi$ at high frequency. The abruptness with which α crosses $-\pi/2$ increases as the resonance gets sharper (i.e. as the "quality factor" Q gets larger; note that Q increases when the damping decreases). In other words, the response $\tilde{A}$ *lags* the driving force by very little at low frequencies, by 90° at resonance, and by nearly 180° at high frequencies. The final step is to substitute back for x_p:

$$\begin{aligned} x_p &= \Re(\tilde{A}e^{i\omega t}) \\ &= \Re(|\tilde{A}|e^{i\alpha}e^{i\omega t}) \\ &= |\tilde{A}|\cos(\omega t + \alpha) \end{aligned} \tag{3.5}$$

with $|\tilde{A}|$ and α as in (3.4).

In some texts, confusion is spread by analyzing the peak in $|\tilde{A}|$ rather than in $\omega^2|\tilde{A}|^2$. (The latter is physically the more meaningful quantity, as it is proportional to the power dissipated in the oscillator.) This confusion leads to messy nonstandard definitions for the resonant frequency and for the quality factor, which are best ignored.

3.3. Homogeneous solutions.

Having found a nontrivial particular solution x_p to the full equation, we turn to the general solution x_h to the homogeneous equation. Since the right-hand side of the homogeneous equation is zero, it is clear that the sum of x_p and x_h will still satisfy the full equation. Taking this sum is the easiest way to obtain a general solution to the full equation.

The homogeneous equation is (3.1) without the driving term:

$$\ddot{x}_h + \gamma\dot{x}_h + \omega_0^2 x_h = 0. \tag{3.6}$$

Using the same complex exponential method, and ignoring the trivial solution $\tilde{A} = 0$, we obtain

$$-\omega^2 + i\gamma\omega + \omega_0^2 = 0.$$

The solutions are given by the quadratic formula:

$$\omega_\pm = \frac{i\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{\gamma^2}{4}} \equiv \frac{i\gamma}{2} \pm \omega_\gamma.$$

The discriminant, which can be written as $\gamma^2(Q^2 - \frac{1}{4})$, distinguishes three cases: (*i*) $Q > \frac{1}{2}$ ("underdamped"); (*ii*) $Q < \frac{1}{2}$ ("overdamped"); (*iii*) $Q = \frac{1}{2}$ ("critically damped").

In the *underdamped* case, the discriminant is positive and ω_γ is real. The general solution is an arbitrarily weighted sum of the two particular solutions involving ω_+ and ω_-:

$$x_h = e^{-\gamma t/2}\Re(\tilde{A}_+ e^{i\omega_\gamma t} + \tilde{A}_- e^{-i\omega_\gamma t}).$$

The last factor can be written as a cosinusoid of $\omega_\gamma t$ within a phase:

$$x_h = Be^{-\gamma t/2}\cos(\omega_\gamma t + \beta), \tag{3.7}$$

where the amplitude B and phase β are adjusted to fit the boundary conditions.

In the *overdamped* case, the discriminant is negative and ω_γ is imaginary. Defining

$$\gamma_\pm \equiv \frac{\omega_\pm}{i} = \frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_0^2},$$

x_h again is an arbitrarily weighted sum of the two particular solutions:

$$x_h = C_+ e^{-\gamma_+ t} + C_- e^{-\gamma_- t}. \tag{3.8}$$

Note that if $\gamma/2 \gg \omega_0$, or $Q \ll 1$ (ultradamping), γ_+ is close to γ while γ_- is much smaller.

Finally, in the *critically damped* case, never exactly achieved in practice, $Q \equiv \frac{1}{2}$, the discriminant vanishes, and

$$\begin{aligned} \omega_+ = \omega_- &= i\gamma/2, \\ x_h &= \Re(\tilde{D}e^{-\gamma t/2}). \end{aligned} \tag{3.9}$$

This is only one solution; to get two adjustable constants, necessary for a second-order differential equation, we must find another. This is accomplished by writing (3.6), with $\gamma = 2\omega_0$, as

$$\left(\frac{d}{dt} + \omega_0\right)\left(\frac{d}{dt} + \omega_0\right)x_h = 0. \tag{3.10}$$

The solution (3.9) is a solution to

$$\left(\frac{d}{dt}+\omega_0\right)x_h = \left(\frac{d}{dt}+\omega_0\right)e^{-\omega_0 t} = 0.$$

If we had another solution $s(t)$ such that

$$\left(\frac{d}{dt}+\omega_0\right)s(t) = e^{-\omega_0 t},$$

then $s(t)$ would also solve (3.10). Substituting the trial solution $s(t) = f(t)\exp(-\omega_0 t)$, we easily find $f(t) = t$. Then the general solution to the critically damped case is

$$x_h = D_1 e^{-\gamma t/2} + D_2 t e^{-\gamma t/2}. \tag{3.11}$$

This is the transient response that designers of mechanical ammeters, shock absorbers, etc., attempt to achieve: it approaches the asymptotic state as quickly as possible without overshoot. Despite the brisk market in heavy-duty shocks, putting them on your car merely moves the suspension from critically damped to overdamped; the ride becomes harsher but not better controlled.

Once the solutions for these three cases are developed, the applications tend to be routine. Most fall in two classes: (*i*) steady-state response to a periodic driving waveform, and (*ii*) transient response to a non-periodic driving waveform, i.e. a switch closing. Only occasionally is one asked to combine (*i*) and (*ii*). If this happens, remember to match boundary conditions using the *sum* of the particular and homogeneous solutions $x_p(t) + x_h(t)$, rather than using only the latter.

3.4. Fourier expansion of the driving term.

In problems of class (*i*), if the driving waveform is (co)sinusoidal the problem is already solved, by (3.5). As an example of the solution of problems of this class, we consider a non-sinusoidal driving force $f(t)$ with period T and zero average value. This problem is solved by expanding $f(t)$ in a *Fourier series* of sines and/or cosines:

$$\begin{aligned} f(t) &= \sum_{n=1}^{\infty} f_n \cos\omega_n t + g_n \sin\omega_n t; \\ \omega_n &\equiv \frac{2\pi n}{T}. \end{aligned} \tag{3.12}$$

The constants f_n and g_n can be found using *Fourier's trick*. For example, to find f_m, where $1 \le m \le \infty$, multiply (3.12) by $(2/T)\cos 2\pi m t/T$, and integrate over one period:

$$\begin{aligned} &\frac{2}{T}\int_0^T dt\, \cos\frac{2\pi m t}{T} f(t) \;= \\ &= \frac{2}{T}\sum_{n=1}^{\infty}\Bigg(f_n \int_0^T dt\, \cos\frac{2\pi m t}{T}\cos\frac{2\pi n t}{T} + \\ &\qquad + g_n \int_0^T dt\, \cos\frac{2\pi m t}{T}\sin\frac{2\pi n t}{T}\Bigg). \end{aligned}$$

Using the orthonormality of the cosines and sines,

$$\begin{aligned} \frac{2}{T}\int_0^T dt\, \cos\frac{2\pi m t}{T}\cos\frac{2\pi n t}{T} &= \delta_{mn} \\ \frac{2}{T}\int_0^T dt\, \sin\frac{2\pi m t}{T}\sin\frac{2\pi n t}{T} &= \delta_{mn} \\ \frac{2}{T}\int_0^T dt\, \cos\frac{2\pi m t}{T}\sin\frac{2\pi n t}{T} &= 0, \end{aligned}$$

all terms on the right-hand side vanish except the $\cos^2$ term with $n = m$. Then (and similarly for g_m),

$$\begin{aligned} f_m &= \frac{2}{T}\int_0^T dt\, \cos\frac{2\pi m t}{T} f(t) \\ g_m &= \frac{2}{T}\int_0^T dt\, \sin\frac{2\pi m t}{T} f(t). \end{aligned} \tag{3.13}$$

Since the oscillator is linear, the solution to a sum of driving terms is the sum of the individual solutions:

$$\begin{aligned} x_p = \sum_{n=1}^{\infty} |\tilde{a}_n| [&f_n \cos(\omega_n t + \alpha_n) + \\ &+ g_n \sin(\omega_n t + \alpha_n)], \end{aligned} \tag{3.14}$$

where

$$\begin{aligned} |\tilde{a}_n| &= \frac{1/m}{[(\omega_0^2 - \omega_n^2)^2 + \gamma^2\omega_n^2]^{1/2}} \\ \alpha_n &= -\arctan\frac{\gamma\omega_n}{\omega_0^2 - \omega_n^2}. \end{aligned}$$

3.5. Response of the underdamped oscillator to a δ-function drive.

As an example of a problem of type (*ii*), we consider the response of an oscillator with $Q > \frac{1}{2}$ to a δ-function driving term. An infinite force F_δ is applied for an infinitesimal time such that its time integral is a constant mv_0:

$$\begin{aligned} F_\delta(t) &= 0 \ (t \neq 0); \\ F_\delta(t) &= \infty \ (t = 0); \\ \int_{-\infty}^{\infty} \frac{F_\delta(t)}{m}\, dt &= v_0 \end{aligned} \tag{3.15}$$

This is equivalent to requiring

$$F_\delta(t) = mv_0\delta(t),$$

where $\delta(t)$ is a one-dimensional Dirac delta function like the three-dimensional type discussed in section 2.5. The action of this force may be simulated by striking a resting mass at $t = 0$ with a hard object so that the mass obtains a velocity v_0. What are the boundary conditions?

Within the infinitesimal time interval that $F_0(t)$ is nonzero, it is so large that all the other forces are negligible in comparison to it. Therefore, after that interval, the mass will acquire the velocity v_0. However, during the same interval, the displacement $\int v\, dt$ of the mass still is infinitesimal, because v is finite. Taking the origin of coordinates at the resting position, the boundary conditions become

$$0 = x(0^+); \quad v_0 = \dot{x}(0^+).$$

Using Eq. (3.7),

$$\begin{aligned} 0 &= B\cos\beta \\ v_0 &= -\frac{\gamma}{2}B\cos\beta - \omega_\gamma B\sin\beta. \end{aligned}$$

Solving for the constants,

$$\begin{aligned} \beta &= \pi/2 \\ B &= \frac{-v_0}{\omega_\gamma}. \end{aligned}$$

Plugging these constants into (3.7) for the solution,

$$\begin{aligned} x(t) &= 0 \ \ (t < 0); \\ &= v_0 \frac{e^{-\gamma t/2}\sin\omega_\gamma t}{\omega_\gamma} \ \ (t > 0); \\ (\omega_\gamma^2 &\equiv \omega_0^2 - \gamma^2/4). \end{aligned} \tag{3.16}$$

3.6. Green function for the underdamped oscillator.

For an underdamped oscillator initially at rest, the solution (3.16) is just the *Green function* $G(t)$ multiplied by v_0. Associated with many homogeneous differential equations and boundary conditions are unique Green functions. If the Green function is known, the solution to the differential equation, in the presence of *any* driving term, may be found by performing a single integration.

Consider a linear differential operator D_t (in the case just considered, $D_t = d^2/dt^2 + \gamma\, d/dt + \omega_0^2$). The Green function $G(t)$ is defined to be the solution to the equation

$$D_t G(t) \equiv \delta(t).$$

Generalizing to a delta-function that peaks at $t = t'$ rather than $t = 0$,

$$D_t G(t, t') \equiv \delta(t - t').$$

For an arbitrary driving term $a(t)$, the solution to the differential equation $D_t x = a(t)$ is the integral

$$x(t) = \int_{-\infty}^{\infty} G(t, t')a(t')\, dt', \tag{3.17}$$

as is easily verified:

$$\begin{aligned} D_t x &= \int_{-\infty}^{\infty} D_t G(t, t')a(t')\, dt' \\ &= \int_{-\infty}^{\infty} \delta(t - t')a(t')\, dt' \\ &= a(t). \end{aligned}$$

In the last step we used the fundamental property of the Dirac δ function

$$\int_{-\infty}^{\infty} \delta(t - t')f(t')\, dt' = f(t)$$

for any function f. That is, the (rest of the) integrand is evaluated where the δ function becomes infinite.

Returning to the underdamped oscillator, for any driving term $a(t)$ the solution obtained with the help of the Green function is

$$x(t) = \int_{-\infty}^{t} \frac{e^{-\gamma(t-t')/2} \sin \omega_\gamma (t-t')}{\omega_\gamma} a(t')\, dt',$$

provided that the mass is at rest at the origin before the driving force starts. We wrote the upper limit of the integral as t rather than ∞ because G vanishes for $t < t'$.

For a practical application in which determining the answer with adequate numerical precision is the main objective, a solution in the form of an integral is wholly acceptable. The integral may be evaluated to arbitrarily high precision using a digital computer.

You already know at least one other Green function. In section 2.5 we found that the gravitational potential U from a point mass m is $U = -Gm/r$. The gravitational force $\mathbf{F} = -m_t \nabla U$, where m_t is a test mass, satisfies

$$\nabla \cdot \mathbf{F} = -4\pi G m m_t \delta^3(\mathbf{r}).$$

Then

$$\begin{aligned} -\nabla^2\big(-m_t \frac{Gm}{r}\big) &= -4\pi G m m_t \delta^3(\mathbf{r}) \\ \nabla^2 \frac{1}{r} &= -4\pi \delta^3(\mathbf{r}) \\ \nabla^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|} &= -4\pi \delta^3(\mathbf{r}-\mathbf{r}'), \end{aligned}$$

where we have generalized the last equation to allow the mass point to be located at any coordinate $\mathbf{r}'$. This demonstrates that the Green function $G(\mathbf{r}, \mathbf{r}')$ for the differential operator ∇^2 is

$$G(\mathbf{r}, \mathbf{r}') = \frac{-1}{4\pi|\mathbf{r}-\mathbf{r}'|},$$

subject to the boundary condition $G(\infty, \mathbf{r}') = 0$. Correspondingly, a differential equation of the general form

$$\nabla^2 f(\mathbf{r}) = a(\mathbf{r}),$$

where a is any driving term, has the Green function solution

$$f(\mathbf{r}) = \iiint \frac{-a(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|} d^3r',$$

where the integral is taken over all space.

3.7. Nonlinear oscillations.

In general, oscillatory motion occurs (for example, in one dimension x) when the total energy E in a conservative system exceeds $U(x)$ only within a finite region. Within this region, say $x_A < x < x_B$, the kinetic energy T is positive. At $x = x_A$ and $x = x_B$, T vanishes; x_A and x_B are the *classical turning points*.

The period of the motion may be determined by a simple integral. Starting from the equation of energy conservation,

$$\begin{aligned} T &= \tfrac{1}{2} m \dot{x}^2 = E - U(x) \\ \frac{dx}{dt} &= \big\{\frac{2}{m}[E - U(x)]\big\}^{1/2} \\ \frac{T_0}{2} &= \int_{t_A}^{t_B} dt \\ &= \int_{x_A}^{x_B} \frac{dx}{\big\{\frac{2}{m}[E-U(x)]\big\}^{1/2}}. \end{aligned} \tag{3.18}$$

For simple harmonic motion, with $U(x) = \frac{1}{2}kx^2$ and $A = -B$, the usual period $T = 2\pi\sqrt{m/k}$ is easily obtained by doing the integral. On the other hand, for a pendulum oscillating through angles that are not negligibly small, the integral for the period is not elementary. With θ substituted for x, I for m, and $U(\theta) = mgl(1-\cos\theta)$, it is an *elliptic integral of the first kind*. This integral may be evaluated by using tables, or by Taylor-series expanding $(1-\cos\theta)$. As expected, to lowest order the Taylor expansion merely recovers the simple-harmonic-motion period. As a practical matter, fast digital computation usually obviates the need for tables or series expansions in evaluating any of these integrals.

In addition to obtaining an integral solution for the period, one may solve the differential equation for nonlinear oscillation if it is

only slightly nonlinear. The procedure used, the *method of perturbations*, is more interesting than this problem alone, because it is a fundamental method of classical and especially of modern theoretical physics.

Consider an undamped linear oscillator under the influence of a small additional nonlinear force $m\lambda x^2$:

$$\ddot{x} + \omega_0^2 x - \lambda x^2 = 0.$$

Here "small" means that $|\lambda x| \ll \omega_0^2$. Since λ is small, we attempt to find a solution of the form

$$\begin{aligned} x(t) &= x_0(t) + \lambda\eta(t) \\ 0 &= \ddot{x}_0 + \omega_0^2 x_0 - \lambda x_0^2 + \lambda\ddot{\eta} + \omega_0^2\lambda\eta + \\ &+ \text{ (terms of higher order in } \lambda). \end{aligned} \tag{3.19}$$

Since λ, though small, is a constant of otherwise arbitrary size, x_0 must solve the simple harmonic equation $\ddot{x}_0 + \omega_0^2 x_0 = 0$. Assuming the boundary condition $\dot{x}_0(0) = 0$, the solution for x_0 is

$$x_0(t) = A\cos\omega_0 t.$$

Plugging x_0 back into (3.19), the first two terms vanish. Neglecting the higher-order terms,

$$\begin{aligned} \lambda(-A^2\cos^2\omega_0 t + \ddot{\eta} + \omega_0^2\eta) &= 0 \\ \ddot{\eta} + \omega_0^2\eta &= \frac{A^2}{2}(\cos 2\omega_0 t + 1), \end{aligned}$$

using the relation $2\cos^2 y = \cos 2y + 1$.

Ignoring the constant driving term on the right-hand side, the last equation is the same as (3.1) with $\gamma = 0$. The particular solution is (3.5) with $\gamma = 0$:

$$\eta = \frac{\frac{A^2}{2}\cos 2\omega_0 t}{\omega_0^2 - (2\omega_0)^2} + \frac{A^2}{2\omega_0^2}.$$

The last constant is added to satisfy the constant driving term. Simplifying,

$$\eta = \frac{A^2}{2\omega_0^2}\big(1 - \tfrac{1}{3}\cos 2\omega_0 t\big).$$

It is characteristic of the solution that η is proportional to a *power* of the unperturbed amplitude A (in this case the square). It is also characteristic that the presence of the nonlinear term causes a perturbation to the response which occurs at a *harmonic* of the fundamental frequency ω_0, in this case the second harmonic. In general, departures from linearity cause an oscillator to exhibit *harmonic distortion*, as is all too obvious in a loudspeaker that is driven too hard.

The perturbation solutions to slightly nonlinear oscillators are rarely as straightforward as in this example. Often η is found to contain terms that increase indefinitely with t, so that $\lambda\eta$ eventually cannot remain small. Nevertheless, the main point of this example is to introduce the method of perturbations. Recapitulating, to solve a standard problem when a small additional force or potential is introduced, add a small new term to the standard solution, plug into the differential equation, retain terms to first order in smallness, take advantage of the cancellations, and solve for the new term.

4. Calculus of variations.

4.1. Euler equation.

Initially we focus on the purely mathematical problem of finding the *shape* of a curve $y(t)$, where t is an independent variable not necessarily equal to the time, such that the quantity

$$J \equiv \int_{t_1}^{t_2} \mathcal{L}(y, \dot{y}, t)\, dt$$

is *stationary*. Here $\mathcal{L}$ is an arbitrary, continuously differentiable function of the indicated variables. In other words, we seek a path $y(t)$ such that the integral of $J = \mathcal{L}(y, \dot{y}, t)$ doesn't vary as the path is varied infinitesimally. Usually this means that J is minimized or maximized.

Consider all possible paths between (t_1, y_1) and (t_2, y_2). To make the problem more specific, consider only the subset of paths which begin at *fixed* $y_1(t_1)$ and end at *fixed* $y_2(t_2)$. We parameterize these paths by a single variable α such that, by convention, $\alpha = 0$ for the "best" path,

i.e. the path which produces a stationary value for the above integral. We label the various paths by α:

$$J_\alpha = \int_{t_1}^{t_2} \mathcal{L}\big(y_\alpha(t), \dot{y}_\alpha(t), t\big)\, dt,$$
$$J(\alpha) = \int_{t_1}^{t_2} \mathcal{L}\big(y(\alpha,t), \dot{y}(\alpha,t), t\big)\, dt.$$

In the latter expression we have chosen to consider J, y, and $\dot{y}$ to be functions of the label α. Both notations have the same meaning; in the following we shall use the latter.

The requirement that α vanish when J is stationary means that

$$\begin{aligned} &\frac{\partial J}{\partial \alpha}\Big|_{\alpha=0} = 0, \text{ where} \\ &\frac{\partial J}{\partial \alpha} = \int_{t_1}^{t_2} \Big[\frac{\partial \mathcal{L}}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\Big]\, dt. \end{aligned} \tag{4.1}$$

Performing a parts integration on the second term,

$$\begin{aligned} \int u dv &= uv - \int v du \\ u = \frac{\partial \mathcal{L}}{\partial \dot{y}} \quad du &= \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} dt \\ dv &= \frac{\partial^2 y}{\partial t \partial \alpha} dt \quad v = \frac{\partial y}{\partial \alpha} \\ \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{y}}\frac{\partial \dot{y}}{\partial \alpha}\, dt &= \frac{\partial \mathcal{L}}{\partial \dot{y}}\frac{\partial y}{\partial \alpha}\Big|_{t_1}^{t_2} - \\ &- \int_{t_1}^{t_2} \frac{\partial y}{\partial \alpha}\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\, dt. \end{aligned}$$

The first term on the right-hand side vanishes because, by assumption, $y(t_1)$ and $y(t_2)$ are the same for every α. Finally,

$$\frac{\partial J}{\partial \alpha} = \int_{t_1}^{t_2} \frac{\partial y}{\partial \alpha}\Big[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\Big]\, dt.$$

The next step is to multiply through by an arbitrary small displacement $\delta\alpha$ and evaluate the derivatives with respect to α at $\alpha = 0$:

$$\frac{\partial J}{\partial \alpha}\Big|_{\alpha=0}\delta\alpha = \int_{t_1}^{t_2} \frac{\partial y}{\partial \alpha}\Big|_{\alpha=0}\delta\alpha\Big[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\Big]\, dt.$$

Defining

$$\begin{aligned} \frac{\partial J}{\partial \alpha}\Big|_{\alpha=0}\delta\alpha &\equiv \delta J \\ \frac{\partial y}{\partial \alpha}\Big|_{\alpha=0}\delta\alpha &\equiv \delta y, \end{aligned}$$

we have

$$\delta J = \int_{t_1}^{t_2} \delta y\Big[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}}\Big]\, dt. \tag{4.2}$$

We are now in a position to make the final argument. Because we are considering all possible paths between (y_1, t_1) and (y_2, t_2), δy is a *completely arbitrary* displacement at each point on the path. δJ can vanish *only if* the part of the integrand that multiplies δy also vanishes, i.e.

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} \\ \mathcal{L} &= \mathcal{L}(y, \dot{y}, t). \end{aligned} \tag{4.3}$$

This is the celebrated *Euler equation*. When $\mathcal{L}$ is equal to the *Lagrangian* $T - U$, J is called the *action* and the Euler equation becomes the *Euler-Lagrange equation*.

4.2. Example using Euler equation.

As an example of the use of the Euler equation, again considering a purely mathematical problem, we *minimize the surface of revolution*. The problem is easier to visualize if we temporarily change the notation. Denote the independent variable, usually t, as the (cylindrical) radius r; and denote the generalized coordinate, usually y, as the distance z along the (cylindrical) axis. In this notation, the Euler equation is

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dr}\frac{\partial \mathcal{L}}{\partial \dot{z}} \\ \mathcal{L} &= \mathcal{L}(z, \dot{z}, r). \end{aligned}$$

In this notation, the problem asks us to find the curve $z(r)$ with fixed endpoints $z(r_1) = z_1$, $z(r_2) = z_2$, such that the surface of revolution about the z axis has minimum area.

Along the curve $z(r)$, the path length is

$$\begin{aligned} ds &= [(dr)^2 + (dz)^2]^{1/2} \\ &= dr[1 + \dot{z}^2]^{1/2}, \end{aligned}$$

where $\dot{z}$ means differentiation of z with respect to the independent variable r. When rotated about the z axis, this element of path length produces an element of surface area

$$dA = 2\pi r[1 + \dot{z}^2]^{1/2} dr.$$

So the problem reduces to minimizing the integral

$$\int_{r_1}^{r_2} r[1 + \dot{z}^2]^{1/2} dr,$$

subject to the condition that the endpoints z_1 and z_2 are fixed. Then the integrand is

$$\mathcal{L}(z, \dot{z}, r) = r[1 + \dot{z}^2]^{1/2}.$$

Applying the Euler equation,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial z} = 0 &= \frac{d}{dr}\frac{\partial \mathcal{L}}{\partial \dot{z}} \\ &= \frac{d}{dr}\frac{\dot{z}r}{[1 + \dot{z}^2]^{1/2}} \\ r_0 = \text{const} &= \frac{\dot{z}r}{[1 + \dot{z}^2]^{1/2}} \\ r_0^2(1 + \dot{z}^2) &= \dot{z}^2 r^2 \\ r_0^2 &= \dot{z}^2(r^2 - r_0^2) \\ \frac{dz}{dr} &= \frac{r_0}{(r^2 - r_0^2)^{1/2}} \\ z &= r_0 \int \frac{dr}{(r^2 - r_0^2)^{1/2}} \\ &= r_0 \cosh^{-1}\frac{r}{r_0} + z_0 \\ r &= r_0 \cosh\frac{z - z_0}{r_0}, \end{aligned}$$

where r_0 and z_0 are constants determined by z_1 and z_2. This is the equation of a *catenary* (the shape of the cables on a suspension bridge).

4.3. Equations of constraint.

An *equation of constraint* is an *additional* equation introduced to constrain the generalized coordinate $y(t)$. Suppose, as before, we consider the problem of making stationary the action J when the integrand $\mathcal{L}$ is a function only of a *single* coordinate y, its time derivative $\dot{y}$, and the independent variable t. In this simplest case it would be foolish to impose an equation of constraint, because the constraint would determine $y(t)$ by itself, and there would be nothing left to make stationary. Therefore, equations of constraint are relevant only when $\mathcal{L}$ is a function of two or more coordinates.

Suppose that $\mathcal{L} = \mathcal{L}(y, z, \dot{y}, \dot{z}, t)$. Temporarily, we will assume that no constraint equations apply. Under these circumstances, the derivation of the Euler equations proceeds much the same way as in section 4.1. The integrand in Eq. (4.1) acquires *four* terms, because $\mathcal{L}$ must be differentiated with respect to y, z, $\dot{y}$, and $\dot{z}$. Equation (4.2) becomes

$$\begin{aligned} \delta J = \int_{t_1}^{t_2} \Big\{ \delta y \Big[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} \Big] + \\ + \delta z \Big[\frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{z}} \Big] \Big\} dt. \end{aligned} \tag{4.4}$$

Since each of the virtual displacements δy and δz are independent, two Euler equations must be satisfied for δJ to vanish:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} \\ 0 &= \frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{z}} \\ \mathcal{L} &= \mathcal{L}(y, z, \dot{y}, \dot{z}, t). \end{aligned} \tag{4.5}$$

For n generalized coordinates, one obtains n equations of the same form.

Now that we understand how to handle the case of two generalized coordinates when there is no equation of constraint, we return to the main problem. We are faced with finding a path $y(t), z(t)$ that makes stationary the path integral of $\mathcal{L}$ when the path's endpoints are fixed – while satisfying an additional equation that constrains y and z. Unfortunately, the constraint equation can take many forms. If it is a *differential* equation that is *not of first order*, the only hope for doing the problem is to solve the constraint

equation, or at least reduce it to first order by changing variables. If the constraint equation is a first-order differential equation

$$g_y \frac{dy}{dt} + g_z \frac{dz}{dt} + g_t = 0, \tag{4.6}$$

where g_y, g_z, and g_t, like $\mathcal{L}$, are functions of y, z, $\dot{y}$, $\dot{z}$, and t, it can be solved by the method of *Lagrange undetermined multipliers*, which is the subject of the next section.

If the constraint equation is an *algebraic* rather than a differential equation, $G(y, z, t) = 0$, it is called *holonomic* and is merely a special case of Eq. (4.6), since

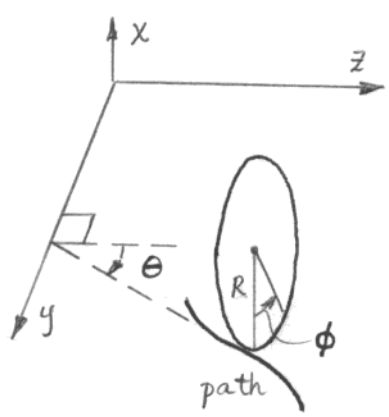
$$g_y \equiv \frac{\partial G}{\partial y} \quad g_z \equiv \frac{\partial G}{\partial z} \quad g_t \equiv \frac{\partial G}{\partial t}.$$

Finally, if the constraint equation takes the form

$$\int_{t_1}^{t_2} \mathcal{G}\, dt = \text{constant}, \tag{4.7}$$

where $\mathcal{G}$ is a function of the same variables as $\mathcal{L}$, the problem can also be solved by using an undetermined multiplier.

Before proceeding to discuss undetermined multipliers, it is instructive to identify the differential (nonholonomic) equation of constraint for a physically meaningful example. Consider a thin coin of radius R rolling upright on the plane $x = 0$, where $\hat{x}$ is up. The C.M. of the coin is at (y, z). Let ϕ be the azimuth of the coin, measured with respect to the point of contact, and let θ be the orientation of the face of the coin, measured with respect to the xz plane. As long as the coin remains upright, these four generalized coordinates define the position and orientation of the coin, and their four time derivatives completely describe its motion.



If the surface $x = 0$ is frictionless, there need be no equation of constraint relating the generalized coodinates. However, if there is friction, we might require that the coin be *rolling without slipping*. In that case, for a particular choice of sign for ϕ and θ, the constraint equations are:

$$-\dot{y} = R\dot{\phi}\sin\theta$$
$$-\dot{z} = R\dot{\phi}\cos\theta.$$

We cannot integrate and solve these equations without solving the whole problem. However, we can identify the coefficients g_i in the notation of Eq. (4.6). For these two equations of constraint, they are, respectively,

$$g_y = 1 \quad g_z = 0 \quad g_\theta = 0 \quad g_\phi = R\sin\theta \quad g_t = 0$$
$$g_y = 0 \quad g_z = 1 \quad g_\theta = 0 \quad g_\phi = R\cos\theta \quad g_t = 0.$$

4.4. Method of Lagrange undetermined multipliers.

In Eq. (4.4), when y and z were two independent generalized coordinates, we argued that δy and δz were independently arbitrary. This yielded the two Euler equations in Eq. (4.5). However, the presence of an equation of constraint which links y and z destroys that independence. The constraint equation (4.6) links the two virtual displacements:

$$g_y \delta y + g_z \delta z = -g_t \delta t = 0.$$

The term on the right-hand side vanishes because the displacements occur at fixed time t for each point on the path $y(t), z(t)$.

Now we introduce the *Lagrange undetermined multiplier* $\lambda(t)$. For any λ it is obvious that

$$(g_y \delta y + g_z \delta_z)\lambda = 0.$$

Inserting this in Eq. (4.4),

$$\delta J = \int_{t_1}^{t_2} \left\{ \delta y \Big[\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} + \lambda g_y \Big] + \right.$$
$$\left. + \delta z \Big[\frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{z}} + \lambda g_z \Big] \right\} dt. \tag{4.8}$$

Since δy and δz cannot be independent of each other, we *choose* δy to be the independent virtual displacement and δz to be the dependent one. This means that the first square bracket in Eq. (4.8) must vanish. Since λ is undetermined, we are free to *choose* λ so that the second square bracket in Eq. (4.8) vanishes too.

To summarize, the effect of introducing the equation of constraint and the Lagrange undetermined multiplier is to increase the number of unknown functions of t from two ($y(t)$ and $z(t)$) to three, with $\lambda(t)$ included; and, correspondingly, to increase the number of differential equations from two to three:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} + \lambda g_y \\ 0 &= \frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{z}} + \lambda g_z \\ 0 &= g_y \dot{y} + g_z \dot{z} + g_t, \end{aligned} \tag{4.9}$$

where the last equation is just a repetition of Eq. (4.6).

4.5. Lagrange multiplier applied to integral constraint.

Suppose the constraint equation is of the integral form (4.7). Since the integral in that equation is constant, as α is varied the virtual displacement of the integral is zero. This means that we can multiply the integrand $\mathcal{G}$ in (4.7) by a *constant* undetermined multiplier Λ and add it to the integrand $\mathcal{L}$ in the virtual displacement of the action:

$$0 = \delta J = \delta \int_{t_1}^{t_2} (\mathcal{L} + \Lambda \mathcal{G})\, dt.$$

The resulting set of Euler equations is the same as Eq. (4.5) with $\mathcal{L}$ replaced by $\mathcal{L} + \Lambda\mathcal{G}$. In their derivation, which we do not elaborate here, the addition of $\Lambda\mathcal{G}$ to $\mathcal{L}$ makes it possible to continue to regard the virtual displacements δy and δz as independent, even in the presence of the integral equation of constraint.

4.6. Alternate form of Euler equation.

Returning to the case in which the integrand is a function of only one generalized coordinate, $\mathcal{L}$ can vary with time both explicitly and also through the time dependence of y or $\dot{y}$:

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \frac{\partial \mathcal{L}}{\partial y}\dot{y} + \frac{\partial \mathcal{L}}{\partial \dot{y}}\ddot{y}.$$

Using the Euler equation (4.3) for $\partial\mathcal{L}/\partial y$,

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \frac{\partial \mathcal{L}}{\partial t} + \dot{y}\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}} + \frac{\partial \mathcal{L}}{\partial \dot{y}}\ddot{y} \\ &= \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt}\big(\dot{y}\frac{\partial \mathcal{L}}{\partial \dot{y}}\big) \\ -\frac{\partial \mathcal{L}}{\partial t} &= \frac{d}{dt}\big(\dot{y}\frac{\partial \mathcal{L}}{\partial \dot{y}} - \mathcal{L}\big) \equiv \frac{d}{dt}\mathcal{H}. \end{aligned} \tag{4.10}$$

If $\mathcal{L}$ is free of *explicit* time dependence, $\mathcal{H} = E$, where E is a constant. This equality can substitute for the Euler equation in solving the same extremization problem. It leads to much simpler algebra in some cases.

If $\mathcal{L}$ is the Lagrangian $T - U$, and if t is the time, $\mathcal{H}$ as defined in (4.10) is the *Hamiltonian*. If $\mathcal{L}$ is free from explicit time dependence, the constant E is the conserved total mechanical energy. If U is independent of $\dot{y}$ (velocity-independent potential) and if T is a quadratic function of $\dot{y}$,

$$\begin{aligned} E = \mathcal{H} &\equiv \dot{y}\frac{\partial \mathcal{L}}{\partial \dot{y}} - \mathcal{L} \\ &= 2T - (T - U) = T + U. \end{aligned}$$

Under these conditions, the conserved total mechanical energy is the sum of kinetic and potential energies as expected.

5. Lagrangian mechanics.

5.1. Hamilton's principle.

The foregoing discussion of the calculus of variations, with or without equations of constraint, acquires physical relevance from a famous postulate by Hamilton. From now on, the independent variable t will represent the time, and the coordinates y_i will be called *generalized coordinates*. They are "generalized" in the sense

that any time-dependent quantity that helps to define the state of a system (Cartesian coordinate, spherical or cylindrical coordinate, Euler angle, etc.) can be chosen as a generalized coordinate. Not all the y_i $(1 \le i \le n)$ in the same problem need to have the same dimension. The path $y_i(t)$ followed by a system is called its *history*.

For systems in which all the forces are conservative, Hamilton's principle states that the history which the system actually will follow is that which makes the action J *stationary*, where

$$J \equiv \int_{t_1}^{t_2} \mathcal{L}(y_i, \dot{y}_i, t)\, dt. \tag{5.1}$$

Here $\mathcal{L}$ is the Lagrangian $T - U$. The calculus of variations supplies n Euler-Lagrange equations that may be solved for the history:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}_i} = \frac{\partial \mathcal{L}}{\partial y_i}. \tag{5.2}$$

As another choice, we may define the Hamiltonian, and, if $\partial\mathcal{L}/\partial t = 0$, we may use the alternate equation

$$\begin{aligned} \mathcal{H}(y_i, \dot{y}_i) &\equiv \dot{y}_i \frac{\partial \mathcal{L}}{\partial \dot{y}_i} - \mathcal{L} \\ &= E = \text{constant}. \end{aligned} \tag{5.3}$$

We accept Hamilton's principle because it is found to reproduce all solutions obtained using Newtonian analysis, and because it agrees with experimental observation for additional classes of problem beyond the reach of Newtonian analysis. Hamiltonian analysis is also the historical path to quantum mechanics.

Often T and U are generalized quadratic functions of the y_i and $\dot{y}_i$ respectively:

$$T = \tfrac{1}{2} T_{ij} \dot{y}_i \dot{y}_j \quad U = \tfrac{1}{2} U_{ij} y_i y_j, \tag{5.4}$$

where the T_{ij} are functions only of the y_i, and the U_{ij} are functions only of the $\dot{y}_i$. In this case $\mathcal{H} = T + U$, and the Euler-Lagrange equations simplify to

$$T_{ij} \ddot{y}_j = -U_{ij} y_j. \tag{5.5}$$

If, for some i, $T_{ij} = m\delta_{ij}$ and $U_{ij} = k\delta_{ij}$, the Euler-Lagrange equation for the ith Cartesian coordinate looks like a component of $m\mathbf{a} = \mathbf{F}$:

$$m\ddot{y}_i = -k y_i.$$

Evidently $(d/dt)(\partial\mathcal{L}/\partial\dot{y}_i)$ plays the role of the rate of change of a momentum, and $\partial\mathcal{L}/\partial y_i$ plays the role of a force. More generally, $(\partial\mathcal{L}/\partial\dot{y}_i)$ is called a *generalized momentum* and $(\partial\mathcal{L}/\partial y_i)$ is called a *generalized force*. For example, if y_i is really an angle, the corresponding generalized momentum is really an angular momentum, and the generalized force is really a torque.

Equation (4.9) added Lagrange multiplier terms to the Euler-Lagrange equations for the case in which two generalized coordinates were mutually constrained. More generally, there may be n generalized coordinates y_i and m constraint equations, $1 \le j \le m$:

$$g_{ji} dy_i + g_{jt} dt = 0. \tag{5.6}$$

Each of the $m \times (n+1)$ g's is a function of the $2n+1$ variables $(y_i, \dot{y}_i, t)$. The problem is solved by adding n Euler-Lagrange equations each containing m Lagrange multipliers λ_j:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{y}_i} = \frac{\partial \mathcal{L}}{\partial y_i} + g_{ji}\lambda_j. \tag{5.7}$$

In addition to the generalized force $\partial\mathcal{L}/\partial y_i$ that is derived from the potential U, the right-hand side contains an additional term $g_{ji}\lambda_j$ (summation implied) which is the *generalized force of constraint* Q_i. This is the physical significance of the Lagrange multipliers: when we solve for the λ's as well as the y_i's, we obtain the forces of contraint as well as the motion. To summarize, $n+m$ unknown functions $y_i(t)$ and $\lambda_j(t)$ are solved by n Euler-Lagrange equations containing Lagrange multipliers, plus m equations of constraint.

Note that the generalized forces of constraint *must do no work* – otherwise we could no longer continue to define the potential energy U. For example, the force of constraint from a wall must be normal rather than frictional.

5.2. The falling ladder.

This interesting problem is solved straightforwardly by using the Euler-Lagrange equations with a Lagrange multiplier. At $t = 0$ a ladder of length h under the influence of gravity is released from rest. The bottom is in contact with a frictionless floor $y = 0$. The top rests on a frictionless wall $x = 0$ with which the ladder makes an initial angle α_0.

Part (a) of the problem is just a warm-up: Assuming $\alpha \ll 1$, find $\alpha(t)$.

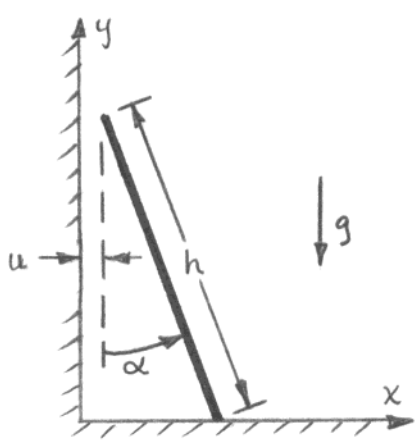
For this part we assume that the ladder remains in contact with the wall. Then its position is specified by only one generalized coordinate, which we take as α. Then the C.M. of the ladder is at $x = (h/2)\sin\alpha,\ y = (h/2)\cos\alpha$. The Lagrangian is:

$$\begin{aligned}
\mathcal{L} &= T - U \\
&= T_{\text{trans}} + T_{\text{rot}} - U \\
&= \tfrac{1}{2}mv^2 + \tfrac{1}{2}I\omega^2 - U \\
&= \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}(\tfrac{1}{12}mh^2\dot{\alpha}^2) - \tfrac{1}{2}mgh\cos\alpha \\
&= \tfrac{1}{8}mh^2\dot{\alpha}^2 + \tfrac{1}{24}mh^2\dot{\alpha}^2 - \tfrac{1}{2}mgh\cos\alpha \\
&= \tfrac{1}{6}mh^2\dot{\alpha}^2 - \tfrac{1}{2}mgh\cos\alpha.
\end{aligned}$$

Applying the Euler-Lagrange equation (5.2),

$$\begin{aligned}
\frac{d}{dt}\tfrac{1}{3}mh^2\dot{\alpha} &= \tfrac{1}{2}mgh\sin\alpha \\
\tfrac{1}{3}mh^2\ddot{\alpha} &= \tfrac{1}{2}mgh\sin\alpha \\
\ddot{\alpha} &= \frac{3g}{2h}\sin\alpha \approx \frac{3g}{2h}\alpha \\
\alpha &\approx \alpha_0\cosh\sqrt{\frac{3g}{2h}}t,
\end{aligned}$$

where in the last equation the boundary conditions have been applied.



Part (b) of the falling ladder problem is more interesting: *Without* assuming $\alpha \ll 1$, find the angle α_1 at which the ladder leaves the wall.

We relax the requirement that the top of the ladder rest against the wall by introducing a second coordinate u, the distance between the top of the ladder and the wall. The equation of constraint imposed by the wall then is $u \geq 0$. Since inequalities are difficult to handle, we instead impose the constraint $u = 0$. Using the method of Lagrange multipliers, we solve for the generalized force of constraint $Q_u = \lambda g_u$. The ladder will leave the wall when the wall exerts no force on it, i.e. when $Q_u = 0$.

Re-expressed in terms of α and u, the C.M. coordinate y is the same, but x acquires the extra term u. This leads to two extra terms

$$\tfrac{1}{2}mh\cos\alpha\dot{\alpha}\dot{u} + \tfrac{1}{2}m\dot{u}^2$$

in T. The constraint equation is

$$\begin{aligned}
g_\alpha d\alpha + g_u du + g_t dt &= 0 \\
g_\alpha = 0 \quad g_u = 1 \quad g_t &= 0.
\end{aligned}$$

When we include the Lagrange multiplier, the Euler-Lagrange equation in α is unchanged, since $\dot{u} = 0$ and $g_\alpha = 0$. The equation in u is

$$\begin{aligned}
\frac{d}{dt}\left[\tfrac{1}{2}m(h\cos\alpha\dot{\alpha} + 2\dot{u})\right] &= \lambda(t) \\
-\sin\alpha\dot{\alpha}^2 + \cos\alpha\ddot{\alpha} &= \frac{2}{mh}\lambda(t).
\end{aligned}$$

In the second equation, we used the fact that $\dot{u} = 0$. When the ladder leaves the wall, $\alpha \equiv \alpha_1$, and $\lambda = 0$. The Euler-Lagrange equation in u becomes

$$\ddot{\alpha}\big|_{\alpha_1} = \dot{\alpha}^2\tan\alpha\big|_{\alpha_1}.$$

Substituting for $\ddot{\alpha}$ from the Euler-Lagrange equation in α,

$$\frac{3g}{2h}\cos\alpha\big|_{\alpha_1} = \dot{\alpha}^2\big|_{\alpha_1}. \tag{5.8}$$

Equation (5.8) would solve the problem if we could obtain a condition expressing $\dot{\alpha}$ in terms of α. Then we would have an equation for α_1 that might be solved. Such a condition is provided by the alternate form of the Euler-Lagrange equation. Returning to the analysis in terms of a single generalized coordinate α, appropriate for the part of the motion in which the ladder remains in contact with the wall, we define

$$\begin{aligned}\mathcal{H} &\equiv \dot{\alpha}\frac{\partial \mathcal{L}}{\partial \dot{\alpha}} - \mathcal{L} \\ &= \dot{\alpha}\tfrac{1}{3}mh^2\dot{\alpha} - \tfrac{1}{6}mh^2\dot{\alpha}^2 + \tfrac{1}{2}mgh\cos\alpha \\ &= \tfrac{1}{6}mh^2\dot{\alpha}^2 + \tfrac{1}{2}mgh\cos\alpha = E,\end{aligned}$$

where in the last equation we used the fact that $\mathcal{L}$ does not depend explicitly on the time. Equating $\mathcal{H}$ above to $\mathcal{H}$ at $(t = 0, \alpha = \alpha_0, \dot{\alpha} = 0)$,

$$\tfrac{1}{6}mh^2\dot{\alpha}^2 + \tfrac{1}{2}mgh\cos\alpha = \tfrac{1}{2}mgh\cos\alpha_0.$$

Setting $\alpha = \alpha_1$ and substituting Eq. (5.8) for $\dot{\alpha}$,

$$\begin{aligned}\tfrac{1}{6}mh^2\frac{3g}{2h}\cos\alpha_1 &= \tfrac{1}{2}mgh(\cos\alpha_0 - \cos\alpha_1) \\ \tfrac{3}{2}\cos\alpha_1 &= \cos\alpha_0 \\ \alpha_1 &= \cos^{-1}(\tfrac{2}{3}\cos\alpha_0).\end{aligned}$$

If the ladder starts at a certain cosine, it leaves the wall when the height of the ladder decreases to $\frac{2}{3}$ of its initial value. For example, if it starts upright (at 0°), it leaves at $\cos^{-1}\frac{2}{3} = 48.2°$.

5.3. Cyclic coordinates and conservation laws.

In section 5.1 we identified $\partial\mathcal{L}/\partial\dot{q}_i$ as a generalized momentum (we have substituted $\dot{q}_i$ for $\dot{y}_i$, following the notation traditionally used for this topic). In fact, this generalized momentum is given a longer name – the *canonical momentum conjugate to* q_i, or (more succinctly) p_i. The reason for the emphasis is that, when there is no constraint equation and $\partial\mathcal{L}/\partial q_i$ vanishes, the Euler-Lagrange equation requires that the canonical momentum be *conserved*. $\mathcal{L}$ is said to be *cyclic* in q_i.

If a system is closed, the homogeneity of spacetime requires the Lagrangian to be invariant to displacements or rotations of the coordinate system, and also to displacements in the zero of time. Using the canonical momenta, each of these invariance principles gives rise to a conservation law.

Consider a displacement δ_x in all the linear coordinates q_x^j in the x direction of the various particles j. The Lagrangian is unchanged under this displacement:

$$\begin{aligned}0 = \frac{\partial\mathcal{L}}{\partial\delta_x} &= \sum_j \frac{\partial\mathcal{L}}{\partial q_x^j}\frac{\partial q_x^j}{\partial\delta_x} \\ &= \sum_j \frac{\partial\mathcal{L}}{\partial q_x^j} \\ &= \frac{d}{dt}\sum_j \frac{\partial\mathcal{L}}{\partial\dot{q}_x^j} \\ &\equiv \frac{d}{dt}\sum_j p_x^j \equiv \frac{d}{dt}P_x.\end{aligned} \tag{5.9}$$

(The first equation does *not* imply summation over x, and the third equation uses the fact that $\partial q_x^j/\partial\delta_x = 1$). Thus P_x, the total linear momentum in the x direction, is conserved.

Consider next a rotation ϵ_x about the x axis. This is equivalent to a displacement ϵ_x in all the angular coordinates θ_x^j about the x axis of the various particles j. By a similar derivation, invariance of the Lagrangian to this rotation leads to the conservation of total angular momentum L_x:

$$\begin{aligned}0 &= \frac{d}{dt}\sum_j \frac{\partial\mathcal{L}}{\partial\theta_x^j} \\ &\equiv \frac{d}{dt}\sum_j l_x^j \equiv \frac{d}{dt}L_x.\end{aligned} \tag{5.10}$$

Finally, we saw in Eq. (5.3) that invariance of $\mathcal{L}$ to a displacement in time t causes the Hamiltonian $\mathcal{H}$ to be conserved. Thus $\mathcal{H}$ must bear a canonical relationship to t similar to that of P_x to δ_x or of L_x to ϵ_x. However, the analogy is not complete, because t plays a special role as the independent variable upon which the q's and p's depend.

6. Hamiltonian mechanics.

6.1. Hamilton's equations.

We have already introduced one equation involving the Hamiltonian $\mathcal{H}$. Equation (4.10) defined $\mathcal{H}$ and obtained its time derivative:

$$\frac{d\mathcal{H}(q_i,\dot{q}_i,t)}{dt} = -\frac{\partial\mathcal{L}(q_i,\dot{q}_i,t)}{\partial t}. \tag{6.1}$$

In Eq. (6.1), the Hamiltonian is (temporarily!) considered to be a function of the same variables as is the Lagrangian.

However, we are advised to put Eq. (6.1) on the back burner, because $\mathcal{H}$ instead is normally considered to be a function of the generalized coordinates, their *canonically conjugate momenta*, and the time: $\mathcal{H} = \mathcal{H}(q_i, p_i, t)$. When expressed in terms of the p_i rather than the $\dot{q}_i$, the Hamiltonian may take on a different functional form. For example, using polar coordinates r and θ, a free particle has the Hamiltonia

$$\begin{aligned}\mathcal{H}(r,\theta,\dot{r},\dot{\theta},t) &= \frac{m}{2}\big(\dot{r}^2 + r^2\dot{\theta}^2\big)\\ \mathcal{H}(r,\theta,p_r,p_\theta,t) &= \frac{1}{2m}\big(p_r^2 + \frac{p_\theta^2}{r^2}\big).\end{aligned}$$

Considering $\mathcal{H}$ to be a function of (q_i, p_i, t), and $\mathcal{L}$ still to be a function of $(q_i, \dot{q}_i, t)$, we examine the total differential of $\mathcal{H}$:

$$\begin{aligned}\mathcal{H} &= \dot{q}_i p_i - \mathcal{L}\\ d\mathcal{H} &= \dot{q}_i dp_i + p_i d\dot{q}_i - \frac{\partial\mathcal{L}}{\partial q_i}dq_i - \frac{\partial\mathcal{L}}{\partial\dot{q}_i}d\dot{q}_i - \frac{\partial\mathcal{L}}{\partial t}dt\\ &= \dot{q}_i dp_i + p_i d\dot{q}_i - \dot{p}_i dq_i - p_i d\dot{q}_i - \frac{\partial\mathcal{L}}{\partial t}dt\\ &= \dot{q}_i dp_i - \dot{p}_i dq_i - \frac{\partial\mathcal{L}}{\partial t}dt.\end{aligned} \tag{6.2}$$

(In the third equality we made use of the Euler-Lagrange equations.)

On the other hand, $\mathcal{H}$ formally is a function of (q_i, p_i, t):

$$d\mathcal{H} \equiv \frac{\partial\mathcal{H}}{\partial q_i}dq_i + \frac{\partial\mathcal{H}}{\partial p_i}dp_i + \frac{\partial\mathcal{H}}{\partial t}dt. \tag{6.3}$$

Since $\mathcal{H}$ depends independently on q_i, p_i, and t, Eqs. (6.2) and (6.3) must be equivalent term by term. The result is *Hamilton's equations*:

$$\begin{aligned}\mathcal{H} &= \mathcal{H}(q_i,p_i,t)\\ \dot{q}_i &= +\frac{\partial\mathcal{H}}{\partial p_i}\\ \dot{p}_i &= -\frac{\partial\mathcal{H}}{\partial q_i}\\ \frac{\partial\mathcal{H}}{\partial t} &= -\frac{\partial\mathcal{L}}{\partial t}.\end{aligned} \tag{6.4}$$

Note that the total and partial time derivatives of $\mathcal{H}(q_i, p_i, t)$ (but not $\mathcal{H}(q_i, \dot{q}_i, t)$!) are equivalent:

$$\begin{aligned}\frac{d\mathcal{H}}{dt} &= \frac{\partial\mathcal{H}}{\partial q_i}\dot{q}_i + \frac{\partial\mathcal{H}}{\partial p_i}\dot{p}_i + \frac{\partial\mathcal{H}}{\partial t}\\ &= -\dot{p}_i\dot{q}_i + \dot{q}_i\dot{p}_i + \frac{\partial\mathcal{H}}{\partial t}\\ &= \frac{\partial\mathcal{H}}{\partial t}.\end{aligned}$$

Therefore, the last of Hamilton's equations is written with equal validity using $d\mathcal{H}/dt$ or $\partial\mathcal{H}/\partial t$. In contrast, Eq. (6.1) considered $\mathcal{H}$ still to be a function of $(q_i, \dot{q}_i, t)$; we do not substitute the partial time derivative of $\mathcal{H}$ there.

6.2. The Poisson bracket.

Applied to straightforward problems that are amenable to Lagrangian analysis, Hamilton's equations substitute two coupled partial differential equations of first order for one Euler-Lagrange equation – an ordinary differential equation of second order. Usually, this is no bargain. Most often, after manipulation, one obtains from Hamiltonian analysis the same differential equations that are found more easily from the Lagrangian.

For more difficult problems, there exists a prescription for making a *canonical transformation* to a new set of generalized coordinates Q_i and canonically conjugate momenta P_i, so that (at least) all the P_i are constants of the motion. Although Hamilton's equations become trivial when expressed in terms of the new variables, finding the right canonical transformation requires solving partial differential equations that

are not trivial. This is the most powerful method for analyzing problems in classical mechanics. It is beyond the scope of a one-semester undergraduate course.

For present purposes, the attraction of Hamilton's equations is that they encourage us to think democratically about coordinates and momenta, and that they reveal essential ideas which led from classical to quantum mechanics. The *Poisson bracket* is a good example of both attractions.

Consider any quantity $\rho(q_i, p_i, t)$ that is a function of the same variables as $\mathcal{H}$. Its total time derivative is

$$\begin{aligned} \frac{d\rho}{dt} &= \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q_i}\dot{q}_i + \frac{\partial\rho}{\partial p_i}\dot{p}_i \\ &= \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i} \\ &\equiv \frac{\partial\rho}{\partial t} + [\rho, \mathcal{H}]. \end{aligned} \tag{6.5}$$

In the second equality we used Hamilton's equations, and in the last we defined the *Poisson bracket* of ρ and $\mathcal{H}$,

$$[\rho, \mathcal{H}] \equiv \frac{\partial\rho}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i}, \tag{6.6}$$

as usual with summation over i implied. Equation (6.5) states that the *implicit* time derivative of any function of (q_i, p_i, t) is given by the Poisson bracket of that function with $\mathcal{H}$: the Hamiltonian is the unique function which controls the time evolution of all other functions. (It is even easier to see why $d\mathcal{H}/dt = \partial\mathcal{H}/\partial t$, as noted in the previous section: the Poisson bracket of $\mathcal{H}$ with itself obviously vanishes.)

In quantum mechanics, if ρ were an *operator*, one would write

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \frac{1}{i\hbar}[\rho, \mathcal{H}],$$

where $[\rho, \mathcal{H}] \equiv \rho\mathcal{H} - \mathcal{H}\rho$ is the *commutator* of ρ with the Hamiltonian operator. The quantity $\hbar$, with the same dimension as pq, is *Planck's constant* h divided by 2π. Thus the Poisson bracket is the quantity in classical mechanics that led to the quantum mechanical commutator.

6.3. Phase space.

Consider a system of n particles in three-dimensional space. This system is described by a set of $6n$ variables $\{q_i, p_i\}$, where, for example, $i = 1\,2\,3$ for $x\,y\,z$ of the first particle, $4\,5\,6$ for $x\,y\,z$ of the second particle, etc. The system is represented as a *single point* in a $6n$ dimensional space called *phase space*.

If we are required to consider an *ensemble* of a large number N of identical systems, each satisfying initial conditions that (in principle) could be unique, the number $(3nN)$ of Euler-Lagrange equations becomes unwieldy. To learn about the average behavior of these systems, we are motivated to consider the *statistical* properties of the ensemble. Suppose that we prepare an ensemble of N systems initially within a closed boundary in $6n$-dimensional phase space. The ensemble can be visualized as a N points within a $6n$-dimensional bag. After a time, each system evolves to a new point in phase space. The bag also evolves to a new shape.

Our first conclusion is that any system that initially lies within the closed boundary will always lie within the (evolving) boundary: *a bounded system remains bounded*. The reason is simple: if at some time a system lies on a phase space point which is part of the boundary, from that time forward it must evolve *with the boundary*. It cannot cross the boundary. Otherwise, we would have two identical systems (one remaining on the boundary, one crossing it) that evolve differently given the same initial conditions.

6.4. Liouville's theorem.

This celebrated theorem has wide-ranging practical implications. It states that the volume in phase space occupied by an ensemble of systems remains constant as the systems evolve, when the Hamiltonian is constant or even when it varies (smoothly) with time. Let ρ be the number of systems per unit phase space volume. (As an aside, if each system were a spin $\frac{1}{2}$ fermion such as an electron, which obeys the

Pauli exclusion principle, the maximum value of ρ would be $2h^{-3n}$, where the factor of 2 allows the electron spin to point up or down.)

To analyze the evolution of ρ we need the $6n$-dimensional velocity $\mathbf{v}_p$ and the $6n$-dimensional gradient operator ∇_p:

$$\mathbf{v}_p \equiv \hat{q}_i\dot{q}_i + \hat{p}_i\dot{p}_i$$
$$\nabla_p \equiv \hat{q}_i\frac{\partial}{\partial q_i} + \hat{p}_i\frac{\partial}{\partial p_i},$$

where summation over i is implied as usual. In the $6n$-dimensonal space, the number N of systems is conserved: an increase in the density of systems in a certain region requires a net flux of systems into the region. This is expressed quantitatively by an equation of continuity:

$$\frac{\partial\rho}{\partial t} = -\nabla_p\cdot(\rho\mathbf{v}_p). \tag{6.7}$$

Using Eq. (6.5), the total time derivative of ρ is

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + [\rho,\mathcal{H}].$$

Substituting for $\partial\rho/\partial t$ from Eq. (6.7),

$$\begin{aligned}\frac{d\rho}{dt} &= [\rho,\mathcal{H}] - \nabla_p\cdot(\rho v_p)\\ &= [\rho,\mathcal{H}] - \Big\{\frac{\partial}{\partial q_i}(\rho\dot{q}_i) + \frac{\partial}{\partial p_i}(\rho\dot{p}_i)\Big\}.\end{aligned} \tag{6.8}$$

Using Hamilton's equations for $\dot{q}_i$ and $\dot{p}_i$, the curly bracket in Eq. (6.8) is

$$\frac{\partial\rho}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial\rho}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i} + \rho\Big[\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i}\Big].$$

The first two terms are just $[\rho,\mathcal{H}]$, cancelling the Poisson bracket in Eq. (6.8). The terms in the square bracket are equal, respectively, to

$$+\frac{\partial^2\mathcal{H}}{\partial q_i\partial p_i} \quad\text{and}\quad -\frac{\partial^2\mathcal{H}}{\partial p_i\partial q_i}.$$

These terms also cancel, provided that $\mathcal{H}$ varies smoothly with q_i and p_i, so that the order of differentiation does not matter.

We have proven Liouville's theorem,

$$\frac{d\rho}{dt} = 0.$$

Since the phase space density ρ and the number of systems N are constant, the volume N/ρ in phase space must also be constant.

It is significant that the phase space volume is proportional to the range in a particular generalized coordinate q_j multiplied by the range in its canonically conjugate momentum p_j. Suppose, at a certain time, that an ensemble of systems occupies a range Δq_j and a range Δp_j. Later, suppose that Δq_j *decreases* by a factor, and there is no change in the range of other coordinates q_i and momenta p_i with $i \neq j$. Then Δp_j must *increase* by the same factor, so that

$$\Delta q_j\Delta p_j = \text{constant}. \tag{6.9}$$

Nothing in classical mechanics suggests that Δq_j or Δp_j should be interpreted as an *uncertainty*. In the present context, each is merely the range of values accessible to an ensemble of identical systems satisfying different initial conditions. But if, for the moment, we *were* to entertain such an interpretation, Eq. (6.9) would resemble an *uncertainty principle* of the type Heisenberg introduced to quantum mechanics. We would be led to conclude that an uncertainty principle can be written for the pair of quantum mechanical operators corresponding to *any* pair of canonically conjugate classical variables. This turns out to be the case.

As a simple application of Liouville's theorem, imagine a truncated cone of half-angle $\ll 1$ that is perfectly reflecting on the inside. Randomly directed, nearly monochromatic photons enter the larger end. Naively, one might hope that the photons would all be funneled down the cone through the small end, yielding an intensified spot of light.

Liouville would reach a different conclusion. At the larger end (area A), the photons entering the cone already occupy a volume in momentum space equal to that of a hemispherical shell with a thickness corresponding to the small range in photon |momentum|. At the smaller end (area a), the emerging photons would occupy this

same volume in momentum space. Since the volume in position ("configuration") space is a/A times smaller there, and the phase space density must remain constant, only a fraction a/A of the photons can emerge from the small end; the remainder must be multiply reflected back to the large end of the cone. (The condition that $\mathcal{H}$ must vary smoothly with the q_i and p_i is equivalent, in this problem, to the requirement that the cone aperture may change significantly only over a distance along the cone corresponding typically to at least a few photon reflections.)

7. Central force motion.

7.1. Reduced mass.

We consider the central force problem with two point masses m_1 and m_2 separated by a vector $\mathbf{r}$ pointing from m_1 to m_2. From the C.M., vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ point to each mass respectively, so that $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. The C.M. is defined by $m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = 0$.

In principle, the two-body system requires six coordinates to describe it – three for the C.M. coordinate $\mathbf{R}$, and three e.g. for $\mathbf{r}$. We want to reduce this number in order to simplify the analysis. By working in the C.M. system, we eliminate $\mathbf{R}$ from further consideration. About the C.M., either of the (parallel) angular momenta $\mathbf{L}_1 \equiv m_1\mathbf{r}_1\times\dot{\mathbf{r}}_1$ and $\mathbf{L}_2 \equiv m_2\mathbf{r}_2\times\dot{\mathbf{r}}_2$ defines the normal to the plane of relative motion. Since the force between the bodies is central, it can exert no torque about the C.M., so that $\mathbf{L}_1 + \mathbf{L}_2$, and therefore the plane of relative motion, must remain invariant. As our two generalized coordinates, we choose r and θ, the azimuthal angle of $\mathbf{r}$ in the plane of relative motion.

In the C.M., the kinetic energy is

$$T = \tfrac{1}{2}m_1\dot{r}_1^2 + \tfrac{1}{2}m_1 r_1^2\dot{\theta}^2 + \tfrac{1}{2}m_2\dot{r}_2^2 + \tfrac{1}{2}m_2 r_2^2\dot{\theta}^2.$$

Substituting for r_1 and r_2 in terms of r, one finds easily

$$\begin{aligned} T &= \tfrac{1}{2}\mu\dot{r}^2 + \tfrac{1}{2}\mu r^2\dot{\theta}^2, \\ \mu &\equiv \frac{m_1 m_2}{m_1 + m_2}. \end{aligned} \tag{7.1}$$

The quantity μ last defined is the *reduced mass*. It is always between 50% and 100% of the *smaller* of m_1 and m_2.

7.2. Constants of the motion.

Expressed in terms of the reduced mass, the Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2) - U(r), \tag{7.2}$$

and the Hamiltonian is

$$\begin{aligned} \mathcal{H} &= \dot{r}\frac{\partial\mathcal{L}}{\partial\dot{r}} + \dot{\theta}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} - \mathcal{L} \\ &= 2T - (T - U) = T + U \\ &= \tfrac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2) + U(r). \end{aligned}$$

In terms of the canonical momenta,

$$\begin{aligned} p_r &\equiv \frac{\partial\mathcal{L}}{\partial\dot{r}} = \mu\dot{r} \\ p_\theta &\equiv \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = \mu r^2\dot{\theta}, \end{aligned}$$

the Hamiltonian is more naturally written

$$\mathcal{H} = \frac{p_r^2}{2\mu} + \frac{p_\theta^2}{2\mu r^2} + U(r).$$

Because the θ coordinate is cyclic ($\partial\mathcal{L}/\partial\theta = 0$), the Euler-Lagrange equation requires the canonically conjugate momentum p_θ (in this case an *angular* momentum) to be conserved:

$$p_\theta = \text{constant} \equiv l. \tag{7.3}$$

Because $\partial\mathcal{L}/\partial t = 0$, the Hamiltonian is constant as well:

$$\begin{aligned} \mathcal{H} &= \text{constant} \equiv E \\ E &= \frac{p_r^2}{2\mu} + \frac{l^2}{2\mu r^2} + U(r) \\ E &= \frac{\mu\dot{r}^2}{2} + \frac{l^2}{2\mu r^2} + U(r). \end{aligned} \tag{7.4}$$

This last equation involving the single coordinate r is the basis of the analysis that follows.

7.3. The repulsive pseudopotential.

We know that the term $l^2/2\mu r^2$ on the right-hand side of Eq. (7.4) is part of the kinetic energy. However, because it depends on r rather than $\dot{r}$, it has the functional form of a *potential* energy, and is called the *pseudopotential* U'. Near the origin, U' is singular and positive, and therefore repulsive.

If the actual potential $U(r)$ is attractive (negative), and is proportional to r^{n+1} with $-3 < n < -1$, the sum of U and U' is $+\infty$ at $r = 0$ and 0 at $r = \infty$, falling to a minimum $-|U_0|$ at some finite radius r_0. If E is negative, the pair of masses is a *bound system*. If $E = -|U_0|$, all of the kinetic energy arises from angular motion $l^2/2\mu r_0^2$ rather than radial motion $\frac{1}{2}\mu\dot{r}^2$, and the masses are in a *stable circular orbit*. If E is larger than $-|U_0|$, but still negative, the radius of the orbit varies between the perigee $r_{\min}$ and the apogee $r_{\max}$. At each of these classical turning points, $E = U + U'$. When $n = -2$, corresponding to a gravitational or Coulomb potential, the orbit is an *ellipse* when $-|U_0| < E < 0$, a *parabola* when $E = 0$, or a *hyperbola* when $E > 0$.

7.4. Period and orbit shape.

Defining

$$\mathcal{E} \equiv \frac{2\mu E}{l^2}; \;\; \mathcal{U} \equiv \frac{2\mu U}{l^2},$$

Eq. (7.4) becomes

$$\begin{aligned} \mathcal{E} &= (\frac{\mu\dot{r}}{l})^2 + r^{-2} + \mathcal{U}(r) \\ \frac{l}{\mu}\frac{dt}{dr} &= (\mathcal{E} - \mathcal{U}(r) - r^{-2})^{-1/2} \\ \frac{\mathcal{T}}{2} &\equiv \int_{r_{\min}}^{r_{\max}} dt \\ &= \frac{\mu}{l}\int_{r_{\min}}^{r_{\max}} \frac{dr}{\sqrt{\mathcal{E} - \mathcal{U}(r) - r^{-2}}}, \end{aligned} \tag{7.5}$$

where $\mathcal{T}$ is the period of radial oscillation, not necessarily equal to the gross orbital period even if the orbit is closed.

Alternatively, setting

$$\frac{d\theta}{dr} = \frac{d\theta}{dt}\frac{dt}{dr} = \frac{l}{\mu r^2}\frac{dt}{dr},$$

a similar integral gives the orbit shape $\theta(r)$:

$$\begin{aligned} \theta_2 - \theta_1 &\equiv \int_{r_1}^{r_2} d\theta \\ &= \int_{r_1}^{r_2} \frac{dr/r^2}{\sqrt{\mathcal{E} - \mathcal{U}(r) - r^{-2}}}. \end{aligned} \tag{7.6}$$

If $\mathcal{U}(r) = kr^{n+1}$, with $n = +5, +3, 0, -4, -5$, or -7, this is an *elliptic integral* found in tables. More importantly, numerical integration can yield an arbitrarily precise orbit for any well-behaved potential.

Conversely, substituting $u \equiv 1/r$, Eq. (7.6) can be re-expressed as

$$\begin{aligned} -d\theta &= (\mathcal{E} - \mathcal{U}(u) - u^2)^{-1/2} du \\ (\frac{du}{d\theta})^2 &= \mathcal{E} - \mathcal{U}(u) - u^2, \end{aligned} \tag{7.7}$$

yielding the potential $\mathcal{U}(u)$, and thus the force law, given the orbit shape $u(\theta)$.

7.5. Orbit for inverse square law force.

For the planetary case ($U = -k/r$), the orbit is easily expressed in closed form. Equation (7.7) becomes

$$(\frac{du}{d\theta})^2 = \mathcal{E} + \frac{2\mu k u}{l^2} - u^2.$$

Differentiating with respect to θ and dividing through by $du/d\theta \neq 0$ (noncircular motion), we obtain

$$\frac{d^2u}{d\theta^2} + u = \frac{\mu k}{l^2}. \tag{7.8}$$

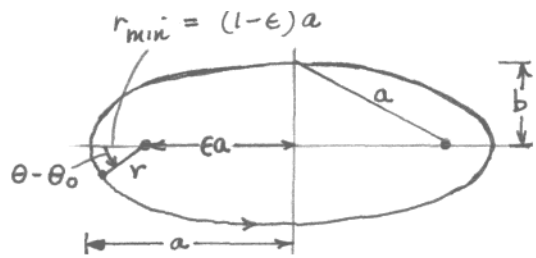
The particular and homogeneous solutions are, respectively,

$$\begin{aligned} u_p &= \frac{\mu k}{l^2} \\ u_h(\theta) &= \frac{\mu k}{l^2}\epsilon\cos(\theta - \theta_0), \end{aligned}$$

where $\dot{u}(\theta_0) \equiv 0$, and $\mu k\epsilon/l^2$ is an adjustable constant. Combining u_p and u_h,

$$u(\theta) \equiv \frac{1}{r(\theta)} = \frac{\mu k}{l^2}(1 + \epsilon\cos(\theta - \theta_0)). \tag{7.9}$$

When $0 < \epsilon < 1$, this is the equation of an *ellipse* with *eccentricity* ϵ. The origin about which θ is measured is one of its two *foci*. These are located within the ellipse, but are well separated from its center if ϵ is significantly greater than zero, i.e. if the ellipse is eccentric.



7.6. Kepler's laws.

We have already proven two of Kepler's laws. The first is that the *areal velocity*,

$$\frac{dA}{dt} = \frac{|\mathbf{r} \times d\mathbf{r}|}{2\,dt} = \frac{l}{2\mu},$$

where dA is an increment of orbit area, is constant. This follows directly from conservation of angular momentum l. The second of Kepler's laws states that the orbits are ellipses with foci at the origin, as shown in Eq. (7.9).

Our remaining objectives are to prove that the total energy is a function only of the major axis of the orbit ellipse, and to establish the third and last of Kepler's laws, which relates the orbit period to the major axis. To proceed, we evaluate Eq. (7.9) at $\theta = \theta_0$ (perigee $r_{\min}$), and at $\theta = \theta_0 + \pi$ (apogee $r_{\max}$):

$$\begin{aligned} r_{\min} &= \frac{l^2}{\mu k(1+\epsilon)}; \quad r_{\max} = \frac{l^2}{\mu k(1-\epsilon)} \\ a &\equiv \frac{r_{\min} + r_{\max}}{2} = \frac{l^2}{\mu k(1-\epsilon^2)}, \end{aligned} \tag{7.10}$$

where a is the *semimajor axis*. In terms of a, we can rewrite

$$r_{\min} = a(1-\epsilon); \quad r_{\max} = a(1+\epsilon). \tag{7.11}$$

At the perigee, where $\dot{r} = 0$, the total energy is

$$\begin{aligned} E &= \frac{l^2}{2\mu r_{\min}^2} - \frac{k}{r_{\min}} \\ &= \frac{l^2}{2\mu a^2(1-\epsilon)^2} - \frac{k}{a(1-\epsilon)} \\ &= \frac{ka(1-\epsilon^2)}{2a^2(1-\epsilon)^2} - \frac{k}{a(1-\epsilon)} \\ &= \frac{k(1+\epsilon)}{2a(1-\epsilon)} - \frac{k}{a(1-\epsilon)} \\ &= \frac{k}{a(1-\epsilon)}\Big(\frac{1+\epsilon}{2} - 1\Big) \\ &= -\frac{k}{2a}. \end{aligned} \tag{7.12}$$

(In the third line we substituted $l^2/\mu = ka(1-\epsilon^2)$ from Eq. (7.10).) Therefore the total energy is a function only of the semimajor axis.

To find the period, it is straightforward to integrate Eq. (7.5). The same result is obtained with less algebra by considering the orbit area A:

$$\mathcal{T} = \frac{A}{dA/dt} = \frac{\pi ab}{l/2\mu}, \tag{7.13}$$

where b is the *semiminor axis*. To evaluate b, consider the right triangle with vertices at the focus, the center of the ellipse, and the tip of the semiminor axis. Its base is $a - r_{\min} = \epsilon a$ and its height is b. The hypotenuse is just a, since the total distance from one focus via the elliptical curve to the other focus is the same for any path. Then

$$\begin{aligned} a^2 &= \epsilon^2 a^2 + b^2 \\ b &= a\sqrt{1-\epsilon^2} \\ \mathcal{T} &= \frac{\pi a^2\sqrt{1-\epsilon^2}}{l/2\mu} \\ &= 2\pi a^2 \mu \sqrt{\frac{1-\epsilon^2}{l^2}} \\ &= \frac{2\pi a^2 \mu}{\sqrt{\mu k a}} \\ &= 2\pi a^{3/2}\sqrt{\frac{\mu}{k}}. \end{aligned} \tag{7.14}$$

(In the fifth equality, we used $l^2/(1-\epsilon^2) = \mu k a$ from Eq. (7.10).)

Equation (7.14), which is Kepler's third law, states that the orbital period is proportional to

the $\frac{3}{2}$ power of the semimajor axis. Neglecting the tiny difference between the reduced mass μ and the actual planetary mass, to which k is proportional, this law predicts a planet's period given its orbit's semimajor axis together with the period and semimajor axis of any other planetary orbit. Its success was one of the experimental cornerstones of Newtonian mechanics.

7.7. Virial theorem.

For a circular orbit with a $1/r^2$ force, the familiar result $T = -E = -U/2$ is easily obtained from Eqs. (7.4) and (7.8) with $\dot{u} = 0$. For an elliptical orbit, both the potential and kinetic energies are functions of θ. To find a simple relation between T, E, and U, we must consider the time average values $\langle T\rangle$ and $\langle U\rangle$ of the kinetic and potential energies. For example, we may integrate

$$\begin{aligned}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} U(t)\,dt &= \frac{1}{2\pi}\int_0^{2\pi} U(\theta)\frac{dt}{d\theta}d\theta \\ &= -\frac{1}{2\pi}\int_0^{2\pi}\frac{k}{r}\frac{\mu r^2}{l}d\theta.\end{aligned}$$

Inserting the elliptical orbit $r(\theta)$, the resulting integral is found in tables, yielding $\langle U\rangle = 2E$ as for the circular orbit.

A more elegant proof of this relation, with extension to a variety of power-law potentials, is provided by the *virial theorem*. Consider a system of N particles indexed by j. Define

$$\begin{aligned}S &\equiv \mathbf{p}_j\cdot\mathbf{r}_j \\ \frac{dS}{dt} &= \dot{\mathbf{p}}_j\cdot\mathbf{r}_j + \mathbf{p}_j\cdot\dot{\mathbf{r}}_j \\ &= \mathbf{F}_j\cdot\mathbf{r}_j + 2T,\end{aligned} \tag{7.15}$$

where $\mathbf{F}_j$ is the force on particle j, and T is the sum of the N kinetic energies.

Define the average value of dS/dt as

$$\Big\langle\frac{dS}{dt}\Big\rangle \equiv \lim_{t\to\infty}\frac{S(t)-S(0)}{t}.$$

This is zero if all the $\mathbf{p}_j$ and $\mathbf{r}_j$ are *bounded*, or if the motion is *periodic* with t chosen to be an integral multiple of the period. Under either condition, Eq. (7.15) requires

$$\langle T\rangle = -\tfrac{1}{2}\langle\mathbf{F}_j\cdot\mathbf{r}_j\rangle.$$

If $\mathbf{F}_j$ is derivable from a potential U_j which is proportional to r_j^{n+1},

$$\begin{aligned}\mathbf{F}_j = -\nabla U_j &= -\frac{(n+1)U_j\hat{r}_j}{|\mathbf{r}_j|} \\ \mathbf{F}_j\cdot\mathbf{r}_j &= -(n+1)U \\ \langle T\rangle &= \frac{n+1}{2}\langle U\rangle,\end{aligned} \tag{7.16}$$

where U is the sum of the N potential energies. With $n = -2$ for the gravitational or Coulomb potential,

$$\begin{aligned}\langle T\rangle &= -\tfrac{1}{2}\langle U\rangle \\ E &= \langle T+U\rangle = \langle T-2T\rangle = -\langle T\rangle,\end{aligned}$$

as asserted above. Equation (7.16) is the *virial theorem*, widely used in classical kinetic theory.

7.8. Stability of circular orbits.

We define the *effective potential* $U_{\rm eff}(r) \equiv U(r)+U'(r)$, where, as in section 7.3, $U'(r)$ is the pseudopotential $l^2/2\mu r^2$. Hamilton's equation in r becomes

$$\begin{aligned}\mathcal{H} &= \frac{p_r^2}{2\mu} + U_{\rm eff}(r) \\ \dot{p}_r &= -\frac{\partial\mathcal{H}}{\partial r} = -\frac{dU_{\rm eff}}{dr}.\end{aligned} \tag{7.17}$$

Therefore a circular orbit, for which $p_r \equiv 0$, is possible only at an extremum of $U_{\rm eff}$. If $U_{\rm eff}(r_0)$ is an extremum, close to that radius we may expand

$$\begin{aligned}\frac{dU_{\rm eff}}{dr} &\approx (r-r_0)\frac{d^2U_{\rm eff}}{dr^2}\Big|_{r=r_0} \\ &\equiv k_{\rm eff}(r-r_0) \\ \mu\ddot{r} = \dot{p}_r &= -k_{\rm eff}(r-r_0).\end{aligned} \tag{7.18}$$

Near the extremum, if the second derivative $k_{\rm eff}$ is *positive* ($U_{\rm eff}$ is a *minimum*), $r-r_0$ experiences simple harmonic motion with angular frequency $\omega_\rho = (k_{\rm eff}/\mu)^{1/2}$. Then the circular orbit is *stable*. Otherwise, $r-r_0$ grows exponentially, and the orbit is *unstable*.

As an alternative to the *method of effective potentials* described above, one may also analyze the stability of circular orbits using the *method of perturbations* as in section 3.7. As an example of both methods, we consider the motion of a particle of mass m in a gravitational field $\mathbf{g} = -g\hat{z}$. The particle is constrained to move on the frictionless surface of an upward-opening cone of half angle α.

In cylindrical coordinates (r, θ, z), the Lagrangian is

$$\mathcal{L} = T - U = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) - mgz.$$

Applying the conical constraint $z = r\cot\alpha$, we may eliminate z:

$$\mathcal{L} = \tfrac{1}{2}m(\dot{r}^2\csc^2\alpha + r^2\dot{\theta}^2) - mgr\cot\alpha.$$

The coordinate θ is cyclic, so that its conjugate momentum $p_\theta = mr^2\dot{\theta} \equiv l$ is conserved. The Euler-Lagrange equation in r is

$$\begin{aligned} \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{r}} &= \frac{\partial\mathcal{L}}{\partial r} \\ m\ddot{r}\csc^2\alpha &= mr\dot{\theta}^2 - mg\cot\alpha \qquad (7.19) \\ \ddot{r} - \frac{l^2}{m^2r^3}\sin^2\alpha &= -g\sin\alpha\cos\alpha. \end{aligned}$$

Since $\partial\mathcal{L}/\partial t = 0$, and T is a quadratic function of $\dot{r}$ and $\dot{\theta}$,

$$\begin{aligned} E = \mathcal{H} &= T + U \\ &= \tfrac{1}{2}m(\dot{r}^2\csc^2\alpha + r^2\dot{\theta}^2) + mgr\cot\alpha \qquad (7.20) \\ &= \tfrac{1}{2}m\dot{r}^2\csc^2\alpha + \frac{l^2}{2mr^2} + mgr\cot\alpha. \end{aligned}$$

The last two terms in the last equation sum to U_{eff}, which is infinite both at $r = 0$ and $r = \infty$. Therefore the intermediate extremum (at $r \equiv r_0$) must be a *minimum*, and a circular orbit at radius r_0 is *stable*.

Proceeding with the method of perturbations, in Eq. (7.19) we substitute $r = r_0 + \lambda\rho$, where λ is a small constant. Retaining only terms of 0th order in λ, Eq. (7.19) becomes

$$-\frac{l^2}{m^2r_0^3}\sin^2\alpha = -g\sin\alpha\cos\alpha. \qquad (7.21)$$

This may be solved for r_0. The remaining terms must also vanish:

$$\begin{aligned} 0 &= \lambda\ddot{\rho} - \frac{l^2\sin^2\alpha}{m^2r_0^3}\big((r_0/r)^3 - 1\big) \\ &= \lambda\ddot{\rho} - \frac{l^2\sin^2\alpha}{m^2r_0^3}\big((1+\lambda\rho/r_0)^{-3} - 1\big) \\ &\approx \lambda\ddot{\rho} + \frac{l^2\sin^2\alpha}{m^2r_0^3}\frac{3\lambda\rho}{r_0} \qquad (7.22) \\ &= \ddot{\rho} + \frac{3l^2\sin^2\alpha}{m^2r_0^4}\rho \\ &\equiv \ddot{\rho} + \omega_\rho^2\rho. \end{aligned}$$

In the third line we expanded $(1+\eta)^{-3} \approx 1 - 3\eta$ for $\eta \ll 1$, and in the last line we identified the angular frequency ω_ρ of radial oscillation.

The angular frequency of the unperturbed orbit is $\Omega = l/mr_0^2$. Therefore

$$\frac{\omega_\rho}{\Omega} = \sqrt{3}\sin\alpha.$$

The orbit is *closed* only for particular cone half angles such that ω_ρ/Ω is equal to a rational number m/n, where m and n are positive integers; n revolutions are required to close the orbit.

In this problem, the frequency of radial oscillation could have been obtained equally well using the method of effective potentials. But the method of perturbations is more general: a similar technique may be used to find the oscillation frequencies of deviations from unperturbed orbits which are *not* circular.

8. Collisions.

8.1. Elastic collisions.

In the absence of external forces, *all* collisions conserve total momentum (see section 2.4). However, not all collisions are *elastic*. Elastic collisions conserve total kinetic energy as well as momentum.

Within the plane of scattering, given the initial momenta, a two-body elastic collision may

be characterized by a single quantity, for example the center of mass (C.M.) scattering angle. Since the transformation between nonrelativistic C.M. and laboratory angles is straightforward, specifying any single final state quantity in the laboratory (e.g. one final angle or energy) determines the whole problem. This can lead to a large set of assigned problems for which the algebra may be tedious but the conceptual basis remains elementary.

It is amusing that the elastic scattering of a nonrelativistic projectile from a stationary target of equal mass always results in an opening angle of $\pi/2$ between the two final state velocities. Label the projectile a, the stationary target b, and the final state particles c and d. According to the conservation laws,

$$\begin{aligned} T_a + (T_b = 0) &= T_c + T_d \\ \Rightarrow \frac{1}{2m}(p_a^2 &= p_c^2 + p_d^2) \\ \mathbf{p}_a + (\mathbf{p}_b = 0) &= \mathbf{p}_c + \mathbf{p}_d \\ \Rightarrow p_a^2 &= p_c^2 + p_d^2 + 2\mathbf{p}_c \cdot \mathbf{p}_d. \end{aligned}$$

For the second and fourth equalities to be consistent, $\mathbf{p}_c$ and $\mathbf{p}_d$ must be mutually perpendicular.

8.2. Rutherford scattering.

By themselves, momentum and energy conservation determine the result of an elastic scattering only if one final state quantity also can be supplied. However, if the force law is known, the final state can be predicted from the initial conditions alone. The classic example is *Rutherford scattering* of two charged particles under the influence of the Coulomb force. In the C.M., consider a nonrelativistic projectile of charge ze impinging with initial relative velocity v_0 upon a target of charge Ze. The electrostatic force between them is centrally directed and of size Zze^2/r^2 in Gaussian units.

This is not enough information to predict the scattered final state, as we have not yet specified whether the collision is "grazing" or "head-on". The missing quantity is the *impact parameter* $b \equiv l/\mu v_0$. It is named for the fact that the undeflected path of the projectile misses the target by the distance b. In terms of these quantities, the constants introduced in section 7.5 become

$$k = -Zze^2 \quad E = \tfrac{1}{2}\mu v_0^2 \quad l = \mu b v_0.$$

Again defining $\theta \equiv \theta_0$ at the perigee $r_{\min}$, Eq. (7.9) requires $\epsilon < -1$ when k is negative, as it is here. Solving the first line of Eq. (7.12) for $r_{\min}$,

$$\begin{aligned} 0 &= \frac{l^2}{2\mu} r_{\min}^{-2} - k r_{\min}^{-1} - E \\ r_{\min}^{-1} &= \frac{k \pm \sqrt{k^2 + 2El^2/\mu}}{l^2/\mu} \\ &= \frac{-k\mu}{l^2}(\sqrt{1+\eta^2} - 1), \end{aligned} \tag{8.1}$$

where

$$\eta \equiv \sqrt{\frac{2El^2}{\mu k^2}} = \frac{\mu b v_0^2}{Zze^2}. \tag{8.2}$$

From Eq. (7.10),

$$r_{\min}^{-1} = \frac{-k\mu}{l^2}(-\epsilon - 1). \tag{8.3}$$

Comparing Eqs. (8.1) and (8.3),

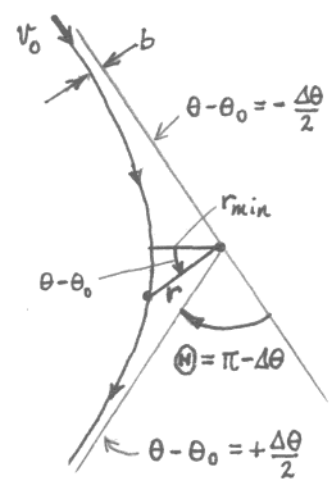
$$-\epsilon = \sqrt{1+\eta^2}. \tag{8.4}$$

To keep r in Eq. (7.9) positive while $k < 0$, we require

$$\begin{aligned} -\epsilon^{-1} &\le \cos(\theta - \theta_0) \\ \cos^{-1}(1+\eta^2)^{-1/2} &\le |\theta - \theta_0| \\ \tan^{-1}\eta &\le |\theta - \theta_0|. \end{aligned} \tag{8.5}$$

Thus $\theta - \theta_0$ varies only within a range

$$\Delta\theta = 2\tan^{-1}\eta. \tag{8.6}$$



The scattering angle Θ is defined as $\Theta \equiv \pi - \Delta\theta$, so that $\Theta = 0$ for a grazing collision and $\Theta = \pi$ for a head-on collision. Then, combining Eqs. (8.6) and (8.2),

$$\begin{aligned} \Theta/2 &= \pi/2 - \Delta\theta/2 \\ &= \pi/2 - \tan^{-1}\eta \\ &= \cot^{-1}\eta \\ &= \tan^{-1}\frac{Zze^2}{\mu b v_0^2} \\ b &= \frac{Zze^2}{\mu v_0^2}\cot\Theta/2. \end{aligned} \tag{8.7}$$

This last equation connects the impact parameter to the scattering angle.

8.3. Scattering cross section.

In Rutherford's experiment, the projectiles were ^{4}He nuclei and the targets were ^{197}Au nuclei in a thin foil. The detectors were dark-adapted students barely observing flashes in scintillating plates. Like more modern instruments, the students could measure the scattering angle Θ, but not the tiny impact parameter b. To interpret his data, Rutherford needed to predict the *distribution* of scattering angles.

Let Γ be the flux of incident projectiles (particles/cm^2-sec), and let dN be the number of incident projectiles/sec with impact parameter between b and $b + db$ and with azimuth between ϕ and $\phi + d\phi$. Using Eq. (8.7),

$$\begin{aligned} dN &= \Gamma b\, db\, d\phi \\ \frac{dN}{\Gamma} &= \Big(\frac{Zze^2}{\mu v_0^2}\Big)^2 \cot\tfrac{\Theta}{2}\, d\cot\tfrac{\Theta}{2}\, d\phi \\ &= \Big(\frac{Zze^2}{\mu v_0^2}\Big)^2 \cot\tfrac{\Theta}{2}\csc^2\tfrac{\Theta}{2}\, d\tfrac{\Theta}{2}\, d\phi \\ &= \Big(\frac{Zze^2}{\mu v_0^2}\Big)^2 \cos\tfrac{\Theta}{2}\sin\tfrac{\Theta}{2}\csc^4\tfrac{\Theta}{2}\, d\tfrac{\Theta}{2}\, d\phi \\ &= \Big(\frac{Zze^2}{2\mu v_0^2}\Big)^2 \csc^4\tfrac{\Theta}{2}\, d\Omega, \end{aligned}$$

where $d\Omega \equiv \sin\Theta\, d\Theta\, d\phi = 4\sin\frac{\Theta}{2}\cos\frac{\Theta}{2}\, d\frac{\Theta}{2}\, d\phi$ is an infinitesimal element of *solid angle*. The quantity $dN/\Gamma d\Omega$ has the dimensions of an *area* and is called the *differential scattering cross section* $d\sigma/d\Omega$:

$$\frac{d\sigma}{d\Omega} = \Big(\frac{Zze^2}{2\mu v_0^2}\Big)^2 \frac{1}{\sin^4\Theta/2}. \tag{8.8}$$

This is the famous differential cross section for Rutherford scattering. Physically, $d\sigma$ is the cross sectional area of the beam which is scattered by the target through angle Θ into an element $d\Omega$ of solid angle. The probability for a single ^{4}He nucleus to scatter into $d\Omega$ is $\Sigma(d\sigma/d\Omega)d\Omega$, where Σ is the number of ^{197}Au nuclei per *square* cm of target foil.

Rutherford scattering has special properties that deserve comment. First, like the gravitational force, the $1/r^2$ Coulomb force is said to have *infinite range* because the *total* scattering cross section σ_T is infinite:

$$\sigma_T \equiv 2\pi\int_0^\pi \frac{d\sigma}{d\Omega}\sin\Theta d\Theta \propto \int_0^\pi \frac{\sin\Theta d\Theta}{\sin^4\Theta/2} = \infty.$$

Second, in the quantum mechanical sense, the individual charges on the projectile and target nuclei act *coherently* to produce the scattering. This follows from the fact that the differential cross section is proportional to the *square* of z and of Z. (In fact, the quantum mechanical calculation for Rutherford scattering yields the identical result.) The cross section is independent of the sign of either charge.

The final comment is that $d\sigma/d\Omega$ is *finite* at $\Theta = \pi$: per unit solid angle, there is a finite probability for *backscattering*. Using this striking prediction, Rutherford was able experimentally to distinguish between a nucleus modeled to be small in extent, relative to $r_{\min}$, and another more extended model for which backscattering could not occur. At ever higher projectile energies, scattering of simple projectiles by ever smaller targets has provided fundamental insights. In the 1950's, elastic scattering of $\approx 10^8$ eV electrons determined the size and shape of nuclei in exquisite detail. In 1968, quarks were discovered in the inelastic scattering of $\approx 10^{10}$ eV electrons from protons.

9. Rotational motion.

9.1. Pseudoforces in rotating systems.

The present discussion builds on the results of section 1. Recall that $\tilde{x}$ is a column vector representing a vector in the body (rotating) system, while $\tilde{x}'$ is the same vector in an inertial (fixed) system. At $t = 0$ the two coordinate systems are coincident, so that $\tilde{x}(0) = \tilde{x}'(0)$. At other times, $\tilde{x}$ and $\tilde{x}'$ are related by a rotation, which is represented by an orthogonal 3×3 matrix Λ:

$$\tilde{x}' = \Lambda \tilde{x}. \tag{9.1}$$

If the body system is rotating relative to the inertial system with angular velocity $\vec{\omega}$, and $\mathbf{P}$ ($\mathbf{P}'$) is the same vector observed in the rotating (fixed) system, recall that $d\mathbf{P}'/dt$ acquires an extra term due to the rotation:

$$\frac{d\mathbf{P}'}{dt} = \frac{d\mathbf{P}}{dt} + \vec{\omega} \times \mathbf{P}. \tag{9.2}$$

Choosing $\mathbf{P}^{(\prime)} = \mathbf{r}^{(\prime)}$, Eq. (9.2) becomes

$$\mathbf{v}' = \mathbf{v} + \vec{\omega} \times \mathbf{r}. \tag{9.3}$$

Alternatively, choosing $\mathbf{P}'$ to be the left hand side of Eq. (9.3), and $\mathbf{P}$ to be the right hand side, Eq. (9.2) becomes

$$\begin{aligned} \frac{d\mathbf{v}'}{dt} &= \big(\frac{d}{dt} + \vec{\omega}\times\big)\big(\mathbf{v} + \vec{\omega} \times \mathbf{r}\big) \\ \frac{\mathbf{F}'}{m} &= \frac{d\mathbf{v}}{dt} + 2\vec{\omega} \times \mathbf{v} + \vec{\omega} \times (\vec{\omega} \times \mathbf{r}) \\ m\frac{d\mathbf{v}}{dt} &= \mathbf{F}' - 2m\vec{\omega} \times \mathbf{v} - m\vec{\omega} \times (\vec{\omega} \times \mathbf{r}) \\ &\equiv \mathbf{F}' + \mathbf{F}_{\text{Coriolis}} + \mathbf{F}_{\text{centrifugal}}. \end{aligned} \tag{9.4}$$

When observed in the body frame, the mass accelerates as though under the influence of *pseudoforces* $\mathbf{F}_{\text{Coriolis}}$ and $\mathbf{F}_{\text{centrifugal}}$, in addition to the *actual* force $\mathbf{F}'$ acting in the inertial frame.

On the surface of the earth, which spins with angular velocity $\vec{\omega}_e$ directed out of the north pole, the centrifugal force points outward from the axis with magnitude $m\omega_e^2 r_e \sin\lambda = 0.003455 mg \sin\lambda$, where r_e is the earth's radius, λ is the colatitude (measured from the north pole) and g is the gravitational acceleration at the surface. The magnitude of the effective (apparent) acceleration $\mathbf{g}_{\text{eff}} \equiv \mathbf{g} + \mathbf{F}_{\text{centrifugal}}/m$ is slightly reduced, especially for colatitudes near $\pi/2$. This causes the earth to bulge near the equator. The direction of $\mathbf{g}_{\text{eff}}$ also shifts slightly. For example, in the northern hemisphere, $\mathbf{g}_{\text{eff}}$ intersects the earth's axis slightly south of its center.

The Coriolis force is proportional to velocity. In the northern hemisphere ($0 < \lambda < \pi/2$), falling bodies experience a force $2m\omega_e v \sin\lambda$ to the *east*, and bodies moving horizontally feel a Coriolis force whose *horizontal* component is $2m\omega_e v \cos\lambda$ to the *right*. In the southern hemisphere, $\cos\lambda$ reverses sign and the horizontal Coriolis force is to the *left*. The Coriolis force accounts *e.g.* for the counterclockwise circulation of storms in the northern hemisphere: as an element of air is sucked into the low-pressure eye, it veers right. As it happened, gunnery tables available to the U.S. Navy at the time of early south Pacific battles in WW II were written only for the northern hemisphere. The seriousness of this blunder was diminished by the emerging dominance of air power.

9.2. Foucault pendulum.

If a freely pivoting pendulum is placed at the north pole, as viewed in an inertial frame it oscillates in a fixed plane while the earth rotates underneath it. Correspondingly, on the earth, its plane of oscillation is observed to precess with angular velocity $-\vec{\omega}_e$. If the pendulum is placed at the equator, its plane is not observed to precess with respect to the earth's surface, because the equator is a symmetry point at which a preferred direction for this precession cannot be identified. More generally, if this *Foucault pendulum* is placed at colatitude λ, its plane of oscillation precesses clockwise ("CW") relative to the earth's surface with angular velocity $\omega_e \cos\lambda$. At least, this guess is consistent with the two limiting cases just considered.

To analyze this problem further, consider a local (unprimed) coordinate system with its ori-

gin on the earth's surface at colatitude λ. We take the $\hat{z}$ direction to be normal to the surface (ignoring centrifugal forces). In this frame, ω_e lies in the z-north plane at angle λ to $\hat{z}$. Including the Coriolis force, the pendulum bob satisfies the equation

$$\begin{aligned}\ddot{\mathbf{r}} &= -g\hat{z} + \frac{\mathbf{T}}{m} - 2\vec{\omega}_e \times \dot{\mathbf{r}} \\ &\approx -g\hat{z} + \frac{\mathbf{T}}{m} - 2\omega_e \cos\lambda\, \hat{z} \times \dot{\mathbf{r}},\end{aligned}$$

where $\mathbf{T}$ is the string tension. In the second equality we used the fact that $\dot{\mathbf{r}}$ lies mainly in the horizontal plane, assuming small-angle oscillations.

The final step is to analyze the same problem in a (starred) system having the same origin, which is precessing CW *with respect to the earth's surface* with angular velocity $\vec{\Omega} = -\hat{z}\omega_e \cos\lambda$. The equation of motion of the pendulum bob becomes

$$\ddot{\mathbf{r}}^* = -g\hat{z} + \frac{\mathbf{T}}{m} - 2\omega_e \cos\lambda\, \hat{z} \times \dot{\mathbf{r}}^* - 2\vec{\Omega} \times \dot{\mathbf{r}}^*.$$

The last two terms cancel, and the resulting equation is that of a simple pendulum. In the starred system, the plane of oscillation does not precess. Therefore, in the *unprimed* system, the plane of oscillation precesses CW with the same angular velocity $\omega_e \cos\lambda$ as does the starred system. This justifies the original assertion.

Large Foucault pendula with slowly precessing planes of oscillation are staples of many science museums. By observing the angle through which the pendulum has precessed, you can monitor the duration of your visit, or, with the help of a watch, measure the museum's latitude.

9.3. Angular velocity from Euler rotations.

Recall from section 1 that the Euler rotation Λ^t is a transformation from the space (inertial) frame to the body frame, by means of three successive transformations:

$$\begin{aligned}\tilde{x} &= \Lambda^t_\psi \tilde{x}''' \\ &= \Lambda^t_\psi \Lambda^t_\theta \tilde{x}'' \\ &= \Lambda^t_\psi \Lambda^t_\theta \Lambda^t_\phi \tilde{x}' \\ &\equiv \Lambda^t \tilde{x}',\end{aligned} \tag{9.5}$$

where Λ^t_ψ, Λ^t_θ, and Λ^t_ϕ are defined in Eq. (1.9). In order to solve problems in the body system, we need to compute the Cartesian components there of the angular velocity $\vec{\omega}$ that results when the Euler angles ϕ, θ, and ψ vary with time.

The angular velocities in question are directed along the $3'$, $1''$, and $3'''$ axes, respectively. Transforming all three to the body system,

$$\begin{aligned}\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix} &= \Lambda^t_\psi \Lambda^t_\theta \Lambda^t_\phi \begin{pmatrix}0\\0\\ \dot{\phi}\end{pmatrix} + \\ &+ \Lambda^t_\psi \Lambda^t_\theta \begin{pmatrix}\dot{\theta}\\0\\0\end{pmatrix} + \Lambda^t_\psi \begin{pmatrix}0\\0\\ \dot{\psi}\end{pmatrix} \\ &= \begin{pmatrix}\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\ \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \\ \dot{\phi}\cos\theta + \dot{\psi}\end{pmatrix}.\end{aligned} \tag{9.6}$$

We shall use this relation in the study of tops.

9.4. Angular momentum of a rigid body.

About the origin, N point masses at *space* coordinates $\mathbf{r}'_i$ $(1 \le i \le N)$ have combined angular momentum

$$\mathbf{L} = \sum_{i=1}^{N} m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i). \tag{9.7}$$

We would like to express $\mathbf{L}$ in the *body* system of coordinates. In that system, if the point masses make up a *rigid body*, we can take advantage of whatever symmetries the body might possess.

If the motion of the body system is limited to a simple *rotation* characterized by angular velocity $\vec{\omega}$, we choose the origins of the space and body axes to be the same. In addition, we choose to consider $\mathbf{L}$ at an instant of time when the two systems are entirely coincident. This means that $\mathbf{r}_i = \mathbf{r}'_i$ for all i at that time, but of course $\dot{\mathbf{r}}_i \neq \dot{\mathbf{r}}'_i$. Assuming that the body is rigid, $\mathbf{r}_i$ is a constant in the body system. Equation (9.3) becomes

$$\dot{\mathbf{r}}'_i = \vec{\omega} \times \mathbf{r}_i. \tag{9.8}$$

The angular momentum may be written

$$\begin{aligned}\mathbf{L} &= \sum_i m_i\big(\mathbf{r}_i \times (\vec{\omega} \times \mathbf{r}_i)\big) \\ &= \sum_i m_i\big(\vec{\omega}(\mathbf{r}_i \cdot \mathbf{r}_i) - \mathbf{r}_i(\mathbf{r}_i \cdot \vec{\omega})\big).\end{aligned} \tag{9.9}$$

Changing to matrix notation in Cartesian coordinates, with $\tilde{L}$, $\tilde{\mathbf{r}}$, and $\tilde{\omega}$ representing column vectors, and with I as the unit matrix,

$$\begin{aligned}\tilde{L} &= \sum_i m_i\big(r_i^2\tilde{\omega} - \tilde{r}_i(\tilde{r}_i^t\tilde{\omega})\big) \\ &= \sum_i m_i\big(r_i^2\mathrm{I}\tilde{\omega} - (\tilde{r}_i\tilde{r}_i^t)\tilde{\omega}\big) \\ &= \big(\sum_i m_i(r_i^2\mathrm{I} - \tilde{r}_i\tilde{r}_i^t)\big)\tilde{\omega} \\ &\equiv \mathcal{I}\tilde{\omega} \\ \mathcal{I} &\equiv \sum_i m_i\big(r_i^2\mathrm{I} - \tilde{r}_i\tilde{r}_i^t\big).\end{aligned} \tag{9.10}$$

In the above we have introduced the *inertia tensor* $\mathcal{I}$, which relates the angular momentum to the angular velocity of a rigid body.

Because the matrix notation in Eq. (9.10) can be cryptic, it is useful to display $\mathcal{I}$ in component form:

$$\begin{aligned}\tilde{r}_i\tilde{r}_i^t &= \begin{pmatrix} x_1^ix_1^i & x_1^ix_2^i & x_1^ix_3^i \\ x_2^ix_1^i & x_2^ix_2^i & x_2^ix_3^i \\ x_3^ix_1^i & x_3^ix_2^i & x_3^ix_3^i \end{pmatrix} \\ \mathcal{I}_{jk} &\equiv \sum_i m_i\big(\delta_{jk}r_i^2 - x_j^ix_k^i\big),\end{aligned} \tag{9.11}$$

where e.g. x_j^i is the jth component ($1 \le j \le 3$) of the coordinate $\mathbf{r}_i$ of the ith mass. For example,

$$\begin{aligned}\mathcal{I}_{11} &= \sum_i m_i\big(r_i^2 - (x_1^i)^2\big) \\ &= \sum_i m_i\big((x_2^i)^2 + (x_3^i)^2\big) \\ \mathcal{I}_{12} &= -\sum_i m_ix_1^ix_2^i.\end{aligned}$$

What makes $\mathcal{I}$ a *tensor* is its transformation property. If $\tilde{x}$ is a vector transforming from the body to the space system according to $\tilde{x}' = \Lambda\tilde{x}$, the inertia tensor transforms according to the *similarity transformation:*

$$\mathcal{I}' = \Lambda\mathcal{I}\Lambda^t. \tag{9.12}$$

9.5. Elementary properties of the inertia tensor.

Obviously, since $\mathcal{I}$ rarely is proportional to the unit matrix, for arbitrary $\hat{\omega}$ the angular momentum $\mathbf{L} = \mathcal{I}\vec{\omega}$ rarely is directed along $\hat{\omega}$. It is true that $\mathbf{L} \parallel \vec{\omega}$ in many elementary problems; this occurs when $\vec{\omega}$ is directed along one of the *principal axes* of the body, defined later on.

A simple example with $\mathbf{L}$ not parallel to $\vec{\omega}$ is a barbell rotating about a vertical axis through its center. The bar is inclined at an angle ψ to the horizontal, so that one weight orbits the axis in an higher plane than does the other. For each weight, $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ must be $\perp$ to the bar, so $\mathbf{L}$ must make the same angle ψ with the axis of rotation.

When $\mathbf{L} = I\vec{\omega}$, where I is the scalar moment of inertia, the kinetic energy of rotation is $T_{\text{rot}} = \frac{1}{2}I\omega^2$. More generally,

$$\begin{aligned}T_{\text{rot}} &= \tfrac{1}{2}\sum_i m_i(v_i')^2 \\ &= \tfrac{1}{2}\sum_i m_i\mathbf{v}_i' \cdot \frac{d\mathbf{r}_i'}{dt} \\ &= \tfrac{1}{2}\sum_i m_i\mathbf{v}_i' \cdot (\vec{\omega} \times \mathbf{r}_i') \\ &= \tfrac{1}{2}\sum_i m_i\vec{\omega} \cdot (\mathbf{r}_i' \times \mathbf{v}_i') \\ &= \tfrac{1}{2}\vec{\omega} \cdot \big(\sum_i m_i\mathbf{r}_i' \times \mathbf{v}_i'\big) \\ &= \tfrac{1}{2}\vec{\omega} \cdot \mathbf{L} \\ &= \tfrac{1}{2}\vec{\omega} \cdot (\mathcal{I}\vec{\omega}) \\ &= \tfrac{1}{2}\tilde{\omega}^t\mathcal{I}\tilde{\omega}.\end{aligned} \tag{9.12}$$

Defining $\tilde{n}$ to be a unit column vector along $\vec{\omega}$, Eq. (9.12) becomes

$$\begin{aligned}T_{\text{rot}} &= \tfrac{1}{2}\omega^2\tilde{n}^t\mathcal{I}\tilde{n} \\ &\equiv \tfrac{1}{2}I\omega^2 \\ I &\equiv \tilde{n}^t\mathcal{I}\tilde{n}.\end{aligned} \tag{9.13}$$

As a bonus, Eq. (9.13) relates the inertia tensor $\mathcal{I}$ to the scalar moment of inertia I about a particular axis of rotation. However, while useful for calculating $T_{\rm rot}$, this scalar I obviously *cannot* be employed to relate $\mathbf{L}$ to $\vec{\omega}$ in the general case.

So far we have been considering a rigid body to be a collection of discrete masses, for which the inertia tensor is given by Eq. (9.11), repeated below. The same equation for a *continuous* mass distribution is obtained by replacing the sum over m_i by an integral over the mass density ρ:

$$\begin{aligned}\mathcal{I}_{jk} &= \sum_i m_i\big(\delta_{jk}\sum_{l=1}^{3}(x_l^i)^2 - x_j^i x_k^i\big)\\ &= \int dx_1\int dx_2\int dx_3\Big(\rho(x_1,x_2,x_3)\times \\ &\times\big(\delta_{jk}\sum_{l=1}^{3}(x_l)^2 - x_j x_k\big)\Big).\end{aligned} \tag{9.14}$$

The integral is taken over the entire volume of the rigid body. It is usually easy to evaluate unless the boundaries of the body are not readily expressible in terms of the Cartesian coordinates x_1, x_2, and x_3. In that event, it may be necessary to transform the integral to cylindrical or spherical coordinates.

9.6. Diagonalization of the inertia tensor.

In this section we prove, for *any* rigid body, that there exist three orthogonal unit vectors $\tilde{s}_r$, $1 \le r \le 3$, such that

$$\mathcal{I}\tilde{s}_r = e_r\tilde{s}_r, \tag{9.15}$$

where e_r is a real constant called the $r^{\rm th}$ *eigenvalue*. The $\tilde{s}_r$ are the *eigenvectors* – the directions of the body's *principal axes*. When the rigid body rotates about any of its principal axes, $\mathbf{L}$ and $\vec{\omega}$ are *parallel*:

$$\tilde{L} = \mathcal{I}\tilde{\omega} = \mathcal{I}(\omega\tilde{s}_r) = e_r(\omega\tilde{s}_r) = e_r\tilde{\omega}.$$

(Summation over r is *never* implied!) This simple relationship between $\mathbf{L}$ and $\vec{\omega}$ is so advantageous that we almost always use the principal axis frame when we work in the body system.

If the eigenvalue equation (9.15) is satisfied,

$$\begin{aligned}0 &= \mathcal{I}\tilde{s} - e\tilde{s} = (\mathcal{I} - e\mathrm{I})\tilde{s}\\ &= (\mathcal{I} - e\mathrm{I})\tilde{s}s\\ &= \det\big((\mathcal{I} - e\mathrm{I})\tilde{s}s\big)\\ &= \det(\mathcal{I} - e\mathrm{I})\det(\tilde{s}s)\\ 0 &= \det(\mathcal{I} - e\mathrm{I}).\end{aligned} \tag{9.16}$$

This *"secular equation"* is a cubic equation with three roots e_1, e_2, and e_3. Once it is solved to yield these eigenvalues e_r, the eigenvectors $\tilde{s}_r$ are obtained from the equations

$$\begin{aligned}(\mathcal{I} - e_r\mathrm{I})\tilde{s}_r &= 0\\ \tilde{s}_r^t\tilde{s}_r &= 1.\end{aligned} \tag{9.17}$$

The latter equality ensures that the eigenvectors are of unit length.

First we prove that the eigenvalues of $\mathcal{I}$ are *real*. Here Eq. (9.11) reminds us that $\mathcal{I}$ is *real* and *symmetric*, and therefore *self-adjoint*: $\mathcal{I}^\dagger = \mathcal{I}$. For brevity dropping the index r,

$$\begin{aligned}\mathcal{I}\tilde{s} &= e\tilde{s}\\ \tilde{s}^\dagger\mathcal{I}\tilde{s} &= e\tilde{s}^\dagger\tilde{s}\\ (\tilde{s}^\dagger\mathcal{I}\tilde{s})^\dagger &= (e\tilde{s}^\dagger\tilde{s})^\dagger\\ \tilde{s}^\dagger(\tilde{s}^\dagger\mathcal{I})^\dagger &= e^*\tilde{s}^\dagger\tilde{s}\\ \tilde{s}^\dagger\mathcal{I}^\dagger\tilde{s} &= e^*\tilde{s}^\dagger\tilde{s}\\ \tilde{s}^\dagger\mathcal{I}\tilde{s} &= e^*\tilde{s}^\dagger\tilde{s}\\ e\tilde{s}^\dagger\tilde{s} &= e^*\tilde{s}^\dagger\tilde{s}\\ e &= e^*.\end{aligned}$$

Obviously the eigenvectors $\tilde{s}_r$ also can be defined to be real, since everything else in the linear Eq. (9.15) is real.

Next we prove that two eigenvectors $\tilde{s}_1$ and $\tilde{s}_2$ corresponding to two different eigenvalues e_1

and e_2 are *orthogonal*:

$$\begin{aligned} e_1 &\neq e_2 \\ \mathcal{I}\tilde{s}_1 &= e_1\tilde{s}_1 \\ \mathcal{I}\tilde{s}_2 &= e_2\tilde{s}_2 \\ \tilde{s}_2^t\mathcal{I}\tilde{s}_1 &= e_1\tilde{s}_2^t\tilde{s}_1 \\ \tilde{s}_1^t\mathcal{I}\tilde{s}_2 &= e_2\tilde{s}_1^t\tilde{s}_2 \\ (\tilde{s}_1^t\mathcal{I}\tilde{s}_2)^t &= (e_2\tilde{s}_1^t\tilde{s}_2)^t \\ \tilde{s}_2^t\mathcal{I}^t\tilde{s}_1 &= e_2\tilde{s}_2^t\tilde{s}_1 \\ \tilde{s}_2^t\mathcal{I}\tilde{s}_1 &= e_2\tilde{s}_2^t\tilde{s}_1 \\ 0 &= (e_2 - e_1)\tilde{s}_2^t\tilde{s}_1 \\ 0 &= \tilde{s}_2^t\tilde{s}_1. \end{aligned}$$

If the same eigenvalue is shared by two or more different eigenvectors, it is easy to form linear combinations of them to construct three mutually orthogonal eigenvectors. Thus *the principal axes of a rigid body are always orthogonal.*

9.7. Principal axis transformation.

For simplicity, arrange the eigenvector signs so that $\mathbf{s}_1 \times \mathbf{s}_2 = \mathbf{s}_3$. The reference frame defined by these three unit vectors is called the *principal axis frame*. Denote it by two primes ($''$). A vector $\tilde{x}''$ in the principal axis frame is related to the same vector $\tilde{x}$ in the (unprimed) body frame, in which $\mathcal{I}$ was originally calculated, by a *principal axis transformation*:

$$\tilde{x}'' = \Lambda\tilde{x}\ .$$

Since $\mathcal{I}$ is a tensor, $\mathcal{I}''$ is related to $\mathcal{I}$ by a similarity transformation:

$$\begin{aligned} \mathcal{I}'' &= \Lambda\mathcal{I}\Lambda^t \\ \Lambda^t\mathcal{I}'' &= \mathcal{I}\Lambda^t \end{aligned}$$

$$\begin{pmatrix} \Lambda_{11}^t & \Lambda_{12}^t & \Lambda_{13}^t \\ \Lambda_{21}^t & \Lambda_{22}^t & \Lambda_{23}^t \\ \Lambda_{31}^t & \Lambda_{32}^t & \Lambda_{33}^t \end{pmatrix} \begin{pmatrix} e_1 & 0 & 0 \\ 0 & e_2 & 0 \\ 0 & 0 & e_3 \end{pmatrix} = \\ = \mathcal{I} \begin{pmatrix} \Lambda_{11}^t & \Lambda_{12}^t & \Lambda_{13}^t \\ \Lambda_{21}^t & \Lambda_{22}^t & \Lambda_{23}^t \\ \Lambda_{31}^t & \Lambda_{32}^t & \Lambda_{33}^t \end{pmatrix}. \tag{9.18}$$

The first column of Eq. (9.19) is just the first eigenvalue equation:

$$e_1 \begin{pmatrix} \Lambda_{11}^t \\ \Lambda_{21}^t \\ \Lambda_{31}^t \end{pmatrix} = \mathcal{I} \begin{pmatrix} \Lambda_{11}^t \\ \Lambda_{21}^t \\ \Lambda_{31}^t \end{pmatrix}. \tag{9.19}$$

Comparing Eq. (9.19) to Eq. (9.15), we see that the column vector on either side of Eq. (9.19) is proportional to the first eigenvector $\tilde{s}_1$. Since the transformation is orthogonal ($\Lambda^t\Lambda = \mathrm{I}$), this constant of proportionality is unity. Therefore the transpose Λ^t of the principal axis transformation matrix is the *matrix which has columns equal to the eigenvectors*. That is, Λ_{ir}^t is the ith component of the rth eigenvector $\tilde{s}_r$.

To recapitulate, the inertia tensor is diagonalized by an orthogonal transformation $\mathcal{I}'' = \Lambda\mathcal{I}\Lambda^t$, where Λ^t is the matrix whose columns are the eigenvectors. The diagonals of the transformed inertia tensor are the eigenvalues.

One may solve inertia tensor problems using various strategies. In one method, calculate $\mathcal{I}$ in a convenient coordinate system, solve the secular equation (9.16) for the eigenvalues, then solve (9.17) for the eigenvectors. This yields the matrix Λ^t, which facilitates a transformation to the principal axis system.

Alternatively, the principal axes of the rigid body may be obvious from symmetry considerations. For example, Eq. (9.11) guarantees that the normal to a plane of symmetry is a principal axis. So is an axis of cylindrical symmetry. Often you can guess the principal axes and calculate the inertia tensor in the principal axis system directly.

9.8. Parallel axis theorem.

Usually it is desirable to know the inertia tensor $\mathcal{I}_{\text{CM}}$ in a system of (body) coordinates with its origin at the C.M. In this system, there is no translational motion of the C.M.; the kinetic energy arises simply from rotation, and is equal to $T = \tilde{\omega}^t\mathcal{I}\tilde{\omega}$, as in Eq. (9.12).

On the other hand, it is often far more convenient to calculate $\mathcal{I}$ in a new (body) system in which the axes have the same orientation, but the origin is displaced from the C.M. Denote by $\mathbf{R}$ the vector *from* the new origin *to* the C.M.

Using $\mathbf{r}_i$ as the coordinate of a mass point in the new system, and $\mathbf{r}_i^*$ as the C.M. coordinate of the same point,

$$\begin{aligned}\mathcal{I} &= \sum_i m_i(r_i^2\mathrm{I} - \tilde{r}_i\tilde{r}_i^t)\\ &= \sum_i m_i\big((\mathbf{r}_i^*+\mathbf{R})^2\mathrm{I} - (\tilde{r}_i^*+\tilde{R})(\tilde{r}_i^{*t}+\tilde{R}^t)\big)\\ &= \sum_i m_i\big((r_i^{*2}+R^2)\mathrm{I} - (\tilde{r}_i^*\tilde{r}_i^{*t}+\tilde{R}\tilde{R}^t)\big)\\ &= \mathcal{I}^{CM} + M(R^2\mathrm{I} - \tilde{R}\tilde{R}^t)\ ,\ \text{or}\\ \mathcal{I}_{jk} &= \mathcal{I}_{jk}^{CM} + M(R^2\delta_{jk} - R_jR_k)\ ,\end{aligned} \tag{9.20}$$

where M is the total mass. The cross terms in the second equality vanished in the third because $\sum_i m_i\mathbf{r}_i^* \equiv 0$. The last two equations say that the inertia tensor at $\mathbf{R}$ is the inertia tensor at the center of mass plus the inertia tensor of a point mass M located at $\mathbf{R}$.

Equation (9.20) is the *parallel axis theorem*. With it one may easily calculate $\mathcal{I}_{CM}$ given $\mathcal{I}$, or *vice versa*.

According to Eq. (9.13), the scalar moment of inertia I about an axis $\hat{n}$ is

$$\begin{aligned}I &= \tilde{n}^t\mathcal{I}\tilde{n}\\ &= \tilde{n}^t\mathcal{I}_{CM}\tilde{n} + \tilde{n}^tM(R^2\mathrm{I} - \tilde{R}\tilde{R}^t)\tilde{n}\\ &= I_{CM} + M(R^2 - \tilde{n}^t\tilde{R}\tilde{R}^t\tilde{n})\\ &= I_{CM} + M(R^2 - \hat{n}\cdot\mathbf{R}\ \mathbf{R}\cdot\hat{n})\\ &= I_{CM} + M(R^2 - R_n^2)\\ &= I_{CM} + M(\mathbf{R}\times\hat{n})^2.\end{aligned} \tag{9.21}$$

Here we have introduced I_{CM}, the scalar moment of inertia for rotation about a parallel axis through the C.M. If $\hat{n}$ is *not* a principal axis, recall that such scalar moments of inertia are useful only for calculating the rotational kinetic energy. If, on the other hand, $\hat{n}$ *is* a principal axis, I and I_{CM} *do* relate $\mathbf{L}$ to $\vec{\omega}$, and Eq. (9.21) is equivalent to the simple form of the parallel axis theorem that is usually found in introductory courses.

10. Euler's equations for rotational motion.

10.1. Evolution of the angular velocity.

Consider the rotation of a free rigid body ($U = 0$) about its C.M., using the body's Euler angles ϕ, θ, and ψ as generalized coordinates. The Lagrangian reduces to

$$\mathcal{L} = T = \tfrac{1}{2}(\omega_1^2\mathcal{I}_{11} + \omega_2^2\mathcal{I}_{22} + \omega_3^2\mathcal{I}_{33})$$

if $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are chosen to be the principal axes. Using Eq. (9.6), it would be possible, though messy, to express T in terms of the Euler angles and their time derivatives. Alternatively, to obtain one Euler-Lagrange equation, it is easier to write

$$\begin{aligned}\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\psi}} &= \frac{\partial\mathcal{L}}{\partial\psi}\\ \frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\omega_3}\frac{\partial\omega_3}{\partial\dot{\psi}}\Big) &= \frac{\partial\mathcal{L}}{\partial\omega_1}\frac{\partial\omega_1}{\partial\psi} + \frac{\partial\mathcal{L}}{\partial\omega_2}\frac{\partial\omega_2}{\partial\psi}.\end{aligned}$$

In the second equality we have used the fact that only ω_3 depends on $\dot{\psi}$, and only ω_1 and ω_2 depend on ψ, according to Eq. (9.6). (We chose to consider the Euler-Lagrange equation in ψ rather than ϕ or θ because of these simplifications.)

Plugging in the partial derivatives, including

$$\frac{\partial\omega_3}{\partial\dot{\psi}} = 1 \quad \frac{\partial\omega_1}{\partial\psi} = \omega_2 \quad \frac{\partial\omega_2}{\partial\psi} = -\omega_1$$

from Eq. (9.6), the Euler-Lagrange equation in ψ becomes

$$\mathcal{I}_{33}\dot{\omega}_3 = \mathcal{I}_{11}\omega_1\omega_2 - \mathcal{I}_{22}\omega_2\omega_1.$$

If an external generalized force Q_ψ in the ψ direction were present, one would add Q_ψ to the right hand side. Since the generalized coordinate ψ is an angle, and $\dot{\psi}$ is a rotation about the $3''' = 3$ axis, Q_ψ is simply the component N_3 of the external torque on the body. Adding this term,

$$\begin{aligned}\mathcal{I}_{33}\dot{\omega}_3 - (\mathcal{I}_{11} - \mathcal{I}_{22})\omega_1\omega_2 &= N_3\\ \mathcal{I}_{11}\dot{\omega}_1 - (\mathcal{I}_{22} - \mathcal{I}_{33})\omega_2\omega_3 &= N_1\\ \mathcal{I}_{22}\dot{\omega}_2 - (\mathcal{I}_{33} - \mathcal{I}_{11})\omega_3\omega_1 &= N_2.\end{aligned} \tag{9.22}$$

Here, recognizing that the $\mathbf{e}_3$ direction is not unique, we have cyclically permuted the indices

to obtain all three equations. These are the famous *Euler equations* prescribing the evolution of the angular velocity of a rigid body. Although we derived them by considering the Euler-Lagrange equation in the Euler angle ψ, Euler's equations involve only the Cartesian components of the angular velocity. These equations can also be derived using only Newtonian mechanics.

10.2. Torque-free symmetrical top.

Consider a rotating body that is *symmetric* ($\mathcal{I}_{11} = \mathcal{I}_{22} \equiv I_0$) and that is free of external torques ($\mathbf{N} = 0$). The first condition is commonly found, as it is satisfied by any object that is cylindrically symmetric about the $\mathbf{e}_3$ axis. However, the latter condition is not often encountered in everyday experience. For the torque to vanish, the top must be located in a "weightless" environment, or be supported on bearings ("gimbals") whose axes intersect its C.M. The earth is one example. As usual, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the principal axes. Two cases are distinguished. If $\mathcal{I}_{33} \equiv I_3 < I_0$, as would be the case for a cigar, the top is called *prolate*. Otherwise, like a pancake, the top is *oblate*.

For either the prolate or oblate top, Euler's equations reduce to

$$\begin{aligned} I_3\dot{\omega}_3 &= 0 \quad \omega_3 = \text{constant} \\ \dot{\omega}_1 &= -\big((I_3/I_0) - 1\big)\omega_2\omega_3 \\ \dot{\omega}_2 &= \big((I_3/I_0) - 1\big)\omega_3\omega_1 \\ \mathbf{e}_1\dot{\omega}_1 + \mathbf{e}_2\dot{\omega}_2 &= \big((I_3/I_0) - 1\big)\omega_3(\omega_1\mathbf{e}_2 - \omega_2\mathbf{e}_1) \\ \frac{d\vec{\omega}}{dt} &= \vec{\Omega} \times \vec{\omega} \\ \vec{\Omega} &\equiv \mathbf{e}_3\omega_3\big((I_3/I_0) - 1\big). \end{aligned} \tag{9.23}$$

The next to last line describes *precession* of $\vec{\omega}$ about the $\mathbf{e}_3$ axis with angular velocity $\Omega = \omega_3((I_3/I_0) - 1)$. The precession vanishes if all three principal moments of inertia are equal. Otherwise it is CCW for an oblate top ($I_3 > I_0$), and CW for a prolate top.

While this precession of $\vec{\omega}$ is fairly straightforward to describe in the body system, the actual motion is complex and quite different from e.g. the familiar precession caused by gravity acting on a top with one point fixed. As viewed in the body system, what is precessing is not the top, but rather the axis about which it is instantaneously rotating. The angular velocity Ω of precession is small compared to ω itself if I_3 and I_0 are nearly equal.

As an example, imagine that the earth is a perfectly rigid body with a slight bulge at the equator, making it oblate. Here the "equator" is defined not by the earth's instantaneous axis of rotation, but rather by a line painted around the earth's circumference at its maximum bulge. Suppose further that, as the result of some cosmic accident, the earth is rotating about an axis at 40° north latitude relative to the equator, e.g. near Denver. Then, as seen by an observer on the earth, this axis of rotation slowly moves east at the same latitude, through Philadelphia, Madrid, etc.

As seen by an observer in the space (inertial) system, the earth's motion is more complicated. This is because the change of $\hat{\omega}$ is slow relative only to the body axes. As observed in the space system, the axis of the earth's rotation changes more rapidly, because the body axes themselves are spinning with angular frequency ω.

Some sense can be made of the motion observed in the inertial system by considering the angular momentum $\mathbf{L}'$, which is conserved in that system owing to the absence of external torques. Since $\vec{\omega}$ is precessing, its magnitude is constant. The kinetic energy $\frac{1}{2}\vec{\omega} \cdot \mathbf{L}'$ (see Eq. (9.12)) also is conserved. Therefore the angle between $\mathbf{L}'$ and $\vec{\omega}$ is fixed.

Temporarily return to the body axes, and consider the plane formed by $\vec{\omega}$ and $\mathbf{e}_3$. This plane must contain the angular momentum $\mathbf{L}$ because the body is symmetric ($\mathcal{I}_{11} = \mathcal{I}_{22}$). In the case of an oblate earth, the angular momentum is closer to $\mathbf{e}_3$ than is the angular velocity because I_3 is larger than I_0. Therefore $\mathbf{L}$ lies in the plane *between* $\mathbf{e}_3$ and $\vec{\omega}$. Both $\vec{\omega}$ and $\mathbf{L}$ precess CCW about the $\mathbf{e}_3$ axis with the same (small) angular velocity Ω.

Return finally to the space axes. Now $\mathbf{L}'$ is fixed. The other two vectors $\mathbf{e}_3$ and $\vec{\omega}$ still lie

in a plane containing $\mathbf{L}'$, one on either side of $\mathbf{L}'$. Each traces a cone around $\mathbf{L}'$, moving with (large) angular velocity $\approx \omega$. Thus, in the space axes, the earth is observed to *wobble*: it rotates rapidly about an axis which itself is precessing rapidly about $\mathbf{L}'$. This is in addition to the slow precession of $\vec{\omega}$ with respect to $\mathbf{e}_3$.

10.3. Stability of force-free rotation.

In the *symmetric* case just considered, any wobbling arises from an initial misalignment of $\vec{\omega}$ and $\mathbf{e}_3$. When these two vectors are aligned initially, no precession of $\vec{\omega}$ occurs.

Here, for an *asymmetric* torque-free rigid body with unequal principal moments of inertia, $I_1 \neq I_2 \neq I_3$, we investigate the stability of rotation along a direction close to that of a principal axis, say $\mathbf{e}_3$. In other words, $\omega_3 \gg \omega_1$ or ω_2. Euler's equations yield

$$\begin{aligned}
I_3\dot{\omega}_3 &= (I_1 - I_2)\omega_1\omega_2 \\
I_2\dot{\omega}_2 &= (I_3 - I_1)\omega_3\omega_1 \\
I_1\dot{\omega}_1 &= (I_2 - I_3)\omega_2\omega_3 \\
I_1\ddot{\omega}_1 &= (I_2 - I_3)(\dot{\omega}_2\omega_3 + \dot{\omega}_3\omega_2) \\
&= (I_2 - I_3)\big(I_2^{-1}(I_3 - I_1)\omega_3^2\omega_1 - \\
&\quad - I_3^{-1}(I_2 - I_1)\omega_2^2\omega_1\big) \\
&\approx (I_2 - I_3)I_2^{-1}(I_3 - I_1)\omega_3^2\omega_1 \\
0 &= \ddot{\omega}_1 + \frac{(I_3 - I_2)(I_3 - I_1)}{I_1 I_2}\omega_3^2\omega_1.
\end{aligned} \tag{9.24}$$

The first equality guarantees that ω_3 is nearly constant, since $\omega_1\omega_2$ is assumed to be small. Then the last equality is nearly a harmonic oscillator equation for ω_1, provided that the coefficient of ω_1 in the last term is positive. This occurs if I_3 is either the largest or the smallest principal moment. But if I_3 is intermediate between I_1 and I_2, the coefficient is negative. Then the solution to Eq. (9.24) is an exponentially growing ω_1, and the rotation is unstable.

11. Heavy symmetric top with one point fixed.

11.1. Constants of the motion.

We pass now from the torque-free top to analysis of a "heavy" top which is influenced both by the force of gravity $m\mathbf{g}$ and by forces exerted upon it to keep its pivot point fixed. The top possesses the same symmetry $\mathcal{I}_{11} = \mathcal{I}_{22}$ as considered previously. Instead of the C.M., here we choose the top's pivot point as the origin, so that its kinetic energy can be considered to be purely rotational. Since the pivot point, like the C.M., lies on the $\mathbf{e}_3$ axis, the parallel axis theorem guarantees that the principal moments $I_1 = I_2 \equiv I$ about that point are also equal.

The Euler angles ϕ, θ, and ψ are ideal for describing the orientation of the top. As usual, $\mathbf{e}_1'$, $\mathbf{e}_2'$, and $\mathbf{e}_3' \equiv -\hat{\mathbf{g}}$ are the fixed axes, with their origin at the pivot point; $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the body axes, with the same origin. According to Euler's convention, the transformation from the space to the body axes consists first of a rotation about $\mathbf{e}_3'$ by ϕ; next a rotation about the line of nodes (the temporary $\mathbf{e}_1$ direction) by θ; and finally a rotation about the $\mathbf{e}_3$ direction by ψ. Therefore ϕ is the azimuth of the top's axis, as viewed in the space system; θ is the polar angle of that axis measured from $\mathbf{e}_3'$; and ψ is the azimuth of the top about the same axis. In other words, spinning of the top about its own axis is represented by $\dot{\psi}$; precession of the top's axis about $\mathbf{e}_3'$ is represented by $\dot{\phi}$; and *nutation*, the (bobbing) variation of the polar angle of the top's axis, is represented by $\dot{\theta}$.

To write the Lagrangian we must express the top's kinetic energy in terms of the Euler angles. Using Eq. (9.6), we evaluate

$$\begin{aligned}
\omega_1^2 + \omega_2^2 &= (\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)^2 + \\
&\quad + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)^2 \\
&= \dot{\phi}^2\sin^2\theta + \dot{\theta}^2 \\
\omega_3^2 &= (\dot{\phi}\cos\theta + \dot{\psi})^2.
\end{aligned}$$

Using these relations, the Lagrangian is

$$\begin{aligned}
\mathcal{L} &= \tfrac{1}{2}(\mathcal{I}_{11}\omega_1^2 + \mathcal{I}_{22}\omega_2^2 + \mathcal{I}_{33}\omega_3^2) - U \\
&= \tfrac{1}{2}I(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) + \\
&\quad + \tfrac{1}{2}I_3(\dot{\phi}\cos\theta + \dot{\psi})^2 - mgh\cos\theta,
\end{aligned} \tag{11.1}$$

where h is the distance along the $\mathbf{e}_3$ axis from the pivot point to the C.M.

The Lagrangian is independent of the two cyclic coordinates ϕ and ψ. The corresponding conjugate momenta are constants of the motion:

$$\begin{aligned} p_\phi &\equiv \frac{\partial \mathcal{L}}{\partial \dot\phi} = \text{constant} \\ &= I\dot\phi \sin^2\theta + I_3(\dot\phi\cos\theta + \dot\psi)\cos\theta \\ p_\psi &\equiv \frac{\partial \mathcal{L}}{\partial \dot\psi} = \text{constant} \\ &= I_3(\dot\phi\cos\theta + \dot\psi). \end{aligned} \tag{11.2}$$

Equation (11.2) can be rearranged to express $\dot\phi$ in terms of p_ϕ and p_ψ:

$$\begin{aligned} p_\phi &= I\dot\phi\sin^2\theta + p_\psi\cos\theta \\ \dot\phi &= \frac{p_\phi - p_\psi\cos\theta}{I\sin^2\theta}. \end{aligned} \tag{11.3}$$

The final constant of the motion is obtained by noting that the Lagrangian has no explicit time dependence, so that the Hamiltonian $\mathcal{H} =$ constant $\equiv E$. Since T is a generalized quadratic function of $\dot\phi$, $\dot\theta$, and $\dot\psi$, $\mathcal{H} = T + U$. Then

$$\begin{aligned} \mathcal{H} &= \tfrac{1}{2}I(\dot\phi^2\sin^2\theta + \dot\theta^2) + \\ &\quad + \tfrac{1}{2}I_3(\dot\phi\cos\theta + \dot\psi)^2 + mgh\cos\theta \\ E &= \tfrac{1}{2}I\dot\phi^2\sin^2\theta + \tfrac{1}{2}I\dot\theta^2 + \frac{p_\psi^2}{2I_3} + mgh\cos\theta \\ &= \frac{(p_\phi - p_\psi\cos\theta)^2}{2I\sin^2\theta} + \frac{I\dot\theta^2}{2} + \frac{p_\psi^2}{2I_3} + mgh\cos\theta. \end{aligned} \tag{11.4}$$

In the last two lines we used Eqs. (11.2) and (11.3) to eliminate $\dot\phi$ and $\dot\psi$. Eq. (11.4) is the starting point for further analysis.

11.2. Equation of motion in a single coordinate.

Introducing the renormalized energy E' and effective potential U',

$$\begin{aligned} E' &\equiv E - p_\psi^2/2I_3 = \text{constant} \\ U' &\equiv \frac{(p_\phi - p_\psi\cos\theta)^2}{2I\sin^2\theta} + mgh\cos\theta, \end{aligned} \tag{11.5}$$

Eq. (11.4) takes the simple form

$$E' = \tfrac{1}{2}I\dot\theta^2 + U'(\theta). \tag{11.6}$$

This is a differential equation for the single coordinate $\theta(t)$.

Before attempting to solve Eq. (11.6), we recognize that U' is infinite both at $\theta = 0$ and at $\theta = \pi$, unless the special condition $|p_\phi| = |p_\psi|$ is satisfied. Therefore U' must reach a minimum at an intermediate polar angle θ_U. When $\theta = \theta_U$ and $\dot\theta = 0$, E' is minimized. There the top moves with the same uniform precession studied in introductory courses. In relation to the effective potential, this minimum energy solution is reminiscent of the circular orbit in the two-body central force problem. As the energy increases (for fixed p_ϕ and p_ψ), the top *nutates* around θ_U between $\theta_{\min}$ and $\theta_{\max}$. This reminds us of the elliptical orbit in the central force problem, except that the frequency of nutation in general is not an integral multiple of the precession frequency.

In the same way as for two-body central force motion, Eq. (11.6) may be rearranged to yield an integral solution to the motion:

$$\begin{aligned} \tfrac{1}{2}I\left(\frac{d\theta}{dt}\right)^2 &= E' - U'(\theta) \\ t &= \int d\theta \sqrt{\frac{I}{2\big(E' - U'(\theta)\big)}}. \end{aligned} \tag{11.7}$$

Equation (11.6) is simplified by substituting $u \equiv \cos\theta$:

$$\begin{aligned} \dot u &= -\dot\theta\sin\theta = -\dot\theta\sqrt{1-u^2} \\ E' &= I\frac{\dot u^2}{2(1-u^2)} + \frac{(p_\phi - p_\psi u)^2}{2I(1-u^2)} + mghu \\ I^2\dot u^2 &= 2I(1-u^2)(E' - mghu) - (p_\phi - p_\psi u)^2. \end{aligned} \tag{11.8}$$

Unfortunately, this equation is cubic in u.

Further simplification is achieved by releasing the top's axis from rest: $\dot\theta(0) = \dot\phi(0) = 0$, with $u(0) \equiv u_0 \equiv \cos\theta_0$. Referring to Eq. (11.3), we see that $p_\phi - p_\psi\cos\theta_0$ also vanishes. Then, in Eq. (11.5), so does the first term in U' when $t = 0$. Since $\dot\theta$ also vanishes at that time,

Eq. (11.6) reduces to $E' = mghu_0$. Equation (11.8) becomes

$$\begin{aligned} I^2\dot{u}^2 &= 2I(1-u^2)mgh(u_0-u) - (p_\phi - p_\psi u)^2 \\ &= 2I(1-u^2)mgh(u_0-u) - p_\psi^2(u_0-u)^2 \\ &= p_\psi^2\big(\alpha(1-u^2)(u_0-u) - (u_0-u)^2\big) \\ \alpha &\equiv \frac{mgh}{p_\psi^2/2I}. \end{aligned} \tag{11.9}$$

In the last line we have introduced the dimensionless constant α, which is a factor of order unity ($I/2I_3$) multiplied by the ratio of the *range in potential energy* ($2mgh$) to the *initial kinetic energy* ($p_\psi^2/2I_3$).

Although its constants are in neater form, Eq. (11.9) is still a cubic in u. One new piece of information falls out easily. Obviously $\dot{u} = 0$ when $u = u_0$, as the initial conditions demand. In addition, $\dot{u} = 0$ when $u_0 - u = \alpha(1-u^2)$. Denote this second turning point by $u = u_0 - \Delta u$. The condition for $\dot{u} = 0$ becomes

$$\begin{aligned} \Delta u &= \alpha\big(1 - (u_0 - \Delta u)^2\big) \\ \Delta u &= \alpha(1 - u_0^2 + 2u_0\Delta u - \Delta u^2). \end{aligned} \tag{11.10}$$

The range Δu within which the top nutates can be obtained by solving this quadratic equation.

11.3. Nutation of a fast top.

Another advantage of Eq. (11.9), relative to earlier versions, is that the right hand side contains two distinct terms. Depending on the value of α, it may be possible to neglect one with respect to the other. For example, a *fast top* has an initial kinetic energy which greatly exceeds its range in potential energy: $\alpha \ll 1$. Then Eq. (11.10) requires Δu to be small, simplifying to

$$\Delta u \approx \alpha(1 - u_0^2) = \alpha \sin^2\theta_0. \tag{11.11}$$

To solve Eq. (11.9) for the special case of the fast top, we apply the method of perturbations. Since u is known to vary between u_0 and $u_0 - \Delta u = u_0 - \alpha\sin^2\theta_0$, we expand u about its central value $u_0 - \frac{1}{2}\Delta u$:

$$\begin{aligned} u &= u_0 - \tfrac{1}{2}\Delta u + \delta \\ &= u_0 - \tfrac{\alpha}{2}\sin^2\theta_0 + \delta. \end{aligned} \tag{11.12}$$

Here the perturbation δ is of the same small order as α.

Now we insert Eq. (11.12) into Eq. (11.9) and retain terms to second order in α and δ:

$$\begin{aligned} &\frac{I^2}{p_\psi^2}\dot{u}^2 = \alpha(1-u^2)(u_0-u) - (u_0-u)^2 \\ &= (u_0-u)\big(\alpha(1-u^2) - (u_0-u)\big) \\ &= (\tfrac{\alpha}{2}\sin^2\theta_0 - \delta)\big(\alpha(1-u^2) - (\tfrac{\alpha}{2}\sin^2\theta_0 - \delta)\big) \\ &\approx (\tfrac{\alpha}{2}\sin^2\theta_0 - \delta)\big(\alpha\sin^2\theta_0 - (\tfrac{\alpha}{2}\sin^2\theta_0 - \delta)\big) \\ &= (\tfrac{\alpha}{2}\sin^2\theta_0 - \delta)(\tfrac{\alpha}{2}\sin^2\theta_0 + \delta) \\ &\frac{I^2}{p_\psi^2}\dot{\delta}^2 = \frac{\alpha^2}{4}\sin^4\theta_0 - \delta^2. \end{aligned} \tag{11.13}$$

Equation (11.13) is merely quadratic in δ and may be solved by taking the time derivative:

$$\begin{aligned} 2\dot{\delta}\ddot{\delta} &= -\frac{p_\psi^2}{I^2}2\delta\dot{\delta} \\ \ddot{\delta} &= -\frac{p_\psi^2}{I^2}\delta \\ \Omega_{\text{nutation}} &= \frac{p_\psi}{I}. \end{aligned} \tag{11.14}$$

For comparison, the average angular velocity of precession is:

$$\begin{aligned} \langle\dot{\phi}\rangle &= \Big\langle\frac{p_\phi - p_\psi\cos\theta}{I\sin^2\theta}\Big\rangle \\ &= \Big\langle\frac{p_\psi(u_0-u)}{I\sin^2\theta}\Big\rangle \\ &\approx \Big\langle\frac{p_\psi\Delta u}{2I\sin^2\theta}\Big\rangle \\ &\approx \frac{p_\psi\alpha\sin^2\theta_0}{2I\sin^2\theta_0} = \frac{\alpha}{2}\frac{p_\psi}{I}. \end{aligned} \tag{11.15}$$

Therefore, independently of θ_0, the fast top nutates $\Omega_{\text{nutation}}/\langle\dot{\phi}\rangle = 2/\alpha$ times for each period of precession.

11.4. Stability of a sleeping top.

Retaining the same initial conditions, we consider finally the special case of an initially upright ("sleeping") top, i.e. $\theta_0 = 0$. Because

of this additional simplification, it will no longer be necessary to assume that the top is *fast*, i.e. that $\alpha \ll 1$. Beginning with the first equality in Eq. (11.13), set $u_0 = 1$ and define $\epsilon \equiv 1 - u$:

$$\begin{aligned} \frac{I^2}{p_\psi^2}\dot{u}^2 &= \alpha(1-u^2)(u_0-u) - (u_0-u)^2 \\ &= \alpha(1-u^2)(1-u) - (1-u)^2 \\ &= (1-u)^2\big(\alpha(1+u)-1\big) \\ \frac{I^2}{p_\psi^2}\dot{\epsilon}^2 &= \epsilon^2\big(\alpha(2-\epsilon)-1\big). \end{aligned} \tag{11.16}$$

Differentiating both sides with respect to time,

$$\begin{aligned} \frac{I^2}{p_\psi^2}\dot{\epsilon}\ddot{\epsilon} &= \epsilon\dot{\epsilon}\big(\alpha(2-\epsilon)-1\big) - \epsilon^2\alpha\dot{\epsilon} \\ \frac{I^2}{p_\psi^2}\ddot{\epsilon} &= \epsilon\big(\alpha(2-\epsilon)-1\big) - \epsilon^2\alpha \\ &= \epsilon\big(2\alpha(1-\epsilon)-1\big) \\ &\approx -(1-2\alpha)\epsilon. \end{aligned} \tag{11.17}$$

When $\alpha < \frac{1}{2}$, perturbations about $\cos\theta = 1$ oscillate stably with angular frequency

$$\Omega_{\text{nutation}} = \frac{p_\psi}{I}\sqrt{1-2\alpha}. \tag{11.18}$$

When $\alpha \ll 1$, this reduces to Eq. (11.14). When $\alpha > \frac{1}{2}$, the sleeping top is unstable. This is a familiar observation. As friction slows $\dot{\psi}$ and raises α, an initially upright top suddenly nutates violently.

12. Coupled oscillatory motion.

12.1. Lagrangian for small oscillations about equilibrium.

Our approach to the *coupled oscillator problem* will be to develop the most general method, usable for any number n of generalized coordinates q_k, $1 \le k \le n$. Here k runs over *both* the number of *dimensions* and the number of *particles* in the problem; for example, a problem with two particles in three dimensions has $n = 6$.

For a velocity-independent potential $U(q_k)$, about a local minimum $U_{\min}$ where all the q's are defined to vanish, we may expand to second order

$$\begin{aligned} U - U_{\min} &= \sum_k q_k \frac{\partial U}{\partial q_k}\Big|_{U_{\min}} + \\ &\quad + \tfrac{1}{2}\sum_{kl} q_k q_l \frac{\partial^2 U}{\partial q_k \partial q_l}\Big|_{U_{\min}} \\ &\equiv 0 + \tfrac{1}{2}\sum_{kl} q_k q_l \mathcal{K}_{kl} \\ &\equiv \tfrac{1}{2} q_k \mathcal{K}_{kl} q_l \\ &\equiv \tfrac{1}{2}\tilde{q}^t \mathcal{K} \tilde{q} \\ \mathcal{K}_{kl} &\equiv \frac{\partial^2 U}{\partial q_k \partial q_l}\Big|_{U_{\min}}. \end{aligned} \tag{12.1}$$

In the second equality we used the fact that all derivatives with respect to the q_k vanish at the local minimum. Notice the summation over k and l implied by the repeated indices in the third equality, and the matrix notation in the fourth.

The elements $\mathcal{K}_{kl}$ of the *spring constant matrix* $\mathcal{K}$ are *constants*. Likewise, we assume that the potential energy T for this system can be written in terms of the constant *mass matrix* $\mathcal{M}$:

$$\begin{aligned} T &\equiv \tfrac{1}{2}\sum_{kl} \dot{q}_k \dot{q}_l \mathcal{M}_{kl} \\ &\equiv \tfrac{1}{2}\dot{q}_k \mathcal{M}_{kl} \dot{q}_l. \end{aligned} \tag{12.2}$$

This is a much stronger assumption than was made in order to obtain Eq. (12.1). For example, the coordinates must be Cartesian; otherwise the elements $\mathcal{M}_{kl}$ of the mass matrix would be functions of the q_k. Then the Lagrangian is

$$\mathcal{L}(\{q_i\},\{\dot{q}_i\}) = \tfrac{1}{2}\dot{q}_k \mathcal{M}_{kl}\dot{q}_l - \tfrac{1}{2} q_k \mathcal{K}_{kl} q_l. \tag{12.3}$$

The oscillator is *coupled* if *any* of the off-diagonal elements of $\mathcal{K}$ *or* $\mathcal{M}$ are nonzero. Otherwise, the Lagrangian merely describes n *uncoupled* oscillators, each affected by none of the others.

Since the matrix elements in the Lagrangian are constant, the Euler-Lagrange equations in each of the n coordinates q_k yield n different equations of the form

$$\mathcal{K}_{kl} q_l + \mathcal{M}_{kl}\ddot{q}_l = 0, \tag{12.4}$$

each with $2n$ terms. In obtaining Eq. (12.4), one makes use of the fact that $\mathcal{K}$ is *symmetric* by definition, and that we *force* $\mathcal{M}$ to be symmetric by associating equal amounts of kinetic energy with $\mathcal{M}_{kl}$ and $\mathcal{M}_{lk}$.

12.2. Harmonic solution to Euler-Lagrange equation.

In analogy with simple harmonic motion of the undamped oscillator, we speculate that there exist solutions to Eq. (12.4) for which each coordinate q_k executes harmonic motion at the *same frequency* and with the *same phase*, give or take π, as every other coordinate. We *do* allow the amplitude of oscillation a_k to vary with k. Such solutions exist; they are called *normal modes of oscillation*. By analogy with the complex exponential method, taking a_k to be a real constant, substitute

$$q_k(t) = \Re\big(a_k e^{i(\omega t+\delta)}\big) \tag{12.5}$$

in Eq. (12.4):

$$\Re\big((\mathcal{K}_{kl} - \omega^2\mathcal{M}_{kl})a_l e^{i(\omega t+\delta)}\big) = 0. \tag{12.6}$$

As usual, we choose to solve the complex equation of which Eq. (12.6) is the real part. Factoring out the common phase,

$$\begin{aligned}(\mathcal{K}_{kl} - \omega^2\mathcal{M}_{kl})a_l &= 0\\ (\mathcal{K} - \omega^2\mathcal{M})\tilde{a} &= 0,\end{aligned} \tag{12.7}$$

where $\tilde{a}$ is a column vector with elements equal to the a_l's. Since everything else in the linear Eq. (12.7) is real, the a_l's may be defined to be real as well. This supports our original speculation.

12.3. Normal mode eigenvalue problem.

Equation (12.7) is similar to the eigenvalue problem encountered when we diagonalized the inertia tensor. In fact, since both $\mathcal{K}$ and $\mathcal{M}$ are real symmetric matrices, it is the *same* problem, except that $\mathcal{K} - \omega^2\mathcal{M}$ replaces $\mathcal{I} - e\mathrm{I}$. By the same arguments used in section (9.6), there exist n *normal frequencies*^{2} ω_r^2, $1 \le r \le n$. These correspond to the real positive eigenvalues of the inertia tensor. Like those eigenvalues, some of the normal frequencies may be the same (these are called "degenerate"). Corresponding to the eigenvectors of the inertia tensor, there exist n real *normal mode vectors* $\tilde{a}^r$.

The normal mode vectors are orthogonal, but only in a special way:

$$\begin{aligned}(\tilde{a}^r)^t\tilde{a}^s &\neq 0 \text{ when } r \neq s, \text{ but}\\ (\tilde{a}^r)^t\mathcal{M}\tilde{a}^s &= 0 \text{ when } r \neq s.\end{aligned} \tag{12.8}$$

This special type of orthogonality is readily understood in the context of the proof in section (9.6). That proof used the fact that the eigenvalue e multiplies the unit matrix I in the eigenvalue equation (9.16). However, in the coupled oscillator problem, the normal frequency^2 ω^2 multiplies $\mathcal{M}$ rather than I.

Equation (12.8) still leaves us free to choose the *lengths* of the normal mode vectors. Our choice produces the most elegant condition:

$$(\tilde{a}^r)^t\mathcal{M}\tilde{a}^s = \delta_{rs}. \tag{12.9}$$

The normal mode vectors are said to be *orthonormal in a space having the metric* $\mathcal{M}$, rather than in a space having the usual unit metric I.

Although the analogy with the inertia tensor eigenvalue problem is complete, we record here the steps required to obtain the normal frequencies and the normal mode vectors. First we identify the spring constant and mass matrices for the particular problem at hand. Then, for a solution to Eq. (12.7) to exist, we demand that the *secular equation* be satisfied:

$$\det(\mathcal{K} - \omega^2\mathcal{M}) = 0. \tag{12.10}$$

This is an nth order polynomial equation for ω^2. Its n roots are the normal frequencies^2 ω_r^2. Next, for *each* of these n frequencies, we solve n simultaneous equations of the form

$$(\mathcal{K} - \omega_r^2\mathcal{M})\tilde{a}^r = 0 \tag{12.11}$$

to obtain each of the normal mode vectors $\tilde{a}^r$. Finally, we adjust the lengths of each of the $\tilde{a}^r$'s

in order to satisfy the orthonormality requirement (12.9).

12.4. Transformation to normal coordinates.

In our study of the inertia tensor $\mathcal{I}$, we benefited by considering the square matrix Λ^t whose columns are its eigenvectors. The untransposed version Λ of that matrix was found to represent the orthogonal transformation that diagonalizes $\mathcal{I}$ by means of the *similarity transformation* $\Lambda\mathcal{I}\Lambda^t$.

Correspondingly, we now form the matrix $\mathcal{A}$ whose columns are the normal mode vectors. This means that the $(kr)^{\text{th}}$ element of $\mathcal{A}$ is the kth component of the rth normal mode vector:

$$\mathcal{A}_{kr} \equiv a_k^r. \tag{12.12}$$

The benefits of considering $\mathcal{A}$ are even more striking. The orthonormality condition (12.9) becomes still more elegant:

$$\begin{aligned} \delta_{rs} &= (\tilde{a}^r)^t \mathcal{M} \tilde{a}^s \\ &= a_k^r \mathcal{M}_{kl} a_l^s \\ &= \mathcal{A}_{kr} \mathcal{M}_{kl} \mathcal{A}_{ls} \\ &= \mathcal{A}^t_{rk} \mathcal{M}_{kl} \mathcal{A}_{ls} \\ &= (\mathcal{A}^t \mathcal{M} \mathcal{A})_{rs} \\ \mathrm{I} &= \mathcal{A}^t \mathcal{M} \mathcal{A}. \end{aligned} \tag{12.13}$$

Equation (12.13) states that, by means of the *congruence transformation* $\mathcal{A}^t\mathcal{M}\mathcal{A}$, the matrix $\mathcal{A}$ of normal mode vectors not only diagonalizes the mass matrix $\mathcal{M}$; it reduces $\mathcal{M}$ to the unit matrix I.

A similar proof using the spring constant matrix yields

$$\begin{aligned} \mathcal{A}^t \mathcal{K} \mathcal{A} &\equiv \Omega^2 \\ &= \begin{pmatrix} \omega_1^2 & 0 & \cdots & 0 \\ 0 & \omega_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_n^2 \end{pmatrix}. \end{aligned} \tag{12.14}$$

Again, by means of the congruence transformation $\mathcal{A}^t\mathcal{K}\mathcal{A}$, the matrix $\mathcal{A}$ of normal mode vectors diagonalizes the spring constant matrix $\mathcal{K}$. The resulting matrix Ω^2 is the *diagonal matrix of normal frequencies*[2] ω_r^2.

How is it possible that the same transformation $\mathcal{A}$ is able to diagonalize two different real symmetric matrices $\mathcal{M}$ and $\mathcal{K}$? Imagine first performing a simple rotation to diagonalize $\mathcal{M}$. Then transform to a renormalized set of generalized coordinates so that $\mathcal{M}$ is proportional to the unit matrix. Finally, perform a second rotation to diagonalize $\mathcal{K}$. This second rotation preserves the diagonal form of $\mathcal{M}$ because it is already the unit matrix.

So far we have been discussing the effect of $\mathcal{A}$ upon the mass and spring constant matrices. Now consider using $\mathcal{A}$ to transform *from* a new set of n generalized coordinates $\{Q_r\}$ *to* the original set of n generalized coordinates $\{q_k\}$:

$$\begin{aligned} q_k(t) &\equiv \sum_r \mathcal{A}_{kr} Q_r(t) \\ \tilde{q} &\equiv \mathcal{A}\tilde{Q} \\ \tilde{Q} &= \mathcal{A}^{-1}\tilde{q} \\ &= \mathcal{A}^t \mathcal{M} \tilde{q}. \end{aligned} \tag{12.15}$$

In the last line we set $\mathcal{A}^{-1} = \mathcal{A}^t\mathcal{M}$ using Eq. (12.13). The $\{Q_r\}$ are called *normal coordinates*.

In the original basis $\{q_k\}$, neither $\mathcal{M}$ nor $\mathcal{K}$ were diagonal. But in the new basis $\{Q_r\}$,

$$\begin{aligned} T &= \tfrac{1}{2}\dot{\tilde{q}}^t \mathcal{M} \dot{\tilde{q}} \\ &= \tfrac{1}{2}(\mathcal{A}\dot{\tilde{Q}})^t \mathcal{M}\mathcal{A}\dot{\tilde{Q}} \\ &= \tfrac{1}{2}\dot{\tilde{Q}}^t \mathcal{A}^t \mathcal{M}\mathcal{A}\dot{\tilde{Q}} \\ &= \tfrac{1}{2}\dot{\tilde{Q}}^t \mathrm{I}\dot{\tilde{Q}} \\ &= \tfrac{1}{2}\dot{\tilde{Q}}^t \dot{\tilde{Q}}. \end{aligned} \tag{12.16}$$

Similarly,

$$U = \tfrac{1}{2}\tilde{Q}^t \Omega^2 \tilde{Q}. \tag{12.17}$$

In the new basis $\{Q_r\}$, called the *normal basis*, both the mass and spring constant matrices are diagonal. Application of the Euler-Lagrange equations yields one simple harmonic equation

for each normal coordinate. In the new basis, Eq. (12.4) is simply

$$\omega_r^2 Q_r + \ddot{Q}_r = 0. \tag{12.18}$$

The oscillation of each normal coordinate is *totally decoupled* from that of any other normal coordinate. The normal modes of oscillation are connected only through the initial conditions. Thereafter, oblivious to the others, each normal mode forever rattles away at its own normal frequency.

When a system of oscillators is excited in *only one* normal mode, what are the motions of the *original* generalized coordinates q_k? The simple answer is provided by Eq. (12.15). For example, if only mode seven is excited, only Q_7 is nonzero:

$$q_k(t) = A_{k7} Q_7(t) = a_k^7 Q_7(t). \tag{12.19}$$

The relative amplitude of the motion of each of the n generalized coordinates is given simply by the relative size of the n components a_k^7 of the seventh normal mode vector. All coordinates q_k execute simple harmonic motion with the same angular frequency ω_7 and the same phase.

Of course, the general solution involves the excitation of all n normal modes. Then each original generalized coordinate executes the sum of n simple harmonic oscillations with n (generally) different angular frequencies and phases. For any coordinate, the amplitude of a particular oscillation frequency is given by the relative amplitude of motion of that coordinate for the particular normal mode which has that frequency, multiplied by the amplitude with which that particular normal mode is excited, as determined by the initial conditions.

12.5. Obtaining normal mode amplitudes from initial conditions.

Let's express the previous paragraph in equations rather than words. The initial conditions require each normal coordinate $Q_r(t)$ to have a unique amplitude and phase:

$$Q_r(t) = \Re(P_r e^{i\omega_r t}), \tag{12.20}$$

where P_r is a complex constant. From Eq. 12.15, the original generalized coordinates q_k become

$$q_k(t) = \Re\Big(\sum_r \mathcal{A}_{kr} P_r e^{i\omega_r t}\Big). \tag{12.21}$$

The two constants $\mathcal{A}_{kr}$ and P_r are different in form and in function. $\mathcal{A}_{kr}$ is *real*, with a magnitude *fixed* by the orthonormality condition (12.9). For oscillation in the rth normal mode, it determines the relative amplitude of the kth generalized coordinate q_k. P_r is *complex*, with a magnitude and phase *adjusted* to fit the initial conditions. It *is* the amplitude with which the rth normal mode is excited.

As for the initial conditions, suppose that

$$q_k(0) \equiv q_{0k}\ ;\ \dot{q}_k(0) \equiv \dot{q}_{0k}, \tag{12.22}$$

where q_{0k} and $\dot{q}_{0k}$ are real constants. (Note that $\dot{q}_{0k} \neq dq_{0k}/dt$!) Suppose also that the complex normal mode amplitude P_r is expressed in terms of its real and imaginary parts:

$$P_r \equiv R_r + \frac{S_r}{i\omega_r}, \tag{12.23}$$

where R_r and S_r are real constants. Then from Eqs. (12.21) and (12.22),

$$\begin{aligned} q_{0k} &= \sum_r \mathcal{A}_{kr} R_r \\ \dot{q}_{0k} &= \sum_r \mathcal{A}_{kr} S_r. \end{aligned} \tag{12.24}$$

Transforming to matrix notation,

$$\begin{aligned} \tilde{q}_0 &= \mathcal{A}\tilde{R}; \quad \tilde{R} = \mathcal{A}^{-1}\tilde{q}_0 \\ \tilde{\dot{q}}_0 &= \mathcal{A}\tilde{S}; \quad \tilde{S} = \mathcal{A}^{-1}\tilde{\dot{q}}_0. \end{aligned} \tag{12.25}$$

Using $\mathcal{A}^{-1} = \mathcal{A}^t\mathcal{M}$, we have finally

$$\begin{aligned} \tilde{R} &= \mathcal{A}^t\mathcal{M}\tilde{q}_0 \\ \tilde{S} &= \mathcal{A}^t\mathcal{M}\tilde{\dot{q}}_0. \end{aligned} \tag{12.26}$$

With $\tilde{R}$ and $\tilde{S}$ specified, so is $\tilde{P}$ from Eq. (12.23), yielding the $q_k(t)$ from Eq. (12.21).

12.6. Energy in normal modes.

Using Eqs. (12.16), (12.17), (12.20), and (12.23), the conserved total energy is

$$\begin{aligned} E &= T + U \\ &= \tfrac{1}{2}\dot{\tilde{Q}}^t(t)\dot{\tilde{Q}}(t) + \tfrac{1}{2}\tilde{Q}^t(t)\Omega^2\tilde{Q}(t) \\ &= \tfrac{1}{2}\dot{\tilde{Q}}^t(0)\dot{\tilde{Q}}(0) + \tfrac{1}{2}\tilde{Q}^t(0)\Omega^2\tilde{Q}(0) \\ &= \tfrac{1}{2}\tilde{S}^t\tilde{S} + \tfrac{1}{2}\tilde{R}^t\Omega^2\tilde{R} \\ &= \tfrac{1}{2}\sum_r (S_r^2 + \omega_r^2 R_r^2) \\ &= \tfrac{1}{2}\sum_r \omega_r^2 P_r^* P_r \\ &= \tfrac{1}{2}\tilde{P}^\dagger\Omega^2\tilde{P}. \end{aligned} \tag{12.27}$$

Although energy *within* a particular normal mode is shifted back and forth between kinetic and potential forms, as occurs in any harmonic oscillator, the sum of potential and kinetic energies in that mode is constant. According to Eq. (12.27), the total energy E is just the sum of energies in the individual modes, which in turn depend only upon the normal frequencies ω_r and the moduli $|P_r|$ of the mode amplitudes.

12.7. Driven coupled oscillator.

So far we have been discussing the general solution to the homogeneous equation, corresponding to free vibrations. If a driving term is present, we need a particular solution of the inhomogenous equation to add to the homogeneous solution. Suppose that each of the original generalized coordinates q_k is subjected to a generalized force

$$f_k \equiv \Re(g_k e^{i\omega t}), \tag{12.28}$$

where g_k is a complex constant. Note that the driving frequency ω is assumed to be the same for all coordinates.

We shall express the generalized force F_r acting on the normal coordinate Q_r as

$$F_r \equiv \Re(G_r e^{i\omega t}). \tag{12.29}$$

Then, recalling from Eq. (12.15) that $q_k = \mathcal{A}_{kr}Q_r$, we require that the work done by the force when calculated in either coordinate system be the same:

$$\begin{aligned} \sum_r F_r\, dQ_r &= \sum_k f_k\, dq_k \\ &= \sum_k f_k \sum_r \mathcal{A}_{kr}\, dQ_r \\ &= \sum_r \Big(\sum_k \mathcal{A}_{kr} f_k\Big) dQ_r \\ F_r &= \sum_k \mathcal{A}_{kr} f_k \\ \tilde{F} &= \mathcal{A}^t \tilde{f} \\ \tilde{G} &= \mathcal{A}^t \tilde{g}. \end{aligned} \tag{12.30}$$

With the addition of the generalized force, the Euler-Lagrange equation satisfied by the normal coordinate Q_r becomes

$$\ddot{Q}_r = -\omega_r^2 Q_r + \Re(G_r e^{i\omega t}). \tag{12.31}$$

Seeking a solution of the form

$$Q_r(t) \equiv \Re(P_r' e^{i\omega t}), \tag{12.32}$$

where P_r' is a complex constant, we easily find

$$P_r' = \frac{G_r}{\omega_r^2 - \omega^2}. \tag{12.33}$$

Introducing

$$\mathcal{T}^2 \equiv \begin{pmatrix} \frac{1}{\omega_1^2-\omega^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\omega_2^2-\omega^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\omega_n^2-\omega^2} \end{pmatrix}, \tag{12.34}$$

the particular solution is

$$\begin{aligned} \tilde{P}' &= \mathcal{T}^2\tilde{G} \\ &= \mathcal{T}^2\mathcal{A}^t\tilde{g} \\ \mathcal{A}\tilde{P}' &= \mathcal{A}\mathcal{T}^2\mathcal{A}^t\tilde{g} \\ \Re(\mathcal{A}\tilde{P}' e^{i\omega t}) &= \Re(\mathcal{A}\mathcal{T}^2\mathcal{A}^t\tilde{g}e^{i\omega t}) \\ \mathcal{A}\tilde{Q}(t) &= \mathcal{A}\mathcal{T}^2\mathcal{A}^t\Re(\tilde{g}e^{i\omega t}) \\ \tilde{q}(t) &= \mathcal{A}\mathcal{T}^2\mathcal{A}^t\tilde{f}(t), \end{aligned} \tag{12.35}$$

where it is understood that the components of $\tilde{f}(t)$ are sinusoidal in ω. Equation (12.35) expresses a simple result: the steady-state response $\tilde{q}(t)$ of the original generalized coordinates to a generalized sinusoidal driving force $\tilde{f}(t)$ is given by the operator $\mathcal{A}\mathcal{T}^2\mathcal{A}^t$, which is just $\mathcal{T}^2$ after a similarity transformation by $\mathcal{A}$.

13. Lagrangian density for continuous systems.

13.1. Calculus of variations for two independent variables.

In section (4.1) we considered the problem of finding the path $y(t)$ such that the action

$$J = \int_{t_1}^{t_2} dt\, \mathcal{L}(y, \tfrac{\partial y}{\partial t}, t)$$

is extremized, where the limits of integration are fixed, and where $\mathcal{L}$ (later called the Lagrangian) is a function of the indicated variables. The solution was given by the Euler equation (4.3):

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial(\frac{\partial y}{\partial t})} = \frac{\partial \mathcal{L}}{\partial y}. \tag{13.1}$$

Here we have chosen to write the first derivative of y with respect to t as $\frac{\partial y}{\partial t}$ rather than as $\dot{y}$. (Since y is a function only of the independent variable t, the partial and total time derivatives of y are equivalent.)

Now we consider a slightly more general problem. The new action to be extremized is a *two*-dimensional integral of a new function $\mathcal{L}'$, later to be called the *Lagrangian density*. The path y which extremizes the action is a function of *two* independent variables s and t. The integral is taken over these same two variables. The Lagrangian density $\mathcal{L}'(y, \frac{\partial y}{\partial s}, \frac{\partial y}{\partial t}, s, t)$ is a function of y and its partial derivatives with respect both to s and t. It may also be a function of s and/or t explicitly. The action is

$$J = \int_{s_1}^{s_2} ds \int_{t_1}^{t_2} dt\, \mathcal{L}'(y, \tfrac{\partial y}{\partial s}, \tfrac{\partial y}{\partial t}, s, t),$$

where the limits of *both* integrals are fixed.

It is not surprising that a derivation of the same type as that in section (4.1) gives a slightly more general Euler equation:

$$\frac{d}{ds}\frac{\partial \mathcal{L}'}{\partial(\frac{\partial y}{\partial s})} + \frac{d}{dt}\frac{\partial \mathcal{L}'}{\partial(\frac{\partial y}{\partial t})} = \frac{\partial \mathcal{L}'}{\partial y}. \tag{13.2}$$

This is the same as Eq. (13.1) except for the added first term, which is equivalent to the second term with t replaced by s. The meaning of $\frac{d}{ds}$ and $\frac{d}{dt}$ requires clarification. For example, $\frac{d}{dt}$ is *total* because it includes the variation of $\mathcal{L}'$ with respect to t both *explicitly* and *implicitly* through the dependence upon t of y and its derivatives. However, $\frac{d}{dt}$ is also *partial* because the other independent variable s must be held fixed throughout the differentiation. To be excessively precise,

$$\begin{aligned}\frac{d}{dt} \equiv &\left(\frac{\partial}{\partial t}\right)_{y,\frac{\partial y}{\partial s},\frac{\partial y}{\partial t},s} + \\ &+ \left(\frac{\partial y}{\partial t}\right)_s \left(\frac{\partial}{\partial y}\right)_{\frac{\partial y}{\partial s},\frac{\partial y}{\partial t},s,t} + \\ &+ \left(\frac{\partial \frac{\partial y}{\partial s}}{\partial t}\right)_s \left(\frac{\partial}{\partial \frac{\partial y}{\partial s}}\right)_{y,\frac{\partial y}{\partial t},s,t} + \\ &+ \left(\frac{\partial \frac{\partial y}{\partial t}}{\partial t}\right)_s \left(\frac{\partial}{\partial \frac{\partial y}{\partial t}}\right)_{y,\frac{\partial y}{\partial s},s,t},\end{aligned} \tag{13.3}$$

where the subscripts are the variables to be held fixed during the differentiation. Note that terms similar to those in the first, second, and last lines of this expression would be present even if we had only one independent variable (t).

The generalization of Eq. (13.2) to n new independent variables s_i, $1 \le i \le n$, is obvious:

$$\frac{d}{ds_i}\frac{\partial \mathcal{L}'}{\partial(\frac{\partial y}{\partial s_i})} + \frac{d}{dt}\frac{\partial \mathcal{L}'}{\partial(\frac{\partial y}{\partial t})} = \frac{\partial \mathcal{L}'}{\partial y}, \tag{13.4}$$

where, as usual, summation over the repeated index i is implied.

13.2. Hamilton's Principle for continuous systems.

The foregoing section was concerned merely with a mathematical problem in variational calculus. As in section (5.1), the connection with physics is recovered by identifying the independent variable t with the *time*. Of what use is the other independent variable s?

One answer is that s may play the role of a *field variable* in one spatial dimension. For example, to describe the state of a string which may be displaced in one transverse direction, we must specify the displacement y as a function both of time t and of the position s along the string. The fixed limits of integration t_1 and t_2 correspond to the fixed time interval over which the action is to be minimized. Correspondingly, the fixed limits of integration s_1 and s_2 correspond to the *fixed endpoints of the string*.

In order to preserve the action's units, we assign to the integrand $\mathcal{L}'$ the units of the Lagrangian $\mathcal{L} = T - U$ divided by those of s. In our example of a string, $\mathcal{L}'$ is the *Lagrangian per unit length* along the string – that is, $T - U$ per unit length. More generally, $\mathcal{L}'$ is called the *Lagrangian density*.

Hamilton's Principle for continuous systems asserts that the system will evolve along the path $y(s,t)$ which minimizes the action – the integral with respect to s and t of the Lagrangian density $\mathcal{L}'$. This means that y satisfies the Euler-Lagrange equation (13.2). Again, the justification of Hamilton's Principle relies on the fact that it reproduces the solutions to problems amenable to Newtonian analysis, while it yields solutions to more complex problems that are confirmed by experiment.

13.3. Transverse wave equation for a string.

Consider an infinitesimal piece Δs of string in a gravity-free region. We assume that any motion is possible only in the single transverse direction y. The kinetic energy of the piece of string is

$$\Delta T = \tfrac{1}{2}\mu\big(\frac{\partial y}{\partial t}\big)^2\Delta s,$$

where μ is the string mass per unit length.

The potential energy requires a bit more discussion. If the string's slope is $\frac{\partial y}{\partial s}$, according to Pythagoras' theorem the length of the piece of string is increased with respect to its equilibrium length by

$$\begin{aligned}\Delta l &= \sqrt{(\Delta s)^2 + (\Delta y)^2} - \Delta s \\ &= \Big(\sqrt{1 + \big(\frac{\partial y}{\partial s}\big)^2} - 1\Big)\Delta s \\ &\approx \tfrac{1}{2}\big(\frac{\partial y}{\partial s}\big)^2\Delta s,\end{aligned}$$

where the approximation is valid for small slopes. If the string is stretched with tension τ, the incremental potential energy associated with this extra length is $\Delta U = \tau\,\Delta l$.

The Lagrangian density corresponding to ΔT and ΔU is

$$\begin{aligned}\mathcal{L}' &= \frac{\Delta T}{\Delta s} - \frac{\Delta U}{\Delta s} \\ &= \tfrac{1}{2}\mu\big(\frac{\partial y}{\partial t}\big)^2 - \tfrac{1}{2}\tau\big(\frac{\partial y}{\partial s}\big)^2.\end{aligned} \tag{13.5}$$

Applying the Euler-Lagrange Eq. (13.2) to $\mathcal{L}'$,

$$\begin{aligned}-\frac{d}{ds}\big(\tau\frac{\partial y}{\partial s}\big) + \frac{d}{dt}\big(\mu\frac{\partial y}{\partial t}\big) &= 0 \\ -\tau\frac{\partial^2 y}{\partial s^2} + \mu\frac{\partial^2 y}{\partial t^2} &= 0 \\ \frac{\partial^2 y}{\partial s^2} - \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} &= 0 \\ c &\equiv \sqrt{\frac{\tau}{\mu}}.\end{aligned} \tag{13.6}$$

In the last line we defined the *phase velocity* c. Equation (13.6), the *wave equation*, is the starting point for our discussion of waves in one dimension.

14. Waves.

14.1. General solution to the wave equation.

The general solution to the wave equation (13.6),

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = 0, \tag{14.1}$$

is

$$y(x,t) = y_+(x-ct) + y_-(x+ct), \tag{14.2}$$

where y_+ and y_- are any twice differentiable functions of their arguments. The first (second) term is an arbitrary smooth shape travelling in the positive (negative) x direction with velocity c. Following the usual convention, in Eq. (14.1) we substituted x for s. Nevertheless, it cannot be emphasized too strongly that x and y play *completely* different roles in the Euler-Lagrange equation (13.2). The former is an *independent* variable like the time; the latter is a variable that *depends* on x and t.

The fact that Eq. (14.2) solves the wave equation is easily verified by substitution. The fact that it is a *general* solution is illustrated by obtaining the $\{y_+, y_-\}$ which satisfy the (fairly) general set of initial conditions

$$y(x,0) \equiv y_0(x); \quad \dot{y}(x,0) \equiv v_0(x). \tag{14.3}$$

Note that the initial conditions for a continuous system are specified as *functions* rather than *numbers*. Substituting the solution (14.2) at $t = 0$,

$$\begin{aligned} y_0(x) &= y_+(x) + y_-(x) \\ v_0(x) &= -cy'_+(x) + cy'_-(x), \end{aligned} \tag{14.4}$$

where the $'$ indicates differentiation of y_+ and y_- with respect to their (different) arguments. Differentiating the first line in Eq. (14.4) with respect to x, dividing the second line by $-c$, and adding,

$$\begin{aligned} 2y'_+(x) &= y'_0(x) - \tfrac{1}{c}v_0(x) \\ y_+(x) &= \tfrac{1}{2}y_0(x) - \tfrac{1}{2c}\int_0^x du\, v_0(u) + C_+. \end{aligned}$$

Similarly,

$$y_-(x) = \tfrac{1}{2}y_0(x) + \tfrac{1}{2c}\int_0^x du\, v_0(u) + C_-,$$

where $C_+ + C_- = 0$. The full solution is then

$$\begin{aligned} y(x,t) = \frac{y_0(x-ct) + y_0(x+ct)}{2} + \\ + \frac{1}{2c}\int_{x-ct}^{x+ct} du\, v_0(u). \end{aligned} \tag{14.5}$$

14.2. Travelling sinusoidal waves.

A special case of the general solution (14.2) occurs when y_+ and y_- are harmonic functions. Allowing the amplitude, angular frequency, and phase of either wave to be arbitrary,

$$\begin{aligned} y_+(x-ct) &\equiv \Re\big(\tilde{A}_+ e^{ik_+(ct-x)}\big) \\ y_-(x+ct) &\equiv \Re\big(\tilde{A}_- e^{ik_-(ct+x)}\big). \end{aligned} \tag{14.6}$$

The constants $\tilde{A}_\pm$ are the *complex wave amplitudes*. The real constants $k_\pm$ have been introduced to make the exponents dimensionless.

Obviously the *time* dependence of Eq. (14.6) is of the form $\exp(i\omega_\pm t)$ with

$$\omega_\pm = ck_\pm. \tag{14.7}$$

The *spatial* dependence is of the similar form $\exp(\mp ik_\pm x)$. In a single (x) dimension, the $k_\pm$ are *wave numbers* $2\pi/\lambda_\pm$, where $\lambda_\pm$ are the wavelengths. In three dimensions, $k_\pm x$ is replaced by $\mathbf{k}_\pm \cdot \mathbf{x}$; the $\mathbf{k}_\pm$ are called *wave vectors*. The relation (14.7) between ω and k is called a *dispersion relation* (because the waves disperse if the relationship between ω and k is nonlinear).

As usual, it is customary and convenient to work directly with the complex displacement $\tilde{A}_\pm \exp\big(ik_\pm(ct \mp x)\big)$, rather than with its real part $y_\pm$.

The phenomenon of *beats* between sinusoidal waves is most simply demonstrated by the case of two waves of equal amplitude and phase but different wave number, travelling in the same ($+x$) direction:

$$y(x,t) = \Re\big(\tilde{A}e^{i(\omega_1 t - k_1 x)} + \tilde{A}e^{i(\omega_2 t - k_2 x)}\big). \tag{14.8}$$

Assuming $\Delta k \ll k_0$, define

$$\begin{aligned} k_0 &\equiv \tfrac{1}{2}(k_1 + k_2) \\ \Delta k &\equiv k_2 - k_1 \\ \omega_1 &\equiv \omega(k_1) \approx \omega_0 - \frac{\Delta k}{2}\frac{d\omega}{dk}\Big|_{k_0} \\ \omega_2 &\equiv \omega(k_2) \approx \omega_0 + \frac{\Delta k}{2}\frac{d\omega}{dk}\Big|_{k_0}. \end{aligned} \tag{14.9}$$

Here we have allowed for the possibility that $\frac{d\omega}{dk}$ is not always equal to a constant value c, as it is for a simple string. This occurs particularly on stiff strings (as on the upper octaves of a piano). Finally, define the *group velocity*

$$v_{\rm gr} \equiv \frac{d\omega}{dk}, \tag{14.10}$$

here not necessarily equal to the *phase velocity*

$$c \equiv \frac{\omega_0}{k_0}. \tag{14.11}$$

With these definitions, Eq. (14.8) becomes

$$\begin{aligned} y(x,t) &= \Re\Big\{\tilde{A}\Big(e^{i(\omega_0 t - k_0 x)}\big(e^{i\frac{\Delta k}{2}(v_{\rm gr}t - x)} + \\ &+ e^{-i\frac{\Delta k}{2}(v_{\rm gr}t - x)}\big)\Big)\Big\} \\ &= \Re\Big\{2\tilde{A}\Big(e^{i(\omega_0 t - k_0 x)}\cos\big(\tfrac{\Delta k}{2}(v_{\rm gr}t - x)\big)\Big)\Big\}. \end{aligned} \tag{14.12}$$

In this result, $\exp\big(i(\omega_0 t - k_0 x)\big)$ is the high-frequency short-wavelength *carrier wave* moving with speed c; $\cos\big(\frac{\Delta k}{2}(v_{\rm gr}t - x)\big)$ is the low-frequency long-wavelength *amplitude modulating wave* moving with speed $v_{\rm gr}$. Because it multiplies the carrier, this modulating wave is an *envelope* that determines the amplitude of the short carrier waves inside it. Since the information content is provided by these amplitude variations, the information is borne by the *modulation*: "the music travels with the group velocity".

14.3. Standing waves and normal modes.

Starting with a pair of sinusoidal waves as in Eq. (14.6), next we investigate the implications of requiring that the string be *fixed* at its two endpoints $x = 0$ and $x = L$. At the first endpoint,

$$\begin{aligned} 0 &= y(x{=}0, t) \\ &= \Re\big(\tilde{A}_+ e^{i\omega_+ t} + \tilde{A}_- e^{i\omega_- t}\big) \\ \Rightarrow \tilde{A}_+ &= -\tilde{A}_- \equiv \frac{\tilde{A}}{2i} \\ \Rightarrow \omega_+ &= \omega_- \equiv \omega; \quad k_+ = k_- \equiv k. \end{aligned}$$

With these substitutions, the displacement y is a *standing wave*

$$y(x,t) = \Re\big(\tilde{A}e^{i\omega t}\big)\sin kx.$$

At the second endpoint $x = L$,

$$\begin{aligned} 0 &= y(x{=}L, t) \\ &= \Re\big(\tilde{A}e^{i\omega t}\big)\sin kL \\ \Rightarrow 0 &= \sin kL \\ \Rightarrow k_n &= \frac{n\pi}{L}, \quad n = 1, 2, 3\ldots \end{aligned}$$

The general solution is a sum of standing waves:

$$y(x,t) = \sum_{n=1}^{\infty} \Re\big(\tilde{A}_n e^{i\omega_n t}\big)\sin\frac{n\pi x}{L}, \tag{14.13}$$

where $\omega_n \equiv \omega(k_n)$ in general, and $\omega_n = n\pi c/L$ for a simple string satisfying the wave equation (14.1).

As is apparent from Eq. (14.13), standing waves are the *product* of a sinusoidal t dependence and a sinusoidal x dependence. Periodically, when $\tilde{A}e^{i\omega t}$ is pure imaginary, y vanishes at all positions x. Likewise, at equally spaced *nodes* where $kx = n\pi$, y vanishes at all times t. Although there is no hint of propagation to the left or right, each standing wave actually is the sum of right- and left-propagating travelling waves having the same frequency and amplitude. All that is required to force any wave to be a standing wave is that the string be fixed at some point.

Equation (14.13) strongly resembles the coupled oscillator Eq. (12.21), which expressed the motion of a generalized coordinate as a sum over normal mode oscillations. In fact, each of the standing waves on a string is a normal mode. There are an infinite number of normal

modes because the string has an infinite number of degrees of freedom (it is equivalent to n equally spaced point masses connected by $(n+1)$ massless springs, as $n \to \infty$).

Aside from their infinite number, the normal modes of the simple string described by the wave equation (14.1) have the same properties as those of a coupled oscillator. Each is excited to a degree required by the initial conditions. Thereafter, each mode continues to oscillate without interference from any other mode. The energy given initially to that mode (on average divided equally between T and U) is retained within that mode forever.

14.4. Fourier expansion in normal modes.

We can use the initial conditions from Eq. (14.3) to evaluate the constants $\tilde{A}_n$ in Eq. (14.13):

$$\begin{aligned} y_0(x) &= \sum_n \Re(\tilde{A}_n) \sin k_n x \\ v_0(x) &= -\sum_n \omega_n \Im(\tilde{A}_n) \sin k_n x. \end{aligned} \tag{14.14}$$

Again we take advantage of *Fourier's trick*, already used in our solution (3.13) to the simple oscillator with a nonlinear driving force. For each line in Eq. (14.14), multiply both sides by $\frac{2}{L}\sin\frac{m\pi x}{L}$, $1 \le m \le \infty$, and integrate over x from 0 to L. Using the identity

$$\frac{2}{L}\int_0^L dx \sin\frac{m\pi x}{L}\sin\frac{n\pi x}{L} = \delta_{mn}, \tag{14.15}$$

all but one term in each right-hand sum vanishes, leaving

$$\begin{aligned} \Re(\tilde{A}_m) &= \frac{2}{L}\int_0^L dx \sin\frac{m\pi x}{L} y_0(x) \\ -\omega_m \Im(\tilde{A}_m) &= \frac{2}{L}\int_0^L dx \sin\frac{m\pi x}{L} v_0(x). \end{aligned} \tag{14.16}$$

14.5. Nodes and antinodes.

For the string we have been considering, *nodes* (zeroes of y) occur by definition at the fixed points $x = 0$ and $x = L$. If the string is vibrating in a single normal mode with $n > 1$, there are additional nodes at intermediate points. For example, when $n = 3$, nodes are found at $x = \frac{L}{3}$ and $\frac{2L}{3}$ as well.

For the same string, *antinodes* (zeroes of $\frac{\partial y}{\partial x}$) are found at $x = \frac{L}{6}$, $\frac{L}{2}$, and $\frac{5L}{6}$. More generally, what physical condition would produce an antinode rather than a node at the boundary $x = L$? An example would be a massive string extending from $x = 0$ to $x = L$, connected to a *massless* string with the same tension extending from $x = L$ to $x = \infty$. Here the slope of the massive string must vanish at $x = L$: otherwise a finite transverse force would be exerted on the massless string, resulting in infinite acceleration.

Since massless strings are hard to find, antinodal boundary conditions are more frequently encountered in the *longitudinal* oscillations of elastic media – for example, sound waves in air. There the phase velocity[2] is $c^2 = E/\rho$, where ρ is the mass density and E, the *elastic modulus*, is the ratio of *stress* to *strain*. (We shall have more to say later about the definition of these quantities.) For a perfect gas, c^2 is equal to the inverse of the *adiabatic compressibility*.

The classic example of an antinodal boundary condition is found in an organ pipe – the closed end is a node, the open end is (approximately) an antinode. Here the fundamental normal mode has only one quarter of its wavelength contained within the pipe, as opposed to one half wavelength for a pipe closed at both ends. The ($\omega/2\pi = 30$ Hz) introductory rumble in *Also Sprach Zarathustra* can be sustained in an open organ pipe only 2.8 m long.

14.6. Reflections of waves at boundaries.

When the boundary conditions are simple (nodal or antinodal), reflections of transverse waves on a string can be determined easily by means of the "virtual string" construction. Consider first a nodal boundary ($y \equiv 0$ at $x = L$). Recall that the wave equation is solved by a smooth shape propagating e.g. along $+\hat{x}$. Consider such a shape incident from the left ($x < L$) upon the nodal boundary.

The construction consists of hypothesizing the existence of a "virtual string" in the forbidden (stringless) region $x > L$, and imagining that the same but *inverted* shape is incident upon the nodal boundary from the right. The two shapes are timed to coincide when they reach $x = L$; the resulting cancellation satisfies the nodal boundary condition $y = 0$ there. Since the boundary condition is satisfied, we have found the correct solution. As the erect shape disappears into the virtual region, the inverted shape leaves the virtual region and appears in the physical region. There the observed waveform is the sum of (what remains of) the incident shape plus (the part of) the inverted shape which has propagated in from the virtual region. After the incident shape has disappeared completely, the reflected shape is fully inverted.

Similarly, the boundary condition $\frac{\partial y}{\partial x} = 0$ at an *antinodal* boundary is satisfied by a *non-inverted* shape on the virtual string; the shape reflected from an antinodal boundary is *erect*.

When the boundary conditions are *not* simple, the reflected and transmitted amplitudes are determined by matching the waveforms on either side of the boundary. For example, when the boundary consists of an interface between two strings of equal tension τ but different mass per unit length μ, we identify two matching conditions. First, the string's *displacement* must be the same on either side, in order for the string to remain continuous. Second, the string's *slope* must also be the same on either side of the boundary. Otherwise there would be a finite net transverse force on the infinitesimal element of string at the boundary, resulting in an infinite acceleration.

Without loss of generality, we place the origin $x = 0$ at the interface between the two strings. The matching conditions are

$$\begin{aligned} y(x{=}0^-,t) &= y(x{=}0^+,t) \\ \frac{\partial y}{\partial x}(x{=}0^-,t) &= \frac{\partial y}{\partial x}(x{=}0^+,t). \end{aligned} \tag{14.17}$$

Denote the string displacement in the region $x < 0$ by y_1 and that in the region $x > 0$ by y_2. For $x < 0$ we must allow for both an incident wave $f_1(t - x/c_1)$ and a reflected wave $g_1(t + x/c_1)$, while for $x > 0$ we need only a transmitted wave $f_2(t - x/c_2)$. Equations (14.17) become

$$\begin{aligned} f_1(t) + g_1(t) &= f_2(t) \\ -\tfrac{1}{c_1} f_1'(t) + \tfrac{1}{c_1} g_1'(t) &= -\tfrac{1}{c_2} f_2'(t). \end{aligned} \tag{14.18}$$

Now differentiate both sides of the first equality with respect to t, multiply the second equality by c_2, and add:

$$\begin{aligned} f_1'(t) + g_1'(t) &= f_2'(t) \\ \tfrac{c_2}{c_1}\big(-f_1'(t) + g_1'(t)\big) &= -f_2'(t) \\ (c_2 + c_1) g_1'(t) &= (c_2 - c_1) f_1'(t) \\ g_1(t) &= \mathcal{R} f_1(t) + \text{const} \\ g_1(t) &= \mathcal{R} f_1(t) \\ \mathcal{R} &\equiv \frac{c_2 - c_1}{c_2 + c_1}. \end{aligned} \tag{14.19}$$

In the next to last line we set the constant of integration equal to zero because it does not represent a wave. In the last line we introduced the *reflected amplitude ratio* $\mathcal{R}$.

Similarly, the transmitted wave is

$$\begin{aligned} f_2(t) &= \mathcal{T} f_1(t) \\ \mathcal{T} &\equiv \frac{2c_2}{c_1 + c_2}, \end{aligned} \tag{14.20}$$

where $\mathcal{T}$ is the *transmitted amplitude ratio*. In terms of the incident wave f_1, the full solution is

$$\begin{aligned} y_1(x,t) &= f_1(t - x/c_1) + \mathcal{R} f_1(t + x/c_1) \\ y_2(x,t) &= \mathcal{T} f_1(t - x/c_2). \end{aligned} \tag{14.21}$$

For the above case in which the string tensions τ_1 and τ_2 are equal, $c_1 \propto \mu_1^{-1/2}$ and $c_2 \propto \mu_2^{-1/2}$ with the same constant of proportionality. Defining

$$Z_1 \equiv (\mu_1 \tau_1)^{-1/2} \quad Z_2 \equiv (\mu_2 \tau_2)^{-1/2}, \tag{14.22}$$

for this case $\tau_1 = \tau_2$ it is also possible to write

$$\mathcal{R} = \frac{Z_2 - Z_1}{Z_2 + Z_1} \quad \mathcal{T} = \frac{2Z_2}{Z_2 + Z_1}. \tag{14.23}$$

Had we worked the most general problem in which the string masses per unit length and the string tensions *both* are allowed to be different on either side of the interface, we would have obtained the results (14.21) and (14.23). The quantity $Z \equiv (\mu\tau)^{-1/2}$ is called the *characteristic impedance* of the string. It is usually more illuminating to characterize a wave medium by its phase velocity c and impedance Z, rather than by the less fundamental properties μ and τ.

It is easy to identify three limiting cases of Eqs. (14.21) and (14.23). When the impedance Z_2 vanishes (right-hand string is a brick wall), $\mathcal{R} = -1$ and $\mathcal{T} = 0$. This is equivalent to the nodal boundary for which we used the virtual string construction. When Z_2 is infinite (right-hand string is a massless filament), $\mathcal{R} = +1$ and $\mathcal{T} = 2$. The transmitted amplitude is *twice* the incident amplitude. This is equivalent to the antinodal boundary. Finally, when the characteristic impedances are the same on either side of the interface, even if the tensions and masses per unit length are different there is *no reflection*.

The reflection formulae (14.21) and (14.23) carry over to waves in other media. For example, considering electromagnetic waves in uniform isotropic media, the electric field $\mathbf{E}$ plays the role of the string displacement; the characteristic impedance is the ratio of $|\mathbf{E}|$ to $|\mathbf{H}|$. In vacuum this ratio is equal to 377 ohms. For electromagnetic waves in a coaxial cable, Z is equal to the same ratio, which is equivalent to $\sqrt{L/C}$, where L and C are the inductance and capacitance per unit length. This ratio is 138 ohms $\times \ln(b/a)$, where a and b are the inner and outer cable radii. Of course, for either of these electromagnetic examples, in vacuum the phase velocity c is the speed of light. When an electromagnetic wave travels in a *refractive* medium, both the phase velocity and the characteristic impedance are reduced by the factor $1/n$, where n is the *index of refraction*.

As a final example, consider a one dimensional Schrödinger wave that is incident on a barrier V which is smaller than its energy E. Here the wave's phase velocity is *directly* proportional to its wave number $k \equiv \sqrt{2m(E-V)}/\hbar$, while the characteristic impedance is *inversely* proportional to k.

15. Mechanics of solids.

15.1. Stress, strain, and waves in a one-dimensional solid.

To set the stage for consideration of a real (three-dimensional) elastic solid, first we consider a hypothetical medium in which the molecules are able to move only in one direction. Let $u(x)$ be the difference between the actual and the equilibrium position x of a molecule. The local distortion $N \equiv \frac{\partial u}{\partial x}$ of the medium is called the *strain*.

Creating a strain requires exerting a force on the medium. Consider a plane of area ΔA with its normal along x. In this one-dimensional case, the force exerted across that plane by one element of the medium upon another is oriented also in the x direction. This force per unit area S is called the *stress*.

The *elastic modulus* E of the medium measures the size of the stress required to produce a strain: $S = EN$. Since the strain is dimensionless, both the stress and the elastic modulus have units of *pressure*. In MKS this is the N/m^2, or *pascal*; one atmosphere is $\approx 10^5$ pascals. Steel, one of the stiffer structural materials, has an elastic modulus of $\approx 2 \times 10^{11}$ pascals.

Consider a volume element $\Delta A \Delta x$ of the medium, to which a stress S is applied to create a strain N. Suppose that the molecules at $x = 0$ are held stationary while the molecules at $x = \Delta x$, already displaced from their equilibrium positions by u, are further displaced by du. Before this extra displacement, the strain is $N = \frac{u}{\Delta x}$. The work done corresponding to du is

$$\begin{aligned} dW &= F\,du = S\Delta A\,du \\ &= EN\Delta A\,du = E\tfrac{u}{\Delta x}\Delta A\,du \\ W &= \tfrac{1}{2}E\tfrac{\Delta A}{\Delta x}u^2 = \tfrac{1}{2}E\,\Delta A\Delta x\,N^2. \end{aligned}$$

Per unit volume, the potential energy U' associated with the strain is $\frac{1}{2}EN^2 = \frac{1}{2}E(\frac{\partial u}{\partial x})^2$.

The kinetic energy T' per unit volume is $\frac{1}{2}\rho(\frac{\partial u}{\partial t})^2$, where ρ is the mass density. Therefore, for this solid in which the molecules move only along x, the Lagrangian per unit volume, or Lagrangian density, is

$$\mathcal{L}' = \tfrac{1}{2}\rho\big(\frac{\partial u}{\partial t}\big)^2 - \tfrac{1}{2}E\big(\frac{\partial u}{\partial x}\big)^2. \tag{15.1}$$

This result is fully analagous to the string's Lagrangian density (13.5). The same application of the Euler-Lagrange equation (13.2) yields a wave equation for this hypothetical solid:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\rho}{E}\frac{\partial^2 u}{\partial t^2} &= 0 \\ \sqrt{\frac{E}{\rho}} &\equiv c, \end{aligned} \tag{15.2}$$

where c is the phase velocity. This is essentially the same as the wave equation for the string. However, in the solid, the molecular displacement is *along* the direction of wave propagation, as is true for a *longitudinal wave*, rather than *transverse* to that direction as for the string.

15.2. Stress and strain tensors in a solid.

All of our remaining discussion of solids parallels that of section 15.1, except that the solid is no longer hypothetical – its molecules will be allowed to move in three dimensions. This greatly complicates the mathematics, but the basic ideas remain the same.

In a solid or a flowing viscous liquid, forces $\Delta\mathbf{F}$ that act across a surface $\Delta\mathbf{A}$ can have components which are parallel as well as perpendicular to the surface. The stress $\mathcal{S}$ relating one to the other must be a *second rank tensor*:

$$\begin{aligned} \Delta\mathbf{F} &\equiv \mathcal{S}\cdot\Delta\mathbf{A} \\ (\Delta F)_i &= \mathcal{S}_{ij}\Delta A_j. \end{aligned} \tag{15.3}$$

Here the first stress index i refers to the Cartesian component of the force, while the second index j refers to the component of the *normal* to the area being considered. As usual, summation over repeated indices is implied.

The stress tensor must be *symmetric* for a static solid. This can be appreciated by considering the four x_1 or x_2 faces of a cube of material within the solid. If $\mathcal{S}_{21}$ were greater than $\mathcal{S}_{12}$, the cube would experience a net torque in the x_3 direction.

In three dimensions, the displacement $\mathbf{u}$ of each molecule from its equilibrium position is a *vector field* depending on x_1, x_2, and x_3. Here the strain must be defined to take into account all three components of $\mathbf{u}$ and of $\mathbf{x}$: it is also a second rank tensor. In analogy to the one-dimensional strain,

$$\mathcal{N}_{ij} \equiv \tfrac{1}{2}\big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\big). \tag{15.4}$$

In Eq. (15.4) we defined the strain tensor to be *manifestly symmetric*. A possible component of the form

$$\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}$$

cannot be allowed to contribute. If it were nonzero but small, such a component would correspond to an infinitesimal rotation. This could not be part of a strain, because it is not a deformation of the solid – no stress would be required to produce it.

15.3. Fourth-rank tensor of elasticity.

In the one-dimensional case, the elastic modulus was the constant of proportionality relating the (scalar) stress to the (scalar) strain. In three dimensions, both the stress and strain are symmetric tensors. Even for isotropic materials, it will turn out that the six independent elements of $\mathcal{S}$ and of $\mathcal{N}$ do not all have the same constant of proportionality to each other. Therefore the two must be related by a *fourth rank tensor, the elasticity* $\mathcal{E}$:

$$\begin{aligned} \mathcal{S} &\equiv \mathcal{E}\mathcal{N} \\ \mathcal{S}_{ij} &= \mathcal{E}_{ijkl}\mathcal{N}_{kl}. \end{aligned} \tag{15.5}$$

In analogy to the one-dimensional potential energy per unit volume,

$$U' = \tfrac{1}{2}NEN\ ,$$

the three-dimensional potential energy density is

$$\begin{aligned} U' &= \tfrac{1}{2}\mathcal{N}_{ij}\mathcal{E}_{ijkl}\mathcal{N}_{kl} \\ &= \tfrac{1}{8}\Big(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\Big)\mathcal{E}_{ijkl}\Big(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\Big). \end{aligned} \tag{15.6}$$

The kinetic energy density is more straightforward:

$$\mathcal{T}' = \tfrac{1}{2}\rho\frac{\partial u_i}{\partial t}\frac{\partial u_i}{\partial t}. \tag{15.7}$$

As for the one-dimensional case, the Lagrangian density $\mathcal{L}'$ is $T' - U'$.

15.4. Elasticity in a homogeneous isotropic solid.

A homogeneous isotropic solid, unlike any crystalline structure, has a fourth-rank tensor of elasticity with components that, by definition, are independent of particular axis directions. The only available building block, aside from scalars, is the unit matrix I with components δ_{ij}. The most general fourth rank tensor that we are able to construct is

$$\mathcal{E}_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{15.8}$$

The last term is manifestly symmetric under interchange of i and j in order to ensure that $\mathcal{S}$ is symmetric. Because of the homogeneity and isotropy, the two values λ and μ, called the *Lamé constants*, are sufficient to determine all 81 elements of $\mathcal{E}$.

Evaluating the stress,

$$\begin{aligned} \mathcal{S}_{ij} &= \mathcal{E}_{ijkl}\mathcal{N}_{kl} \\ &= \lambda\delta_{ij}\delta_{kl}\mathcal{N}_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\mathcal{N}_{kl} \\ &= \lambda\delta_{ij}\,\mathrm{tr}\,\mathcal{N} + 2\mu\mathcal{N}_{ij}. \end{aligned} \tag{15.9}$$

For example, comparing the first and third equalities in Eq. (15.9), one has

$$\begin{aligned} \mathcal{S}_{11} &= \lambda(\mathcal{N}_{11} + \mathcal{N}_{22} + \mathcal{N}_{33}) + 2\mu\mathcal{N}_{11} \\ \Rightarrow \mathcal{E}_{1111} &= \lambda + 2\mu \\ \Rightarrow \mathcal{E}_{1122} &= \lambda \\ \mathcal{S}_{12} &= 2\mu\mathcal{N}_{12} = \mathcal{E}_{1212}\mathcal{N}_{12} + \mathcal{E}_{1221}\mathcal{N}_{21} \\ \Rightarrow \mathcal{E}_{1212} &= \mathcal{E}_{1221} = \mu. \end{aligned} \tag{15.10}$$

It is instructive to visualize the distortion that, for example, is controlled by the elasticity element $\mathcal{E}_{1111}$. Consider a solid cube of volume l^3 with faces normal to $\hat{x}_1$, $\hat{x}_2$, or $\hat{x}_3$. Suppose that the the cube is stretched in the $\hat{x}_1$ direction, so that the *diagonal* strain element $\mathcal{N}_{11}$ is positive. This stretching is caused by a stress $\mathcal{S}_{11}$. Further assume that *no* other distortion is present: *all* the other elements of $\mathcal{N}$ vanish. This means that *additional* stresses must be exerted in order to prevent the cube from shrinking in the $\hat{x}_2$ and $\hat{x}_3$ directions, as would be its natural tendency. In other words, forces must be exerted to hold the sides $x_2 =$ constant and $x_3 =$ constant at their equilibrium separation l. The fact that $\mathcal{E}_{1111} = \lambda + 2\mu$ (Eq. (15.10)) means that the effective elastic modulus Y_{eff} for this particular stretching mode is equal to $\lambda + 2\mu$.

Similarly, for the same cube, consider a *shear displacement* of the four faces normal to $\hat{x}_1$ or $\hat{x}_2$ so that the face normal to $\hat{x}_3$ changes in shape from a square to a rhombus. The four sides of the rhombus make angles $\pm\Delta\phi/2$ with the edges of the original square. This describes a distortion corresponding to nonzero *off-diagonal* strain elements $\mathcal{N}_{12} = \mathcal{N}_{21} = \Delta\phi/2$. All other possible distortions (and strain elements) are assumed to vanish. Because $\mathcal{E}_{1212} = 2\mu$, the stress $\mathcal{S}_{12} = \mathcal{S}_{21}$ which must be applied to create this shear displacement is

$$\begin{aligned} \mathcal{S}_{12} &= \mathcal{E}_{1212}\mathcal{N}_{12} + \mathcal{E}_{1221}\mathcal{N}_{21} \\ &= \mu\frac{\Delta\phi}{2} + \mu\frac{\Delta\phi}{2} = \mu\Delta\phi. \end{aligned}$$

15.5. Young's modulus and Poisson's ratio.

The strains and stresses described in the last two paragraphs may be straightforward to visualize, but they are not the easiest quantities to measure. A more practical approach to determining the Lamé constants is to experiment with a solid rectangular bar, for example having relaxed square cross section w^2 and length l. One such experiment measures *Young's modulus* Y, the ratio of the pressure F/w^2 with which the ends are pulled apart to the fractional

increase $\Delta l/l$ in the bar's length. Another measures *Poisson's ratio* σ, the (negative) ratio of the fractional change in w to that in l for the same experiment:

$$\begin{aligned} Y &\equiv \frac{F/w^2}{\Delta l/l} \\ \sigma &\equiv -\frac{\Delta w/w}{\Delta l/l}. \end{aligned} \tag{15.11}$$

For both experiments, *no* constraint is imposed on the side walls of the bar.

The price paid for simplifying these measurements is the task of finding λ and μ, given Y and σ. Taking $\hat{x}_1$ to be the long axis of the bar, we know that the only stress applied is $\mathcal{S}_{11}$. Only the diagonal elements of the strain are nonzero: since the bar is not twisted, by symmetry the off-diagonal elements must vanish. By definition of the Poisson ratio σ, $\mathcal{N}_{22} = \mathcal{N}_{33} \equiv -\sigma\mathcal{N}_{11}$. And by definition of Young's modulus Y, $\mathcal{S}_{11} \equiv Y\mathcal{N}_{11}$. Starting from Eq. (15.9),

$$\begin{aligned} \mathcal{S}_{11} &= \lambda(\mathcal{N}_{11} + \mathcal{N}_{22} + \mathcal{N}_{33}) + 2\mu\mathcal{N}_{11} \\ Y &= \lambda(1 - \sigma - \sigma) + 2\mu \\ \mathcal{S}_{22} &= \lambda(\mathcal{N}_{11} + \mathcal{N}_{22} + \mathcal{N}_{33}) + 2\mu\mathcal{N}_{22} \\ 0 &= \lambda(1 - \sigma - \sigma) - 2\mu\sigma. \end{aligned}$$

Subtracting the fourth from the second equality,

$$\begin{aligned} Y &= 2\mu(1 + \sigma) \\ \mu &= \frac{Y}{2(1 + \sigma)}. \end{aligned} \tag{15.12}$$

With Eq. (15.12), the second equality alone yields

$$\begin{aligned} \lambda &= \frac{2\mu\sigma}{1 - 2\sigma} \\ &= \frac{\sigma Y}{(1 + \sigma)(1 - 2\sigma)}. \end{aligned} \tag{15.13}$$

Equations (15.12) and (15.13) express the Lamé constants in terms of Young's modulus and Poisson's ratio. The inverse equations are

$$\begin{aligned} \sigma &= \frac{\lambda}{2(\lambda + \mu)} \\ Y &= \mu\frac{3\lambda + 2\mu}{\lambda + \mu}. \end{aligned} \tag{15.14}$$

Note that $0 \le \sigma \le \frac{1}{2}$. The upper limit is obtained either if $\lambda = \infty$ (medium is incompressible) or if $\mu = 0$ (medium cannot support a shear stress). Mechanical engineers often *assume* that $\sigma = \frac{1}{2}$ when better data are not available. If the medium is incompressible, Young's modulus is *not* infinite; shear displacement still allows the bar to elongate, with modulus $Y = 3\mu$.

15.6. Waves in solids.

So far we have been discussing the *statics* of three-dimensional solids. As an introduction to the *dynamics*, we consider wave propagation in a solid medium. As was the case for waves on a string, we neglect the damping forces that cause the wave eventually to die out. For a cast bell with $Q \approx 10^3$, able to ring for seconds at a frequency of hundreds of Hz, this is a good approximation.

The Lagrangian density is given by the difference of Eqs. (15.7) and (15.6). Since the elasticity tensor $\mathcal{E}_{ijkl}$ is symmetric under the interchanges $i \leftrightarrow j$ and $k \leftrightarrow l$ for a solid that is homogeneous and isotropic, we may replace $\mathcal{N}_{ij}$ and $\mathcal{N}_{kl}$ in Eq. (15.6) by $\partial u_i/\partial x_j$ and $\partial u_k/\partial x_l$. The potential energy density becomes

$$\begin{aligned} U' &= \tfrac{1}{2}\frac{\partial u_i}{\partial x_j}\big(\lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\big)\frac{\partial u_k}{\partial x_l} \\ &= \frac{\lambda}{2}\frac{\partial u_i}{\partial x_i}\frac{\partial u_j}{\partial x_j} + \frac{\mu}{2}\big(\frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j} + \frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\big). \end{aligned} \tag{15.15}$$

The Lagrangian density is independent of $\mathbf{u}$. Since its dependence on $\partial u/\partial t$ is confined to T', and its dependence upon $\partial u/\partial x$ is confined to U', the Euler-Lagrange equation in the nth component of $\mathbf{u}$ becomes

$$\frac{d}{dt}\frac{\partial T}{\partial(\frac{\partial u_n}{\partial t})} = \frac{d}{dx_k}\frac{\partial U}{\partial(\frac{\partial u_n}{\partial x_k})}. \tag{15.16}$$

Using Eq. (15.7), the left hand side is

$$\frac{d}{dt}\big(\rho(\mathbf{x}, t)\frac{\partial u_n}{\partial t}\big) = \rho\frac{\partial^2 u_n}{\partial t^2} + \frac{\partial\rho}{\partial t}\frac{\partial u_n}{\partial t}.$$

Since the time-varying component of ρ is assumed to be a small fraction of its average, the second term is negligible compared to the first.

The right hand side of Eq. (15.16) is

$$\begin{aligned}
&\lambda\frac{d}{dx_n}\Big(\frac{\partial u_j}{\partial x_j}\Big)+\mu\frac{d}{dx_k}\Big(\frac{\partial u_n}{\partial x_k}+\frac{\partial u_k}{\partial x_n}\Big)\\
=&\lambda\frac{\partial^2 u_j}{\partial x_n\partial x_j}+\mu\Big(\frac{\partial^2 u_n}{\partial x_k\partial x_k}+\frac{\partial^2 u_k}{\partial x_k\partial x_n}\Big)\\
=&(\lambda+\mu)\frac{\partial}{\partial x_n}\frac{\partial u_k}{\partial x_k}+\mu\frac{\partial^2 u_n}{\partial x_k\partial x_k}\\
=&(\lambda+\mu)\frac{\partial}{\partial x_n}(\nabla\cdot\mathbf{u})+\mu\nabla^2 u_n.
\end{aligned}$$

Considered together, the Euler-Lagrange equations for $n = 1$, 2, and 3 are equivalent to the vector equation

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u})+\mu\nabla^2\mathbf{u}. \qquad (15.17)$$

Suppose first that the displacement field $\mathbf{u}$ is *divergenceless*, so that there can be no compression. Then any wave still present is a *shear wave*. Equation (15.17) becomes

$$\mu\nabla^2\mathbf{u}-\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=0, \qquad (15.18)$$

satisfied by a shear wave with phase velocity $\sqrt{\mu/\rho}$.

Conversely, suppose that $\mathbf{u}$ is *curlless* so that there can be no shear. Using the "*bac cab*" rule,

$$\begin{aligned}
0&=\nabla\times(\nabla\times\mathbf{u})\\
&=\nabla(\nabla\cdot\mathbf{u})-\nabla^2\mathbf{u},
\end{aligned}$$

Eq. (15.17) becomes

$$(\lambda+2\mu)\nabla^2\mathbf{u}-\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=0. \qquad (15.19)$$

This describes a *compression wave* with the larger phase velocity $\sqrt{(\lambda+2\mu)/\rho}$. An earthquake including both compression and shear waves might be felt first as a sharp compressive jolt, followed by a rolling shear motion.

16. Mechanics of fluids.

16.1. Static fluids.

A fluid (liquid or gas) differs from a solid in that it cannot support a shear stress if it is static (i.e. if the fluid is not moving). In terms of the Lamé constants, $\mu\equiv 0$, so that Young's modulus $Y = 0$ and Poisson's ratio $\sigma=\frac{1}{2}$.

Consider an infinitesimal cube of side l with one corner at the origin. The $\hat{x}_1$ component of the force $\mathbf{F}$ exerted on the cube is

$$\begin{aligned}
F_1=l^2\big\{&\big(\mathcal{S}_{11}(l,x_2,x_3)-\mathcal{S}_{11}(0,x_2,x_3)\big)+\\
&+\big(\mathcal{S}_{12}(x_1,l,x_3)-\mathcal{S}_{12}(x_1,0,x_3)\big)+\\
&+\big(\mathcal{S}_{13}(x_1,x_2,l)-\mathcal{S}_{13}(x_1,x_2,0)\big)\big\}\\
&=l^3\frac{\partial}{\partial x_j}\mathcal{S}_{1j}.
\end{aligned}$$

Defining $\mathbf{f}$ to be the force per unit volume,

$$\begin{aligned}
f_i&=\frac{\partial}{\partial x_j}\mathcal{S}_{ij}\\
\tilde{f}^t&=\tilde{\partial}^t\mathcal{S},
\end{aligned} \qquad (16.1)$$

where $\tilde{\partial}^t$ is a row vector with elements equal to $\partial/\partial x_j$, and, as usual, a sum is taken over the repeated index j.

For a static fluid, the stress tensor has no off-diagonal (shear) components. Also, the fluid's homogeneity requires the diagonal elements to be equal. Then

$$\mathcal{S}\equiv -p\,\mathrm{I},$$

where p is the *pressure* and I is the unit matrix. Equation (16.1) reduces to

$$\mathbf{f}=-\nabla p.$$

We add a (conservative) external force $-\rho\nabla\phi$ derived from ϕ, a *potential per unit mass*. Since ρ is the mass per unit volume, this additional term is also a force per unit volume. Insisting that the total force on any static element of fluid vanish, we obtain the basic equation of fluid statics

$$0=\mathbf{f}=-\nabla p-\rho\nabla\phi. \qquad (16.2)$$

The two most common applications of Eq. (16.2) involve an external force due to gravity, for which the gravitational potential is $\phi = gz$, where z is the vertical coordinate. For an *incompressible* fluid (ρ = constant), integrating Eq. (16.2),

$$\begin{aligned} 0 &= -\nabla p - \rho\nabla(gz) \\ -p_0 &= -p - \rho g(z - z_0) \\ p &= p_0 - \rho g(z - z_0). \end{aligned}$$

This leads directly to *Archimedes' Principle*, which states that the buoyant force on a (partially or totally) submerged object is the weight of the water that it displaces.

For a *perfect gas* at constant temperature, $\rho = \rho_0 p/p_0$. Equation (16.2) yields

$$\begin{aligned} 0 &= -\nabla p - \frac{\rho_0 p}{p_0}\nabla(gz) \\ \frac{\partial p/\partial z}{p} &= -\frac{\rho_0 g}{p_0} \\ p &= p_0 e^{-\rho_0 g(z-z_0)/p_0}. \end{aligned}$$

Applying this result to the earth's atmosphere, where $\rho_0 \approx 1$ kg/m^3, $p_0 \approx 10^5$ pascals, and $g \approx 10$ m/sec^2, we estimate that the atmosphere's pressure is reduced by a factor e after an elevation gain of $p_0/\rho_0 g \approx 10^4$ m.

16.2. Flow of a nonviscous fluid.

Fluids that are not static *do* support a shear stress. Temporarily we choose to neglect this fact, focusing on "dry" or runny liquids. Now that it accelerates the fluid, we no longer require the force per unit volume in Eq. (16.2) to vanish:

$$\begin{aligned} -\nabla p - \rho\nabla\phi &= \rho\frac{d\mathbf{v}}{dt} \\ -\frac{\nabla p}{\rho} - \nabla\phi &= \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}. \end{aligned} \tag{16.3}$$

This is the *Navier-Stokes equation*. On the right hand side of the last equality we have written the total time derivative of $\mathbf{v}$ as a *convective derivative*. The first term describes the explicit change of velocity with time ("Spring approaches and the river flows ever faster"). The second term describes the change in $\mathbf{v}$ due to flow of the fluid ("water moves from a pool to the rapids").

We are faced with two unknown scalar fields (p and ρ) and one unknown vector field ($\mathbf{v}$), but we have derived only one vector equation (16.3). We need two additional scalar equations. The first is an *equation of state* relating ρ to p, as in the two examples of section (16.1). The second is the *equation of continuity* expressing the conservation of fluid molecules:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{16.4}$$

16.3. *Steady* flow of an *incompressible* nonviscous fluid.

Under the assumption of *steady flow*, the fluid can be moving, but its velocity field, and every other characteristic of the fluid, possesses no *explicit* time dependence. That is, $\partial/\partial t$ of *anything* vanishes. Also in this section we assume that the fluid is incompressible (ρ = constant). The Navier-Stokes equation (16.3) becomes

$$-\nabla\big(\frac{p}{\rho} + \phi\big) = (\mathbf{v}\cdot\nabla)\mathbf{v}.$$

Using the vector identity

$$\tfrac{1}{2}\nabla(\mathbf{v}\cdot\mathbf{v}) = (\mathbf{v}\cdot\nabla)\mathbf{v} + \mathbf{v}\times(\nabla\times\mathbf{v}),$$

the above equation may be rewritten as

$$\nabla\big(\frac{p}{\rho} + \phi + \frac{v^2}{2}\big) = \mathbf{v}\times(\nabla\times\mathbf{v}). \tag{16.5}$$

Taking the dot product of $\mathbf{v}$ with Eq. (16.5),

$$\begin{aligned} 0 &= (\mathbf{v}\cdot\nabla)\big(\frac{p}{\rho} + \phi + \frac{v^2}{2}\big) \\ &= \big(\frac{d}{dt} - \frac{\partial}{\partial t}\big)\big(\frac{p}{\rho} + \phi + \frac{v^2}{2}\big) \\ &= \frac{d}{dt}\big(\frac{p}{\rho} + \phi + \frac{v^2}{2}\big). \end{aligned} \tag{16.6}$$

This is the first form of *Bernoulli's equation*. It says that the quantity $p/\rho + \phi + v^2/2$ for an element of fluid is constant as the fluid moves – it

is *constant along streamlines*. Bernoulli's equation is an expression of *energy conservation* for the fluid, with p/ρ and ϕ the *potential* terms, and $\frac{1}{2}v^2$ the *kinetic*. It is responsible for the lift on airplane wings and the operation of spray nozzles.

16.3. Steady *irrotational* flow of an incompressible nonviscous fluid.

In this section we make the still stronger assumption that the flow is *irrotational*, $\nabla \times \mathbf{v} \equiv 0$. Then the right-hand side of Eq. (16.5) vanishes altogether, and

$$\frac{p}{\rho} + \phi + \frac{v^2}{2} = \text{constant} \tag{16.7}$$

throughout the entire fluid. This is the second form of Bernoulli's equation.

With this strong set of assumptions, we can solve for $\mathbf{v}$ without even considering the Navier-Stokes equation. For steady flow the first term in the continuity equation (16.4) vanishes, and for an incompressible fluid the second term reduces to

$$\nabla \cdot \mathbf{v} = 0. \tag{16.8}$$

If the flow is irrotational, we may write $\mathbf{v} = -\nabla\xi$, where ξ is the *velocity potential*. Equation (16.8) is equivalent to *Laplace's equation*

$$\nabla^2 \xi = 0. \tag{16.9}$$

Solving Laplace's equation is a standard problem in applied mathematics. A unique solution exists whenever ξ or its normal derivative is specified over an entire closed surface, part of which may be at infinity. For fluid flow problems, it is more common to specify the normal derivative of ξ (velocity normal to the boundary) than it is to specify ξ itself. Analytic solutions to Laplace's equation may be obtained by *series expansions* involving harmonic and hyperbolic functions in Cartesian coordinates, Bessel and Neumann functions in cylindrical coordinates, or spherical harmonics in spherical coordinates. For solutions of Laplace's equations in two dimensions, *conformal transformations* are useful. Numerical solutions to Laplace's equation can be obtained by a variety of methods. A simple procedure is to set up a square or cubic grid and demand (by iterative solution) that ξ at each grid point be equal to the average of its nearest neighbors.

16.5. Flow of a viscous fluid.

At last we no longer neglect the shear stress in a flowing fluid. Consider two parallel planes of fluid with area A separated by Δz along their normal. Suppose that the top plane is moving with velocity $\Delta\mathbf{v} = \Delta v_y \hat{y}$ while the bottom plane is stationary. The *first coefficient of viscosity* η is defined by

$$\frac{F}{A} \equiv \eta \frac{\Delta v_y}{\Delta z}, \tag{16.10}$$

where F is the force in the y direction exerted by the top plane on the bottom plane. Identifying F/A with a stress $\mathcal{S}_{yz}$, for this simple example Eq. (16.10) may be written

$$\mathcal{S}_{yz} = \eta \frac{\partial v_y}{\partial z}.$$

Symmetrizing,

$$\mathcal{S}_{ij} \equiv \eta\Big(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\Big). \tag{16.11}$$

This is the formal definition of η for an *incompressible* fluid, in which viscous forces are caused only by shear flow.

Considering once more the strain $\mathcal{N}$ defined by Eq. (15.4), it is clear that Eq. (16.11) may alternatively be written

$$\mathcal{S}_{ij} = 2\eta \frac{\partial \mathcal{N}_{ij}}{\partial t}. \tag{16.12}$$

What happens if the fluid is both viscous and compressible? Resistance to rapid compression will be exerted as a result of its viscosity, in addition to that from its Lamé constant λ. Equation (16.12) must be extended:

$$\mathcal{S}_{ij} = 2\eta \frac{\partial \mathcal{N}_{ij}}{\partial t} + \eta' \delta_{ij} \frac{\partial}{\partial t} \mathrm{tr}\, \mathcal{N}, \tag{16.13}$$

where η' is the *second coefficient of viscosity*. Notice the similarity between Eq. (15.9) for a

homogeneous isotropic solid and Eq. (16.13) for a compressible liquid: apart from redefinition of constants, the right hand side of (16.13) is just the partial time derivative of that in (15.9). Note also that the coefficient of $\eta'\delta_{ij}$ in (16.13) is just $\nabla\cdot\mathbf{v}$, which vanishes for an incompressible fluid according to Eq. (16.8).

The above are the mathematical definitions. The physical picture is that thermal motion of molecules causes them to be exchanged between the strata of relatively moving fluid. These exchanged molecules retain their original average velocity in the direction of slippage, tending to equalize the velocities of the strata. One analogy is that of two open coal trains passing each other with workers shoveling coal back and forth.

In a perfect gas at fixed temperature, η is independent of pressure p. This is because the mean free path is inversely proportional to p, while the density is directly proportional to p; the viscosity is proportional to their product.

16.6. Vorticity.

The force $\mathbf{f}$ per unit volume on a viscous fluid is the usual term $-\nabla p-\rho\nabla\phi$ from Eq. (16.2) plus a term $\mathbf{f}^{\rm v}$ from the viscosity. Combining Eqs. (16.1) and (16.13), and neglecting possible spatial variation of the viscosity,

$$\begin{aligned} f_i^{\rm v} &= \frac{\partial}{\partial x_j}\big(\eta(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i})+\eta'\delta_{ij}(\nabla\cdot\mathbf{v})\big)\\ &= \eta\frac{\partial}{\partial x_j}(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i})+\eta'\frac{\partial}{\partial x_i}(\nabla\cdot\mathbf{v})\\ &= \eta\nabla^2 v_i+(\eta+\eta')\frac{\partial}{\partial x_i}(\nabla\cdot\mathbf{v})\\ \mathbf{f}^{\rm v} &= \eta\nabla^2\mathbf{v}+(\eta+\eta')\nabla(\nabla\cdot\mathbf{v}). \end{aligned} \tag{16.14}$$

Adding $\mathbf{f}^{\rm v}$ to the Navier-Stokes equation,

$$\begin{aligned} -\frac{\nabla p}{\rho}-\nabla\phi+\frac{\eta}{\rho}\nabla^2\mathbf{v}+\frac{\eta+\eta'}{\rho}\nabla(\nabla\cdot\mathbf{v}) &= \\ = \frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}. \end{aligned} \tag{16.15}$$

Exploiting the same vector identity used to prove Eq. (16.5), the last term in Eq. (16.15) may be rewritten as $\frac{1}{2}\nabla v^2-\mathbf{v}\times(\nabla\times\mathbf{v})$. Our plan is take the *curl* of Eq. (16.15) after this modification. All the terms proportional to gradients will vanish. Again assuming that the fluid is incompressible so that ρ is constant, we are left with the curl of

$$\frac{\eta}{\rho}\nabla^2\mathbf{v}=\frac{\partial\mathbf{v}}{\partial t}-\mathbf{v}\times(\nabla\times\mathbf{v}). \tag{16.16}$$

Taking said curl, and defining the *vorticity* $\vec{\Omega}\equiv\nabla\times\mathbf{v}$, we obtain

$$\begin{aligned} \frac{\eta}{\rho}\nabla^2\vec{\Omega} &= \frac{\partial\vec{\Omega}}{\partial t}-\nabla\times(\mathbf{v}\times\vec{\Omega})\\ &= \frac{\partial\vec{\Omega}}{\partial t}+\nabla\times(\vec{\Omega}\times\mathbf{v}). \end{aligned} \tag{16.17}$$

In a nonviscous fluid for which $\eta=0$, the left hand side vanishes. It will be seen that the right hand side can lead to *persistent vortices* in the fluid.

16.7. Diffusion of the vorticity: Reynolds number.

Invoking a similar vector identity,

$$\begin{aligned} \nabla\times(\vec{\Omega}\times\mathbf{v}) = (\mathbf{v}\cdot\nabla)\vec{\Omega}-(\vec{\Omega}\cdot\nabla)\mathbf{v}+\\ +\vec{\Omega}(\nabla\cdot\mathbf{v})-\mathbf{v}(\nabla\cdot\vec{\Omega}), \end{aligned}$$

we see that the last term in the identity vanishes by definition of $\vec{\Omega}$, and the second last term vanishes according to the continuity equation (16.4) when the fluid is incompressible. Equation (16.17) becomes

$$\begin{aligned} \frac{\eta}{\rho}\nabla^2\vec{\Omega} &= \frac{\partial\vec{\Omega}}{\partial t}+(\mathbf{v}\cdot\nabla)\vec{\Omega}-(\vec{\Omega}\cdot\nabla)\mathbf{v}\\ &= \frac{d\vec{\Omega}}{dt}-(\vec{\Omega}\cdot\nabla)\mathbf{v}\\ \frac{d\vec{\Omega}}{dt} &= \frac{\eta}{\rho}\nabla^2\vec{\Omega}+(\vec{\Omega}\cdot\nabla)\mathbf{v}. \end{aligned} \tag{16.18}$$

In some cases the velocity has no spatial dependence along the direction of the vorticity. For example, $\mathbf{v}$ may be oriented in the x-y plane and may depend only upon x and y, in which case $\vec{\Omega}$ is oriented along $\hat{z}$. In these special instances, Eq. (16.18) reduces to

$$\frac{d\vec{\Omega}}{dt}=\frac{\eta}{\rho}\nabla^2\vec{\Omega}. \tag{16.19}$$

This is a *diffusion equation* for each component of the vorticity, in which the diffusion constant $D \equiv \eta/\rho$. Unlike more traditional applications of this equation, e.g. the diffusion of impurity gas molecules, Eq. (16.19) describes the diffusion of *vorticity*. The small whirlpool formed when a rower's oar is lifted from the water of a still pond diffuses outward into a larger, slower vortex and eventually disappears. On the other hand, Eq. (16.19) requires the vorticity to *persist* if η/ρ may be neglected.

Since η/ρ has units of $\mathrm{m}^2/\mathrm{sec}$, we must construct another quantity of the same dimensions with which to compare it. Consider a pipe of circular cross section with diameter d, in which fluid flows with average velocity V. Then Vd has the same dimensions as η/ρ. The ratio of the two is called the *Reynolds number* R_e:

$$R_e \equiv \frac{\rho V d}{\eta}. \tag{16.20}$$

The Reynolds number is large when η/ρ is small, and vice versa.

With the help of Eq. (16.19), we can guess what happens in the pipe. If R_e is small, the velocity $\eta/\rho d$ characterizing diffusion of the vorticity is much faster than the flow velocity. If a small vortex is present, it spreads out quickly along the pipe and dissipates. The absence of vortices results in *laminar flow*, for which the velocity field is easy to calculate. In smooth pipes, laminar flow occurs for Reynolds numbers up to surprisingly large values, of order a few hundred. But when R_e is large (above 10,000), an element of fluid flows so fast that the vorticity cannot spread out rapidly enough to escape it. This is *turbulent flow*. It is so difficult to model that research in this area is a frontier of mathematical physics. For example, it is virtually impossible to calculate and predict the eddies that cause a flag to flutter.

– End –