# Level 2 Classical Mechanics 2019

V. Eke (Dept. of Physics, Durham University)

## Contents

**1 Synopsis** **1**

**2 Motivation** **4**

**3 Generalised Description of Mechanical Systems** **5**

**4 The Lagrangian** **7**

**5 Variational Calculus** **11**

**6 Linear Oscillators** **14**

**7 Driven Oscillators** **16**

**8 Coupled Small Oscillations** **19**

**9 Normal Modes** **22**

**10 Central Forces** **24**

**11 Gravitational Attraction** **27**

**12 Noether's Theorem and Hamiltonian Mechanics** **30**

**13 Canonical Transformations and Poisson Brackets** **33**

**14 Rotating Reference Frames** **37**

**15 Inertial Forces on Earth** **40**

**16 Rotational Inertia, Angular Momentum and Kinetic Energy** **43**

**17 The Parallel and Principal Axis Theorems** **47**

**18 Rigid Body Dynamics and Stability** **50**

WORKSHOP 1: APPLICATIONS OF THE LAGRANGIAN APPROACH **52**

WORKSHOP 2: APPLICATIONS OF VARIATIONAL METHODS **54**

WORKSHOP 3: TOSSING THE CABER **55**

WORKSHOP 4: PARTICLES ON A STRETCHED STRING **56**

WORKSHOP 5: CENTRAL FORCES **57**

WORKSHOP 6: HAMILTONIAN MECHANICS **58**

WORKSHOP 7: CANONICAL TRANSFORMATION TO A ROTATING FRAME **59**

WORKSHOP 8: INERTIAL FORCES **60**

WORKSHOP 9: RIGID BODY MOTION **61**

**A Second-Order Ordinary Linear Differential Equations with Constant Coefficients I**
A.1 Definition . . . . . I
A.2 Solution . . . . . I
A.2.1 Auxiliary Equation . . . . . I
A.2.2 General Solution (Roots not Repeated) . . . . . I
A.3 Possible Complications . . . . . II
A.3.1 Repeated Roots . . . . . II
A.3.2 Additional Constant . . . . . II
A.4 Representing Position in a Plane by a Single Complex Number . . . . . II

**B Taylor Series III**
B.1 Taylor's Theorem of the Mean . . . . . III
B.2 Taylor Series in One Variable . . . . . III
B.3 Taylor Series in Two Variables . . . . . IV
B.4 Taylor Series in Three Variables . . . . . IV

**C Determinants, Eigenvalues and Eigenvectors V**
C.1 Symmetric Matrices . . . . . V
C.2 Determinants . . . . . V
C.3 The Conventional Eigenvalue Problem . . . . . V
C.4 Generalized Eigenvalue Problem . . . . . VII

# 1 Synopsis

## Description of Mechanical Systems

SEE ALSO HAND AND FINCH, 1.1–1.5
Dynamical variables. Degrees of freedom. Constraints. Generalized coordinates.

## The Lagrangian Approach

SEE ALSO HAND AND FINCH, 1.6–1.11
D'Alembert's principle. Generalised equations of motion. The Lagrangian. Ignorable coordinates.

## Variational Calculus

SEE ALSO HAND AND FINCH, 2.1–2.8
Euler equation. Hamilton's principle. Lagrange multipliers and constraints.

## Linear Oscillators

SEE ALSO HAND AND FINCH, 3.1–3.3
Equilibrium. Lagrangian near to static equilibrium. Simple harmonic oscillator. Damped oscillator and the quality factor, Q.

## Driven Oscillators

SEE ALSO HAND AND FINCH, 3.4–3.8
Equation of motion. Dirac $\delta$ function. Impulsive driving force. Green's function, G. General driving force. Resonance.

## Coupled Small Oscillations

SEE ALSO HAND AND FINCH, 9.1–9.2
Two Coupled Pendulums. Problem solving recipe. Matrix form of L.

## Normal Modes

SEE ALSO HAND AND FINCH, 9.3–9.5
Normal coordinates. Mode orthogonality. Linear triatomic molecule.

## Central Forces

SEE ALSO HAND AND FINCH, 4.3–4.4
Two interacting bodies. Translational symmetry and conservation of momentum. Ignorable centre of mass motion. Rotational invariance and conservation of angular momentum. Kepler's 2nd law. Equivalent 1D problem.

## Gravitational Attraction

SEE ALSO HAND AND FINCH, 4.1, 4.5–4.6
Solving 1D systems by quadrature. Gravitational potential and the $u = 1/r$ transformation. Kepler's 1st and 3rd laws.

## Noether's Theorem and Hamiltonian Mechanics

SEE ALSO HAND AND FINCH, 5.1–5.5
Invariance of $L$ under continuous transformations. Noether's theorem. Legendre transformations. The Hamiltonian. Hamilton's equations of motion.

## Canonical Transformations and Poisson Brackets

SEE ALSO HAND AND FINCH, 6.1–6.3
Canonical Transformations. Equivalence of Lagrangians. Generating functions. The transformed Hamiltonian. Poisson brackets.

## Rotating Reference Frames

SEE ALSO HAND AND FINCH, 7.1–7.7
Accelerating reference frames. Rotated reference frames in 2D. Small rotations in 3D. Velocity and acceleration. Inertial forces in a rotating frame.

## Inertial Forces on Earth

SEE ALSO HAND AND FINCH, 7.7–7.10
Procedure for determining a local coordinate system. Inertial forces. Centrifugal force. Effective gravity on Earth's surface. Coriolis Force. Euler force.

## Rotational Inertia, Angular Momentum and Kinetic Energy

SEE ALSO HAND AND FINCH, 8.1–8.3
Displacement of a rigid body. Moment of inertia. Angular momentum and rotational kinetic energy. The inertia tensor. Generalized angular momentum and rotational kinetic energy.

## The Parallel and Principal Axis Theorems

SEE ALSO HAND AND FINCH, 8.2,8.7
Displaced axis theorem. Parallel axis theorem. Symmetry of the inertia tensor. Principal axis transformation. Principal axis theorem.

## Rigid Body Dynamics and Stability

SEE ALSO HAND AND FINCH, 8.4–8.5
Euler's equations of motion. Torque-free motion.

## Books

**Hand & Finch**, *Analytical Mechanics*, CUP. Appropriate level and style.
**Thornton & Marion**, *Classical Dynamics of Particles and Systems*, Thomson. Second half of book pretty much covers this course.
**Kibble & Berkshire**, *Classical Mechanics*, Imperial College Press. Covers more than needed for this course. Sometimes terse treatment. Lots of questions and answers.
**Brizard**, *An Introduction to Lagrangian Mechanics*, World Scientific. Quite a mathematical approach. Covers almost all of the course, and not much else.
**Goldstein**, *Classical Mechanics*, Addison-Wesley. More formal treatment. Goes far beyond this course.

# 2 Motivation

## Newton's formulation

Formally, the system comprises $N$ point particles. For particle $i$,

$$m_i \ddot{\underline{r}}_i = \underline{F}_i(\underline{r}_i, \underline{r}_j, ..., \dot{\underline{r}}_i, \dot{\underline{r}}_j, ..., t).$$

The scheme is conceptually simple if $\underline{F}_i$ is explicitly known, e.g. gravitation, Lorentz forces. Microscopic systems are 'simple' in this sense (provided that classical physics still applies at the relevant scales). Macroscopic systems include forces due to constraints. All Newton says about such forces is that

$$\underline{F}_i \text{ due to } j = -\underline{F}_j \text{ due to } i.$$

## Features of this scheme

1. Valid in an inertial (non-accelerating) frame for objects moving with speeds much less than that of light and sizes much larger than those of atoms.

2. Conceptually simple, with direct solutions in many elementary cases.

3. Totally deterministic in principle, like any formulation of classical mechanics (but might need to know initial conditions with infinite precision).

4. Constraints need to be handled explicitly.

5. Conservation of energy, momentum and angular momentum are *derived*, and thus not *fundamental* within this approach. For electromagnetism, a common approach is to take Maxwell's Equations and the Lorentz Force as fundamental. Then conservation of energy follows from Poynting's theorem and a page of algebra. Surely energy and momentum are more central than this?

6. Classical physics is only approximate and applies to the limiting case of large bodies and low velocities. While force is central to the Newtonian approach, it doesn't even appear in Schrödinger's equation.

The aim of the present course is to explore the alternative formulations of classical mechanics of Lagrange and Hamilton. These are important in the development of modern physics (e.g. quantum mechanics, chaotic behaviour, ...). Each formulation is a new set of 'first principles'; provided they predict the correct behaviour, they need not be derived from any other set of first principles.

# 3 Generalised Description of Mechanical Systems

SEE ALSO HAND AND FINCH, 1.1–1.5

## Dynamical variables

These are any set of variables that describe completely the *configuration* (positions of parts) of a mechanical system. In 3D space, possible sets of dynamical variables include $(x, y, z)$, $(r, \phi, z)$ and $(r, \theta, \phi)$.

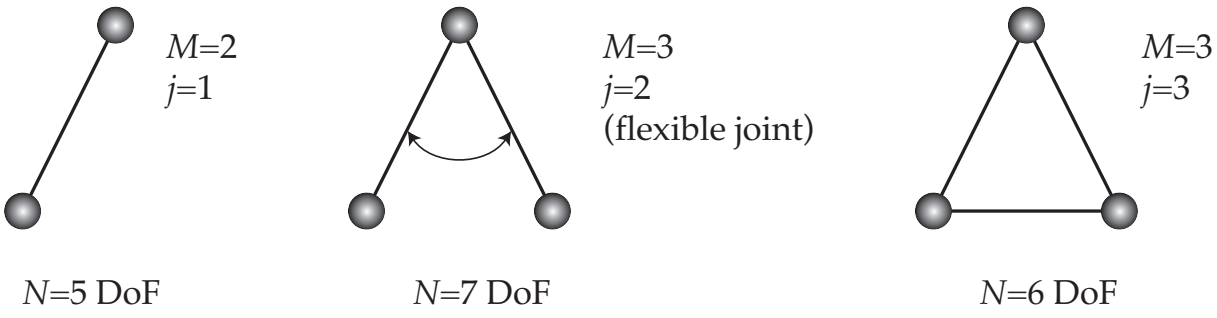
They can change under the action of forces — the *equation of motion* of a system specifies the dynamical variables as functions of time.

Newton's laws imply that these functions are found by solving *second order* differential equations — hence the initial positions/angles and velocities/angular velocities (or equivalent information) must be known.

## Degrees of Freedom (DoF)

The motion of a point mass $\underline{r}(t)$ is described by 3 *independent* dynamical variables, i.e., the point mass has $N = 3$ DoF.

A *system* of $M$ point masses has $N = 3M$ DoF, but the existence of $j$ *independent* constraints reduces this number to $N = 3M - j$ DoF.




Any rigid body containing more than 2 point masses connected by rigid rods has 6 DoF (or fewer, if there are additional constraints).

## Types of constraints

If constraints on a system of $M$ point masses can be expressed in the form

$$f(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_M, t) = 0,$$

then the constraints are called *holonomic*.

*Rheonomic* (running law) constraints have an explicit time dependence (e.g. bead on a rotating wire). *Scleronomic* (rigid law) constraints do not, i.e., $f(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_M) = 0$ (e.g. pendulum).

*Nonholonomic* constraints involve either differential equations or inequalities (rather than algebraic equations), e.g.: velocities (coin rolling without slipping on a 2D surface), inequalities (particle sliding down the outside of a bowling ball), most velocity dependent forces (friction).

## Generalised coordinates $q_k$

The existence of $j$ constraints means that the coordinates of e.g. $M$ point masses are *no longer independent*.

We require only as many coordinates $q_k$ as there are DoF ($N = 3M - j$).

If the constraints are holonomic then we can express the position of the $i$th part of a system as

$$\underline{r}_i = \underline{r}_i(q_1, \ldots, q_k, \ldots, q_N, t).$$

There may be an *explicit time dependence* particularly if the constraints are *rheonomic*.

The generalised velocities are $\dot{q}_k = dq_k/dt$.

We regard $q_k$ and $\dot{q}_k$ as *independent* variables (this is *very* important, and means for example that $\partial q_k/\partial \dot{q}_k = \partial \dot{q}_k/\partial q_k = 0$).

By the chain rule, one can derive the following identity, which can be a useful check:

$$\underline{v}_i = \dot{\underline{r}}_i = \sum_k \frac{\partial \underline{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \underline{r}_i}{\partial t} \quad \Rightarrow \quad \frac{\partial \dot{\underline{r}}_i}{\partial \dot{q}_k} = \frac{\partial \underline{r}_i}{\partial q_k} \qquad \text{“cancelling the dots.”}$$

This last step comes from taking the partial derivative with respect to $\dot{q}_k$.

# 4 The Lagrangian

SEE ALSO HAND AND FINCH, 1.6–1.11

## Lagrange's formulation of mechanics

We will assume:

1. Holonomic constraints

2. Constraining forces do no virtual work

3. Applied forces are conservative, such that a scalar potential energy function exists.

These assumptions imply: no friction or viscous drag, reaction forces are normal to surfaces, strings are inextensible and always tight. Note that the potential function may change with time.

## D'Alembert's Principle

Newtons 2nd law for particle $i$ is

$$\underline{\dot{p}}_i = \underline{F}_i.$$

This implies that

$$\sum_i (\underline{F}_i - \underline{\dot{p}}_i) \cdot \delta \underline{r}_i = 0,$$

where $\delta \underline{r}_i$ represents an arbitrary virtual displacement. A virtual displacement is a hypothetical change of coordinates at one instant in time that is compatible with the constraints. This force includes both the constraint and applied forces, such that

$$\underline{F}_i = \underline{F}_i^{(c)} + \underline{F}_i^{(a)}.$$

Given that we have supposed the constraining forces do no virtual work, this gives us D'Alembert's Principle:

$$\sum_i (\underline{F}_i^{(a)} - \underline{\dot{p}}_i) \cdot \delta \underline{r}_i = 0. \tag{1}$$

We have removed the constraint forces from this equation, but have yet to rewrite the equation in terms of the generalised coordinates that are *independent* of one another.

## Generalised Equations of Motion (EoM)

We would like to rewrite D'Alembert's Principle ($\sum_i (\underline{F}_i^{(a)} - \underline{\dot{p}}_i) \cdot \delta \underline{r}_i = 0$) in terms of the independent, generalised coordinates, $q_k$. Recall

$$\underline{v}_i = \frac{d}{dt}(\underline{r}_i) = \sum_k \frac{\partial \underline{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \underline{r}_i}{\partial t}. \tag{2}$$

Taking the partial derivative of this with respect to $\dot{q}_k$ yields

$$\frac{\partial \underline{v}_i}{\partial \dot{q}_k} = \frac{\partial \underline{r}_i}{\partial q_k}. \tag{3}$$

By analogy with Eq. (2), we can write

$$\frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_j}\right) = \sum_k \frac{\partial^2 \underline{r}_i}{\partial q_j \partial q_k}\dot{q}_k + \frac{\partial^2 \underline{r}_i}{\partial q_j \partial t}. \tag{4}$$

Taking the partial derivative of Eq. (2) with respect to $q_j$ gives

$$\frac{\partial \underline{v}_i}{\partial q_j} = \sum_k \frac{\partial^2 \underline{r}_i}{\partial q_j \partial q_k}\dot{q}_k + \frac{\partial^2 \underline{r}_i}{\partial q_j \partial t}. \tag{5}$$

Comparing Eqs. (4) and (5) leads to

$$\frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_j}\right) = \frac{\partial \underline{v}_i}{\partial q_j}. \tag{6}$$

The 1st term in Eq. (1), which represents the virtual work done by the applied force (now without the $^{(a)}$, as we no longer need to worry about the constraint force) can be written as

$$\delta W = \sum_i \underline{F}_i \cdot \delta \underline{r}_i = \sum_i \underline{F}_i \cdot \sum_k \frac{\partial \underline{r}_i}{\partial q_k}\delta q_k = \sum_k \underbrace{\left(\sum_i \underline{F}_i \cdot \frac{\partial \underline{r}_i}{\partial q_k}\right)}_{\text{Generalised forces } \mathcal{F}_k = \frac{\delta W}{\delta q_k}} \delta q_k = \sum_k \mathcal{F}_k \delta q_k.$$

The 2nd term in Eq. (1), can be written as

$$\sum_i \underline{\dot{p}}_i \cdot \delta \underline{r}_i = \sum_i m_i \underline{\dot{v}}_i \cdot \delta \underline{r}_i = \sum_i \sum_k m_i \underline{\dot{v}}_i \cdot \frac{\partial \underline{r}_i}{\partial q_k}\delta q_k.$$

If we note that

$$\underline{\dot{v}}_i \cdot \frac{\partial \underline{r}_i}{\partial q_k} = \frac{d}{dt}\left(\underline{v}_i \cdot \frac{\partial \underline{r}_i}{\partial q_k}\right) - \underline{v}_i \cdot \frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_k}\right)$$

then we can write

$$\sum_i \underline{\dot{p}}_i \cdot \delta \underline{r}_i = \sum_i \sum_k m_i \left[\frac{d}{dt}\left(\underline{v}_i \cdot \frac{\partial \underline{r}_i}{\partial q_k}\right) - \underline{v}_i \cdot \frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_k}\right)\right] \delta q_k.$$

Using Eqs. (3) and (6), we can write

$$\sum_i \underline{\dot{p}}_i \cdot \delta \underline{r}_i = \sum_i \sum_k \left[ \frac{d}{dt} \left( m_i \underline{v}_i \cdot \frac{\partial \underline{v}_i}{\partial \dot{q}_k} \right) - m_i \underline{v}_i \cdot \frac{\partial \underline{v}_i}{\partial q_k} \right] \delta q_k.$$

Changing the order of summation gives

$$\sum_i \underline{\dot{p}}_i \cdot \delta \underline{r}_i = \sum_k \left[ \frac{d}{dt} \left( \frac{\partial}{\partial \dot{q}_k} \left( \sum_i \frac{1}{2} m_i \underline{v}_i^2 \right) \right) - \frac{\partial}{\partial q_k} \left( \sum_i \frac{1}{2} m_i \underline{v}_i^2 \right) \right] \delta q_k.$$

Therefore, D'Alembert's Principle can be written as

$$\sum_k \left[ \mathcal{F}_k - \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_k} \right) + \frac{\partial T}{\partial q_k} \right] \delta q_k = 0,$$

where the kinetic energy, $T = \sum_i \frac{1}{2} m_i \underline{v}_i^2 = T(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N, t)$.
The $\delta q_k$ are independent, so the generalised equations of motion are

$$\mathcal{F}_k - \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_k} \right) + \frac{\partial T}{\partial q_k} = 0. \tag{7}$$

We are assuming that the applied forces are conservative, so $\underline{F}_i = -\nabla_i V(\underline{r}_1, \ldots, \underline{r}_i, \ldots, \underline{r}_M)$.
Correspondingly,

$$\mathcal{F}_k = -\sum_i \nabla_i V \cdot \frac{\partial \underline{r}_i}{\partial q_k} = -\frac{\partial V}{\partial q_k} \qquad \text{(follows from the chain rule).}$$

Substituting this into Eq. (7), and defining the Lagrangian by

$$L = T - V,$$

gives the *Euler-Lagrange equations*:

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0.$$

These are the *fundamental equations* of Lagrangian mechanics. Note that $L$ is a function of all of the $q_k$, $\dot{q}_k$ and $t$.

## Ignorable coordinates

If the time derivative of a coordinate appears in the Lagrangian, but the coordinate itself does not, then this is an *ignorable coordinate*. As

$$\frac{\partial L}{\partial q_k} = 0,$$

the Euler-Lagrange equation implies that

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) = 0.$$

Thus, the *canonically conjugate momentum* to the coordinate, defined as

$$p_k = \frac{\partial L}{\partial \dot{q}_k},$$

is a constant of the motion.

Example: a free particle in 1D with displacement $x$, has a Lagrangian $L = \frac{1}{2}m\dot{x}^2$. Thus, the coordinate $x$ is ignorable, and $p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$ is constant.

Having constants of motion can greatly simplify finding the equation of motion of a system. The canonically conjugate momentum plays a central role in the Hamiltonian approach to quantum mechanics.

# 5 Variational Calculus

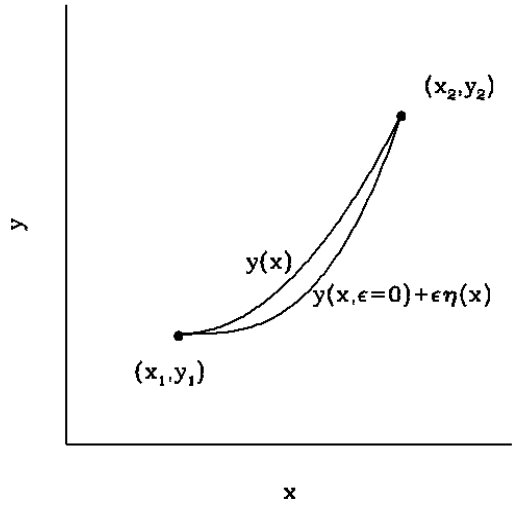
SEE ALSO HAND AND FINCH, 2.1–2.8

## Calculus of Variations

What path $y(x)$ leads to an extreme value of the functional

$$I[y] = \int_{x_1}^{x_2} f(y, y', x)dx, \tag{8}$$

where $y' = dy/dx$ and the trajectory goes through points $(x_1, y_1)$ and $(x_2, y_2)$?




Consider a small variation such that the path becomes $y(x, \epsilon = 0) + \epsilon\eta(x)$, where $\eta(x_1) = \eta(x_2) = 0$. For $I[y]$ to be stationary, we want $(dI/d\epsilon)_{\epsilon=0} = 0$. Note that $\partial y/\partial\epsilon = \eta$ and $\partial y'/\partial\epsilon = \eta'$.

$$\begin{aligned} \frac{dI}{d\epsilon} &= \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y}\frac{\partial y}{\partial \epsilon} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \epsilon} \right\} dx \\ &= \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta' \right\} dx \end{aligned}$$

Integrating the second term by parts gives

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\frac{d\eta}{dx}dx = \left[\eta\frac{\partial f}{\partial y'}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) dx.$$

The first term is zero because $\eta = 0$ at the ends of the path. Therefore

$$\frac{dI}{d\epsilon} = 0 \Rightarrow \int_{x_1}^{x_2} \left[\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y}\right] \eta(x)dx = 0 \qquad \forall\eta(x).$$

Hence we arrive at Euler's equation

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y} = 0.$$

Recall the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0.$$

For the case where $f$ does not explicitly depend upon $x$, ie $\partial f/\partial x = 0$, one can use

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial y'}\frac{dy'}{dx}$$

in conjunction with the Euler equation to show that

$$\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) = 0. \tag{9}$$

This is a useful alternative form of the Euler equation.

## Hamilton's Principle of Least Action

The Euler-Lagrange equations are the solution to the calculus of variations problem: for what route $\underline{q}(t)$ between points $\underline{q}_1$ and $\underline{q}_2$ at $t_1$ and $t_2$ is the integral

$$S[\underline{q}(t)] = \int_{t_1}^{t_2} L(\underline{q}(t), \dot{\underline{q}}(t), t)dt$$

an extrumum? Hamilton's Principle states that a mechanical system moves in just such a way as to minimise this integral, which is also known as the *action functional*, $S[\underline{q}(t)]$. The argument of a *function* is a number, and a number is returned — the argument of a *functional* is a *function*, and a number is returned.

Hamilton's Principle can be written as $\delta S = 0$, where the $\delta$ represents a change with respect to the path between the end points of the integral. The variational derivative of $L$ with respect to $q$ is written

$$\frac{\delta L}{\delta q} \equiv \frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right).$$

## Lagrange multipliers

If a mechanical system has an applied constraint and the number of DoF reduces, then this is usually dealt with by eliminating a dynamical variable and finding the equation of motion for a smaller set of generalised coordinates. An alternative method involves using Lagrange multipliers.
Consider a pendulum of length $l$. The natural generalised coordinate to choose to describe the motion is the angle $\theta$. Suppose you 'stubbornly and foolishly' insist on using $x$ and $y$, then you also have the constraint $x^2 + y^2 = l^2$. The variation of the action is

$$\delta S = \int \left[\frac{\delta L}{\delta x}\delta x + \frac{\delta L}{\delta y}\delta y\right] dt.$$

Hamilton's Principle implies that $\delta S = 0$, but the non-independence of $x(t)$ and $y(t)$ means that we cannot simply say that each variational derivative is separately zero. Our constraint can be written as

$$G(x, y) = x^2 + y^2 = c \quad \text{a constant.}$$

Now

$$\delta G = \frac{\partial G}{\partial x}\delta x + \frac{\partial G}{\partial y}\delta y = 0$$

is how the variations in $x$ and $y$ are connected. The variation of the action can then be written as

$$\delta S = \int \left[ \left( \frac{\delta L}{\delta x} - \lambda \frac{\partial G}{\partial x} \right) \delta x + \left( \frac{\delta L}{\delta y} - \lambda \frac{\partial G}{\partial y} \right) \delta y \right] dt.$$

where $\lambda$ is an arbitrary function of the independent variable ($t$ in this case) called a Lagrange multiplier. If we choose $\lambda$ such that the coefficient of $\delta x$ is set to zero, then $\delta S = 0$ implies that the coefficient of $\delta y$ must also vanish. It is as if we had independent variables $x$ and $y$.

More generally, if we wish to find the extremum of a functional $F(x, y)$ with respect to variations in the path $y(x)$, with an additional imposed constraint $G(x, y) = c$, a constant, then the path providing the extreme value of the functional can be found from

$$\delta(F - \lambda G) = 0,$$

where $\lambda$ is the Lagrange multiplier.

## An example of Lagrange multipliers

How should a fixed length of rope, $L$, be arranged in the half-plane $y > 0$ such that it encloses a maximal area when the two ends are placed on the line $y = 0$?

We wish to find an extreme value of $A = \int_{x_1}^{x_2} y dx$ subject to the constraint $\int ds = L$, where $ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + y'^2}\, dx$ represents a length along the rope. Using a Lagrange multiplier, we require $\delta(A - \lambda L) = 0$, i.e.

$$\delta \int_{x_1}^{x_2} \left( y - \lambda \sqrt{1 + y'^2} \right) dx = 0\,.$$

Comparison with equation (8) gives $f(y, y', x) = y - \lambda\sqrt{1 + y'^2}$, from which Euler's equation implies

$$\frac{d}{dx}\left[ \frac{-\lambda y'}{\sqrt{1 + y'^2}} \right] - 1 = 0\,.$$

Integrating w.r.t. $x$ and using the substitution $t = (x - a)/\lambda$, where $a$ is the integration constant, this can be rewritten as

$$\frac{-\dot{y}}{\sqrt{\lambda^2 + \dot{y}^2}} = t\,,$$

where $\dot{y} = dy/dt$. Rearranging for $\dot{y}$ and integrating w.r.t. $t$ yields

$$(x - a)^2 + (y - b)^2 = \lambda^2,$$

a semi-circle. The three constants, $a$, $b$ and $\lambda$ are set by $x_1$, $x_2$ and $L$.

# 6 Linear Oscillators

SEE ALSO HAND AND FINCH, 3.1–3.3

## Equilibrium

A mechanical system that remains at rest is in *equilibrium*. This occurs at points in configuration space where all generalised forces $\mathcal{F}_k$ *vanish*.

For a conservative system, this corresponds to configurations where the potential energy $V(q_1, \ldots, q_N)$ is *stationary*, i.e., its first derivatives with respect to the $q_k$ vanish.

## Lagrangian Near Static Equilibrium (1 DoF)

Carry out a 2nd order Taylor series expansion of some $L(q, \dot{q})$ around $q = q_S$ [where $q_S$ is a stationary point, i.e., $\partial V/\partial q|_{q=q_S} = 0$] and $\dot{q} = \dot{q}_S = 0$. Choose $q_S = 0$, then

$$\begin{aligned} L \approx L_{\text{approx}} =& L(q_S, \dot{q}_S) + q \left.\frac{\partial L}{\partial q}\right|_{q_S,\dot{q}_S} + \dot{q} \left.\frac{\partial L}{\partial \dot{q}}\right|_{q_S,\dot{q}_S} \\ &+ \frac{1}{2}\left( q^2 \left.\frac{\partial^2 L}{\partial q^2}\right|_{q_S,\dot{q}_S} + 2q\dot{q} \left.\frac{\partial^2 L}{\partial q \partial \dot{q}}\right|_{q_S,\dot{q}_S} + \dot{q}^2 \left.\frac{\partial^2 L}{\partial \dot{q}^2}\right|_{q_S,\dot{q}_S} \right) \\ &= A + Bq + C\dot{q} + Dq^2 + Eq\dot{q} + F\dot{q}^2. \end{aligned}$$

Substitute the resulting $L_{\text{approx}}$ into the *Euler-Lagrange equation* (n.b., $B = 0$ by the definition of equilibrium)

$$\frac{d}{dt}(\overbrace{C + Eq + 2F\dot{q}}^{\frac{\partial L_{\text{approx}}}{\partial \dot{q}}}) - \overbrace{(2Dq + E\dot{q})}^{\frac{\partial L_{\text{approx}}}{\partial q}} = 0.$$

This yields $\ddot{q} - (D/F)q = 0$, which is identical to the EoM for a particle of mass $m$ derived from

$$L = \frac{m}{2}\left(\dot{q}^2 + \frac{D}{F}q^2\right). \tag{10}$$

## Simple Harmonic Oscillator (SHO)

The Lagrangian for a simple harmonic oscillator, such as a mass $m$ attached to a spring with spring constant $k$, can be written as

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2.$$

If $D/F < 0$, then Eq. (10) is the Lagrangian for a simple harmonic oscillator, with $k = -m(D/F)$.

Close to equilibrium one therefore observes *simple harmonic motion* with characteristic frequency $\omega = \sqrt{-D/F}$. Hence, the motion is described as *stable* (if $D/F > 0$, it is unstable).

$\ddot{q} + \omega^2 q = 0$ is a *linear, homogeneous* differential equation, and can be solved to give

$$q(t) = q(0)\cos(\omega t) + \frac{\dot{q}(0)}{\omega}\sin(\omega t).$$

## Damping Force

A frictional force $F_d$, acts to *suppress* motion, i.e., it must act in opposition to motion, and vanish when the motion ceases.

The simplest such force is $F_d = -\gamma\dot{q} = -(m\omega/Q)\dot{q}$ (expressed in convenient form for a damped SHO), where $Q$ is the dimensionless *quality factor* of the oscillator.

## Damped Simple Harmonic Oscillator

We insert the SHO Lagrangian $L = (m/2)(\dot{q}^2 - \omega^2 q^2)$ into the Euler-Lagrange equation and incorporate the (non-conservative) damping force separately

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = F_d.$$

This corresponds to taking Eq. (7), grouping the conservative forces (in the form of a potential energy function $V$) with $T$ to make $L$, and keeping the non-conservative forces separate.

The result is a *linear, homogeneous* differential equation, $\ddot{q} + (\omega/Q)\dot{q} + \omega^2 q = 0$.

This is solved by setting $q = e^{\lambda t}$ and determining the allowed values of $\lambda$ from the resulting auxiliary equation. The solutions differ qualitatively depending on the value of $Q$.

(See linosc.py in the duo course documents.)

# 7 Driven Oscillators

SEE ALSO HAND AND FINCH, 3.4–3.8

## Oscillator Driven by an External Force

We consider an undamped oscillator, driven by a time-dependent external force $F(t)$ with no spatial dependence.

Whatever supplies the driving force is *not considered to be part of the dynamical system*. In particular we do not consider any effect of forces the oscillator exerts on the source of the external force.

The corresponding Lagrangian is given by

$$L = \underbrace{\frac{m\dot{q}^2}{2}}_{T} - \underbrace{\left[\frac{m\omega^2 q^2}{2} - F(t)q\right]}_{V},$$

which yields a second order *linear, inhomogeneous* differential equation:

$$\ddot{q} + \omega^2 q = F(t)/m. \tag{11}$$

We will split the driving force into an infinite number of impulsive forces and sum up the displacements caused by each of these impulsive forces to find $q(t)$. This method is exactly analogous to solving Poisson's equation in electrostatics ($\nabla^2\phi(\mathbf{r}) = -\rho(\mathbf{r})/\epsilon$) to find the potential, $\phi(\mathbf{r})$, caused by a charge distribution $\rho(\mathbf{r})$ ($\phi(\mathbf{r}) \triangleq q(t)$ and $\rho(\mathbf{r}) \triangleq F(t)$).

## Dirac $\delta$-Functions

To express an *impulsive* (instantaneously applied and infinitely intense) force mathematically we make use of Dirac *$\delta$-functions*.

$\delta(t-t')$ is the *zero-width limit* of a function/distribution of $t$, sharply peaked at $t = t'$, with unit area: $\lim_{\epsilon\to 0}\int_{t'-\epsilon}^{t'+\epsilon}\delta(t-t')dt = 1$ (dimensionless). If $t$ has the dimension of *time*, then $\delta(t-t')$ must have the dimension of inverse time, i.e., *frequency*.

The effect of $\delta(t-t')$ is defined (necessarily under an integral) as

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

and we can define a single impulsive force at time $t'$ as $F(t) = K\delta(t-t')$, where $K$ is the total impulse provided.

## Response of an oscillator to a $q$-Independent Impulsive Force

With $F(t) = K\delta(t-t')$, integrating the differential equation for the driven oscillator around $t = t'$ yields

$$\int_{t'-\epsilon}^{t'+\epsilon}(\ddot{q}+\omega^2 q)dt = \int_{t'-\epsilon}^{t'+\epsilon}\frac{K}{m}\delta(t-t')dt,$$

$$\Rightarrow \dot{q}(t'+\epsilon) - \dot{q}(t'-\epsilon) + \underbrace{\omega^2\int_{t'-\epsilon}^{t'+\epsilon}q(t)dt}_{\to 0 \text{ as } \epsilon\to 0,\ \text{unless } q(t')\to\infty} = \frac{K}{m}.$$

Hence, there is an instantaneous change in the *velocity*: $\dot{q}(t'_+) = \dot{q}(t'_-) + K/m$, where $t'_\pm = \lim_{\epsilon\to 0} t' \pm \epsilon$ (this is the definition of an impulsive force).

Note that for a finite velocity, $q(t'_+) = q(t'_-)$, i.e., there is *no* instantaneous change in *position* when the impulsive force is applied.

## Time-Evolution Sequence of a SHO with an Impulsive Force

1. $t < t'$: ($F(t) = 0$) system evolves as a SHO, to position $q(t'_-)$, velocity $\dot{q}(t'_-)$.
2. $t = t'$: ($F(t) \neq 0$) instantaneous jump in velocity: $q(t') = q(t'_-)$, $\dot{q}(t') = \dot{q}(t'_-) + K/m$.
3. $t > t'$: ($F(t) = 0$) system evolves as a SHO:

$$q(t) = q(t'_-)\cos(\omega[t-t']) + \frac{\dot{q}(t'_-) + K/m}{\omega}\sin(\omega[t-t']). \tag{12}$$

## (Causal) Green's Function $G$

The Green's function for this system is the solution to the differential equation (c.f. Eq. (11)) $\ddot{G}(t-t') + \omega^2 G(t-t') = \delta(t-t')$. It represents the response of the oscillator to a single unit-sized impulsive force, scaled as $G(t-t') = q(t)m/K$, so that the differential equation only has a $\delta$-function on the right hand side.

Considering the oscillator to be at rest prior to the application of the impulsive force, the solution for $G$ is, by comparison with Eq. (12),

$$\begin{aligned} G(t-t') &= 0, \quad t-t' \leq 0, \\ G(t-t') &= \frac{1}{\omega}\sin(\omega[t-t']), \quad t-t' \geq 0. \end{aligned} \tag{13}$$

(See green.py in the duo course documents.)

## General Driving Force $F(t)$

Using the definition of the $\delta$-function, the equation of motion for the driven oscillator (Eq. (11)) can be written as

$$\ddot{q} + \omega^2 q = (1/m)\int_{-\infty}^{\infty} dt' F(t')\delta(t-t').$$

This is effectively just splitting up the general driving force into an infinite number of impulsive forces. We can then eliminate $\delta(t-t')$ by substituting in the differential equation for the Green's function $G$:

$$\ddot{q} + \omega^2 q = (1/m)\int_{-\infty}^{\infty} dt' F(t')[\ddot{G}(t-t') + \omega^2 G(t-t')].$$

Comparing terms, and substituting in the solution for $G(t-t')$ (Eq. (13)) it follows that

$$q(t) = \frac{1}{m}\int_{-\infty}^{t} F(t')G(t-t')dt' = \frac{1}{m\omega}\int_{-\infty}^{t} F(t')\sin(\omega[t-t'])dt'.$$

Hence, the displacement at time $t$ is found by summing over the displacements caused by all the preceding impulsive forces.

(See driven_undamped.py and driven_damped.py in the duo course documents for animations of the behaviour of undamped and damped, driven oscillators, including the splitting of the solution for the damped, driven oscillator into transient and steady-state components.)

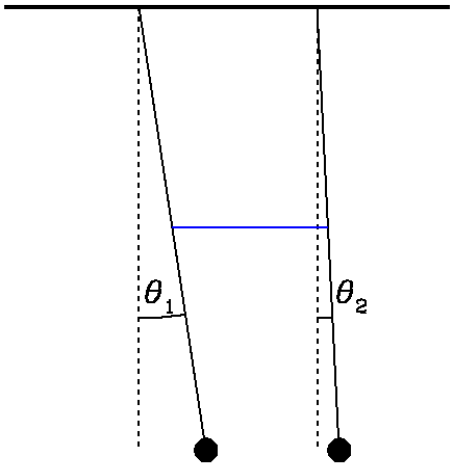
# 8 Coupled Small Oscillations

SEE ALSO HAND AND FINCH, 9.1–9.2

Lagrangian formulation particularly useful for small oscillations of a set of coupled oscillators about equilibrium positions. Important in acoustics, molecular spectroscopy, mechanical systems and coupled electrical circuits.

We consider small oscillations around a stable equilibrium point of a mechanical system. If this motion does not depart too far from a stable equilibrium point, then the system resembles a set of *coupled oscillators*.

## Two coupled pendulums




Consider a system with two pendulums of length $l$, connected half way down by a spring with constant $k$, and unstretched length equal to the horizontal separation of the pivot points. There are 2 DoF - use $\theta_1$ and $\theta_2$ as generalised coordinates. If the rigid pendulums have all of their mass, $m$, contained in the bob, and the spring is massless, then the kinetic energy is

$$T = \frac{ml^2}{2}\left(\dot{\theta}_1^2 + \dot{\theta}_2^2\right).$$

Without coupling, and assuming small oscillations,

$$V = \frac{mgl}{2}(\theta_1^2 + \theta_2^2).$$

The spring provides an extra potential energy,

$$V_{coupling} = \frac{k}{2}\left(\frac{l}{2}\right)^2(\theta_2 - \theta_1)^2.$$

Hence

$$L = \frac{ml^2}{2}(\dot{\theta}_1^2 + \dot{\theta}_2^2) - \frac{1}{2}\left(mgl(\theta_1^2 + \theta_2^2) + k\left(\frac{l}{2}\right)^2(\theta_2 - \theta_1)^2\right).$$

The $\theta_1\theta_2$ term makes life awkward, so transform to *centre of mass* $\theta_c = (\theta_1 + \theta_2)/2$ and *relative* $\theta_r = \theta_2 - \theta_1$ coordinates. Defining $\omega_0 = \sqrt{g/l}$, and the coupling constant $\eta = \frac{kl}{4mg}$,

$$L = ml^2 \left[ \left( \dot{\theta}_c^2 - \omega_0^2 \theta_c^2 \right) + \frac{1}{4} \left( \dot{\theta}_r^2 - \omega_0^2 (1 + 2\eta) \theta_r^2 \right) \right].$$

The Euler-Lagrange equations yield two *separated* harmonic oscillator EoMs:

$$\ddot{\theta}_C + \omega_0^2 \theta_C = 0$$

and

$$\ddot{\theta}_r + \omega_0^2 (1 + 2\eta) \theta_r = 0.$$

In a system of $N$ coupled oscillators there *always* exist $N$ generalised coordinates that separate, yielding $N$ independent, homogeneous, linear differential equations.

Those generalised coordinates undergoing simple harmonic motion are called *normal coordinates*.

(See cpend.py in the duo course documents.)

## A more general recipe

1. Find an equilibrium configuration, i.e., values of the generalised coordinates where all generalised forces $\mathcal{F}_k = -\partial V / \partial q_k$ are $= 0$.

2. Taylor expand the Lagrangian $L$ to *second order* in the generalised coordinates and velocities, around values for the generalised coordinates $q_k$ given by the equilibrium configuration, with the values of the generalised velocities $\dot{q}_k$ set $= 0$.

3. For $N$ DoF, the Taylor expansion must in principal be in $2N$ variables (all generalised coordinates and velocities). In practice, expansions in fewer variables may be adequate, e.g. if some terms are already in quadratic form.

## Matrix form of $L$

If $a_k$ are the equilibrium values of an initially used set of generalised coordinates $q'_k$, then we can define a new set of generalised coordinates $q_k = q'_k - a_k$.

These $q_k$ will have equilibrium values $= 0$, and $L$ can then take the form

$$L = \underbrace{\underline{\dot{q}}^T \hat{\tau} \underline{\dot{q}}}_{T} - \underbrace{\underline{q}^T \hat{v} \underline{q}}_{V} + c \text{ (a constant)},$$

where $\underline{q}$ is a column vector, and its transpose $\underline{q}^T = (q_1, \ldots, q_N)$ is a row vector.

The matrices $\hat{\tau}$ and $\hat{v}$ are *symmetric* (i.e., equal to their transposes, so that e.g. $\tau_{kj} = \tau_{jk}$), with matrix elements given by

$$\tau_{jk} = \frac{1}{2} \frac{\partial^2 T}{\partial \dot{q}_j \partial \dot{q}_k} \bigg|_{\dot{q}_j, \dot{q}_k = 0}, \qquad v_{jk} = \frac{1}{2} \frac{\partial^2 V}{\partial q_j \partial q_k} \bigg|_{q_j, q_k = 0}.$$

## Equations of Motion

Note that partial differentiation of the Lagrangian $L$ with respect to the generalised coordinates and velocities yields

$$\frac{\partial L}{\partial \dot{q}_j} = 2\sum_{k=1}^{N} \tau_{jk}\dot{q}_k = 2(\hat{\tau}\underline{\dot{q}})_j, \qquad \frac{\partial L}{\partial q_j} = -2\sum_{k=1}^{N} v_{jk}q_k = -2(\hat{v}\underline{q})_j,$$

and that the Euler-Lagrange equations therefore yield a *linear system* of differential equations: $\hat{\tau}\underline{\ddot{q}} + \hat{v}\underline{q} = 0$.

# 9 Normal Modes

SEE ALSO HAND AND FINCH, 9.3–9.5

The Euler-Lagrange equations for a system of coupled oscillators can be written compactly in matrix notation as $\hat{\tau}\underline{\ddot{q}} + \hat{v}\underline{q} = 0$.

## Working Towards a Solution

Inserting a *trial solution* $\underline{q} = \underline{b}e^{i\omega t}$ into the Euler-Lagrange equations above yields

$$(\hat{v} - \omega^2\hat{\tau})\underline{b} = 0. \tag{14}$$

We also know (due to a theorem from linear algebra) that, unless $\underline{b} = \underline{0}$, this can only be fulfilled if $|-\hat{\tau}\omega^2 + \hat{v}| = 0$.

(This strictly only represents an eigenvalue problem if $\hat{\tau}$ is a multiple of the identity matrix. If different masses are associated with the different generalised velocities, then a variable transformation $\xi_i = \sqrt{m_i}q_i$ rectifies this. In practice, we can treat the solutions to Eq. (14) as *eigenvectors* with their associated *eigenvalues*.)

Calculating the determinant will yield an $N$th order polynomial in $\omega^2$. Hence, there are $N$ generalised values $\omega_j^2$ (possible values of $\omega^2$), each with its corresponding vector $\underline{b}_j$ that will solve Eq. (14).

Returning to the example of the coupled pendulums, before we used our intuition to convert from $(\theta_1, \theta_2)$ to $(\theta_c, \theta_r)$, we had

$$\begin{aligned} L =& \frac{ml^2}{2}(\dot{\theta}_1^2 + \dot{\theta}_2^2) - \frac{1}{2}\left(mgl(\theta_1^2 + \theta_2^2) + k\left(\frac{l}{2}\right)^2(\theta_2 - \theta_1)^2\right) \\ =& \frac{mgl}{2}\left(\frac{1}{\omega_0^2}(\dot{\theta}_1^2 + \dot{\theta}_2^2) - ((1+\eta)(\theta_1^2 + \theta_2^2) - \eta(\theta_1\theta_2 + \theta_2\theta_1))\right), \end{aligned}$$

where $\omega_0 = \sqrt{g/l}$ and $\eta = \frac{kl}{4mg}$. We can thus infer that

$$\omega^2\hat{\tau} = \frac{mgl}{2}\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

where $\lambda = \frac{\omega^2}{\omega_0^2}$, and

$$\hat{v} = \frac{mgl}{2}\begin{pmatrix} 1+\eta & -\eta \\ -\eta & 1+\eta \end{pmatrix}$$

The values of $\omega$ for which the system oscillates at a single frequency are thus

$$\left|-\hat{\tau}\omega^2 + \hat{v}\right| = (1 + \eta - \lambda)^2 - \eta^2 = 0,$$

i.e. $\lambda_{(1)} = 1$ and $\lambda_{(2)} = 1 + 2\eta$, which correspond to $\omega_{(1)} = \omega_0$ and $\omega_{(2)} = \omega_0\sqrt{1+2\eta}$. To find the mode vectors associated with these modes of oscillation, substitute $\omega_{(j)}$ back into Eq. (14). This gives

$$\frac{b_{(j)2}}{b_{(j)1}} = \frac{1 + \eta - \lambda}{\eta},$$

from which, $b_{(1)1} = b_{(1)2}$ and $b_{(2)1} = -b_{(2)2}$.

## Normal Coordinates

It is common to normalise the mode vectors $\underline{b}_{(j)}$, i.e., to have $\underline{b}_{(j)}^T = (b_{(j)1}, b_{(j)2}, \ldots, b_{(j)N})$ such that $\sqrt{\sum_{k=1}^{N} b_{(j)k}^2} = 1$. Using the vector elements, we may define a new set of *separated* generalised normal coordinates $r_j$ via

$$r_j = \sum_{k=1}^{N} b_{(j)k} q_k.$$

Note that, for the coupled pendulum example, $r_1 \propto (\theta_1 + \theta_2)$ and $r_2 \propto (\theta_2 - \theta_1)$.

If $\omega_j^2 > 0$, the equilibrium configuration is *stable*, and the normal coordinate $r_j$ evolves as $r_j(t) = r_j(0)\cos(\omega_j t) + (\dot{r}_j(0)/\omega_j)\sin(\omega_j t)$. Such motion is called a *normal mode*, where all parts of the system oscillate with the same frequency. Motion in general can be described by a superposition of normal modes.

If $\omega_j^2 < 0$, the equilibrium is *unstable*, and if $\omega_j^2 = 0$, the motion is also not oscillatory, usually corresponding to e.g. rotation or translation of the centre of mass.

A *normal mode* is the motion of a normal coordinate.

## Mode orthogonality

As we have effectively been finding eigenvectors of a real symmetric matrix, they will be orthogonal, ie

$$\underline{b}_i^T \hat{\tau} \underline{b}_j = 0 \quad \text{if} \quad i \neq j.$$

1. If $\omega_i$'s are all different, then $b_i$'s are unique.
2. If some $\omega_i$'s are the same, then there is a choice of $b_i$'s, but can always find orthogonal $b_i$'s (if $b_i$ and $b_j$ are eigenvectors with the same $\omega$, then so is any linear combination $\alpha b_i + \beta b_j$).

The $b_i$'s are the 'natural axes' in N-dimensional space. With respect to them, the transformed potential energy, $\hat{v}'$, is diagonal (i.e. contains no cross terms). In these coordinates, the system undergoes simple harmonic oscillations.

(See triatomic.py in the duo course documents for an example of normal modes in a system with three degrees of freedom.)

# 10 Central Forces

SEE ALSO HAND AND FINCH, 4.3–4.4

## Two Interacting Bodies

Consider two bodies (with position vectors $\underline{r}_1$ and $\underline{r}_2$) interacting via a *central force*, defined as a force depending only on the distance $r$ between the two bodies, directed along the line between them.

The potential energy $V_{12}$ from which such a force is derived is only dependent on $r$, i.e., $V_{12} = V_{12}(\underline{r}_1, \underline{r}_2) = V_{12}(|\underline{r}_1 - \underline{r}_2|) = V(r)$ for such a force.

If we consider the general case of 2 point masses interacting via a central force, with additional external forces, then the Lagrangian is given by

$$L = \frac{1}{2}(m_1\dot{\underline{r}}_1^2 + m_2\dot{\underline{r}}_2^2) - \underbrace{V_1(\underline{r}_1) - V_2(\underline{r}_2)}_{\text{external potentials}} - \underbrace{V_{12}(|\underline{r}_1 - \underline{r}_2|)}_{\text{interaction potential}}.$$

In the *absence* of external potentials ($V_1 = V_2 = 0$), this Lagrangian is *translationally invariant*: $L(\underline{r}_1, \underline{r}_2, \dot{\underline{r}}_1, \dot{\underline{r}}_2) = L(\underline{r}_1 + \underline{a}, \underline{r}_2 + \underline{a}, \dot{\underline{r}}_1, \dot{\underline{r}}_2)$.

## Translational Symmetry and Conservation of Momentum

Consider an *infinitesimal translation* $\epsilon\underline{a}$, and Taylor expand the Lagrangian $L$ to first order in $\epsilon$:

$$L(\underline{r}_1 + \epsilon\underline{a}, \underline{r}_2 + \epsilon\underline{a}, \dot{\underline{r}}_1, \dot{\underline{r}}_2) = L(\underline{r}_1, \underline{r}_2, \dot{\underline{r}}_1, \dot{\underline{r}}_2) + \epsilon\underline{a} \cdot (\nabla_1 + \nabla_2)L + O(\epsilon^2).$$

We neglect the $O(\epsilon^2)$ term. Since $\underline{a}$ is arbitrary, if $L$ is to be translationally invariant, then $(\nabla_1 + \nabla_2)L$ must $= 0$, i.e.,

$$\frac{\partial L}{\partial x_1} + \frac{\partial L}{\partial x_2} = 0, \qquad \frac{\partial L}{\partial y_1} + \frac{\partial L}{\partial y_2} = 0, \qquad \frac{\partial L}{\partial z_1} + \frac{\partial L}{\partial z_2} = 0.$$

From the Euler-Lagrange equations for $x_1$ and $x_2$ (and equivalently for $y$ and $z$):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_1}\right) + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_2}\right) = \frac{\partial L}{\partial x_1} + \frac{\partial L}{\partial x_2} = 0.$$

$\partial L/\partial \dot{x}_1 = p_{1x}$ defines the *canonically conjugate momentum* $p_{1x}$ to $x_1$ (here this is equal to the mechanical momentum $m\dot{x}_1$, but this need not in general be so). Hence, $\dot{\underline{P}} = \dot{\underline{p}}_1 + \dot{\underline{p}}_2 = 0$, and so the *total* momentum $\underline{P}$ is *conserved*.

## Ignorable Centre of Mass (CoM) Motion and Conservation of Momentum

Conservation of the total momentum $\underline{P}$ implies that the CoM (coordinate $\underline{R} = (m_1\underline{r}_1 + m_2\underline{r}_2)/(m_1 + m_2)$) moves like a free particle.

Rewrite $L$ in terms of $\underline{R}$ and the relative coordinate $\underline{r} = \underline{r}_1 - \underline{r}_2$:

$$L = \frac{1}{2}M\dot{\underline{R}}^2 + \frac{1}{2}\mu\dot{\underline{r}}^2 - V_{12}(r),$$

where $M = m_1 + m_2$ (*total* mass) and $\mu = m_1 m_2/(m_1 + m_2)$ (the *reduced* mass).

$\underline{R}$ does not appear in the Lagrangian $L$. Such coordinates are called *ignorable*. For any ignorable generalised coordinate $q$, it follows from the Euler-Lagrange equation that $d(\partial L/\partial \dot{q})/dt \equiv \dot{p}_q = 0$, and hence that $\underline{P}$ is conserved in this case.

## Rotational Invariance and Conservation of Angular Momentum

If we transform to the centre of mass frame (a frame moving with the CoM), then $\underline{R} = \dot{\underline{R}} = 0$, and $L = (1/2)\mu \dot{\underline{r}}^2 - V_{12}(r)$. Note that $V_{12}(r)$ is dependent on *distance* only, not direction, and hence that $L$ is *rotationally invariant*.

In terms of spherical coordinates ($x = r\cos\phi\sin\theta$, $y = r\sin\phi\sin\theta$, $z = r\cos\theta$), we deduce that $\dot{\underline{r}}^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta$. Hence,

$$L = \frac{1}{2}\mu[\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta] - V_{12}(r). \tag{15}$$

$\phi$ is ignorable, as it does not appear in $L$. Therefore

$$p_\phi \equiv \frac{\partial L}{\partial \dot{\phi}} = \mu r^2\dot{\phi}\sin^2\theta = \text{ constant } \equiv \underbrace{J_z}_{\text{can show to be the } z \text{ component of } \underline{J} = \underline{r}\times\mu\dot{\underline{r}}}$$

For any choice of z axis, the angular momentum in that direction is constant. Therefore the total angular momentum, $\underline{J}$, is *always* constant. Without loss of generality, we can choose the coordinates such that $\underline{J}$ always lies along the z axis.

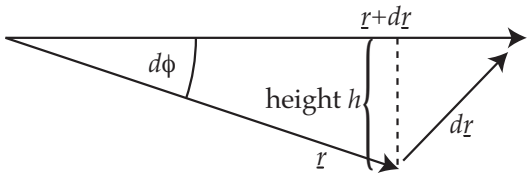
## Motion Takes Place in a Plane

As $\underline{J} = \underline{r}\times\mu\dot{\underline{r}}$, both $\underline{p} \equiv \mu\dot{\underline{r}}$ and $\underline{r}$ must be perpendicular to $\underline{J}$.

Hence, $\underline{p}$ and $\underline{r}$ must be in a plane perpendicular to $\underline{J}$, and remain in that plane if $\underline{J}$ is to remain constant.

Choose $\underline{J}$ to lie along the $z$ direction, so that the angle of elevation $\theta = \pi/2$, $\dot{\theta} = 0$, and $J = |\underline{J}| = J_z = \mu r^2\dot{\phi}$.

## Equal Areas are Swept out in Equal Times (Kepler's Second Law)




Consider motion in the $\theta = \pi/2$ plane. If the position vector changes incrementally from $\underline{r}$ to $\underline{r} + d\underline{r}$, with the angle changing by $d\phi$, then the area swept out, $dA$, is given by the area of the triangle formed by $\underline{r}$, $d\underline{r}$, and $\underline{r} + d\underline{r}$.

By straightforward trigonometry $dA = (1/2)|\underline{r} + d\underline{r}||\underline{r}|\sin(d\phi)$. Keeping only first-order terms, $dA = (1/2)r^2 d\phi$.

It follows that

$$\frac{dA}{dt} = \frac{1}{2}r^2\dot{\phi} = \frac{J}{2\mu}. \tag{16}$$

$J$ is conserved, so the rate at which area is swept out is constant.

Kepler's second law is therefore a consequence of angular momentum conservation, itself a consequence of considering a central force, and has nothing specifically to do with gravity as such.

## Equivalent 1D Problem

We can remove $\theta$ and $\phi$ entirely from the Lagrangian.

Following Eq. (15), we can write $E = T + V = (1/2)\mu[\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2(\theta)] + V_{12}(r)$, which, if we make use of $\dot{\theta} = 0$, $\theta = \pi/2$, $\dot{\phi} = J/\mu r^2$, simplifies to

$$E = \frac{1}{2}\mu\dot{r}^2 + \underbrace{\frac{J^2}{2\mu r^2} + V_{12}(r)}_{V_{\text{eff}}(r)}. \tag{17}$$

The effective potential $V_{\text{eff}}(r)$ combines the interaction potential with the so-called *angular momentum barrier* (responsible for the centrifugal force).

The total energy $E$ is the same as that corresponding to an effective Lagrangian $L_{\text{eff}} = (1/2)\mu\dot{r}^2 - V_{\text{eff}}$, from which the Euler-Lagrange equation for $r$ yields

$$\mu\ddot{r} = \frac{J^2}{\mu r^3} - \frac{\partial V_{12}}{\partial r}. \tag{18}$$

(See central_force.py in the duo course documents for an example of motion of a particle in a central potential.)

# 11 Gravitational Attraction

SEE ALSO HAND AND FINCH, 4.1, 4.5–4.6

## Gravitational Interaction Potential

The potential energy between two gravitating bodies of masses $m_1$ and $m_2$ is $V_{12}(r) = -k/r$, where $k = Gm_1m_2$ and $G$ is the gravitational constant. The differential equation to be solved is then

$$\mu\ddot{\vec{r}} = \frac{J^2}{\mu r^3} - \frac{k}{r^2}$$

We would like to solve this equation to determine the orbits of the bodies.

## The $u = 1/r$ Transformation

It is helpful to change the dependent variable from $r$ to $u = 1/r$. From $J = \mu r^2(d\phi/dt)$, we can determine that $dt = [\mu/(Ju^2)]d\phi$ and hence that

$$\frac{d}{dt} = \frac{Ju^2}{\mu}\frac{d}{d\phi} \Rightarrow \frac{dr}{dt} = -\frac{1}{u^2}\frac{du}{dt} = -\frac{J}{\mu}\frac{du}{d\phi}.$$

Expressing the energy in the CoM frame [Eq. (17)] in terms of $u$ and $\phi$, gives

$$E = \frac{J^2}{2\mu}\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] - ku. \tag{19}$$

Differentiating with respect to $\phi$ (note that $E$ is constant, i.e., $dE/d\phi = 0$),

$$\frac{J^2}{2\mu}\left(2\frac{du}{d\phi}\frac{d^2u}{d\phi^2} + 2u\frac{du}{d\phi}\right) - k\frac{du}{d\phi} = 0,$$

from which

$$\frac{d^2u}{d\phi^2} + u = \frac{\mu k}{J^2}. \tag{20}$$

This can be solved to give $(1/r) \equiv u = (\mu k/J^2) + B\cos\phi$ ($B$ is an integration constant, and we have set a second integration constant, equivalent to an arbitrary phase added to $\phi$, to 0).

## Elliptical Orbits (Kepler's First Law)

Define $p \equiv J^2/\mu k$ and $\epsilon \equiv pB$. Then putting the solution for $u$ into Eq. (19) yields

$$E = \frac{\mu k^2}{2J^2}(\epsilon^2 - 1) = \frac{k}{2p}(\epsilon^2 - 1).$$

If $0 \le \epsilon < 1$, the energy $E$ is negative, implying a *bound state*.

The solution to Eq. (20) can be written as

$$p = r + \epsilon r\cos\phi.$$

Expressed in cartesian coordinates, this becomes $p = \sqrt{x^2 + y^2} + \epsilon x$, or equivalently

$$(1 - \epsilon^2)x^2 + 2\epsilon p x + y^2 - p^2 = 0. \tag{21}$$

Restricting ourselves to $0 \leq \epsilon < 1$, Eq. (21) describes an *ellipse* (Kepler's first law):

$$\frac{1}{a^2}\left(x + \frac{\epsilon p}{1 - \epsilon^2}\right)^2 + \frac{y^2}{b^2} = 1, \tag{22}$$

where the semimajor axis $a = p/(1 - \epsilon^2)$ and the seminor axis $b = p/\sqrt{1 - \epsilon^2}$ (with these quantities defined, the energy takes the simple form $E = -k/2a$).

$\epsilon$ is the eccentricity of the elliptical orbit. $\epsilon = 0$ for a circular orbit, $\epsilon = 1$ describes a parabolic trajectory and $\epsilon > 1$ for a hyperbola.

## Kepler's Third Law

From Kepler's second law $dA/dt = J/(2\mu)$ (Eq. (16)), the period $\tau$ of an elliptical orbit follows from the total area $A = \pi ab$ of that orbit , i.e., $\tau = A2\mu/J$. A little algebra yields

$$\tau = 2\pi\sqrt{\frac{1}{G(m_1 + m_2)}}a^{3/2}$$

Kepler's third law states that $\tau^2 \propto a^3$ with the same proportionality constant for all the planets. As the mass of the sun dominates the masses of the individual planets, this is approximately true.

## Bertrand's Theorem

For force laws of the form

$$F \propto -cr^{\alpha},$$

only $\alpha = -2$ or $\alpha = 1$ give rise to closed, non-circular orbits. (Note that a circular orbit is possible for any $\alpha$.) A closed orbit is one for which the radius at a fixed azimuthal angle, $\theta$, is described by a function $r(\theta)$. $\alpha = -2$ corresponds to the familiar inverse square law, where the orbit is an ellipse and the central force originates from one focus. $\alpha = 1$ is Hooke's law, which gives rise to elliptical orbits that are centred on the source of the force.
The precession of the perihelion of Mercury was one hint that the inverse square law of gravity was not the whole story. The General Theory of Relativity accurately predicts the size of the effect.

## Solving 1D Systems by Quadrature

Consider a 1D system such as a point mass in an external potential. The Lagrangian $L = (1/2)m\dot{q}^2 - V(q)$, and it has energy $E = (1/2)m\dot{q}^2 + V(q)$. This implies

$$\dot{q} \equiv \frac{dq}{dt} = \pm\sqrt{\frac{2[E - V(q)]}{m}} \Rightarrow dt = dq\left\{\pm\sqrt{\frac{m}{2[E - V(q)]}}\right\}.$$

The time $t$ taken to move from $q = 0$ to $q(t)$ is then

$$t = \sqrt{\frac{m}{2}} \int_0^q dq' \sqrt{\frac{1}{E - V(q')}}. \tag{23}$$

The elapsed time as a function of the coordinate, $t(q)$, must then be inverted if we wish to obtain $q(t)$ (this may have to be carried out numerically).

A problem "solved" as an integral (i.e., put into integral form, as in Eq. (23)) is said to be *solved by quadrature*.

# 12 Noether's Theorem and Hamiltonian Mechanics

SEE ALSO HAND AND FINCH, 5.1–5.5

## Invariance of $L$ under Continuous Transformations

Consider a set of continuous transformations, i.e. depending on a continuously variable parameter, $s$ (e.g. translations along or rotations about the $z$ axis are single-variable families of transformations).
When $s = 0$, this should yield the *identity transformation*, i.e., no change.
If $q_k(t)$ is a solution of the original EoM for a Lagrangian $L$, then $Q_k(s,t)$ will be a solution for the Euler-Lagrange equations evaluated within a transformed set of coordinates, where $Q_k(0,t) = q_k(t)$.
If the Lagrangian $L$ is *invariant* (independent of $s$) under this transformation, then

$$\begin{aligned} L' \equiv & L(Q_1(s,t), \ldots, Q_N(s,t), \dot{Q}_1(s,t), \ldots, \dot{Q}_N(s,t), t) \\ = & L(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N, t). \end{aligned}$$

If the Lagrangian is independent of $s$, $dL'/ds = 0$, and dropping the time variation for simplicity, then

$$\frac{dL'}{ds} = \sum_{k=1}^{N} \left( \underbrace{\frac{\partial L'}{\partial Q_k}}_{= \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{Q}_k}\right)} \frac{dQ_k}{ds} + \frac{\partial L'}{\partial \dot{Q}_k} \underbrace{\frac{d\dot{Q}_k}{ds}}_{= \frac{d}{dt}\left(\frac{dQ_k}{ds}\right)} \right) = \frac{d}{dt}\left( \sum_{k=1}^{N} \frac{\partial L'}{\partial \dot{Q}_k} \frac{dQ_k}{ds} \right) = 0.$$

This implies the existence of a *constant of the motion*

$$I(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N) \equiv \sum_{k=1}^{N} p_k (dQ_k/ds)|_{s=0},$$

which is evaluated at $s = 0$ for convenience.
For multiple parameters $s_1, s_2, \ldots$, the procedure can be repeated for each to find associated constants $I_1, I_2, \ldots$.
This procedure effectively describes *Noether's theorem*.

## Noether's Theorem

**If the Lagrangian is invariant under a continuous symmetry transformation, then there are conserved quantities associated with that symmetry, one for each parameter of the transformation. These can be found by differentiating each coordinate with respect to the parameters of the transformation in the immediate neighbourhood of the identity transformation, multiplying by the conjugate momentum, and summing over the degrees of freedom.**

Take the example of a point mass, $m$, moving in free space with $L = \frac{1}{2}m\dot{q}^2$. For transformed coordinates $Q = q + s$, this gives

$$\left.\frac{dQ}{ds}\right|_{s=0} = 1 \quad \Rightarrow I(q) = m\dot{q}.$$

## Legendre Transformations

Consider a function $A(x,y)$ of two variables. Introduce a third variable $z$ and define $B(x,y,z) \equiv yz - A(x,y)$. By application of the chain rule,

$$dB = zdy + ydz - \left.\frac{\partial A}{\partial x}\right|_y dx - \left.\frac{\partial A}{\partial y}\right|_x dy = \underbrace{\left(z - \left.\frac{\partial A}{\partial y}\right|_x\right)}_{\text{vanishes, as we define } z = z(x,y) \equiv \left.\frac{\partial A}{\partial y}\right|_x} dy + ydz - \left.\frac{\partial A}{\partial x}\right|_y dx. \tag{24}$$

For this choice of $z$, $B$ has no explicit $y$ dependence. Thus $y$ can be written in terms of $x$ and $z$ giving

$$B = B(x, y(x,z), z) = B(x,z) = yz - A(x,y)$$

This is a Legendre transformation (cf thermodynamics). Furthermore, it follows from Eq. (24) that

$$\left.\frac{\partial B}{\partial z}\right|_x = y, \qquad \left.\frac{\partial B}{\partial x}\right|_z = -\left.\frac{\partial A}{\partial x}\right|_y.$$

Note that the same Legendre transformation takes $A(x,y)$ to $B(x,z)$ as takes $B(x,z)$ back to $A(x,y)$. $x$ is called a passive variable, and $y$ an active variable.

## The Hamiltonian

A Legendre transformation of a time-independent, 1 DoF Lagrangian $L(q,\dot{q})$, where the active variable is the velocity $\dot{q}$, the passive variable is the coordinate $q$, and the third variable is the canonically conjugate momentum $p$, yields the *Hamiltonian* $H$:

$$H(q,p) \equiv p\dot{q} - L(q,\dot{q}), \quad \text{where} \quad p \equiv \left.\frac{\partial L}{\partial \dot{q}}\right|_{\text{constant } q}. \tag{25}$$

Differentiating this equation and using the Euler-Lagrange equation, we find *Hamilton's equations*.

$$\left.\frac{\partial H}{\partial p}\right|_q = \dot{q}, \qquad \left.\frac{\partial H}{\partial q}\right|_p = -\left.\frac{\partial L}{\partial q}\right|_{\dot{q}} = -\frac{d}{dt}\left.\frac{\partial L}{\partial \dot{q}}\right|_q = -\dot{p}.$$

Note that, while the Lagrangian formulation of mechanics produces $N$ second order ordinary differential equations (in the $q_k$), the Hamiltonian treatment yields $2N$ first order ordinary differential equations (in $q_k$ and $p_k$).

## Hamilton's Principle

From Hamilton's principle, variations in the action $S = \int L dt$ are given by $\delta S = S[q + \delta q] - S[q] = \int \delta L dt = 0$. Using the relationship between $L$ and $H$ given by Eq. (25),

$$\begin{aligned}\delta L &= \dot{q}\delta p + p\delta\dot{q} - \underbrace{\left(\frac{\partial H}{\partial q}\delta q + \frac{\partial H}{\partial p}\delta p\right)}_{\delta H} \\ &= \dot{q}\delta p + \frac{d}{dt}(p\delta q) - \dot{p}\delta q - \frac{\partial H}{\partial q}\delta q - \frac{\partial H}{\partial p}\delta p = \left(\dot{q} - \frac{\partial H}{\partial p}\right)\delta p - \left(\dot{p} + \frac{\partial H}{\partial q}\right)\delta q + \frac{d}{dt}(p\delta q).\end{aligned}$$

Note that $d(p\delta q)/dt$ vanishes on integration, as $\delta q = 0$ at the endpoints of the physical path, and that Hamilton's equations imply that the $\delta q$, $\delta p$ terms also vanish. Hence, $\int \delta L dt = 0$ and independent infinitesimal variations in $\delta q$, $\delta p$ from the physical path in coordinate-momentum space (called *phase space* as opposed to the configuration space of Lagrangian mechanics) do not change the action, $S$.

## Hamilton's Equations of Motion

For $N$ DoF, the Hamiltonian is defined using the Legendre transformation

$$H \equiv \left(\sum_{k=1}^{N} p_k \dot{q}_k\right) - L,$$

where $L$ may now be explicitly time-dependent. Therefore include a time differential to give

$$\begin{aligned}dH &= \sum_{k=1}^{N}\left(p_k d\dot{q}_k + \dot{q}_k dp_k - \frac{\partial L}{\partial \dot{q}_k}d\dot{q}_k - \frac{\partial L}{\partial q_k}dq_k\right) - \frac{\partial L}{\partial t}dt \\ &= \sum_{k=1}^{N}(\dot{q}_k dp_k - \dot{p}_k dq_k) - \frac{\partial L}{\partial t}dt.\end{aligned}$$

We can also consider time as a passive variable in the Legendre transformation, which yields

$$\left.\frac{\partial H}{\partial t}\right|_{q_1,\ldots,p_1,\ldots} = -\left.\frac{\partial L}{\partial t}\right|_{q_1,\ldots,\dot{q}_1,\ldots} \Rightarrow \frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

If there is no explicit time dependence in $L$, $H$ is a constant of the motion. If $T$ is also a quadratic form in the $\dot{q}_k$, $H$ is the *total energy* $E = T + V$.

Finally, *Hamilton's equations of motion* in complete generality are:

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \qquad \frac{dH}{dt} = -\frac{\partial L}{\partial t}.$$

# 13 Canonical Transformations and Poisson Brackets

SEE ALSO HAND AND FINCH, 6.1–6.3

## Canonical Transformations

The Lagrangian formulation of mechanics makes coordinate transformations (i.e. $Q = Q(q,t)$) fairly straightforward. These types of change to new generalised coordinates and their associated generalised velocities are commonly called *point transformations*.

Coordinate transformations take place in configuration space. In *phase space* one can consider new coordinates $Q$ and momenta $P$ that are functions of the original coordinates $q$ and their momenta $p$ (and possibly the time $t$):

$$Q = Q(q,p,t), \qquad P = P(q,p,t).$$

These are commonly called *contact transformations*.

**A transformation is by definition *canonical*, if it preserves the structure of Hamilton's equations for all dynamical systems.** Hence, for a specific dynamical system, a new Hamiltonian $H'$ that is a function of $Q$ and $P$ is produced, and Hamilton's equations for $H'$, $P$, and $Q$ must describe the correct transformed dynamics.

## Equivalence of Lagrangians

Consider two Lagrangians $L$ and $L'$ such that $L'(q,\dot{q},t) = L(q,\dot{q},t) - dF(q,t)/dt$ ($F$ is independent of $\dot{q}$). Substituting $L'$ into the Euler-Lagrange equation yields

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} &= \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}}\left(\frac{dF}{dt}\right)\right] - \frac{\partial}{\partial q}\left(\frac{dF}{dt}\right) \\ &= \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}}\left(\dot{q}\frac{\partial F}{\partial q} + \frac{\partial F}{\partial t}\right)\right] - \frac{\partial}{\partial q}\left(\dot{q}\frac{\partial F}{\partial q} + \frac{\partial F}{\partial t}\right) \\ &= \frac{d}{dt}\left(\frac{\partial F}{\partial q}\right) - \left(\dot{q}\frac{\partial^2 F}{\partial q^2} + \frac{\partial^2 F}{\partial q \partial t}\right) = 0,\end{aligned}$$

i.e., the Euler-Lagrange equation derived from $L$. Hence, the descriptions of a physical system provided by $L$ and $L'$ are *equivalent*.

We now consider instead a starting Lagrangian $L(q,\dot{q})$ and a transformed Lagrangian $L'(Q,\dot{Q})$ related by a similar $F$ that is a function of both $q$ and $Q$, and perhaps $t$, such that

$$L'(Q,\dot{Q},t) = \lambda L(q,\dot{q},t) - \frac{dF(q,Q,t)}{dt} \tag{26}$$

($\lambda \neq 1$ is associated with a change of units or *scale transformation*. We will always consider $\lambda = 1$)

We next integrate Eq. (26) with respect to $t$ (and with $\lambda = 1$):

$$\int_{t_1}^{t_2} L'dt = \int_{t_1}^{t_2} Ldt + F(q(t_1),Q(t_1),t_1) - F(q(t_2),Q(t_2),t_2) \tag{27}$$

Recalling Hamilton's principle, if we take variations of Eq. (27) [assuming that arbitrary variations $\delta q(t)$ imply arbitrary variations $\delta Q(t)$] the variation of $\int_{t_1}^{t_2} L dt$ vanishes. If we further assume that $\delta F = 0$ at the end points, then we can say that the variation of $\int_{t_1}^{t_2} L' dt$ vanishes, and it follows that the two descriptions are equivalent.

## The Generating Function

$F(q, Q, t)$ is called a *generating function*. From the chain rule, the time derivative for the generating function is $dF/dt = (\partial F/\partial q)\dot{q} + (\partial F/\partial Q)\dot{Q} + \partial F/\partial t$

By construction, $\dot{q}$ does not appear explicitly in $L'$, and so, from Eq. (26)

$$0 = \frac{\partial L'}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{q}} - \frac{\partial}{\partial \dot{q}}\left(\frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q} + \frac{\partial F}{\partial t}\right) = \frac{\partial L}{\partial \dot{q}} - \frac{\partial F}{\partial q} \Rightarrow p = \frac{\partial F}{\partial q}. \tag{28}$$

Similarly

$$\frac{\partial L'}{\partial \dot{Q}} = -\frac{\partial F}{\partial Q} \Rightarrow P = -\frac{\partial F}{\partial Q}. \tag{29}$$

These are known as the *implicit transformation equations*.

With the implicit transformation equations we have two equations and two unknowns, $P(p, q)$ and $Q(p, q)$. To find an explicit form for the transformation, solve Eq. (28) to express $Q = Q(q, p, t)$, and then insert this relation into Eq. (29) to get $P = P(p, q, t)$.

Extension to $N > 1$ DoF is in principle straightforward, although in practice the task of determining $2N$ unknowns from $2N$ equations may be formidable.

For example, consider the generating function $F = qQ$, then

$$p = \frac{\partial F}{\partial q} = Q, \quad P = -\frac{\partial F}{\partial Q} = -q.$$

This particular generating function essentially interchanges the coordinates and momenta.

## The Transformed Hamiltonian

To find the new Hamiltonian $H'(Q, P)$, return to the definition of $H'$ in terms of a Legendre transformation (Eq. (25)). Hence,

$$\begin{aligned} H'(Q, P, t) \equiv & P\dot{Q} - L' = -\frac{\partial F}{\partial Q}\dot{Q} - L + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial Q}\dot{Q} + \frac{\partial F}{\partial t} = p\dot{q} - L + \frac{\partial F}{\partial t} \\ = & H(q, p, t) + \frac{\partial F(q, Q, t)}{\partial t} \end{aligned} \tag{30}$$

where we have made use of both Eq. (29) and Eq. (28). Finally, $p$ and $q$ appearing in the RHS of Eq. (30) must be expressed in terms of $P$ and $Q$.

If there is no explicit time dependence to $F$, then the new Hamiltonian $H'$ is simply the original Hamiltonian $H$, rewritten by inserting the inverse of the transformation equations expressing $P$ and $Q$ in terms of $p$ and $q$.

## Four Forms of Generating Function

The generating function need not be written in terms of $q$ and $Q$. Any combination of one original and one new dynamical variable may be employed, in conjunction with appropriately modified implicit transformation equations:

$$\begin{aligned} F_1(q,Q,t), &\qquad p=\frac{\partial F_1}{\partial q}, \quad P=-\frac{\partial F_1}{\partial Q};\\ F_2(q,P,t), &\qquad p=\frac{\partial F_2}{\partial q}, \quad Q=\frac{\partial F_2}{\partial P};\\ F_3(p,Q,t), &\qquad q=-\frac{\partial F_3}{\partial p}, \quad P=-\frac{\partial F_3}{\partial Q};\\ F_4(p,P,t), &\qquad q=-\frac{\partial F_4}{\partial p}, \quad Q=\frac{\partial F_4}{\partial P}. \end{aligned} \tag{31}$$

The different forms of generating function may be related via Legendre transformations, or the implicit transformation equations can simply be determined independently in a similar manner to that used for $F=F_1$ above.

## Poisson Brackets

For $N$ DoF, a *Poisson bracket* for two arbitrary functions $F$ and $G$ is defined as

$$\{F,G\}\equiv\sum_{k=1}^{N}\left(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k}-\frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k}\right)$$

For the case where $N=1$,

$$\{F,G\}\equiv\frac{\partial F}{\partial q}\frac{\partial G}{\partial p}-\frac{\partial F}{\partial p}\frac{\partial G}{\partial q}.$$

This is just the Jacobian defined by

$$J\left(\frac{F,G}{q,p}\right)\equiv\frac{\partial(F,G)}{\partial(q,p)}\equiv\begin{vmatrix}\frac{\partial F}{\partial q} & \frac{\partial G}{\partial q}\\ \frac{\partial F}{\partial p} & \frac{\partial G}{\partial p}\end{vmatrix}.$$

Clearly, $\{q,p\}=1$. If we consider a Poisson bracket of $q$ and $p$ defined with respect to the canonically transformed dynamical variables $Q$ and $P$, then (using Eq. (31) and the chain rule)

$$\frac{\partial q}{\partial Q}\frac{\partial p}{\partial P}-\frac{\partial q}{\partial P}\frac{\partial p}{\partial Q}=\frac{\partial q}{\partial Q}\frac{\partial^2 F_2}{\partial P\partial q}+\frac{\partial^2 F_4}{\partial P\partial p}\frac{\partial p}{\partial Q}=\frac{\partial Q}{\partial q}\frac{\partial q}{\partial Q}+\frac{\partial Q}{\partial p}\frac{\partial p}{\partial Q}=\frac{\partial Q}{\partial Q}=1,$$

i.e., the value of the Poisson bracket for $q$ and $p$ is independent of the representation. Conversely, $\{Q,P\}=1$ is a necessary and sufficient condition for a transformation to be canonical.

There is an equivalent chain rule for Jacobians whereby

$$\frac{\partial(F,G)}{\partial(Q,P)}=\frac{\partial(F,G)}{\partial(q,p)}\frac{\partial(q,p)}{\partial(Q,P)},$$

so

$$\frac{\partial F}{\partial Q}\frac{\partial G}{\partial P} - \frac{\partial F}{\partial P}\frac{\partial G}{\partial Q} = \frac{\partial F}{\partial q}\frac{\partial G}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial G}{\partial q}.$$

We therefore do not need to specify a particular set of coordinates and conjugate momenta when discussing Poisson brackets.

## Poisson Bracket Formulation of Hamilton's Equations

Hamilton's equations of motion can be reformulated in terms of Poisson brackets:

$$\dot{q} = \{q, H\}, \quad \dot{p} = \{p, H\}. \quad \text{For an arbitrary function } F(q, p, t), \quad \dot{F} = \{F, H\} + \frac{\partial F}{\partial t}.$$

# 14 Rotating Reference Frames

SEE ALSO HAND AND FINCH, 7.1–7.7

## Accelerating Reference Frames

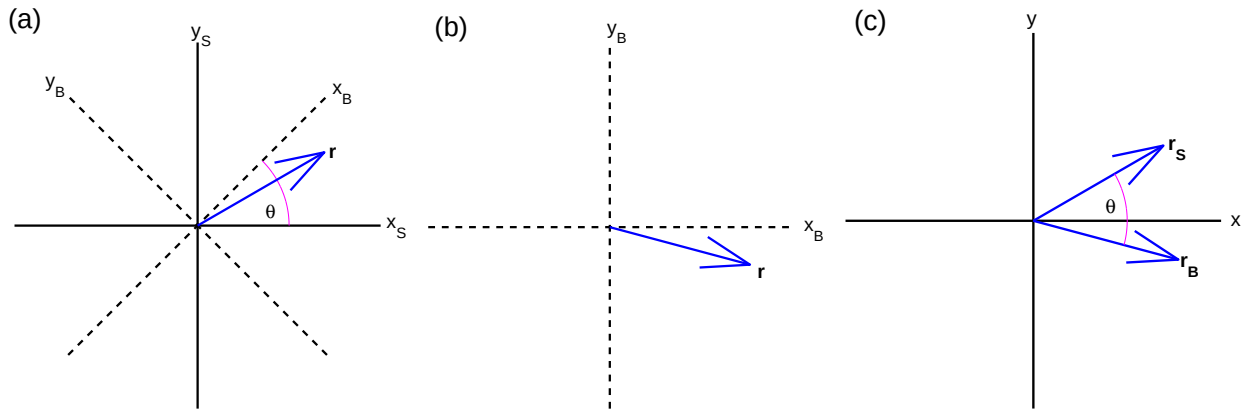
One can only specify the position $\underline{r}_S$ of a point mass with respect to a chosen *reference frame* $S$.

One can equally choose another reference frame $B$ to specify the position $\underline{r}_B$ of the same point mass, displaced by $\underline{R}$ from $S$. The positions of the point mass relative to the two frames are then related by $\underline{r}_B = \underline{r}_S - \underline{R}$.

If $S$ is an inertial frame where Newton's laws hold (fixed in Space), but $\ddot{\underline{R}} \neq 0$, then $B$ is *not* an inertial frame (attached to an accelerating Body). Hence, in frame $S$, $m\ddot{\underline{r}}_S = \underline{F}$, but in frame $B$,

$$m\ddot{\underline{r}}_B = m(\ddot{\underline{r}}_S - \ddot{\underline{R}}) = \underbrace{\underline{F}}_{\text{true force}} - \underbrace{m\ddot{\underline{R}}.}_{\text{fictitious force}}$$

## Rotated Reference Frames in 2D




(a) In frame $S$, the position vector $\underline{r}$ has positive $x$ and $y$ components. A new fixed reference frame $B$ is defined with respect to an original reference frame $S$ by an *anticlockwise* rotation through angle $\theta$.

(b) In frame $B$, $\underline{r}$ has a positive $x$ component and a *negative* y component.

(c) An observer who changes their observation frame from $S$ to $B$ finds that $\underline{r}$ appears to have rotated *clockwise* through an angle $\theta$, from $\underline{r}_S$ to $\underline{r}_B$.

## Rotations in 3D

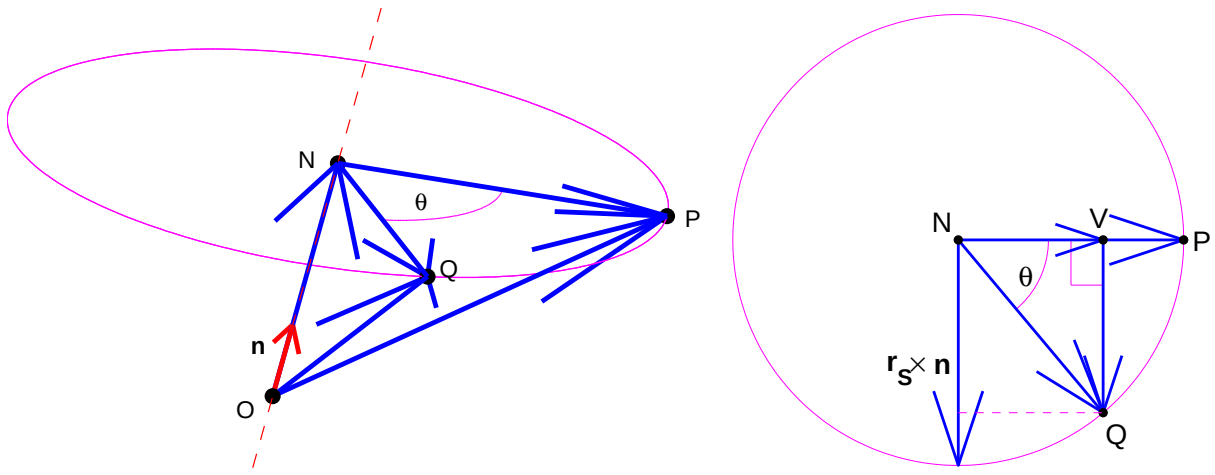



Figure 1: **Left**: Overall and **Right:** top views of a general *clockwise* rotation of a vector $\underline{OP} = \underline{r}_S$ through an angle $\theta$ about an axis defined by the unit vector $\underline{n}$, to become $\underline{OQ} = \underline{r}_B$. This is what we would observe if the reference frame we were in rotated in an *anticlockwise* direction.

We wish to determine $\underline{OQ} = \underline{ON} + \underline{NV} + \underline{VQ}$. Firstly,

$$\underline{ON} = (\underline{n} \cdot \underline{r}_S)\underline{n} \quad \text{(component of } \underline{r}_S \text{ in } \underline{n} \text{ direction)}.$$

Noting that $\underline{OP} = \underline{ON} + \underline{NP} \Rightarrow \underline{NP} = \underline{r}_S - (\underline{n} \cdot \underline{r}_S)\underline{n}$, and that $\underline{NQ}$ has the same *magnitude* as $\underline{NP}$, it follows that

$$\underline{NV} = \frac{\underline{NP}}{|\underline{NP}|}|\underline{NQ}|\cos(\theta) = [\underline{r}_S - (\underline{n} \cdot \underline{r}_S)\underline{n}]\cos(\theta)$$

We note that $\underline{r}_S \times \underline{n}$ has *direction* perpendicular to $\underline{r}_S = \underline{OP}$ and $\underline{n}$ (as does $\underline{VQ}$), and *magnitude* $|\underline{r}_S \times \underline{n}| = |\underline{NQ}|$. Hence,

$$\underline{VQ} = \frac{(\underline{r}_S \times \underline{n})}{|\underline{r}_S \times \underline{n}|}|\underline{NQ}|\sin(\theta) = (\underline{r}_S \times \underline{n})\sin(\theta).$$

Gathering the terms calculated above, we have

$$\underbrace{\underline{OQ}}_{\underline{r}_B} = \overbrace{(\underline{n} \cdot \underline{r}_S)\underline{n}}^{\underline{ON}} + \overbrace{[\underline{r}_S - (\underline{n} \cdot \underline{r}_S)\underline{n}]\cos(\theta)}^{\underline{NV}} + \overbrace{(\underline{r}_S \times \underline{n})\sin(\theta)}^{\underline{VQ}},$$

which is generally rearranged slightly to obtain the *rotation formula*:

$$\underline{r}_B = \underline{r}_S\cos(\theta) + (\underline{n} \cdot \underline{r}_S)\underline{n}[1 - \cos(\theta)] + (\underline{r}_S \times \underline{n})\sin(\theta). \tag{32}$$

## Infinitesimal Rotations

Consider a point fixed in frame $S$ at $\underline{r}_S = \underline{r}$. If frame $B$ is infinitesimally rotated w.r.t. $S$ by $d\theta$, then $\underline{r}_B = \underline{r} + d\underline{r}$ ($\Rightarrow d\underline{r} = \underline{r}_B - \underline{r}_S$). Taking the small-angle limit (i.e. $\theta = d\theta$), the rotation formula (Eq. (32)) transforms to $\underline{r} + d\underline{r} = \underline{r} + (\underline{r} \times \underline{n})d\theta$.

Hence, (noting that for any two vectors $\underline{U} \times \underline{V} = -\underline{V} \times \underline{U}$) the velocity of the point in reference frame $B$ relative to that in frame $S$ is given by:

$$\left[\frac{d\underline{r}}{dt}\right]_{\text{in } B} = -\left(\underline{n}\frac{d\theta}{dt}\right) \times \underline{r} = -\underbrace{\underline{\omega}}_{\textit{angular velocity}} \times \underline{r}.$$

Note that even if $\underline{r}$ is constant in reference frame $B$, it is time dependent if viewed from reference frame $S$.

## Velocity and Acceleration

If the vector $\underline{r}$ *does* have a time-dependence in reference frame $S$ (e.g. represents the position of a moving particle), then this must also be accounted for:

$$\underbrace{\left[\frac{d\underline{r}}{dt}\right]_{\text{in } B}}_{\text{velocity in } B} = \underbrace{\left[\frac{d\underline{r}}{dt}\right]_{\text{in } S}}_{\text{velocity in } S} - \underbrace{\underline{\omega} \times \underline{r}.}_{\text{velocity due to motion of } B \text{ relative to } S}$$

This equation is completely general for *any* vector $\underline{A}$, so

$$\left[\frac{d\underline{A}}{dt}\right]_{\text{in } B} = \left[\frac{d\underline{A}}{dt}\right]_{\text{in } S} - \underline{\omega} \times \underline{A}.$$

We can use the operator

$$\left[\frac{d}{dt}\right]_{\text{in } B} = \left(\left[\frac{d}{dt}\right]_{\text{in } S} - \underline{\omega}\times\right)$$

in conjunction with $\underline{v}_B = \underline{v}_S - \underline{\omega} \times \underline{r}$ to determine the acceleration in the non-inertial frame $B$.

$$\begin{aligned}
\left[\frac{d\underline{v}_B}{dt}\right]_{\text{in } B} &= \left(\left[\frac{d}{dt}\right]_{\text{in } S} - \underline{\omega}\times\right)\underline{v}_B \\
&= \left[\frac{d}{dt}\right]_{\text{in } S}(\underline{v}_S - \underline{\omega} \times \underline{r}) - \underline{\omega} \times \underline{v}_B \\
&= \underline{a}_S - \dot{\underline{\omega}} \times \underline{r} - \underline{\omega} \times \underline{v}_S - \underline{\omega} \times \underline{v}_B \\
&= \underline{a}_S - \dot{\underline{\omega}} \times \underline{r} - \underline{\omega} \times (\underline{v}_B + \underline{\omega} \times \underline{r}) - \underline{\omega} \times \underline{v}_B \\
&= \underline{a}_S - 2\underline{\omega} \times \underline{v}_B - \underline{\omega} \times (\underline{\omega} \times \underline{r}) - \dot{\underline{\omega}} \times \underline{r}
\end{aligned}$$

## Inertial Forces in a Rotating Frame

If considering motion of a point mass, then multiplying $\left[\frac{d\underline{v}_B}{dt}\right]_{\text{in } B}$ by $m$ and noting that $\underline{F} = m(d\underline{v}_S/dt)_S$, we deduce that

$$m\ddot{\underline{r}} = \underline{F} - \underbrace{2m\underline{\omega} \times \dot{\underline{r}}}_{\text{coriolis force}} - \underbrace{m\underline{\omega} \times (\underline{\omega} \times \underline{r})}_{\text{centrifugal force}} - \underbrace{m\dot{\underline{\omega}} \times \underline{r}.}_{\text{Euler force}}$$

# 15 Inertial Forces on Earth

SEE ALSO HAND AND FINCH, 7.7–7.10

## Procedure for Determining a Local Coordinate Systems

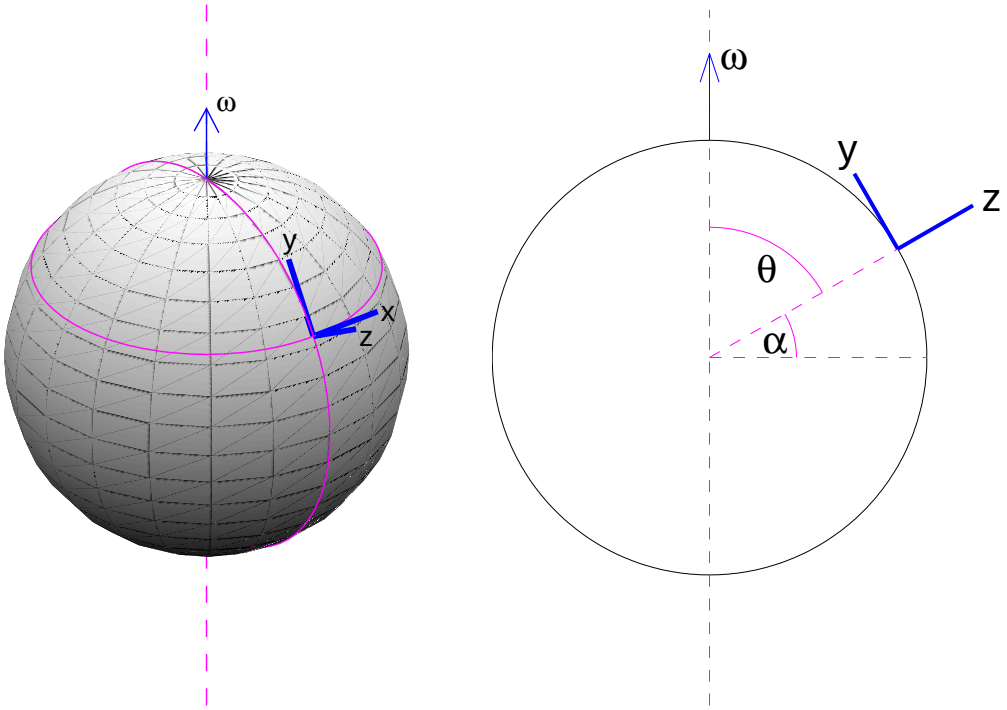



Figure 2: **Left:** Overall view of a local coordinate system in the region of a point on the surface of the spherical earth. The positive $x$ direction is east, the positive $y$ direction is north, the positive $z$ direction is out from the centre of the earth. **Right:** Cross-sectional view, showing the local $y$ and $z$ axes only (the $x$ axis points perpendicularly into the page). $\alpha$ is the *latitude*, and $\theta$ the *colatitude* (in spherical coordinates, the angle of elevation).

1. Model the Earth as a perfect, solid sphere, and require an appropriate coordinate system to describe motion near a point on its surface. Start with a cartesian coordinate system with the $z$ axis running through the poles, and the $x$ and $y$ axes emerging through the equator (in this frame $\underline{\omega} = (0, 0, \omega)$).

2. Rotate this coordinate system around its $x$ axis so that the $z$ axis emerges at a chosen point on the Earth's surface. In these new coordinates, this looks like a rotation of $-\theta$ about the x axis, thus the angular velocity in the local frame is

$$\underline{\omega} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix} = \begin{pmatrix} 0 \\ \omega\sin\theta \\ \omega\cos\theta \end{pmatrix}$$

3. Finally, we displace the origin of the coordinate system by $\underline{R}$ to the point where the local $z$ axis emerges from the Earth's surface (where $R = |\underline{R}|$ is the radius of the Earth).

## Inertial Forces in the Local Coordinate System

Note that the derivation of centrifugal, coriolis, and Euler force terms assumes $\underline{r}$ to be defined with respect to a point on the *axis of rotation*.

The final displacement of the local coordinate system origin means the associated position vector $\underline{r}'$ does not begin on the axis of rotation.

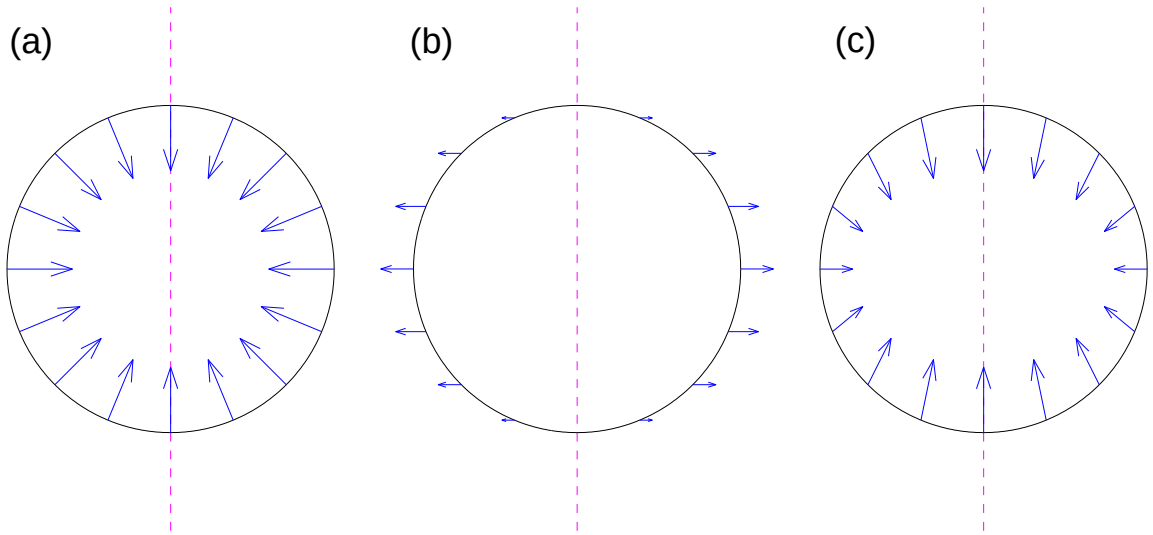
One must therefore substitute $\underline{r} = \underline{r}' + \underline{R}$ into the inertial force terms, where $\underline{r}'$ is the position vector associated with the *local* coordinate system.

## Centrifugal Force $-m\underline{\omega} \times (\underline{\omega} \times \underline{r})$

The direction of the centrifugal force can be determined quite generally: $\underline{\omega} \times \underline{r}$ is perpendicular to $\underline{\omega}$, $\underline{r}$, and $\underline{\omega} \times (\underline{\omega} \times \underline{r})$ is perpendicular to both $\underline{\omega}$ and $\underline{\omega} \times \underline{r}$.

Hence, the centrifugal force $-m\underline{\omega} \times (\underline{\omega} \times \underline{r})$ points *directly away* from the axis of rotation, meaning it can *counteract gravity*.

## Effective Gravity on Earth's Surface




(a) Direction and magnitude of the force due to earth's gravitation on the surface of a spherical earth model.

(b) Direction of the centrifugal force on the surface of a spherical earth model (magnitude relative to the gravitational force greatly exaggerated for effect).

(c) Effect of adding the gravitational and centrifugal forces together. The net force for a static object *does not* point exactly to the centre of the earth except at the poles and the equator.

## Coriolis Force $-2m\underline{\omega} \times \underline{\dot{r}}$

In the local coordinate system defined for the surface of the Earth, the coriolis force is given by

$$-2m\underline{\omega} \times \underline{\dot{r}}' = -2m\omega(\dot{z}\sin\theta - \dot{y}\cos\theta, \dot{x}\cos\theta, -\dot{x}\sin\theta)$$

Note that it is proportional to *velocities*, and therefore only comes into play when there is *motion*.

One often considers phenomena when motion is effectively constrained to be "in plane," when we ignore the $z$ dimension. In this case the coriolis force $x$ component $\rightarrow 2m\omega\dot{y}\cos\theta$, and the coriolis force $y$ component $\rightarrow -2m\omega\dot{x}\cos\theta$.

## Loose Ends

The Euler force is not zero on Earth, but is commonly taken to be negligible.

In a rotating reference frame that is also accelerating translationally with acceleration $\underline{\ddot{R}}$, the general EOM is given by

$$m\underline{\ddot{r}} = \underline{F} - 2m\underline{\omega} \times \underline{\dot{r}} - m\underline{\omega} \times (\underline{\omega} \times \underline{r}) - m\underline{\dot{\omega}} \times \underline{r} - m\underline{\ddot{R}}.$$

# 16 Rotational Inertia, Angular Momentum and Kinetic Energy

SEE ALSO HAND AND FINCH, 8.1–8.3

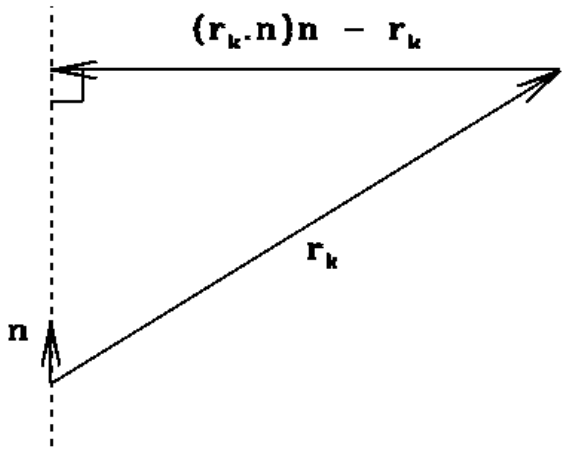
## DoF of Rigid Bodies

A rigid body is a system of mass points (or a continuum) constrained so that distances between all points are constant (e.g. connected by massless rigid rods). A system of 2 connected point masses has 5 DoF, and a system of 3 connected point masses has 6 DoF.

Connecting a further point mass (e.g. by massless rigid rods) to 3 "original" point masses completely specifies its position relative to the rigid body, and contributes 3 coordinates $-3$ constraints $= 0$ additional DoF. **Any rigid body more complex than two connected point masses has 6 DoF.**

These 6 DoF correspond to a translation (3 DoF) and a rotation (2 DoF to define a rotation axis and 1 DoF for the rotation).

## Moment of Inertia $I$ about an Axis of Rotation




Consider $N$ mass points, rotating with angular velocity $\omega$ anticlockwise about an axis of rotation (assumed to pass through the origin), where the direction of the axis of rotation is defined by the unit vector $\underline{n}$ (i.e., the angular velocity vector $\underline{\omega} = \omega\underline{n}$).

Each mass point has a moment of inertia that is the square of its distance from the axis of rotation multiplied by its mass $m$. For $N$ mass points, $I$ is the sum of such terms. Hence,

$$I = \sum_{k=1}^{N} m_k |(\underline{r}_k \cdot \underline{n})\underline{n} - \underline{r}_k|^2 = \sum_{k=1}^{N} m_k [r_k^2 - (\underline{r}_k \cdot \underline{n})^2].$$

## Angular Momentum $\underline{J}$ and Rotational Kinetic Energy

The angular momentum (about the origin) of a system of mass points is given by $\underline{J} = \sum_{k=1}^{N} \underline{r}_k \times \underline{p}_k = \sum_{k=1}^{N} m_k \underline{r}_k \times \dot{\underline{r}}_k$, (if the point about which we calculate $\underline{J}$ is $\underline{R} \neq \underline{0}$, then we replace $\underline{r}_k$, $\dot{\underline{r}}_k$ by $\underline{r}'_k$, $\dot{\underline{r}}'_k$, where $\underline{r}'_k = \underline{r}_k - \underline{R}$).
If rigid body constraints are imposed, any velocity $\dot{\underline{r}}_k$ is due to *rotation* only.
$\underline{\omega} \times \underline{r}_k$ is the velocity of the $k$th mass point due to its anticlockwise rotation.
Using the vector identity $\underline{a} \times (\underline{b} \times \underline{c}) = (\underline{a} \cdot \underline{c})\underline{b} - (\underline{a} \cdot \underline{b})\underline{c}$ and recalling $\underline{\omega} = \omega \underline{n}$,

$$\underline{J} = \sum_{k=1}^{N} m_k [\underline{r}_k \times (\underline{\omega} \times \underline{r}_k)] = \sum_{k=1}^{N} m_k [r_k^2 \underline{\omega} - (\underline{r}_k \cdot \underline{\omega})\underline{r}_k]$$

$$\Rightarrow \underline{J} \cdot \underline{n} = \omega \sum_{k=1}^{N} m_k [r_k^2 - (\underline{r}_k \cdot \underline{n})^2] = \omega I.$$

Similarly, the kinetic energy reduces to $T = \frac{1}{2} \sum_{k=1}^{N} m_k \dot{r}_k^2 = \frac{\omega^2}{2} \sum_{k=1}^{N} m_k |\underline{n} \times \underline{r}_k|^2$. Noting that $|\underline{n} \times \underline{r}_k| = |(\underline{r}_k \cdot \underline{n})\underline{n} - \underline{r}_k|$, the rotational kinetic energy takes a very simple form:

$$T = \frac{\omega^2}{2} \sum_{k=1}^{N} m_k |(\underline{r}_k \cdot \underline{n})\underline{n} - \underline{r}_k|^2 = \frac{I\omega^2}{2}.$$

## Definition of the Inertia Tensor $\hat{I}$

From $\underline{J} = \sum_{k=1}^{N} m_k [r_k^2 \underline{\omega} - (\underline{r}_k \cdot \underline{\omega})\underline{r}_k]$, we infer that

$$J_x = \sum_{k=1}^{N} m_k (r_k^2 - x_k^2)\omega_x - m_k x_k y_k \omega_y - m_k x_k z_k \omega_z,$$
$$J_y = \sum_{k=1}^{N} m_k (r_k^2 - y_k^2)\omega_y - m_k y_k x_k \omega_x - m_k y_k z_k \omega_z,$$
$$J_z = \sum_{k=1}^{N} m_k (r_k^2 - z_k^2)\omega_z - m_k z_k x_k \omega_x - m_k z_k y_k \omega_y,$$

and use the linear transformation relating the angular velocity vector $\underline{\omega}$ to the angular momentum vector $\underline{J}$ to define the *inertia tensor* $\hat{I}$:

$$\underbrace{\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix}}_{\underline{J}} = \underbrace{\begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}}_{\hat{I}} \underbrace{\begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}}_{\underline{\omega}}$$

$$\text{with} \quad I_{\alpha\beta} = \sum_{k}^{N} m_k (r_k^2 \delta_{\alpha\beta} - r_{k,\alpha} r_{k,\beta}), \qquad \text{where } \delta_{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}$$

For *continuous* rigid bodies $I_{\alpha\beta} = \int_{\text{volume}} dxdydz\rho(x,y,z)(r^2\delta_{\alpha\beta} - r_\alpha r_\beta)$, where $\rho$ is a mass density.

## Generalised Description of the Rotational Kinetic Energy

Using the vector identity $\underline{a} \cdot (\underline{b} \times \underline{c}) = \underline{b} \cdot (\underline{c} \times \underline{a}) = \underline{c} \cdot (\underline{a} \times \underline{b})$, the rotational kinetic energy can be phrased more generally in terms of the moment of inertia tensor as

$$T = \frac{1}{2}\sum_{k=1}^{N} m_k \dot{r}_k^2 = \frac{1}{2}\sum_{k=1}^{N} m_k \dot{\underline{r}}_k \cdot (\underline{\omega} \times \underline{r}_k) = \frac{1}{2}\underline{\omega} \cdot \underbrace{\left[\sum_{k=1}^{N} m_k (\underline{r}_k \times \dot{\underline{r}}_k)\right]}_{\underline{J} = \hat{I}\underline{\omega}}.$$

Hence $T = (1/2)\underline{\omega} \cdot (\hat{I}\underline{\omega}) = (\omega^2/2)\underline{n}^T \hat{I}\underline{n}$, where $\underline{n}^T \hat{I}\underline{n}$ is the moment of inertia $I$ about the axis defined by $\underline{n}$.

## Centre of Mass (CoM)

The centre of mass position and velocity for a system of $N$ mass points (mass of $k$th point $= m_k$) are defined by (total mass $M = \sum_{k=1}^{N} m_k$):

$$\underline{R}_C = \frac{\sum_{k=1}^{N} m_k \underline{r}_k}{M}, \qquad \dot{\underline{R}}_C = \frac{\sum_{k=1}^{N} m_k \dot{\underline{r}}_k}{M}.$$

It immediately follows that the *total* momentum $\underline{P} = \sum_{k=1}^{N} m_k \dot{\underline{r}}_k = M\dot{\underline{R}}_C$ is equivalent to the momentum of a single particle of mass $M = \sum_{k=1}^{N} m_k$ moving with the CoM.

## Separating the CoM Angular Momentum and Kinetic Energy

The total angular momentum about some point (situated at the origin of the coordinate system) is $\underline{J} = \sum_{k=1}^{N} m_k (\underline{r}_k \times \dot{\underline{r}}_k)$.
Let $\underline{r}_k = \underline{r}'_k + \underline{R}_C$. Hence,

$$\begin{aligned}
\underline{J} &= \sum_{k=1}^{N} m_k (\underline{r}'_k + \underline{R}_C) \times (\dot{\underline{r}}'_k + \dot{\underline{R}}_C) \\
&= \sum_{k=1}^{N} m_k (\underline{r}'_k \times \dot{\underline{r}}'_k + \underline{R}_C \times \dot{\underline{r}}'_k + \underline{r}'_k \times \dot{\underline{R}}_C + \underline{R}_C \times \dot{\underline{R}}_C) \\
&= \sum_{k=1}^{N} m_k (\underline{r}'_k \times \dot{\underline{r}}'_k) + M\underline{R}_C \times \underbrace{\frac{\sum_{k=1}^{N} m_k \dot{\underline{r}}'_k}{M} + \frac{\sum_{k=1}^{N} m_k \underline{r}'_k}{M}}_{\text{CoM velocity and position with respect to CoM, i.e., = 0}} \times M\dot{\underline{R}}_C + M\underline{R}_C \times \dot{\underline{R}}_C.
\end{aligned}$$

$\underline{J} = \sum_{k=1}^{N} m_k (\underline{r}'_k \times \dot{\underline{r}}'_k) + M\underline{R}_C \times \dot{\underline{R}}_C$ can be expressed as the angular momentum *about* the CoM plus the angular momentum of the CoM about the *origin*.
Similarly, the CoM kinetic energy, and that for motion with respect to the CoM can be separated:

$$
\begin{aligned}
T =& \frac{1}{2}\sum_{k=1}^{N} m_k \dot{r}_k^2 = \frac{1}{2}\sum_{k=1}^{N} m_k \underline{\dot{r}}_k \cdot \underline{\dot{r}}_k = \frac{1}{2}\sum_{k=1}^{N} m_k (\underline{\dot{r}}'_k + \underline{\dot{R}}_C) \cdot (\underline{\dot{r}}'_k + \underline{\dot{R}}_C) \\
=& \frac{1}{2}\sum_{k=1}^{N} m_k \dot{r}_k'^2 + \underbrace{\underline{\dot{R}}_C \cdot \sum_{k=1}^{N} m_k \underline{\dot{r}}'_k}_{=0} + \frac{1}{2} M \dot{R}_C^2.
\end{aligned}
$$

In general, if the external forces acting on a rigid body act through the CoM (e.g. gravity), then the CoM motion can be considered separately.

# 17 The Parallel and Principal Axis Theorems

SEE ALSO HAND AND FINCH, 8.2,8.7

## Displaced Axis Theorem

In a similar fashion to the angular momentum and rotational kinetic energy, one can insert the CoM position $\underline{R}_C$ into the elements of the Inertia tensor $I_{\alpha\beta}$. With $\underline{r}_k = \underline{r'}_k + \underline{R}_C$,

$$\begin{aligned} I_{\alpha\beta} &= \sum_{k=1}^{N} m_k(r_k^2\delta_{\alpha\beta} - r_{k,\alpha}r_{k,\beta}) = \sum_{k=1}^{N} m_k[|\underline{r'}_k + \underline{R}_C|^2\delta_{\alpha\beta} - (r'_{k,\alpha} + R_{C,\alpha})(r'_{k,\beta} + R_{C,\beta})] \\ &= \sum_{k=1}^{N} m_k[(r_k'^2 + 2\underline{r'}_k \cdot \underline{R}_C + R_C^2)\delta_{\alpha\beta} - (r'_{k,\alpha}r'_{k,\beta} + r'_{k,\alpha}R_{C,\beta} + R_{C,\alpha}r'_{k,\beta} + R_{C,\alpha}R_{C,\beta})] \\ &= \sum_{k=1}^{N} m_k(r_k'^2\delta_{\alpha\beta} - r'_{k,\alpha}r'_{k,\beta}) + M(R_C^2\delta_{\alpha\beta} - R_{C,\alpha}R_{C,\beta}) \\ &\quad + \underbrace{2\underline{R}_C \cdot \sum_{k=1}^{N} m_k\underline{r'}_k\delta_{\alpha\beta}}_{=0} - \underbrace{R_{C,\alpha}\sum_{k=1}^{N} m_k r'_{k,\beta}}_{=0} - \underbrace{R_{C,\beta}\sum_{k=1}^{N} m_k r'_{k,\alpha}}_{=0} \end{aligned}$$

We have used the fact that the CoM position with respect to the CoM is $= 0$.

Hence, the inertia tensor of a rigid body, defined with respect to rotations about the origin can be related to the inertia tensor of the same rigid body, defined with respect to rotations about its CoM, through

$$\hat{I} = \hat{I}_{\text{CoM}} + M\hat{A},$$

where $\hat{A}$ can be represented by a matrix, the elements of which are determined by the elements of the CoM position vector: $A_{\alpha\beta} = R_C^2\delta_{\alpha\beta} - R_{C,\alpha}R_{C,\beta}$. **This is the displaced axis theorem.**

## Parallel Axis Theorem

Consider a moment of inertia $I_{\text{COM}}$ about an axis of rotation passing through the CoM. This is equivalent to considering a diagonal element of some representation of the inertia tensor.

From the *displaced axis theorem*, the moment of inertia $I$ about a parallel axis is

$$I = I_{\text{CoM}} + Md^2,$$

where $d$ is the distance of parallel axis from the axis of rotation passing through the CoM. **This is the parallel axis theorem**.

## Symmetric Nature of the Inertia Tensor

The inertia tensor $\hat{I}$ is defined for a particular rigid body, and for rotations of that body about a particular point (frequently the centre of mass).

In any representation (e.g. for some set of cartesian axes, conveniently chosen with the origin at the point of rotation) the inertia tensor is represented by a *symmetric* matrix, i.e., $I_{\alpha\beta} = I_{\beta\alpha}$. We say $\hat{I}$ is a *symmetric tensor*.

The *matrix* elements depend on the chosen representation (the *tensor* is defined representation free). We wish to know if there is a *diagonal* matrix representation of the inertia tensor.

## Orthogonal Matrices and Orthogonal Diagonalization

The inverse $\hat{P}^{-1}$ of a matrix $\hat{P}$ is defined such that $\hat{P}^{-1}\hat{P} = \hat{P}\hat{P}^{-1} = \hat{1}$, the identity matrix. A matrix $\hat{P}$ for which an inverse $\hat{P}^{-1}$ exists is called invertible.
A real invertible square matrix $\hat{P}$ such that its inverse $\hat{P}^{-1} = \hat{P}^T$, its transpose, is said to be an *orthogonal matrix*. (So-called because both its rows and columns form orthonormal sets).
A general $2 \times 2$ orthogonal matrix has a single free parameter $\theta$, and describes a 2-dimensional rotation when acting on a 2-dimensional column vector:

$$\hat{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The inverse matrix describes a rotation back in the opposite direction.
Similarly, in 3 dimensions, one can consider separate rotations about each axis:

$$\hat{P}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \hat{P}_y = \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix},$$
$$\hat{P}_z = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$(\hat{P}_z\hat{P}_y\hat{P}_x)^T\hat{P}_z\hat{P}_y\hat{P}_x = \hat{P}_x^T\hat{P}_y^T\hat{P}_z^T\hat{P}_z\hat{P}_y\hat{P}_x = \hat{1}$, and the product $\hat{P}_z\hat{P}_y\hat{P}_x$ is a way of formulating a general $3 \times 3$ orthogonal matrix, describing a general rotation in 3-dimensions.
A square matrix $\hat{A}$ is called *orthogonally diagonalizable* if there is an invertible matrix $\hat{P}$ such that $\hat{P}^{-1}\hat{A}\hat{P} = \hat{P}^T\hat{A}\hat{P} = \hat{D}$ is diagonal; $\hat{P}$ *orthogonally diagonalizes* $\hat{A}$. **If $\hat{A}$ is an $N \times N$ matrix, then the following are equivalent. (a) $\hat{A}$ is orthogonally diagonalizable; (b) $\hat{A}$ has an orthonormal set of $N$ eigenvectors; (c) $\hat{A}$ is symmetric.**
Consider the $3 \times 3$ case. Let us say that a $3 \times 3$ matrix $\hat{A}$ has 3 orthonormal eigenvectors $\underline{p}_1 = (p_{11}, p_{21}, p_{31})$, $\underline{p}_2 = (p_{12}, p_{22}, p_{32})$, $\underline{p}_3 = (p_{13}, p_{23}, p_{33})$ with corresponding eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$.
Let $\hat{P}$ be the matrix whose columns are these eigenvectors:

$$\hat{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}$$

This implies that

$$\hat{A}\hat{P} = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \lambda_3 p_{13} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \lambda_3 p_{23} \\ \lambda_1 p_{31} & \lambda_2 p_{32} & \lambda_3 p_{33} \end{pmatrix} = \underbrace{\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}}_{\hat{P}} \underbrace{\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}}_{\hat{D}}.$$

Since the columns of $\hat{P}$ are orthonormal, $\hat{P}$ is an orthogonal matrix. Hence, $\hat{A}\hat{P} = \hat{P}\hat{D} \Rightarrow \hat{P}^{-1}\hat{A}\hat{P} = \hat{P}^T\hat{A}\hat{P} = \hat{D}$, i.e., $\hat{A}$ is orthogonally diagonalizable.
To help see that $\hat{A}$ must be symmetric, i.e., that $\hat{A}^T = \hat{A}$, we use $\hat{D} = \hat{P}^{-1}\hat{A}\hat{P} = \hat{P}^T\hat{A}\hat{P} \Rightarrow \hat{P}\hat{D}\hat{P}^{-1} = \hat{P}\hat{D}\hat{P}^T = \hat{A}$. Hence,

$$\hat{A}^T = (\hat{P}\hat{D}\hat{P}^T)^T = \hat{P}\hat{D}^T\hat{P}^T = \hat{P}\hat{D}\hat{P}^T = \hat{A}.$$

## Principal Axis Theorem

$\hat{I}$ is a symmetric tensor, and is always represented by a $3 \times 3$ symmetric matrix. Therefore, $\hat{I}$ has three eigenvalues (which may be repeated). These are the *principal moments of inertia*, called $I_1$, $I_2$, $I_3$.

$\hat{I}$ must also have three orthonormal eigenvectors (in the case of repeated eigenvalues, there is some freedom in choosing what form the corresponding eigenvectors should take). The directions the eigenvectors point in determine the *principal axes*, labelled 1, 2, 3.

It is always possible to find a rotated coordinate system such that the representation of $\hat{I}$ is diagonal. The diagonal elements are the principal moments of inertia (eigenvalues of $\hat{I}$), the coordinate system is set by the principal axes (eigenvectors of $\hat{I}$), and a principal moment of inertia $I_k$ is the moment of inertia associated with rotation about the principal axis $k$.

Rotating the coordinate system such that the representation of $\hat{I}$ is diagonal is a *principal axis transformation*, and that this is always possible is the *principal axis theorem*.

## Deducing the Principal Axes

For highly symmetric objects, it is often possible to guess a coordinate system for which all off-diagonal elements of $\hat{I}$ are $= 0$.

One can then deduce the principal moments of inertia directly, without having to carry out a principal axis transformation.

Prolate ellipsoids (e.g. rugby ball or an egg) have $I_3 < I_1 = I_2$, whereas oblate ellipsoids (smartie or to a small extent, the Earth) have $I_3 > I_1 = I_2$.

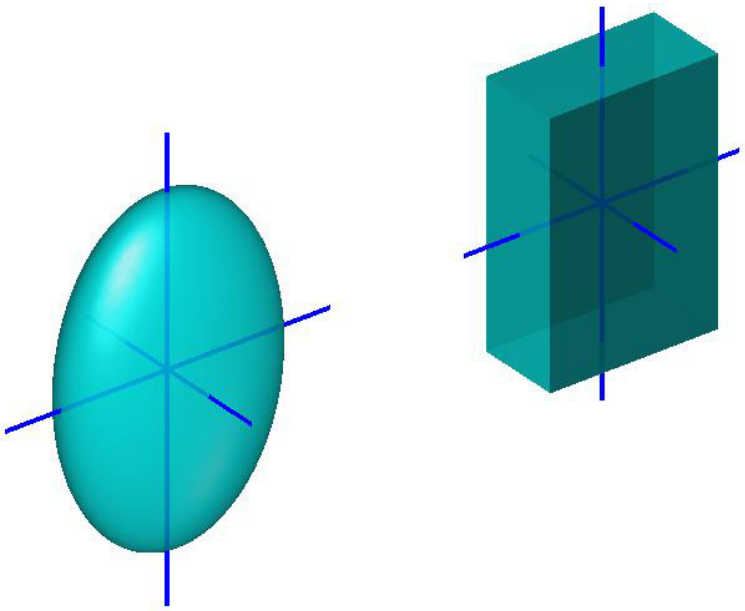


Figure 3: Principal axes of an asymmetric ellipsoid, and an asymmetric parallelepiped, for rotation about the centre of mass. The principal moments of inertia have been chosen to be identical for the two objects.

# 18 Rigid Body Dynamics and Stability

SEE ALSO HAND AND FINCH, 8.4–8.5

## Euler's Equations of Motion

We consider $N$ mass points subject to rigid-body constraints, and take the time derivative of the angular momentum vector $\underline{J} = \sum_{k=1}^{N} m_k(\underline{r}_k \times \dot{\underline{r}}_k)$, to get

$$\dot{\underline{J}} = \underbrace{\sum_{k=1}^{N} m_k(\dot{\underline{r}}_k \times \dot{\underline{r}}_k)}_{=0,\text{ as } \underline{a}\times\underline{a}=0 \text{ for any vector } \underline{a}} + \sum_{k=1}^{N} m_k(\underline{r}_k \times \ddot{\underline{r}}_k) = \underbrace{\sum_{k=1}^{N} \underline{r}_k \times \underline{F}_k}_{\text{total torque } \underline{N}}.$$

Hence, we can say, viewing the rigid body from outside (*space* axes/coordinate system) $(d\underline{J}/dt)_S = \underline{N}$.

Recalling that, for any vector $\underline{a}$, $(d\underline{a}/dt)_B = (d\underline{a}/dt)_S - \underline{\omega} \times \underline{a}$, viewing the rigid body while travelling with its rotational motion (*body* axes/coordinate system) we can say that $(d\underline{J}/dt)_B + \underline{\omega} \times \underline{J} = \underline{N}$.

We assume a coordinate system defined by the principal axes, i.e., the decomposition $\underline{J} = (I_1\omega_1, I_2\omega_2, I_3\omega_3)$, and thus derive *Euler's equations of motion*:

$$\begin{aligned} I_1\dot{\omega}_1 - \omega_2\omega_3(I_2 - I_3) &= N_1, \\ I_2\dot{\omega}_2 - \omega_3\omega_1(I_3 - I_1) &= N_2, \\ I_3\dot{\omega}_3 - \omega_1\omega_2(I_1 - I_2) &= N_3. \end{aligned}$$

## Torque-Free Motion

If there are no external torques (e.g. a freely spinning object moving through the Earth's gravitational field), then we can set $N_1 = N_2 = N_3 = 0$.

Note that the total angular momentum (about the CoM) and the total kinetic energy (associated with motion about the CoM) are both conserved.

$J^2 = J_1^2 + J_2^2 + J_3^2$ describes a *sphere* in angular momentum space, and $T = J_1^2/2I_1 + J_2^2/2I_2 + J_3^2/2I_3$ describes an *ellipsoid*.

Where the sphere and ellipsoid intersect shows the possible trajectories for the angular momentum components $J_k = \omega_k I_k$.

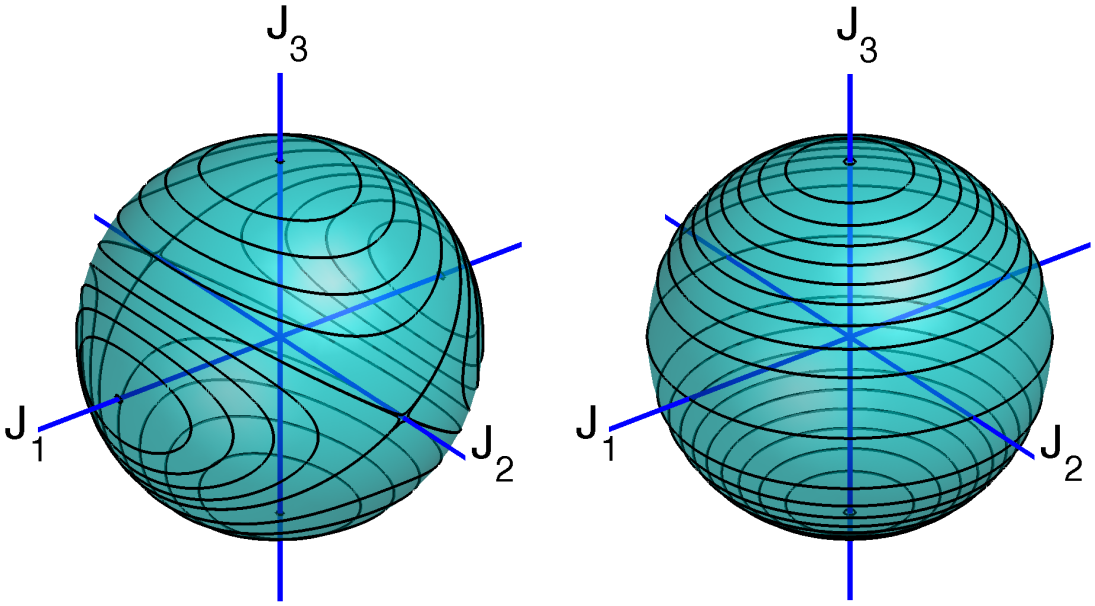



Figure 4: **Left:** Contours of constant $T$ on a sphere described by $J^2 = 1$, for a system where $I_1 > I_2 > I_3$. Rotation about principal axis 2 is *unstable*, rotation about principal axes 1 and 3 is *stable*. **Right:** Equivalent diagram for the case where $I_1 = I_2 > I_3$. Principal axes 1 and 2 can be rotated freely about principal axis 3, and there are no points on the $J^2 = 1$ sphere about which the rotation is unstable.

# Workshop 1: Applications of the Lagrangian Approach

1. Consider a point mass (of mass $m$) falling vertically in the Earth's gravitational field (take the acceleration $= g$ to be constant, i.e., independent of the distance of the point mass from the Earth's surface, and ignore air resistance).

   (a) What is the kinetic energy $T$ of the point mass? What is the potential energy $V$? Hence, determine the Lagrangian $L = T - V$.

   (b) Use the Euler-Lagrange equation to determine a second order differential equation for the vertical motion ($z$ direction) of the point mass.

   (c) Find the equation of motion, and express the solution in terms of the point mass's initial position $z(0)$ and velocity $\dot{z}(0)$.

2. Consider a point mass (of mass $m$) attached to a spring (spring constant $k$), constrained somehow to move only in the $x$ direction, where $x = 0$ is taken to be the equilibrium position of the point mass. This is just a one-dimensional simple harmonic oscillator (undamped), with characteristic oscillation frequency $\omega = \sqrt{k/m}$.

   (a) What is the kinetic energy $T$ of the point mass? What is the potential energy $V$? Hence, determine the Lagrangian $L = T - V$.

   (b) Use the Euler-Lagrange equation to determine an equation for $\ddot{x}$.

   (c) Solve this linear, homogeneous, second-order differential equation using the trial solution $x \propto e^{\lambda t}$.

   (d) Find the general solution for the equation of motion, in terms of the point mass's initial position $x(0)$ and velocity $\dot{x}(0)$. Note that the position and velocity are both real, not complex.

   (e) The two independent real constant prefactors in the general oscillator solution, $A$ and $B$, can equally well be represented in terms of two other real constants $A'$ and $\phi'$, through $A = A' \sin\phi'$, and $B = A' \cos\phi'$. Use this, along with the trigonometric identity $\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$, to show your solution can also be expressed in the form

   $$x(t) = A' \sin(\omega t + \phi'),$$

   and express $A'$ and $\phi'$ in terms of the initial position and velocity.

3. Consider the situation of a small block (SB) of mass $m$ sliding on a frictionless inclined plane (IP) of height $h$, which itself has a mass $M$ and rests on a flat surface, without any friction between this surface and the inclined plane. Consider a set of cartesian axes such that the $Y$ axis points up along the side of the initial location of the inclined plane, and the $X$ axis points along the underside of the inclined plane (we ignore the $Z$ direction). Take $(X_{\text{SB}}, Y_{\text{SB}})$ and $(X_{\text{IP}}, Y_{\text{IP}})$ to be the locations of the centre of mass of the SB, and the right-angled corner of the IP in this coordinate system, respectively.

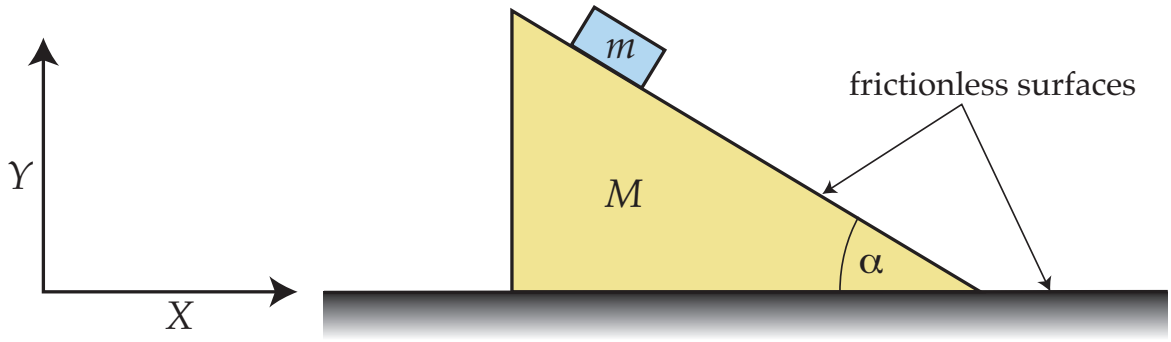
   (a) Because the motion of the SB is constrained to remain on the upper surface of the IP, the value for the dynamical variable $Y_{\text{SB}}$ can be determined exactly if the values for the dynamical variables $X_{\text{SB}}$, $X_{\text{IP}}$ are known. Derive an equation (a constraint equation) describing this relationship.

   (b) Consider an alternative dynamical variable $d$, the distance of the SB centre of mass from the top of the IP. Determine expressions for $X_{\text{SB}}$, $Y_{\text{SB}}$ in terms of $X_{\text{IP}}$, $d$, $h$ and $\alpha$. Taking the time-derivative, determine equivalent expressions for $\dot{X}_{\text{SB}}$, $\dot{Y}_{\text{SB}}$.

   (c) Trivially, the IP is constrained not to move in the $Y$ direction, and so $Y_{\text{IP}}$ and $\dot{Y}_{\text{IP}}$ can be neglected. Determine an expression for the total kinetic energy $T$ of the SB and IP system in terms of $\dot{X}_{\text{IP}}$ and $\dot{d}$. Determine an expression for the total potential energy $V$ in terms of $X_{\text{IP}}$, $h$ and $d$.

   (d) Write down the Lagrangian $L$ using your expressions for $T$ and $V$. From the Lagrangian, derive equations for $\ddot{d}(t)$ and $\ddot{X}_{\text{IP}}(t)$. Solve these differential equations, writing the solutions in terms of the initial positions and velocities $d(0)$, $X_{\text{IP}}(0)$, $\dot{d}(0)$, $\dot{X}_{\text{IP}}(0)$.

   (e) Assuming that the SB is initially at rest at the top of the IP, which is also at rest, derive an expression for the distance the IP has moved once the SB has reached the bottom of the IP, and hence (assuming $\alpha = \pi/4$ radians) determine how large $m$ must be relative to $M$ for this distance to be $= h/2$.



# Workshop 2: Applications of Variational Methods

1. Use Euler's equation to find the shortest route between 2 points $(x_1, y_1)$ and $(x_2, y_2)$ in the x-y plane. [Hint: start by writing the total distance travelled in the form of a calculus of variations problem.]

2. A rope of length $L$, with uniform density, hangs under gravity between points $(x_1, y_1)$ and $(x_2, y_2)$, where $y$ is measured in the vertical direction. Using the method of Lagrange multipliers (see the example on page 13), minimise the gravitational potential energy to determine the function $y(x)$ that describes the path of the rope between these end points.

# Workshop 3: Tossing the Caber

The sport of tossing the caber entails trying to flip a caber (or uniform density thin log) so that it rotates through at least 180° about an axis perpendicular to its length prior to landing. This endeavour can approximately be split into the following phases:

1. the competitor supports the caber from beneath with their hands a height $h$ above the ground, holding it almost vertically, before lumbering forwards;
2. the competitor stops, while still holding the base of the caber, and the caber topples forward;
3. once the caber has toppled an angle $\theta_i$ away from the vertical, the competitor applies a vertical impulsive force to its base;
4. the caber experiences no further torques and subsequently moves under the influence of gravity until landing.

We would like to determine if a launch angle, $\theta_i$, will produce a successfully flipped caber.

(a) Using $\theta$ as the generalised coordinate, integrate along the caber to determine the kinetic energy for a caber of mass $m$ and length $2l$ during phase 2. Hence, find the Lagrangian for the caber and use the Euler-Lagrange equation to show that

$$\ddot{\theta} = \frac{3g}{4l} \sin\theta.$$

(b) Using the small angle approximation, which is sufficiently accurate for the situation under consideration here, and assuming that the initial angular velocity of the caber is $\dot{\theta}_0$ when $\theta = 0$, find the equation of motion for $\theta$ during phase 2.

(c) Calculate the moment of inertia, $I$, for rotations about an axis through the centre of mass and perpendicular to its length, of a uniform rod with mass $m$ and length $2l$.

(d) If the vertical impulsive force applied through the base of the caber is of the form $F = K\delta(t - t_i)$, then determine the torque being applied to the caber, $H$. Using your result for part (c) and

$$\int H dt = I \int \ddot{\theta} dt,$$

show that the angular velocity of the caber for $t > t_i$ can be written as

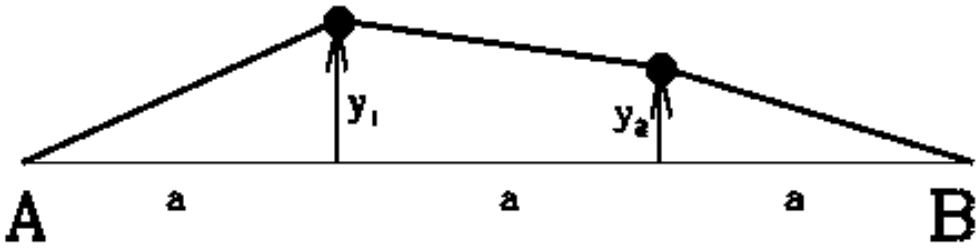
$$\dot{\theta}(t > t_i) = \dot{\theta}(t_i^-) + \frac{3K}{ml} \sin\theta_i,$$

where $\dot{\theta}(t_i^-)$ represents the angular velocity of the caber just before the impulsive force is applied.

(e) Determine the equations of motion for $\theta$ and the height of the centre of mass of the caber, $z_c$, for $t' = t - t_i > 0$. Under the assumption that a successfully tossed caber involves the caber turning through more than 180° before the centre of mass is within a distance $l$ of the ground, write down an inequality that should be satisfied for a successfully tossed caber.

(f) A caber is $\sim$ 5 m long and has a mass of $\sim$ 60 kg. Estimate values for the various other quantities and determine if an angle exists for which the caber is successfully tossed. (You may wish to write a computer code to solve the problem, or have a play with the one in the duo workshop solutions folder.)

# Workshop 4: Particles on a stretched string




Two particles of mass $m$ are attached $1/3$ and $2/3$ of the way along a uniform light elastic string of unstretched length less than $3a$. When fixed at the end points A and B as shown in the plan view above, the string has a tension $F$ along it. We wish to investigate small horizontal displacements in the transverse direction, denoted by the generalised coordinates $y_1$ and $y_2$ for the two masses. You should assume that the particles move on a smooth horizontal surface.

1. Write down the kinetic and potential energies of this system, and hence determine its Lagrangian. The expression $V = (1/2)kx^2$ represents the potential energy stored in a string stretched by $x$ from its unstretched length, so cannot be used in this case because the unstretched length is not specified. Instead, determine the potential energy by considering the work done stretching a string under a tension $F$.

2. Use the Euler-Lagrange equation to find coupled second order differential equations for $y_1$ and $y_2$, and rewrite these as a single matrix equation using the notation

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

3. Solve this equation to determine the frequencies for the normal modes of oscillation.

4. Calculate the mode vectors corresponding to these frequencies and sketch the motion corresponding to each mode.

5. Write down, in terms of $F$ and $m$, the solutions for $y_1$ and $y_2$ at $t \geq 0$ if the system is disturbed from rest at equilibrium by an impulsive force at $t = 0$ that imparts a velocity $\dot{y}_2 = v$.

# Workshop 5: Central Forces

A particle of mass $m$ moves in a central force field with potential energy given by $V(r)$. In spherical polar coordinates, the Lagrangian of the particle can be written as

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) - V(r).$$

(a) Using the fact that the azimuthal angle, $\phi$, is an ignorable coordinate, determine the associated constant of the motion, $J$. What does this constant represent? Express the total energy of the particle as a function of the single variable, $r$, in a way that incorporates the effective potential energy

$$V_{\text{eff}} = V(r) + \frac{J^2}{2mr^2}$$

and explain why no $\theta$ dependence is required.

(b) Assuming that $V(r) = Ar^{n+1}/(n+1)$ where $A > 0$, for each of the cases $n+1 < -2$, $-2 < n+1 < 0$ and $n+1 > 0$, sketch the radial variation of the effective potential. In each case, mark the radius, $r_0$, where the potential is stationary with respect to radius and a circular orbit exists, and state whether or not the orbit is stable or unstable. Calculate an expression for the radius $r_0$.

(c) By performing a Taylor series expansion of the effective potential about the point $r = r_0$, show that, for small perturbations away from a stable circular orbit, the radius of the particle performs simple harmonic motion with an angular frequency

$$\omega = \sqrt{n+3}\frac{J}{mr_0^2}.$$

(d) By comparing this angular frequency with that for a circular orbit at radius $r_0$, determine which values of $n$ give rise to closed orbits. Describe the orbits for the cases $n = -2$ and $n = 1$.

# Workshop 6: Hamiltonian Mechanics

1. Two particles of different mass, $m_1$ and $m_2$, are connected by a massless spring of spring constant $k$ and equilibrium length $d$. The system lies on a horizontal, frictionless, table and may both oscillate and rotate. Use the definitions of the centre of mass position and relative coordinate,

$$\underline{R} = \frac{m_1\underline{r}_1 + m_2\underline{r}_2}{M}, \quad \underline{r} = \underline{r}_2 - \underline{r}_1$$

respectively, where $M = m_1 + m_2$, to show that

$$\frac{1}{2}M\dot{\underline{R}}^2 + \frac{1}{2}\mu\dot{\underline{r}}^2 = \frac{1}{2}m_1\dot{\underline{r}}_1^2 + \frac{1}{2}m_2\dot{\underline{r}}_2^2,$$

where $\mu = m_1m_2/M$ is the reduced mass.
Hence show that the Lagrangian can be written as

$$L = \frac{\mu}{2}\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - \frac{k}{2}(r-d)^2,$$

where $r$ and $\phi$ represent the relative separation of the two masses in polar coordinates. Find the Hamiltonian of the system and Hamilton's equations of motion.

2. Suppose a bug of mass $m$ is crawling on a turntable rotating arbitrarily around an axis perpendicular to its plane. The bug's polar coordinates relative to the turntable are $r$ and $\phi$, whereas in the inertial, lab frame, they are $r_{\text{lab}} = r$ and $\phi_{\text{lab}} = \phi + \theta(t)$, where $\theta(t)$ is the angle between the two coordinate systems and the turntable rotates anticlockwise when viewed from above.

   In mixed coordinates, the Lagrangian of the bug is

$$L = \frac{m}{2}v_{\text{lab}}^2 - V(r,\phi),$$

   where $v_{\text{lab}}$ is the speed of the bug in the lab frame and $V(r,\phi)$ represents an arbitrary potential, expressed in terms of its polar coordinates.

   (a) Why should we use $v_{\text{lab}}^2$ and not $v^2$ in the Lagrangian?

   (b) Substitute the rotating coordinates into the expression for the lab kinetic energy in the Lagrangian then find the canonically conjugate momenta $p_r$ and $p_\phi$.

   (c) Calculate the bug's Hamiltonian in terms of $r, \phi, p_r, p_\phi$. Prove that, for arbitrary variation of $\theta$ with time,

$$H = H_{\text{lab}} - \dot{\theta}p_\phi,$$

   where $H_{\text{lab}}$ is the bug's Hamiltonian if $\dot{\theta} = 0$.

# Workshop 7: Canonical Transformation to a Rotating Frame

Consider a point particle moving in 2 dimensions held within a rotating (with angular frequency $\omega$) anisotropic harmonic potential $V(x,y) = (m/2)\{\omega_1^2[x\cos(\omega t) - y\sin(\omega t)]^2 + \omega_2^2[y\cos(\omega t) + x\sin(\omega t)]^2\}$ (such a potential is sometimes used to generate superfluid vortices in Bose-Einstein condensed atomic gases).

1. Determine the Hamiltonian $H$ from the Lagrangian $L$ for this system, using a Legendre transformation.

2. Consider a generating function of the form $F_2(x,y,P_X,P_Y) = P_X[x\cos(\omega t) - y\sin(\omega t)] + P_Y[y\cos(\omega t) + x\sin(\omega t)]$. Determine expressions for $p_x$, $p_y$, $X$, $Y$, from the implicit transformation equations ($X$ and $Y$ are the "new" coordinates, and $P_X$, $P_Y$ their canonically conjugate momenta).

3. Determine the transformed Hamiltonian $H'(X,Y,P_X,P_Y)$ (which should be time-independent), in terms of the new coordinates and momenta, using the expressions you have calculated for $p_x$, $p_y$, $X$, $Y$.

4. Use Hamilton's equations to determine equations for $\dot{X}$, $\dot{P}_X$. In the case where $\omega = 0$, find the coupled equations of motion for $X$, $P_X$, in terms of their initial values.

5. What is the value of the Poisson bracket $\{X, P_X\}$? Justify your answer.

# Workshop 8: Inertial Forces

1. A stone is dropped from a helicopter hovering a height $h$ over a point on the equator. $h$ is small in comparison with the radius of the Earth, $R$, so that the acceleration due to gravity can be taken as a constant, $g$. Ignore air resistance.

   (a) The operator relating time derivatives of a vector viewed in an inertial frame, S, to those measured in a frame B, rotating with angular velocity $\underline{\omega}$ (for example that associated with the rotation of the Earth), is

   $$\left[\frac{d}{dt}\right]_{\text{in } B} = \left(\left[\frac{d}{dt}\right]_{\text{in } S} - \underline{\omega}\times\right).$$

   Use this to show that the force measured in the rotating frame satisfies

   $$m\ddot{\underline{r}} = \underline{F} - m\underline{\omega} \times (\underline{\omega} \times \underline{r}) - 2m\underline{\omega} \times \dot{\underline{r}} - m\dot{\underline{\omega}} \times \underline{r},$$

   where $\dot{\underline{r}}$ and $\ddot{\underline{r}}$ represent the velocity and acceleration measured in frame B, and $\underline{F}$ is the force in the inertial frame. (Note that $\underline{r}$ is the instantaneous position measured relative to an origin on the axis of rotation, and does not depend upon the frame of reference.)

   In which directions do the three inertial forces point for the falling stone?

   The rate at which the Earth's rotation is decreasing is very small, so you should now assume $\dot{\omega} = 0$.

   (b) Ignoring inertial forces, find expressions for the speed and height of the stone as a function of time, and the time taken before the stone hits the ground, $t_0$. What condition must be satisfied for the inertial forces to be unimportant in determining $t_0$?

   (c) Using the Coriolis force and the approximation that it always acts in the same direction, how far, in terms of $t_0$, does the stone land from the point beneath the helicopter, and in which direction?

   (d) Solve the same problem by using the conservation of angular momentum, and compare your answer with that from (c).

2. What angular velocity would be required for your effective weight at the equator to be half that at the north pole? (Take the earth's radius to be 6378 km). How many minutes would a full rotation of the earth take?

# Workshop 9: Rigid Body Motion

1. A rigid body is turning about a fixed point $O$, and $Oxyz$ are Cartesian axes. The components of velocity of the particle with coordinates $(1,0,0)$ are $(0,2,5)$. Find the component in the direction of the $x$ axis of the velocity of the particle with coordinates $(0,0,1)$.

2. (a) For a uniform density cylinder of height $2h$, radius $R$ and total mass $M$, the inertia tensor is defined as

$$I_{\alpha\beta} = \int_{\text{volume}} dxdydz\rho(x,y,z)(|\underline{r}|^2\delta_{\alpha\beta} - r_\alpha r_\beta),$$

where $\rho$ is the mass density, $|\underline{r}|^2 = x^2+y^2+z^2$, and $\alpha$ and $\beta$ run over the three Cartesian coordinates of vector $\underline{r}$ (i.e. $r_1 \equiv x$, $r_2 \equiv y$ and $r_3 \equiv z$). Show that, with the $z$ axis lined up with the symmetry axis of the cylinder, the inertia tensor about the centre of mass can be written as

$$\hat{I}_{\text{CoM}} = \begin{pmatrix} I_{xx} & 0 & 0 \\ 0 & I_{xx} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix},$$

where $I_{xx} = M(R^2/4 + h^2/3)$ and $I_{zz} = MR^2/2$.

A victorious table tennis player celebrated by throwing their bat into the air. The subsequent rotational motion of the table tennis bat can be understood by approximating it as two connected cylinders - a long, thin cylinder for the handle, and a short, fat cylinder for the blade with which the ball is typically hit - and calculating the inertia tensor of this compound object. Assume that the blade has radius $R$, height $2h \ll R$ and mass $M$, while the handle has radius $u \ll R$, height $R$ and mass $M$. The bat is constructed by attaching one end of the handle to the circular edge of the blade such that the axis of symmetry of the handle passes through the centre of mass of the blade.

For the remainder of the question, use 'body' coordinates such that axis 1 is along the axis of symmetry of the handle and axis 2 is parallel to the symmetry axis of the blade.

(b) Where is the centre of mass of the bat?

Using the displaced axis theorem,

$$\hat{I} = \hat{I}_{\text{CoM}} + M\hat{A},$$

where $A_{\alpha\beta} = |\underline{R}_C|^2\delta_{\alpha\beta} - R_{C,\alpha}R_{C,\beta}$, and $\underline{R}_C$ is the vector representing the centre of mass position vector of either the handle or the blade with respect to the centre of mass position of the whole bat, determine the inertia tensors for the handle and the blade about the centre of mass of the bat. Hence find the inertia tensor for the bat with respect to rotations about its centre of mass.

(c) Rank the principal moments of inertia of the bat about the 1, 2 and 3 axes ($I_1$, $I_2$ and $I_3$ respectively) in order of increasing size.

The initial angular velocity given to the bat is $\underline{\omega} = (0, \omega_2, \omega_3)$, where $\omega_2 \gg \omega_3 > 0$. Describe the subsequent rotational motion of the bat.

# A Second-Order Ordinary Linear Differential Equations with Constant Coefficients

## A.1 Definition

A second-order homogeneous ordinary linear differential equation with constant coefficients has the form:

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx = 0 \tag{33}$$

This has been written to appear to be a differential equation of a position variable with respect to time, as this is most obviously relevant for classical mechanics.

## A.2 Solution

### A.2.1 Auxiliary Equation

The recipe for solving such a linear differential equation begins by substituting $e^{\lambda t}$ for $x$. Thus

$$a\lambda^2 e^{\lambda t} + b\lambda e^{\lambda t} + ce^{\lambda t} = 0 \tag{34}$$

which, cancelling $e^{\lambda t}$ everywhere, becomes

$$a\lambda^2 + b\lambda + c = 0. \tag{35}$$

This is the *auxiliary equation*. One can then solve the auxiliary equation by finding the roots from the quadratic formula:

$$\lambda_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{36}$$

Thus

$$(x - \lambda_+)(x - \lambda_-) = 0, \tag{37}$$

i.e., assuming the roots are not repeated, there are 2 distinct values of $\lambda$ solving the auxiliary equation, which may be complex or imaginary.

### A.2.2 General Solution (Roots not Repeated)

The general solution is then a linear superposition of the two particular solutions $e^{\lambda_+ t}$ and $e^{\lambda_- t}$, i.e.,

$$x(t) = A_+ e^{\lambda_+ t} + A_- e^{\lambda_- t} \tag{38}$$

where the $A_+, A_-$ are arbitrary constants (because we are dealing with a *second-order* differential equation, there must be *two* integration constants in its general solution), often determined by the initial position:

$$x(0) = A_+ + A_- \tag{39}$$

and the initial velocity, through

$$\dot{x}(t) = A_+\lambda_+ e^{\lambda_+ t} + A_-\lambda_- e^{\lambda_- t} \Rightarrow \dot{x}(0) = A_+\lambda_+ + A_-\lambda_-. \tag{40}$$

The roots $\lambda_+$ and $\lambda_-$ may be imaginary — if this is the case and $x$ is a real quantity (such as position) then $A_+$ and $A_-$ must have (complex) values such that the overall imaginary contribution is zero.

## A.3 Possible Complications

### A.3.1 Repeated Roots

If $\lambda_+ = \lambda_- = \lambda$, then $e^{\lambda t}$ solves the differential equation, and a second linearly independent solution is given by $te^{\lambda t}$. The general solution is thus given by

$$x(t) = Ae^{\lambda t} + Bte^{\lambda t}. \tag{41}$$

### A.3.2 Additional Constant

If the differential equation has the form

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx + k = 0 \tag{42}$$

then this can be rephrased as

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + c\left(x + \frac{k}{c}\right) = 0. \tag{43}$$

One can then define a new variable $\tilde{x} = x + k/c$, for which it follows

$$a\frac{d^2\tilde{x}}{dt^2} + b\frac{d\tilde{x}}{dt} + c\tilde{x} = 0. \tag{44}$$

One then finds the general solution for $\tilde{x}(t)$, as shown above, and the general solution for $x(t)$ from it, by $x(t) = \tilde{x}(t) - k/c$.

## A.4 Representing Position in a Plane by a Single Complex Number

If we have a situation where, for example, we have coupled motion in the $x, y$ plane, i.e., the $x$ and $y$ variables cannot be separated:

$$\frac{d^2x}{dt^2} + a_x\frac{dx}{dt} + b_x\frac{dy}{dt} + c_xx + d_xy + k_x = 0, \tag{45}$$

$$\frac{d^2y}{dt^2} + a_y\frac{dx}{dt} + b_y\frac{dy}{dt} + c_yx + d_yy + k_y = 0, \tag{46}$$

then we can try to simplify matters by considering the complex variable $\xi = x + iy$, and adding the bottom equation multiplied by $i$ to the top equation. We then consider the differential equation given by

$$\frac{d^2\xi}{dt^2} + \cdots = \frac{d^2x}{dt^2} + \cdots + i\left(\frac{d^2y}{dt^2} + \cdots\right). \tag{47}$$

In general, this will produce an equation containing terms proportional to both $\xi$ and its complex conjugate $\xi^*$, as well as their first derivatives, which in no way simplifies anything! However, in a number of cases of physical interest, the result takes the form

$$\frac{d^2\xi}{dt^2} + \alpha\frac{d\xi}{dt} + \beta\xi + \gamma = 0, \tag{48}$$

where $\alpha$, $\beta$, $\gamma$ are (in general) *complex* constant coefficients, i.e., there is *no* coupling to the complex conjugate variable $\xi^*$. If this is the case, then the general solution for $\xi(t)$ is determined using exactly the same methodology as described on the first two pages. Having determined the general

solution for $\xi(t)$, one derives the general solution for $x(t)$ by taking the real part of $\xi(t)$, and the general solution for $y(t)$ by taking the imaginary part.

The very substantial simplification of the mathematics offered by such an approach means it is at least worth considering. Otherwise it is necessary to use methods developed for linear systems of equations; this will also be necessary if motion in the $x$, $y$, and $z$ directions cannot be separated. Extension to higher-order linear differential equations is straightforward. See essentially any textbook on ordinary differential equations for further details, if interested.

# B Taylor Series

## B.1 Taylor's Theorem of the Mean

We consider a function of $N$ variables $f(x_1, x_2, \ldots, x_N)$. This could, for example, be a Lagrangian describing the motion of $n$ objects, which would be considered a function of $n$ generalized coordinates and $n$ generalized velocities (which are considered *independent* variables), making for a total of $N = 2n$ variables. Or it could be just the potential energy term of such a Lagrangian, with $N = n$ variables.

If, within a region (i.e., some volume of $n$-dimensional space) all possible partial derivatives of $f$ exist, Taylor's theorem of the mean implies

$$\begin{aligned} f(x_{01} + a_1, \ldots, x_{0N} + a_N) =& f(x_{01}, \ldots, x_{0N}) \\ &+ \sum_{k=1}^{n} \frac{1}{k!} \left( a_1 \frac{\partial}{\partial x_1} + \cdots + a_N \frac{\partial}{\partial x_N} \right)^k f(x_1, \ldots, x_N) \Bigg|_{x_{01}, \ldots, x_{0N}} + O(x_k^{n+1}) \end{aligned} \tag{49}$$

(Taylor's theorem of the mean is in fact a more general result — see essentially any moderately advanced calculus text if interested). It is also assumed that the function $f$ is well-defined over the region of interest (in particular, there should be no infinities), i.e., the above series *converges* within this region. The vertical line at the right-hand-side of the derivative terms means that the derivative is first calculated, and then evaluated at $x_1 = x_{01}$, $x_2 = x_{02}$, etc..

## B.2 Taylor Series in One Variable

If we restrict ourselves to $f$ being a function of just one variable $f(x)$, then

$$f(x_0 + a) = f(x_0) + \sum_{k=1}^{n} \frac{1}{k!} a^k \frac{\partial^k}{\partial x^k} f(x) \Bigg|_{x_0} + O(x^{n+1}). \tag{50}$$

If we write the above in a form where $a = x - x_0$, then we get the usual form of a single-variable Taylor series:

$$f(x) = f(x_0) + \sum_{k=1}^{n} \frac{1}{k!} \left[ \frac{\partial^k}{\partial x^k} f(x) \Bigg|_{x_0} \right] (x - x_0)^k + O(x^{n+1}). \tag{51}$$

Note that, for example, the function $f(x) = x^{-1}$ cannot be expanded around $x_0 = 0$, as this point is *singular* [$f(x) \to \infty$ as $x \to 0$]. Any value of $x_0$ slightly different from zero is acceptable.

## B.3 Taylor Series in Two Variables

It is not so easy to write down such a compact formulation for a Taylor series in two or more variables, but the procedure is essentially the same. Bearing in mind that we are frequently only interested in (at most) the first- and second-order derivative terms, we write down

$$\begin{aligned} f(x_0+a, y_0+b) =& f(x_0,y_0) + \left( a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} \right) f(x,y)\Big|_{x_0,y_0} \\ &+ \frac{1}{2}\left( a^2\frac{\partial^2}{\partial x^2} + 2ab\frac{\partial^2}{\partial x \partial y} + b^2\frac{\partial^2}{\partial y^2} \right) f(x,y)\Big|_{x_0,y_0} + \cdots . \end{aligned} \tag{52}$$

Similarly to the case for a single-variable Taylor series, we set $a = x - x_0$ and $b = y - y_0$, yielding

$$\begin{aligned} f(x,y) =& f(x_0,y_0) + \left\{ \left[ \frac{\partial}{\partial x} f(x,y)\Big|_{x_0,y_0} \right] (x-x_0) + \left[ \frac{\partial}{\partial y} f(x,y)\Big|_{x_0,y_0} \right] (y-y_0) \right\} \\ &+ \frac{1}{2}\left\{ \left[ \frac{\partial^2}{\partial x^2} f(x,y)\Big|_{x_0,y_0} \right] (x-x_0)^2 + 2\left[ \frac{\partial^2}{\partial x \partial y} f(x,y)\Big|_{x_0,y_0} \right] (x-x_0)(y-y_0) \right. \\ &\left. + \left[ \frac{\partial^2}{\partial y^2} f(x,y)\Big|_{x_0,y_0} \right] (y-y_0)^2 \right\} + \cdots . \end{aligned} \tag{53}$$

## B.4 Taylor Series in Three Variables

Proceeding in an exactly analogous fashion, we determine for a function of three variables, $f(x,y,z)$, that

$$\begin{aligned} f(x,y,z) =& f(x_0,y_0,z_0) \\ &+ \left\{ \left[ \frac{\partial}{\partial x} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (x-x_0) + \left[ \frac{\partial}{\partial y} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (y-y_0) \right. \\ &\left. + \left[ \frac{\partial}{\partial z} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (z-z_0) \right\} \\ &+ \frac{1}{2}\left\{ \left[ \frac{\partial^2}{\partial x^2} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (x-x_0)^2 + \left[ \frac{\partial^2}{\partial y^2} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (y-y_0)^2 \right. \\ &+ \left[ \frac{\partial^2}{\partial z^2} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (z-z_0)^2 + 2\left[ \frac{\partial^2}{\partial x \partial y} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (x-x_0)(y-y_0) \\ &+ 2\left[ \frac{\partial^2}{\partial x \partial z} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (x-x_0)(z-z_0) \\ &\left. + 2\left[ \frac{\partial^2}{\partial y \partial z} f(x,y,z)\Big|_{x_0,y_0,z_0} \right] (y-y_0)(z-z_0) \right\} + \cdots . \end{aligned} \tag{54}$$

This can be extended to functions of more than three variables, and to include higher-order derivatives, as desired, in a reasonably straightforward fashion.

# C Determinants, Eigenvalues and Eigenvectors

## C.1 Symmetric Matrices

In classical mechanics, we can usually restrict ourselves to considering real, symmetric matrices, i.e., matrices with real entries that are equal to their own transpose. Formally, we say a matrix $\hat{A}$ is symmetric if

$$\hat{A} = \hat{A}^T. \tag{55}$$

In the case of a general $2 \times 2$ matrix $\hat{B}$

$$\hat{B}^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}. \tag{56}$$

Hence, if $\hat{B}$ is to equal $\hat{B}^T$, it follows that $b$ must equal $c$, i.e., the most general $2 \times 2$ symmetric matrix has the form

$$\hat{B} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}. \tag{57}$$

## C.2 Determinants

The determinant of the matrix $\hat{B}$ is given by

$$\det(\hat{B}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{58}$$

There is considerable freedom in the choice of steps to take when calculating the determinant of a $3 \times 3$ matrix $\hat{C}$. The most conventional formulation is

$$\det(\hat{C}) = \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix} = c_{11} \begin{vmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{vmatrix} - c_{12} \begin{vmatrix} c_{21} & c_{23} \\ c_{31} & c_{33} \end{vmatrix} + c_{13} \begin{vmatrix} c_{21} & c_{22} \\ c_{31} & c_{32} \end{vmatrix}. \tag{59}$$

We thus have a sum of 3 terms involving determinants of $2 \times 2$ matrices, which we already know how to calculate.

Similarly, the determinant of an $n \times n$ matrix $\hat{D}$ is given by a sum of $n$ terms involving determinants of $(n-1) \times (n-1)$ matrices. These $(n-1) \times (n-1)$ matrices are conventionally formed by removing the 1st row and the $k$th column from the matrix $\hat{D}$. In this case, the determinant is formulated by

$$\det(\hat{D}) = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix} = \sum_{k=1}^{n} (-1)^{1+k} d_{1k} \begin{vmatrix} \cdots & d_{2,k-1} & d_{2,k+1} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & d_{n,k-1} & d_{n,k+1} & \cdots \end{vmatrix}. \tag{60}$$

## C.3 The Conventional Eigenvalue Problem

If we take a matrix $\hat{A}$, a vector $\underline{v}$, and a scalar $\lambda$, such that

$$\hat{A}\underline{v} = \lambda \underline{v}, \tag{61}$$

we say that $\underline{v}$ is an *eigenvector* of the matrix $\hat{A}$, with *eigenvalue* $\lambda$. It obviously follows that

$$\hat{A}\underline{v} - \lambda\underline{v} = \underline{0} \Rightarrow (\hat{A} - \lambda\hat{1})\underline{v} = \underline{0}, \tag{62}$$

where $\hat{1}$ is the *identity* matrix [i.e., the matrix with ones down the main (top-left-to-bottom-right) diagonal and zeros everywhere else]. It so happens that this can only be possible if the determinant of the matrix $(\hat{A} - \lambda\hat{1})$ is zero, discounting the trivial solution $\underline{v} = \underline{0}$.

The procedure for finding the eigenvalues associated with the matrix $\hat{A}$ then begins by calculating the determinant of $(\hat{A} - \lambda\hat{1})$, and setting it equal to zero. If $\hat{A}$ is a real, symmetric $n \times n$ matrix, this yields an $n$th-degree polynomial in $\lambda$ which always has $n$ real roots (although roots may be repeated, or *degenerate*).

We will work this through for a real $2 \times 2$ symmetric matrix

$$\hat{B} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}. \tag{63}$$

To determine the eigenvalues of $\hat{B}$, we calculate

$$\begin{vmatrix} a-\lambda & b \\ b & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - b^2 = \lambda^2 - (a+d)\lambda + (ad - b^2) = 0. \tag{64}$$

The values of $\lambda$ that solve this equation can in this case be generally determined from the quadratic formula. There are two solutions, $\lambda_+$ and $\lambda_-$, given by

$$\lambda_\pm = \frac{(a+d) \pm \sqrt{(a-d)^2 + 4b^2}}{2}. \tag{65}$$

The eigenvector $\underline{v}_+$ associated with the eigenvalue $\lambda_+$ can then be determined from

$$\begin{pmatrix} a-\lambda_+ & b \\ b & d-\lambda_+ \end{pmatrix} \begin{pmatrix} v_{1+} \\ v_{2+} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{66}$$

which implies

$$(\lambda_+ - a)v_{1+} = bv_{2+} \tag{67}$$
$$(\lambda_+ - d)v_{2+} = bv_{1+}. \tag{68}$$

It is common to choose $v_{1+} = 1$, and then normalize the result. Hence, from Eq. (68)

$$\underline{v}_+ = \frac{1}{\sqrt{1 + (\lambda_+ - d)^2/b^2}} \begin{pmatrix} 1 \\ (\lambda_+ - d)/b \end{pmatrix}, \tag{69}$$

and equivalently, for the determining of $\underline{v}_-$. Because the polynomial $(a - \lambda_+)(d - \lambda_+) - b^2 = 0$ fixes a relationship between the matrix components $a$, $b$, $d$, and the eigenvalue $\lambda_+$, Eq. (67) and Eq. (68) are not independent. This is why we have ignored Eq. (67) when determining $\underline{v}_+$.

When considering 3- and higher-dimensional vectors, some more involved algebraic manipulation is usually required to produce isolated expressions for the eigenvector components. For an $n \times n$ matrix, we will end up with $n$ equations like Eq. (67) and Eq. (68), and we can always ignore one of them as being not independent.

If, for example two roots in the characteristic polynomial are identical (repeated roots), then we can determine two independent (i.e., orthogonal) eigenvectors with the same eigenvalue. As any linear combination of the two will also be an eigenvector with the same eigenvalue, there is some freedom in choosing what form the two orthogonal eigenvectors should take. See essentially any linear algebra text for more details and proofs of the above results.

## C.4 Generalized Eigenvalue Problem

It can happen (for example when determining normal modes) that we have a situation which is expressed more like

$$(\hat{A} - \lambda \hat{B})\underline{v} = \underline{0}. \tag{70}$$

where $\hat{B}$ is a matrix which is not necessarily proportional to the identity (although it will still be symmetric, in the case of determining normal modes). It is still the case that the determinant of $(\hat{A} - \lambda \hat{B})$ must be zero for this equation to be nontrivially solvable.

This is a generalized eigenvalue problem, and the eigenvalues and eigenvectors are determined in much the same way, i.e., setting the determinant $= 0$, finding the roots of the resulting polynomial in order to determine the eigenvalues, then using the calculated values of the eigenvalues to determine the relationships between the components of their corresponding eigenvectors, and finally (usually) normalizing the eigenvectors.