SEMESTER – I

19PHP103 CLASSICAL MECHANICS AND RELATIVITY 4H – 4C

Instruction Hours / week: L: 4 T: 0 P: 0 **Marks:** Internal: **40** External: **60** Total: **100**
End Semester Exam: 3 Hours

Course objective

- Studying Classical Mechanics will gives an idea about how classical physics deal with matter and energy.
- Have a deep understanding of Newton's laws
- Be able to solve the Newton equations for simple configurations using various methods,
- To give an insight into the classical methods of physics.

Course Outcomes (Cos)

1. Students can understand the classical laws of motion.
2. Competency in using the essential mathematical skills needed for describing mechanics and special relativity
3. They can develop problem solving skills.

Unit I: Lagrangian Dynamics

Mechanics of a system of particles - Generalized Co-ordinates - Constraints - D' Alembert's Principle and Lagrange's Equations - Velocity-Dependent Potentials and the Dissipation Function – Simple applications of Lagrangian Formulation - Hamilton's Principle - Derivation of Lagrange's Equations from Hamilton's Principle - Conservation Theorems and Symmetry Properties - Energy Function and the Conservation of Energy.

Unit II: Central Force Problem and Classical Scattering

Reduction to the Equivalent One-Body Problem - The Equations of Motion and First Integrals – Classification of Orbits – The Viral Theorem – Kepler's Problem: Inverse Square Law of Force –The Motion in time in the Kepler's problem - Scattering in a Central Force Field - Transformation of the Scattering Problem to Laboratory Coordinates- Three body problem.

Unit III: Rigid body Dynamics and Small Oscillations

The Independent coordinates of a Rigid Body - Orthogonal Transformations - The Euler Angles - Angular Momentum and Kinetic Energy of Motion about a Point – Tensors - the Inertia Tensor and the Moment of Inertia - The Eigen values of the Inertia Tensor and the Principal Axis Transformation - Solving Rigid Body Problems and the Euler Equations of Motion. Small Oscillations - Frequencies of Free Vibration, and Normal Coordinates, Linear Tri atomic Molecule.

Unit IV: Hamiltonian dynamics and Canonical Transformations

Legendre Transformations and the Hamilton Equations of Motion - Cyclic Coordinates and Conservation Theorems – Derivation of Hamilton's Equation from Variational principle – Principle of Least Action - Equations of Canonical Transformation - Examples of Canonical Transformations - The Harmonic Oscillator - The Symplectic Approach to Canonical Transformations - Poisson Brackets and Other Canonical Invariants - The Angular Momentum Poisson Bracket Relations.

Unit V: The Special Theory of Relativity

Basic Postulates of the Special Theory – Newton's Law of Motion – Non-Variance Maxwell's Equation - Michelson Morley Experiment – Null results - Lorentz Transformations – Concept of Inertial frame – Velocity Addition and Thomas Precession – Length Contraction – Vectors and Metric Tensor – Relativistic Kinematics – Relativistic Angular Momentum – Introduction to the General theory of Relativity – Gravitation and acceleration and their relation to non-inertial frames of reference – Minkowski space and Lorenz transformation.

SUGGESTED READINGS

1. Goldstein, Poole and Safko, *Classical Mechanics*, 3 edition, Pearson Publication (2001)
2. Hartle B. James, 1st edition, 2009, Gravity, An Introduction to Einstein's General Relativity, Darling Kindersley (India) Pvt.Ltd., Delhi.
3. John R. Taylor, *Classical Mechanics*, University Science Books, (2004)
4. Louis N. Hand and Janet D. Finch, *Analytical mechanics*, Cambridge University Press, (1998)

J.C. Upadyaya, *Classical Mechanics*, Himalayan Publishing House, New Delhi (2009)

Lagrangian Dynamics

Mechanics of a system of particles - Generalized Co-ordinates - Constraints - D' Alembert's Principle and Lagrange's Equations - Velocity-Dependent Potentials and the Dissipation Function – Simple applications of Lagrangian Formulation - Hamilton's Principle - Derivation of Lagrange's Equations from Hamilton's Principle - Conservation Theorems and Symmetry Properties - Energy Function and the Conservation of Energy.

MECHANICS OF A PARTICLE

Let $\mathbf{r}$ be the radius vector of a particle from some given origin and $\mathbf{v}$ its vector velocity:

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}. \tag{1.1}$$

The *linear momentum* $\mathbf{p}$ of the particle is defined as the product of the particle mass and its velocity:

$$\mathbf{p} = m\mathbf{v}. \tag{1.2}$$

In consequence of interactions with external objects and fields, the particle may experience forces of various types, e.g., gravitational or electrodynamic; the vector sum of these forces exerted on the particle is the total force $\mathbf{F}$. The mechanics of the particle is contained in *Newton's second law of motion*, which states that there exist frames of reference in which the motion of the particle is described by the differential equation

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \equiv \dot{\mathbf{p}}, \tag{1.3}$$

or

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}). \tag{1.4}$$

In most instances, the mass of the particle is constant and Eq. (1.4) reduces to

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = m\mathbf{a}, \tag{1.5}$$

where $\mathbf{a}$ is the vector acceleration of the particle defined by

$$\mathbf{a} = \frac{d^2\mathbf{r}}{dt^2}. \tag{1.6}$$

The equation of motion is thus a differential equation of second order, assuming $\mathbf{F}$ does not depend on higher-order derivatives.

A reference frame in which Eq. (1.3) is valid is called an *inertial* or *Galilean system*. Even within classical mechanics the notion of an inertial system is something of an idealization. In practice, however, it is usually feasible to set up a coordinate system that comes as close to the desired properties as may be required. For many purposes, a reference frame fixed in Earth (the "laboratory system") is a sufficient approximation to an inertial system, while for some astronomical purposes it may be necessary to construct an inertial system by reference to distant galaxies.

Many of the important conclusions of mechanics can be expressed in the form of conservation theorems, which indicate under what conditions various mechanical quantities are constant in time. Equation (1.3) directly furnishes the first of these, the

Conservation Theorem for the Linear Momentum of a Particle: If the total force, $\mathbf{F}$*, is zero, then* $\dot{\mathbf{p}} = 0$ *and the linear momentum,* $\mathbf{p}$*, is conserved.*

The angular momentum of the particle about point O, denoted by $\mathbf{L}$, is defined as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{1.7}$$

where $\mathbf{r}$ is the radius vector from O to the particle. Notice that the order of the factors is important. We now define the *moment of force* or *torque* about O as

$$\mathbf{N} = \mathbf{r} \times \mathbf{F}. \tag{1.8}$$

The equation analogous to (1.3) for $\mathbf{N}$ is obtained by forming the cross product of $\mathbf{r}$ with Eq. (1.4):

$$\mathbf{r} \times \mathbf{F} = \mathbf{N} = \mathbf{r} \times \frac{d}{dt}(m\mathbf{v}). \tag{1.9}$$

Equation (1.9) can be written in a different form by using the vector identity:

$$\frac{d}{dt}(\mathbf{r} \times m\mathbf{v}) = \mathbf{v} \times m\mathbf{v} + \mathbf{r} \times \frac{d}{dt}(m\mathbf{v}), \tag{1.10}$$

where the first term on the right obviously vanishes. In consequence of this identity, Eq. (1.9) takes the form

$$\mathbf{N} = \frac{d}{dt}(\mathbf{r} \times m\mathbf{v}) = \frac{d\mathbf{L}}{dt} \equiv \dot{\mathbf{L}}. \tag{1.11}$$

Note that both $\mathbf{N}$ and $\mathbf{L}$ depend on the point O, about which the moments are taken.

As was the case for Eq. (1.3), the torque equation, (1.11), also yields an immediate conservation theorem, this time the

Conservation Theorem for the Angular Momentum of a Particle: If the total torque, $\mathbf{N}$, is zero then $\dot{\mathbf{L}} = 0$, and the angular momentum $\mathbf{L}$ is conserved.

Next consider the work done by the external force $\mathbf{F}$ upon the particle in going from point 1 to point 2. By definition, this work is

$$W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{s}. \tag{1.12}$$

For constant mass (as will be assumed from now on unless otherwise specified), the integral in Eq. (1.12) reduces to

$$\int \mathbf{F} \cdot d\mathbf{s} = m \int \frac{d\mathbf{v}}{dt} \cdot \mathbf{v}\, dt = \frac{m}{2} \int \frac{d}{dt}(v^2)\, dt,$$

and therefore

$$W_{12} = \frac{m}{2}(v_2^2 - v_1^2). \tag{1.13}$$

The scalar quantity $mv^2/2$ is called the kinetic energy of the particle and is denoted by T, so that the work done is equal to the change in the kinetic energy:

$$W_{12} = T_2 - T_1. \tag{1.14}$$

If the force field is such that the work W_{12} is the same for any physically possible path between points 1 and 2, then the force (and the system) is said to be *conservative*. An alternative description of a conservative system is obtained by imagining the particle being taken from point 1 to point 2 by one possible path and then being returned to point 1 by another path. The independence of W_{12} on the particular path implies that the work done around such a closed circuit is zero, i.e.:

$$\oint \mathbf{F} \cdot d\mathbf{s} = 0. \tag{1.15}$$

Physically it is clear that a system cannot be conservative if friction or other dissipation forces are present, because $F \cdot ds$ due to friction is always positive and the integral cannot vanish.

By a well-known theorem of vector analysis, a necessary and sufficient condition that the work, W_{12}, be independent of the physical path taken by the particle is that **F** be the gradient of some scalar function of position:

$$\mathbf{F} = -\nabla V(\mathbf{r}), \tag{1.16}$$

where V is called the *potential*, or *potential energy*. The existence of V can be inferred intuitively by a simple argument. If W_{12} is independent of the path of integration between the end points 1 and 2, it should be possible to express W_{12} as the change in a quantity that depends only upon the positions of the end points. This quantity may be designated by $-V$, so that for a differential path length we have the relation

$$\mathbf{F} \cdot d\mathbf{s} = -dV$$

or

$$F_s = -\frac{\partial V}{\partial s},$$

which is equivalent to Eq. (1.16). Note that in Eq. (1.16) we can add to V any quantity constant in space, without affecting the results. Hence *the zero level of V is arbitrary.*

For a conservative system, the work done by the forces is

$$W_{12} = V_1 - V_2. \tag{1.17}$$

Combining Eq. (1.17) with Eq. (1.14), we have the result

$$T_1 + V_1 = T_2 + V_2, \tag{1.18}$$

which states in symbols the

Energy Conservation Theorem for a Particle: If the forces acting on a particle are conservative, then the total energy of the particle, $T + V$, is conserved.

MECHANICS OF A SYSTEM OF PARTICLES

In generalizing the ideas of the previous section to systems of many particles, we must distinguish between the *external forces* acting on the particles due to sources outside the system, and *internal forces* on, say, some particle i due to all other particles in the system. Thus, the equation of motion (Newton's second law) for the ith particle is written as

$$\sum_j \mathbf{F}_{ji} + \mathbf{F}_i^{(e)} = \dot{\mathbf{p}}_i, \tag{1.19}$$

where $\mathbf{F}_i^{(e)}$ stands for an external force, and $\mathbf{F}_{ji}$ is the internal force on the ith particle due to the jth particle ($\mathbf{F}_{ii}$, naturally, is zero). We shall assume that the $\mathbf{F}_{ij}$ (like the $\mathbf{F}_i^{(e)}$) obey Newton's third law of motion in its original form: that the forces two particles exert on each other are equal and opposite. This assumption (which does not hold for all types of forces) is sometimes referred to as the *weak law of action and reaction*.

Summed over all particles, Eq. (1.19) takes the form

$$\frac{d^2}{dt^2}\sum_i m_i \mathbf{r}_i = \sum_i \mathbf{F}_i^{(e)} + \sum_{\substack{i,j \\ i\neq j}} \mathbf{F}_{ji}. \tag{1.20}$$

The first sum on the right is simply the total external force $\mathbf{F}^{(e)}$, while the second term vanishes, since the law of action and reaction states that each pair $\mathbf{F}_{ij} + \mathbf{F}_{ji}$ is zero. To reduce the left-hand side, we define a vector $\mathbf{R}$ as the average of the radii vectors of the particles, weighted in proportion to their mass:

$$\mathbf{R} = \frac{\sum m_i \mathbf{r}_i}{\sum m_i} = \frac{\sum m_i \mathbf{r}_i}{M}. \tag{1.21}$$

The vector $\mathbf{R}$ defines a point known as the *center of mass*, or more loosely as the center of gravity, of the system (cf. Fig. 1.1). With this definition, (1.20) reduces to

$$M\frac{d^2\mathbf{R}}{dt^2} = \sum_i \mathbf{F}_i^{(e)} \equiv \mathbf{F}^{(e)}, \tag{1.22}$$

which states that the center of mass moves as if the total external force were acting on the entire mass of the system concentrated at the center of mass. Purely internal forces, if the obey Newton's third law, therefore have no effect on the

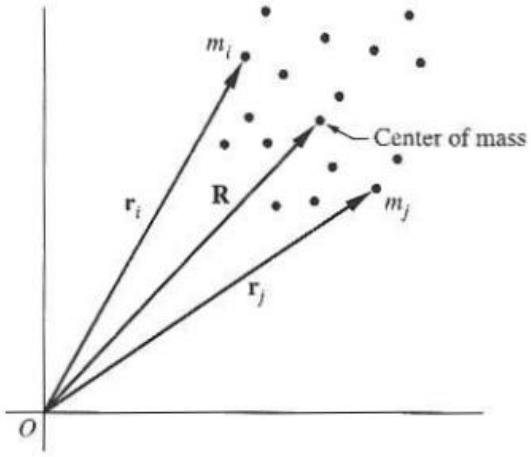


FIGURE 1.1 The center of mass of a system of particles.

motion of the center of mass. An oft-quoted example is the motion of an exploding shell—the center of mass of the fragments traveling as if the shell were still in a single piece (neglecting air resistance). The same principle is involved in jet and rocket propulsion. In order that the motion of the center of mass be unaffected, the ejection of the exhaust gases at high velocity must be counterbalanced by the forward motion of the vehicle at a slower velocity.

By Eq. (1.21) the total linear momentum of the system,

$$\mathbf{P} = \sum m_i \frac{d\mathbf{r}_i}{dt} = M\frac{d\mathbf{R}}{dt}, \tag{1.23}$$

is the total mass of the system times the velocity of the center of mass. Consequently, the equation of motion for the center of mass, (1.23), can be restated as the

Conservation Theorem for the Linear Momentum of a System of Particles: If the total external force is zero, the total linear momentum is conserved.

We obtain the total angular momentum of the system by forming the cross product $\mathbf{r}_i \times \mathbf{p}_i$ and summing over i. If this operation is performed in Eq. (1.19), there results, with the aid of the identity, Eq. (1.10),

$$\sum_i (\mathbf{r}_i \times \dot{\mathbf{p}}_i) = \sum_i \frac{d}{dt}(\mathbf{r}_i \times \mathbf{p}_i) = \dot{\mathbf{L}} = \sum_i \mathbf{r}_i \times \mathbf{F}_i^{(e)} + \sum_{\substack{i,j \\ i \neq j}} \mathbf{r}_i \times \mathbf{F}_{ji}. \quad (1.24)$$

The last term on the right in (1.24) can be considered a sum of the pairs of the form

$$\mathbf{r}_i \times \mathbf{F}_{ji} + \mathbf{r}_j \times \mathbf{F}_{ij} = (\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ji}, \quad (1.25)$$

using the equality of action and reaction. But $\mathbf{r}_i - \mathbf{r}_j$ is identical with the vector $\mathbf{r}_{ij}$ from j to i (cf. Fig. 1.2), so that the right-hand side of Eq. (1.25) can be written as

$$\mathbf{r}_{ij} \times \mathbf{F}_{ji}.$$

If the internal forces between two particles, in addition to being equal and opposite, also lie along the line joining the particles—a condition known as the *strong law of action and reaction*—then all of these cross products vanish. The sum over pairs is zero under this assumption and Eq. (1.24) may be written in the form

$$\frac{d\mathbf{L}}{dt} = \mathbf{N}^{(e)}. \quad (1.26)$$

The time derivative of the total angular momentum is thus equal to the moment of the external force about the given point. Corresponding to Eq. (1.26) is the

Conservation Theorem for Total Angular Momentum: $\mathbf{L}$ *is constant in time if the applied (external) torque is zero.*

CONSTRAINTS

From the previous sections one might obtain the impression that all problems in mechanics have been reduced to solving the set of differential equations (1.19):

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i^{(e)} + \sum_j \mathbf{F}_{ji}.$$

One merely substitutes the various forces acting upon the particles of the system, turns the mathematical crank, and grinds out the answers! Even from a purely physical standpoint, however, this view is oversimplified. For example, it may be necessary to take into account the *constraints* that limit the motion of the system. We have already met one type of system involving constraints, namely rigid bodies, where the constraints on the motions of the particles keep the distances r_{ij} unchanged. Other examples of constrained systems can easily be furnished. The beads of an abacus are constrained to one-dimensional motion by the supporting wires. Gas molecules within a container are constrained by the walls of the vessel to move only *inside* the container. A particle placed on the surface of a solid sphere is subject to the constraint that it can move only on the surface or in the region exterior to the sphere.

Constraints may be classified in various ways, and we shall use the following system. If the conditions of constraint can be expressed as equations connecting the coordinates of the particles (and possibly the time) having the form

$$f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, t) = 0, \tag{1.37}$$

then the constraints are said to be *holonomic*. Perhaps the simplest example of holonomic constraints is the rigid body, where the constraints are expressed by equations of the form

$$(\mathbf{r}_i - \mathbf{r}_j)^2 - c_{ij}^2 = 0.$$

A particle constrained to move along any curve or on a given surface is another obvious example of a holonomic constraint, with the equations defining the curve or surface acting as the equations of a constraint.

Constraints not expressible in this fashion are called nonholonomic. The walls of a gas container constitute a nonholonomic constraint. The constraint involved in the example of a particle placed on the surface of a sphere is also nonholonomic, for it can be expressed as an inequality

$$r^2 - a^2 \geq 0$$

In the case of holonomic constraints, the first difficulty is solved by the introduction of *generalized coordinates*. So far we have been thinking implicitly in terms of Cartesian coordinates. A system of N particles, free from constraints, has $3N$ independent coordinates or *degrees of freedom*. If there exist holonomic constraints, expressed in k equations in the form (1.37), then we may use these equations to eliminate k of the $3N$ coordinates, and we are left with $3N - k$ independent coordinates, and the system is said to have $3N - k$ degrees of freedom. This elimination of the dependent coordinates can be expressed in another way, by the introduction of new, $3N - k$, independent variables $q_1, q_2, \ldots, q_{3N-k}$ in terms of which the old coordinates $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$ are expressed by equations of the form

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_1(q_1, q_2, \ldots, q_{3N-k}, t) \\ &\vdots \\ \mathbf{r}_N &= \mathbf{r}_N(q_1, q_2, \ldots, q_{3N-k}, t) \end{aligned} \tag{1.38}$$

containing the constraints in them implicitly. These are *transformation* equations from the set of $(\mathbf{r}_l)$ variables to the (q_l) set, or alternatively Eqs. (1.38) can be considered as parametric representations of the $(\mathbf{r}_l)$ variables. It is always assumed that we can also transform back from the (q_l) to the $(\mathbf{r}_l)$ set, i.e., that Eqs. (1.38) combined with the k equations of constraint can be inverted to obtain any q_i as a function of the $(\mathbf{r}_l)$ variable and time.

D'ALEMBERT'S PRINCIPLE AND LAGRANGE'S EQUATIONS

A virtual (infinitesimal) displacement of a system refers to a change in the configuration of the system as the result of any arbitrary infinitesimal change of the coordinates $\delta\mathbf{r}_i$, *consistent with the forces and constraints imposed on the system at the given instant* t. The displacement is called virtual to distinguish it from an actual displacement of the system occurring in a time interval dt, during which

the forces and constraints may be changing. Suppose the system is in equilibrium; i.e., the total force on each particle vanishes, $\mathbf{F}_i = 0$. Then clearly the dot product $\mathbf{F}_i \cdot \delta\mathbf{r}_i$, which is the virtual work of the force $\mathbf{F}_i$ in the displacement $\delta\mathbf{r}_i$, also vanishes. The sum of these vanishing products over all particles must likewise be zero:

$$\sum_i \mathbf{F}_i \cdot \delta\mathbf{r}_i = 0. \tag{1.40}$$

As yet nothing has been said that has any new physical content. Decompose $\mathbf{F}_i$ into the applied force, $\mathbf{F}_i^{(a)}$, and the force of constraint, $\mathbf{f}_i$,

$$\mathbf{F}_i = \mathbf{F}_i^{(a)} + \mathbf{f}_i, \tag{1.41}$$

so that Eq. (1.40) becomes

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta\mathbf{r}_i + \sum_i \mathbf{F}_i \cdot \delta\mathbf{r}_i = 0. \tag{1.42}$$

We now restrict ourselves to systems for which *the net virtual work of the forces of constraint is zero*. We have seen that this condition holds true for rigid bodies and it is valid for a large number of other constraints. Thus, if a particle is constrained to move on a surface, the force of constraint is perpendicular to the surface, while the virtual displacement must be tangent to it, and hence the virtual work vanishes. This is no longer true if sliding friction forces are present, and we must exclude such systems from our formulation. The restriction is not unduly hampering, since the friction is essentially a macroscopic phenomenon. On

We therefore have as the condition for equilibrium of a system that the virtual work of the *applied forces* vanishes:

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta\mathbf{r}_i = 0. \tag{1.43}$$

Equation (1.43) is often called the *principle of virtual work*. Note that the coefficients of $\delta\mathbf{r}_i$ can no longer be set equal to zero; i.e., in general $\mathbf{F}_i^{(a)} \neq 0$, since the $\delta\mathbf{r}_i$ are not completely independent but are connected by the constraints. In order to equate the coefficients to zero, we must transform the principle into a form involving the virtual displacements of the q_i, which are independent. Equation (1.43) satisfies our needs in that it does not contain the $\mathbf{f}_i$, but it deals only with statics; we want a condition involving the general motion of the system.

To obtain such a principle, we use a device first thought of by James Bernoulli and developed by D'Alembert. The equation of motion,

$$\mathbf{F}_i = \dot{\mathbf{p}}_i,$$

can be written as

$$\mathbf{F}_i - \dot{\mathbf{p}}_i = 0,$$

which states that the particles in the system will be in equilibrium under a force equal to the actual force plus a "reversed effective force" $-\dot{\mathbf{p}}_i$. Instead of (1.40), we can immediately write

$$\sum_i (\mathbf{F}_i - \dot{\mathbf{p}}_i) \cdot \delta\mathbf{r}_i = 0, \tag{1.44}$$

and, making the same resolution into applied forces and forces of constraint, there results

$$\sum_i (\mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i) \cdot \delta\mathbf{r}_i + \sum_i \mathbf{f}_i \cdot \delta\mathbf{r}_i = 0.$$

We again restrict ourselves to systems for which the virtual work of the forces of constraint vanishes and therefore obtain

$$\sum_i (\mathbf{F}_i^{(a)} - \dot{\mathbf{p}}_i) \cdot \delta\mathbf{r}_i = 0, \tag{1.45}$$

which is often called *D'Alembert's principle*. We have achieved our aim, in that the forces of constraint no longer appear, and the superscript $^{(a)}$ can now be dropped without ambiguity. It is still not in a useful form to furnish equations of motion for the system. We must now transform the principle into an expression involving virtual displacements of the generalized coordinates, which are then independent of each other (for holonomic constraints), so that the coefficients of the δq_i can be set separately equal to zero.

The translation from $\mathbf{r}_i$ to q_j language starts from the transformation equations (1.38),

$$\mathbf{r}_i = \mathbf{r}_i(q_1, q_2, \ldots, q_n, t) \tag{1.45'}$$

(assuming n independent coordinates), and is carried out by means of the usual "chain rules" of the calculus of partial differentiation. Thus, $\mathbf{v}_i$ is expressed in terms of the $\dot{q}_k$ by the formula

$$\mathbf{v}_i \equiv \frac{d\mathbf{r}_i}{dt} = \sum_k \frac{\partial \mathbf{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \mathbf{r}_i}{\partial t}. \tag{1.46}$$

Similarly, the arbitrary virtual displacement $\delta\mathbf{r}_i$ can be connected with the virtual displacements δq_i by

$$\delta\mathbf{r}_i = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j. \tag{1.47}$$

Note that no variation of time, δt, is involved here, since a virtual displacement by definition considers only displacements of the coordinates. (Only then is the virtual displacement perpendicular to the force of constraint if the constraint itself is changing in time.)

In terms of the generalized coordinates, the virtual work of the $\mathbf{F}_i$ becomes

$$\begin{aligned}\sum_i \mathbf{F}_i \cdot \delta\mathbf{r}_i &= \sum_{i,j} \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j \\ &= \sum_j Q_j \delta q_j,\end{aligned} \tag{1.48}$$

where the Q_j are called the components of the *generalized force*, defined as

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}. \tag{1.49}$$

Note that just as the q's need not have the dimensions of length, so the Q's do not necessarily have the dimensions of force, but $Q_j \delta q_j$ must always have the dimensions of work. For example, Q_j might be a torque N_j and dq_j a differential angle $d\theta_j$, which makes $N_j\, d\theta_j$ a differential of work.

We turn now to the other other term involved in Eq. (1.45), which may be written as

$$\sum_i \dot{\mathbf{p}}_i \cdot \delta\mathbf{r}_i = \sum_i m_i \ddot{\mathbf{r}}_i \cdot \delta\mathbf{r}_i.$$

Expressing $\delta\mathbf{r}_i$ by (1.47), this becomes

$$\sum_{i,j} m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j.$$

Consider now the relation

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \left[\frac{d}{dt}\left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}\right) - m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right)\right]. \tag{1.50}$$

In the last term of Eq. (1.50) we can interchange the differentiation with respect

$$\frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_j}\right) = \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j} = \sum_k \frac{\partial^2 \mathbf{r}_i}{\partial q_j \partial q_k}\dot{q}_k + \frac{\partial^2 \mathbf{r}_i}{\partial q_j \partial t},$$
$$= \frac{\partial \mathbf{v}_i}{\partial q_j},$$

by Eq. (1.46). Further, we also see from Eq. (1.46) that

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j}. \tag{1.51}$$

Substitution of these changes in (1.50) leads to the result that

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \left[\frac{d}{dt}\left(m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j}\right) - m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_j}\right],$$

and the second term on the left-hand side of Eq. (1.45) can be expanded into

$$\sum_j \left\{\frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\sum_i \frac{1}{2} m_i v_i^2\right)\right] - \frac{\partial}{\partial q_j}\left(\sum_i \frac{1}{2} m_i v_i^2\right) - Q_j\right\} \delta q_j.$$

Identifying $\sum_i \frac{1}{2} m_i v_i^2$ with the system kinetic energy T, D'Alembert's principle (cf. Eq. (1.45)) becomes

$$\sum \left\{\left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j}\right] - Q_j\right\} \delta q_j = 0. \tag{1.52}$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j. \tag{1.53}$$

There are n such equations in all.

When the forces are derivable from a scalar potential function V,

$$\mathbf{F}_i = -\nabla_i V.$$

Then the generalized forces can be written as

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum_i \nabla_i V \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

which is exactly the same expression for the partial derivative of a function $-V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t)$ with respect to q_j:

$$Q_j = -\frac{\partial V}{\partial q_i}. \tag{1.54}$$

Equations (1.53) can then be rewritten as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial (T - V)}{\partial q_i} = 0. \tag{1.55}$$

The equations of motion in the form (1.55) are not necessarily restricted to conservative systems; only if V is not an explicit function of time is the system conservative (cf. p. 4). As here defined, the potential V does not depend on the generalized velocities. Hence, we can include a term in V in the partial derivative with respect to $\dot{q}_j$:

$$\frac{d}{dt}\left(\frac{\partial (T - V)}{\partial \dot{q}_j}\right) - \frac{\partial (T - V)}{\partial q_j} = 0.$$

Or, defining a new function, the *Lagrangian* L, as

$$L = T - V, \tag{1.56}$$

the Eqs. (1.53) become

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0, \tag{1.57}$$

expressions referred to as "Lagrange's equations."

Note that for a particular set of equations of motion there is no unique choice of Lagrangian such that Eqs. (1.57) lead to the equations of motion in the given generalized coordinates. Thus, in Derivations 8 and 10 it is shown that if $L(q, \dot{q}, t)$ is an approximate Lagrangian and $F(q, t)$ is *any* differentiable function of the generalized coordinates and time, then

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{dF}{dt} \tag{1.57'}$$

VELOCITY-DEPENDENT POTENTIALS AND THE DISSIPATION FUNCTION

Lagrange's equations can be put in the form (1.57) even if there is no potential function, V, in the usual sense, providing the generalized forces are obtained from a function $U(q_j, \dot{q}_j)$ by the prescription

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_j}\right). \tag{1.58}$$

In such case, Eqs. (1.57) still follow from Eqs. (1.53) with the Lagrangian given by

$$L = T - U. \tag{1.59}$$

Here U may be called a "generalized potential," or "velocity-dependent potential." The possibility of using such a "potential" is not academic; it applies to one very important type of force field, namely, the electromagnetic forces on moving charges. Considering its importance, a digression on this subject is well worthwhile.

Consider an electric charge, q, of mass m moving at a velocity, $\mathbf{v}$, in an otherwise charge-free region containing both an electric field, $\mathbf{E}$, and a magnetic field, $\mathbf{B}$, which may depend upon time and position. The charge experiences a force, called the Lorentz force, given by

$$\mathbf{F} = q[\mathbf{E} + (\mathbf{v} \times \mathbf{B})]. \tag{1.60}$$

Both $\mathbf{E}(t, x, y, z)$ and $\mathbf{B}(t, x, y, z)$ are continuous functions of time and position derivable from a scalar potential $\phi(t, x, y, z)$ and a vector potential $\mathbf{A}(t, x, y, z)$ by

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{1.61a}$$

and

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{1.61b}$$

The force on the charge can be derived from the following velocity-dependent potential energy

$$U = q\phi - q\mathbf{A} \cdot \mathbf{v}, \tag{1.62}$$

so the Lagrangian, $L = T - U$, is

Considering just the x-component of Lagrange's equations gives

$$m\ddot{x} = q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x}\right) - q\left(\frac{\partial \phi}{\partial x} + \frac{dA_x}{dt}\right). \tag{1.64}$$

The total time derivative of A_x is related to the particle time derivative through

$$\begin{aligned}\frac{dA_x}{dt} &= \frac{\partial A_x}{\partial t} + \mathbf{v}\cdot\nabla A_x \\ &= \frac{\partial A_x}{\partial t} + v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_x}{\partial y} + v_z\frac{\partial A_x}{\partial z}.\end{aligned} \tag{1.65}$$

Equation (1.61b) gives

$$(\mathbf{v}\times\mathbf{B})_x = v_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + v_z\left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\right).$$

Combining these expressions gives the equation of motion in the x-direction

$$m\ddot{x} = q\left[\mathrm{E}_x + (\mathbf{v}\times\mathbf{B})_x\right]. \tag{1.66}$$

On a component-by-component comparison, Eqs. (1.66) and (1.60) are identical, showing that the Lorentz force equation is derivable from Eqs. (1.61) and (1.62).

Note that if not all the forces acting on the system are derivable from a potential, then Lagrange's equations can always be written in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = Q_j,$$

where L contains the potential of the conservative forces as before, and Q_j represents the forces *not* arising from a potential. Such a situation often occurs when frictional forces are present. It frequently happens that the frictional force is proportional to the velocity of the particle, so that its x-component has the form

$$F_{fx} = -k_x v_x.$$

Frictional forces of this type may be derived in terms of a function $\mathcal{F}$, known as *Rayleigh's dissipation function*, and defined as

$$\mathcal{F} = \frac{1}{2}\sum_i\left(k_x v_{ix}^2 + k_y v_{iy}^2 + k_z v_{iz}^2\right), \tag{1.67}$$

where the summation is over the particles of the system. From this definition it is clear that

$$F_{fx} = -\frac{\partial \mathcal{F}}{\partial v_x},$$

or, symbolically,

$$\mathbf{F}_f = -\nabla_v \mathcal{F}. \tag{1.68}$$

We can also give a physical interpretation to the dissipation function. The work done *by* the system *against* friction is

$$dW_f = -\mathbf{F}_f \cdot d\mathbf{r} = -\mathbf{F}_f \cdot \mathbf{v}\, dt = \left(k_x v_x^2 + k_y v_y^2 + k_z v_z^2\right) dt.$$

Hence, $2\mathcal{F}$ is the rate of energy dissipation due to friction. The component of the generalized force resulting from the force of friction is then given by

$$\begin{aligned} Q_j = \sum_i \mathbf{F}_{fi} \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} &= -\sum \nabla_v \mathcal{F} \cdot \frac{\partial \mathbf{r}_i}{\partial q_i} \\ &= -\sum \nabla_v \mathcal{F} \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j}, \qquad \text{by (1.51),} \\ &= -\frac{\partial \mathcal{F}}{\partial \dot{q}_j}. \end{aligned} \tag{1.69}$$

An example is Stokes' law, whereby a sphere of radius a moving at a speed v, in a medium of viscosity η experiences the frictional drag force $\mathbf{F}_f = 6\pi\eta\mathbf{a}v$. The Lagrange equations with dissipation become

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} + \frac{\partial \mathcal{F}}{\partial \dot{q}_j} = 0, \tag{1.70}$$

so that two scalar functions, L and $\mathcal{F}$, must be specified to obtain the equations of motion.

SIMPLE APPLICATIONS OF THE LAGRANGIAN FORMULATION

Let us now consider simple examples of this procedure:

1. Single particle in space
   (a) Cartesian coordinates
   (b) Plane polar coordinates
2. Atwood's machine
3. Time-dependent constraint—bead sliding on rotating wire

1. (a) *Motion of one particle: using Cartesian coordinates.* The generalized forces needed in Eq. (1.53) are obviously F_x, F_y, and F_z. Then

$$T = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right),$$

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0,$$

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x}, \qquad \frac{\partial T}{\partial \dot{y}} = m\dot{y}, \qquad \frac{\partial T}{\partial \dot{z}} = m\dot{z},$$

and the equations of motion are

$$\frac{d}{dt}(m\dot{x}) = F_x, \quad \frac{d}{dt}(m\dot{y}) = F_y, \quad \frac{d}{dt}(m\dot{z}) = F_z. \tag{1.74}$$

We are thus led back to the original Newton's equations of motion.

(b) *Motion of one particle: using plane polar coordinates.* Here we must express T in terms of $\dot{r}$ and $\dot{\theta}$. The equations of transformation, i.e., Eqs. (1.38), in this case are simply

$$x = r\cos\theta$$

$$y = r\sin\theta.$$

By analogy to (1.46), the velocities are given by

$$\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta,$$

$$\dot{y} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta.$$

The kinetic energy $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$ then reduces formally to

$$T = \tfrac{1}{2}m\left[\dot{r}^2 + \left(r\dot{\theta}\right)^2\right]. \tag{1.75}$$

An alternative derivation of Eq. (1.75) is obtained by recognizing that the plane polar components of the velocity are $\dot{r}$ along $\mathbf{r}$, and $r\dot{\theta}$ along the direction perpendicular to r, denoted by the unit vector $\mathbf{n}$. Hence, the square of the velocity expressed in polar coordinates is simply $\dot{r}^2 + (r\dot{\theta})^2$. With the aid of the expression

$$d\mathbf{r} = \hat{\mathbf{r}}\,dr + r\hat{\boldsymbol{\theta}}\,d\theta + \hat{\mathbf{k}}\,dz$$

for the differential position vector, $d\mathbf{r}$, in cylindrical coordinates, where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors in the $\mathbf{r}$ and $\boldsymbol{\theta}$-directions, respectively, the components of the generalized force can be obtained from the definition, Eq. (1.49),

$$Q_r = \mathbf{F}\cdot\frac{\partial \mathbf{r}}{\partial r} = \mathbf{F}\cdot\hat{\mathbf{r}} = F_r,$$

$$Q_\theta = \mathbf{F}\cdot\frac{\partial \mathbf{r}}{\partial \theta} = \mathbf{F}\cdot r\hat{\boldsymbol{\theta}} = rF_\theta,$$

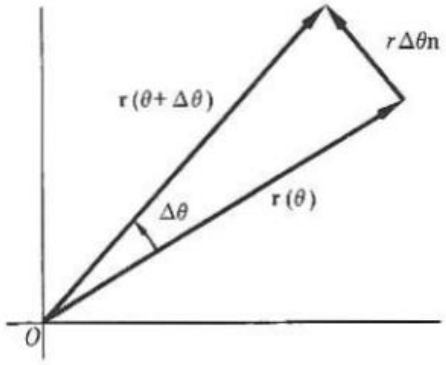


FIGURE 1.6 Derivative of r with respect to θ.

since the derivative of $\mathbf{r}$ with respect to θ is, by the definition of a derivative, a vector in the direction of $\hat{\theta}$ (cf. Fig. 1.6). There are two generalized coordinates, and therefore two Lagrange equations. The derivatives occurring in the r equation are

$$\frac{\partial T}{\partial r} = mr\dot{\theta}^2, \qquad \frac{\partial T}{\partial \dot{r}} = m\dot{r}, \qquad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}}\right) = m\ddot{r},$$

and the equation itself is

$$m\ddot{r} - mr\dot{\theta}^2 = F_r,$$

the second term being the centripetal acceleration term. For the θ equation, we have the derivatives

$$\frac{\partial T}{\partial \theta} = 0, \qquad \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta}, \qquad \frac{d}{dt}\left(mr^2\dot{\theta}\right) = mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta},$$

so that the equation becomes

$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) = mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} = rF_\theta.$$

Note that the left side of the equation is just the time derivative of the angular momentum, and the right side is exactly the applied torque, so that we have simply rederived the torque equation (1.26), where $L = mr^2\dot{\theta}$ and $N^{(\ell)} = rF_\theta$.

2. *Atwood's machine*—(See Fig. 1.7) an example of a conservative system with holonomic, scleronomous constraint (the pulley is assumed frictionless and massless). Clearly there is only one independent coordinate x, the position of the other weight being determined by the constraint that the length of the rope between them is l. The potential energy is

$$V = -M_1gx - M_2g(l - x),$$

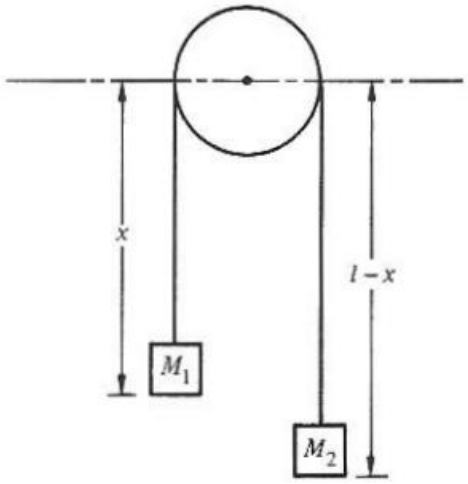


FIGURE 1.7 Atwood's machine.

while the kinetic energy is

$$T = \tfrac{1}{2}(M_1 + M_2)\dot{x}^2.$$

Combining the two, the Lagrangian has the form

$$L = T - V = \tfrac{1}{2}(M_1 + M_2)\dot{x}^2 + M_1 g x + M_2 g(l - x).$$

There is only one equation of motion, involving the derivatives

$$\frac{\delta L}{\partial x} = (M_1 - M_2)\, g,$$

$$\frac{\partial L}{\partial \dot{x}} = (M_1 + M_2)\,\dot{x},$$

so that we have

$$(M_1 + M_2)\,\ddot{x} = (M_1 - M_2)\, g,$$

or

$$\ddot{x} = \frac{M_1 - M_2}{M_1 + M_2} g,$$

3. *A bead (or ring) sliding on a uniformly rotating wire in a force-free space.* The wire is straight, and is rotated uniformly about some fixed axis perpendicular to the wire. This example has been chosen as a simple illustration of a constraint being time dependent, with the rotation axis along z and the wire in the xy plane. The transformation equations explicitly contain the time.

$$x = r\cos\omega t, \qquad (\omega = \text{angular velocity of rotation})$$

$$y = r\sin\omega t. \qquad (r = \text{distance along wire from rotation axis})$$

While we could then find T (here the same as L) by the same procedure used to obtain (1.71), it is simpler to take over (1.75) directly, expressing the constraint by the relation $\dot{\theta} = \omega$:

$$T = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\omega^2\right).$$

Note that T is not a homogeneous quadratic function of the generalized velocities, since there is now an additional term not involving $\dot{r}$. The equation of motion is then

$$m\ddot{r} = mr\omega^2 = 0$$

or

$$\ddot{r} = r\omega^2,$$

which is the familiar simple harmonic oscillator equation with a change of sign.

HAMILTON'S PRINCIPLE

The integral *Hamilton's principle* describes the motion of those mechanical systems for which all forces (except the forces of constraint) are derivable from a generalized scalar potential that may be a function of the coordinates, velocities, and time. Such systems will be denoted as *monogenic*. Where the potential is an explicit function of position coordinates only, then a monogenic system is also conservative (cf. Section 1.2).

For monogenic systems, Hamilton's principle can be stated as

> *The motion of the system from time t_1 to time t_2 is such that the line integral (called the* action *or the* action integral*),*

$$I = \int_{t_1}^{t_2} L\,dt, \tag{2.1}$$

> *where $L = T - V$, has a stationary value for the actual path of the motion.*

We can summarize Hamilton's principle by saying that the motion is such that the *variation* of the line integral I for fixed t_1 and t_2 is zero:

$$\delta I = \delta \int_{t_1}^{t_2} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\,dt = 0. \tag{2.2}$$

Where the system constraints are holonomic, Hamilton's principle, Eq. (2.2), is both a necessary and sufficient condition for Lagrange's equations, Eqs. (1.57). Thus, it can be shown that Hamilton's principle follows directly from Lagrange's equations. Instead, however, we shall prove the converse, namely, that Lagrange's equations follow from Hamilton's principle, as being the more important theorem.

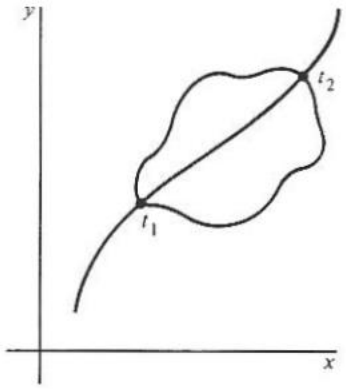


FIGURE 2.1 Path of the system point in configuration space.

3. *The brachistochrone problem.* (See Fig. 2.4a.) This well-known problem is to find the curve joining two points, along which a particle falling from rest under the influence of gravity travels from the higher to the lower point in the least time.

If v is the speed along the curve, then the time required to fall an arc length ds is ds/v, and the problem is to find a minimum of the integral

$$t_{12} = \int_1^2 \frac{ds}{v}.$$

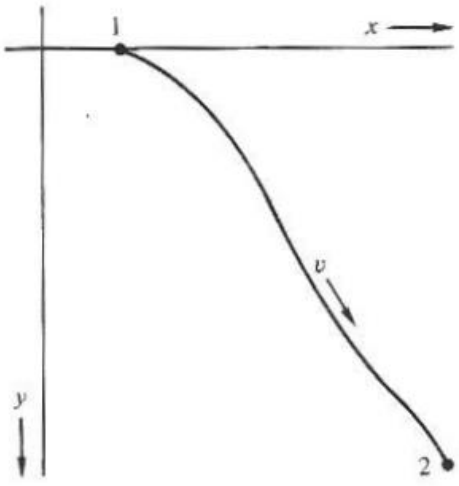


FIGURE 2.4a The brachistochrone problem.

If y is measured down from the initial point of release, the conservation theorem for the energy of the particle can be written as

$$\tfrac{1}{2}mv^2 = mgy$$

or

$$v = \sqrt{2gy}.$$

Then the expression for t_{12} becomes

$$t_{12} = \int_1^2 \frac{\sqrt{1+\dot{y}^2}}{\sqrt{2gy}}\,dx,$$

and f is identified as

$$f = \sqrt{\frac{1+\dot{y}^2}{2gy}}.$$

The integration of Eq. (2.11) with this form for f is straightforward and is left as an exercise.

The solution in terms of its one parameter, a, given by

$$\frac{y}{a} = 1 - \cos\left[\frac{x+\sqrt{y(2a-y)}}{a}\right],$$

is sketched in Fig. 2.4b for the first cycle ($0 \le x \le 2\pi a$) and the beginning of the second cycle. Three cases of solutions are indicated. A power-series expansion of the solution for the limit $y \ll a$ gives

$$y = \frac{x^2}{2}a.$$

The brachistochrone problem is famous in the history of mathematics, for it was the analysis of this problem by John Bernoulli that led to the formal foundation of the calculus of variations.

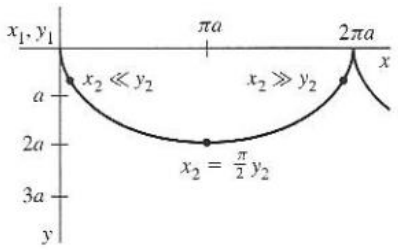


FIGURE 2.4b Catenary solution to the brachistochrone problem showing positions on the curve for the three cases $x_2 \ll y_2$, $x_2 = \frac{\pi}{2} y_2$, and $x_2 \gg y_2$.

DERIVATION OF LAGRANGE'S EQUATIONS FROM HAMILTON'S PRINCIPLE

The fundamental problem of the calculus of variations is easily generalized to the case where f is a function of many independent variables y_i, and their derivatives $\dot{y}_i$. (Of course, all these quantities are considered as functions of the parametric variable x.) Then a variation of the integral J,

$$\delta J = \delta \int_1^2 f(y_1(x); y_2(x), \ldots, \dot{y}_1(x); \dot{y}_2(x), \ldots, x)\, dx, \tag{2.14}$$

is obtained, as before, by considering J as a function of parameter α that labels a possible set of curves $y_1(x, \alpha)$. Thus, we may introduce α by setting

$$\begin{aligned} y_1(x, \alpha) &= y_1(x, 0) + \alpha\eta_1(x), \\ y_2(x, \alpha) &= y_2(x, 0) + \alpha\eta_2(x), \\ &\vdots \qquad\quad \vdots \qquad\quad \vdots \end{aligned} \tag{2.15}$$

where $y_1(x, 0)$, $y_2(x, 0)$, etc., are the solutions of the extremum problem (to be obtained) and η_1, η_2, etc., are independent functions of x that vanish at the end points and that are continuous through the second derivative, but otherwise are completely arbitrary.

The calculation proceeds as before. The variation of J is given in terms of

$$\frac{\partial J}{\partial \alpha}\, d\alpha = \int_1^2 \sum_i \left(\frac{\partial f}{\partial y_i} \frac{\partial y_i}{\partial \alpha}\, d\alpha + \frac{\partial f}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial \alpha}\, d\alpha \right) dx. \tag{2.16}$$

Again we integrate by parts the integral involved in the second sum of Eq. (2.16):

$$\int_1^2 \frac{\delta f}{\partial \dot{y}_i} \frac{\partial^2 y_i}{\partial \alpha\, \partial x}\, dx = \frac{\partial f}{\partial \dot{y}_i} \frac{\partial y_i}{\partial \alpha}\bigg|_1^2 - \int_1^2 \frac{\partial y_i}{\partial \alpha} \frac{d}{dx} \left(\frac{\partial f}{\partial \dot{y}_i} \right) dx,$$

where the first term vanishes because all curves pass through the fixed end points. Substituting in (2.16), δJ becomes

$$\delta J = \int_1^2 \sum_i \left(\frac{\partial f}{\partial y_i} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}_i} \right) \delta y_i\, dx, \tag{2.17}$$

where, in analogy with (2.12), the variation δy_i is

$$\delta y_i = \left(\frac{\partial y_i}{\partial \alpha} \right)_0 d\alpha.$$

Since the y variables are independent, the variations δy_i are independent (e.g., the functions $\eta_i(x)$ will be independent of each other). Hence, by an obvious extension of the fundamental lemma (cf. Eq. (2.10)), the condition that δJ is zero

requires that the coefficients of the δy_i separately vanish:

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial \dot{y}_i} = 0, \qquad i = 1, 2, \ldots, n. \tag{2.18}$$

Equations (2.18) represent the appropriate generalization of (2.11) to several variables and are known as the *Euler-Lagrange differential equations*. Their so-

For present purposes, what we have derived here suffices, for the integral in Hamilton's principle,

$$I = \int_1^2 L(q_i, \dot{q}_i, t)\, dt, \tag{2.19}$$

has just the form stipulated in (2.14) with the transformation

$$x \to t$$

$$y_i \to q_i$$

$$f(y_i, \dot{y}_i, x) \to L(q_i, \dot{q}_i, t).$$

In deriving Eqs. (2.18), we assumed that the y_i variables are independent. The corresponding condition in connection with Hamilton's principle is that the generalized coordinates q_i be independent, which requires that the constraints be holonomic. The Euler-Lagrange equations corresponding to the integral I then become the Lagrange equations of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \ldots, n,$$

and we have accomplished our original aim, to show that Lagrange's equations follow from Hamilton's principle—for monogenic systems with holonomic constraints.

CONSERVATION THEOREMS AND SYMMETRY PROPERTIES

In many problems a number of first integrals of the equations of motion can be obtained immediately; by this we mean relations of the type

$$f(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots, t) = \text{constant}, \tag{2.43}$$

which are first-order differential equations. These first integrals are of interest because they tell us something physically about the system. They include, in fact, the conservation laws obtained in Chapter 1.

Let us consider as an example a system of mass points under the influence of forces derived from potentials dependent on position only. Then

$$\frac{\partial L}{\partial \dot{x}_i} \equiv \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum \frac{1}{2} m_i \left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2 \right)$$
$$= m_i \dot{x}_i = p_{ix},$$

which is the x component of the linear momentum associated with the ith particle. This result suggests an obvious extension to the concept of momentum. The generalized momentum associated with the coordinate q_j shall be defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j}. \tag{2.44}$$

The terms *canonical momentum* and *conjugate momentum* are often also used for p_j. Notice that if q_j is not a Cartesian coordinate, p_j does not necessarily have the dimensions of a linear momentum. Further, if there is a velocity-dependent potential, then even with a Cartesian coordinate q_j the associated *generalized* momentum will not be identical with the usual *mechanical* momentum. Thus, in the case of a group of particles in an electromagnetic field, the Lagrangian is (cf. 1.63)

$$L = \sum_i \frac{1}{2} m_i \dot{r}_i^2 - \sum_i q_i \phi(x_i) + \sum_i q_i \mathbf{A}(x_i) \cdot \dot{\mathbf{r}}_i$$

(q_i here denotes charge) and the generalized momentum conjugate to x_i is

$$p_{ix} = \frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i + q_i A_x, \tag{2.45}$$

i.e., mechanical momentum plus an additional term.

If the Lagrangian of a system does not contain a given coordinate q_j (although it may contain the corresponding velocity $\dot{q}_j$), then the coordinate is said to be *cyclic* or *ignorable*. This definition is not universal, but it is the customary one and will be used here. The Lagrange equation of motion,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0,$$

reduces, for a cyclic coordinate, to

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = 0$$

or

$$\frac{dp_j}{dt} = 0,$$

which mean that

$$p_j = \text{constant}. \tag{2.46}$$

Hence, we can state as a general conservation theorem that *the generalized momentum conjugate to a cyclic coordinate is conserved.*

We first consider a generalized coordinate q_j, for which a change dq_j represents a translation of the system as a whole in some given direction. An example would be one of the Cartesian coordinates of the center of mass of the system. Then clearly q_j cannot appear in T, for velocities are not affected by a shift in the origin, and therefore the partial derivative of T with respect to q_j must be zero. Further, we will assume conservative systems for which V is not a function of the velocities, so as to eliminate such complications as electromagnetic forces. The Lagrange equation of motion for a coordinate so defined then reduces to

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} \equiv \dot{p}_j = -\frac{\partial V}{\partial q_j} \equiv Q_j. \tag{2.48}$$

We will now show that (2.48) is the equation of motion for the total linear momentum, i.e., that Q_j represents the component of the total force along the direction of translation of q_j, and p_j is the component of the total linear momentum along this direction. In general, the generalized force Q_j is given by Eq. (1.49):

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}.$$

Since dq_j corresponds to a translation of the system along some axis, the vectors $\mathbf{r}_i(q_j)$ and $\mathbf{r}_i(q_j + dq_j)$ are related as shown in Fig. 2.7. By the definition of a derivative, we have

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \lim_{dq_j \to 0} \frac{\mathbf{r}_i\left(q_j + dq_j\right) - \mathbf{r}_i\left(q_j\right)}{dq_j} = \frac{dq_j}{dq_j}\mathbf{n} = \mathbf{n}, \tag{2.49}$$

where $\mathbf{n}$ is the unit vector along the direction of the translation. Hence,

$$Q_j = \sum \mathbf{F}_i \cdot \mathbf{n} = \mathbf{n} \cdot \mathbf{F},$$

which (as was stated) is the component of the total force in the direction of $\mathbf{n}$. To prove the other half of the statement, note that with the kinetic energy in the form

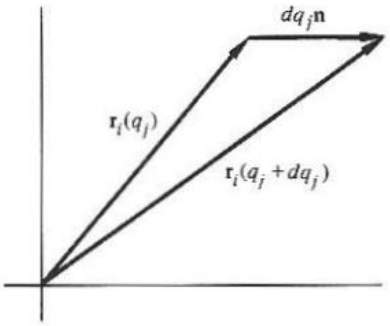


FIGURE 2.7 Change in a position vector under translation of the system.

$$T = \frac{1}{2}\sum m_i \dot{\mathbf{r}}_i^2,$$

the conjugate momentum is

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j}$$

$$= \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

using Eq. (1.51). Then from Eq. (2.49)

$$p_j = \mathbf{n} \cdot \sum_i m_i \mathbf{v}_i,$$

which again, as predicted, is the component of the total system linear momentum along **n**.

Suppose now that the translation coordinate q_j that we have been discussing is cyclic. Then q_j cannot appear in V and therefore

$$-\frac{\partial V}{\partial q_j} \equiv Q_j = 0.$$

But this is simply the familiar conservation theorem for linear momentum—that if a given component of the total applied force vanishes, the corresponding component of the linear momentum is conserved.

The generalized force Q_j is again given by

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

only the derivative now has a different meaning. Here the change in q_j must correspond to an infinitesimal rotation of the vector $\mathbf{r}_i$, keeping the magnitude of the vector constant. From Fig. 2.8, the magnitude of the derivative can easily be obtained:

$$|d\mathbf{r}_i| = r_i \sin\theta \, dq_j$$

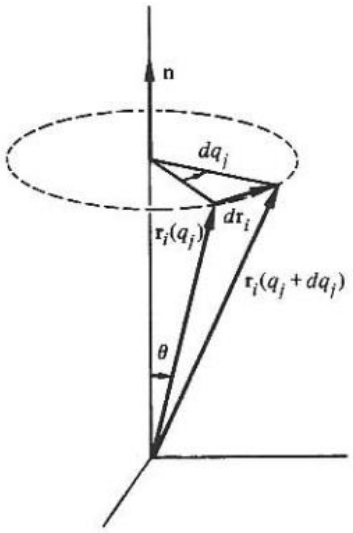


FIGURE 2.8 Change of a position vector under rotation of the system.

and

$$\left|\frac{\partial \mathbf{r}_i}{\partial q_j}\right| = r_i \sin\theta,$$

and its direction is perpendicular to both $\mathbf{r}_i$ and $\mathbf{n}$. Clearly, the derivative can be written in vector form as

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \mathbf{n} \times \mathbf{r}_i. \tag{2.50}$$

With this result, the generalized force becomes

$$\begin{aligned} Q_j &= \sum_i \mathbf{F}_i \cdot \mathbf{n} \times \mathbf{r}_i \\ &= \sum_i \mathbf{n} \cdot \mathbf{r}_i \times \mathbf{F}_i, \end{aligned}$$

reducing to

$$Q_j = \mathbf{n} \cdot \sum_i \mathbf{N}_i = \mathbf{n} \cdot \mathbf{N},$$

which proves the first part. A similar manipulation of p_j with the aid of Eq. (2.50) provides proof of the second part of the statement:

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \mathbf{n} \cdot \mathbf{r}_i \times m_i \mathbf{v}_i = \mathbf{n} \cdot \sum \mathbf{L}_i = \mathbf{n} \cdot \mathbf{L}.$$

ENERGY FUNCTION AND THE CONSERVATION OF ENERGY

Another conservation theorem we should expect to obtain in the Lagrangian formulation is the conservation of total energy for systems where the forces are derivable from potentials dependent only upon position. Indeed, it is possible to demonstrate a conservation theorem for which conservation of total energy represents only a special case. Consider a general Lagrangian, which will be a function of the coordinates q_j and the velocities $\dot{q}_j$ and may also depend explicitly on the time. (The explicit time dependence may arise from the time variation of external potentials, or from time-dependent constraints.) Then the total time derivative of L is

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j}\frac{dq_j}{dt} + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}. \tag{2.51}$$

From Lagrange's equations,

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right),$$

and (2.51) can be rewritten as

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}$$

or

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j\frac{\partial L}{\partial \dot{q}_j}\right) + \frac{\partial L}{\partial t}.$$

It therefore follows that

$$\frac{d}{dt}\left(\sum_j \dot{q}_j\frac{\partial L}{\partial \dot{q}_j} - L\right) + \frac{\partial L}{\partial t} = 0. \tag{2.52}$$

The quantity in parentheses is oftentimes called the *energy function** and will be denoted by h:

$$h(q_1, \ldots, q_n; \quad \dot{q}_1, \ldots, \dot{q}_n; \quad t) = \sum_j \dot{q}_j\frac{\partial L}{\partial \dot{q}_j} - L, \tag{2.53}$$

and Eq. (2.52) can be looked on as giving the total time derivative of h:

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t}. \tag{2.54}$$

If the Lagrangian is not an explicit function of time, i.e., if t does not appear in L explicitly but only implicitly through the time variation of q and $\dot{q}$, then Eq. (2.54) says that h is conserved. It is one of the first integrals of the motion and is sometimes referred to as Jacobi's integral.†

Under certain circumstances, the function h is the total energy of the system. To determine what these circumstances are, we recall that the total kinetic energy of a system can always be written as

$$T = T_0 + T_1 + T_2, \tag{1.73}$$

where T_0 is a function of the generalized coordinates only, $T_1(q, \dot{q})$ is linear in the generalized velocities, and $T_2(q, \dot{q})$ is a quadratic function of the $\dot{q}$'s. For a very wide range of systems and sets of generalized coordinates, the Lagrangian can be similarly decomposed as regards its functional behavior in the $\dot{q}$ variables:

$$L(q, \dot{q}, t) = L_0(q, t) + L_1(q, \dot{q}, t) + L_2(q, \dot{q}, t). \tag{2.55}$$

$$\sum_i x_i \frac{\partial f}{\partial x_i} = nf. \tag{2.56}$$

Applied to the function h, Eq. (2.53), for the Lagrangians of the form (2.55), this theorem implies that

$$h = 2L_2 + L_1 - L = L_2 - L_0. \tag{2.57}$$

If the transformation equations defining the generalized coordinates, Eqs. (1.38), do not involve the time explicitly, then by Eqs. (1.73) $T = T_2$. If, further, the potential does not depend on the generalized velocities, then $L_2 = T$ and $L_0 = -V$, so that

$$h = T + V = E, \tag{2.58}$$

and the energy function is indeed the total energy. Under these circumstances, if V does not involve the time explicitly, neither will L. Thus, by Eq. (2.54), h (which is here the total energy), will be conserved.

Finally, note that where the system is not conservative, but there are frictional forces derivable from a dissipation function $\mathcal{F}$, it can be easily shown that $\mathcal{F}$ is related to the decay rate of h. When the equations of motion are given by Eq. (1.70), including dissipation, then Eq. (2.52) has the form

$$\frac{dh}{dt} + \frac{\partial L}{\partial t} = \sum_j \frac{\partial \mathcal{F}}{\partial \dot{q}_j} \dot{q}_j.$$

By the definition of $\mathcal{F}$, Eq. (1.67), it is a homogeneous function of the $\dot{q}$'s of degree 2. Hence, applying Euler's theorem again, we have

$$\frac{dh}{dt} = -2\mathcal{F} - \frac{\partial L}{\partial t}. \tag{2.59}$$

If L is not an explicit function of time, *and* the system is such that h is the same as the energy, then Eq. (2.59) says that $2\mathcal{F}$ is the rate of energy dissipation,

$$\frac{dE}{dt} = -2\mathcal{F}, \tag{2.60}$$

a statement proved above (cf. Sec. 1.5) in less general circumstances.

****

Unit II: Central Force Problem and Classical Scattering

Central Force Problem and Classical Scattering
Reduction to the Equivalent One-Body Problem - The Equations of Motion and First Integrals – Classification of Orbits – The Viral Theorem – Kepler's Problem: Inverse Square Law of Force –The Motion in time in the Kepler's problem - Scattering in a Central Force Field - Transformation of the Scattering Problem to Laboratory Coordinates- Three body problem.

REDUCTION TO THE EQUIVALENT ONE-BODY PROBLEM

Consider a monogenic system of two mass points, m_1 and m_2 (cf. Fig. 3.1), where the only forces are those due to an interaction potential U. We will assume at first that U is any function of the vector between the two particles, $\mathbf{r}_2 - \mathbf{r}_1$, or of their relative velocity, $\dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1$, or of any higher derivatives of $\mathbf{r}_2 - \mathbf{r}_1$. Such a system has six degrees of freedom and hence six independent generalized coordinates. We choose these to be the three components of the radius vector to the center of mass, $\mathbf{R}$, plus the three components of the difference vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. The Lagrangian will then have the form

$$L = T(\dot{\mathbf{R}}, \dot{\mathbf{r}}) - U(\mathbf{r}, \dot{\mathbf{r}}, \ldots). \tag{3.1}$$

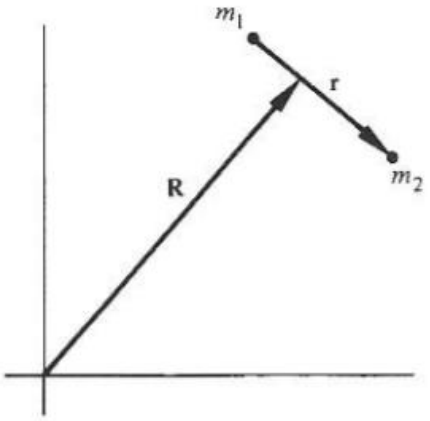


FIGURE 3.1 Coordinates for the two-body problem.

The kinetic energy T can be written as the sum of the kinetic energy of the motion of the center of mass, plus the kinetic energy of motion about the center of mass, T':

$$T = \tfrac{1}{2}(m_1 + m_2)\dot{\mathbf{R}}^2 + T'$$

with

$$T' = \tfrac{1}{2}m_1\dot{\mathbf{r}}_1'^2 + \tfrac{1}{2}m_2\dot{\mathbf{r}}_2'^2.$$

Here $\mathbf{r}_1'$ and $\mathbf{r}_2'$ are the radii vectors of the two particles relative to the center of mass and are related to $\mathbf{r}$ by

$$\begin{aligned}\mathbf{r}_1' &= -\frac{m_2}{m_1 + m_2}\mathbf{r},\\ \mathbf{r}_2' &= \frac{m_1}{m_1 + m_2}\mathbf{r}\end{aligned} \tag{3.2}$$

Expressed in terms of $\mathbf{r}$ by means of Eq. (3.2), T' takes on the form

$$T' = \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\dot{\mathbf{r}}^2$$

and the total Lagrangian (3.1) is

$$L = \frac{m_1 + m_2}{2}\dot{\mathbf{R}}^2 + \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\dot{\mathbf{r}}^2 - U(\mathbf{r}, \dot{\mathbf{r}}, \ldots). \tag{3.3}$$

It is seen that the three coordinates $\mathbf{R}$ are cyclic, so that the center of mass is either at rest or moving uniformly. None of the equations of motion for $\mathbf{r}$ will contain terms involving $\mathbf{R}$ or $\dot{\mathbf{R}}$. Consequently, the process of integration is particularly simple here. We merely drop the first term from the Lagrangian in all subsequent discussion.

The rest of the Lagrangian is exactly what would be expected if we had a fixed center of force with a single particle at a distance $\mathbf{r}$ from it, having a mass

$$\mu = \frac{m_1 m_2}{m_1 + m_2}, \tag{3.4}$$

where μ is known as the *reduced mass*. Frequently, Eq. (3.4) is written in the form

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}. \tag{3.5}$$

Thus, the central force motion of two bodies about their center of mass can always be reduced to an equivalent one-body problem.

THE EQUATIONS OF MOTION AND FIRST INTEGRALS

Since the problem is spherically symmetric, the total angular momentum vector,

$$\mathbf{L} = \mathbf{r} \times \mathbf{p},$$

is conserved. It therefore follows that $\mathbf{r}$ is always perpendicular to the fixed direction of $\mathbf{L}$ in space. This can be true only if $\mathbf{r}$ always lies in a plane whose normal is parallel to $\mathbf{L}$. While this reasoning breaks down if $\mathbf{L}$ is zero, the motion in that case must be along a straight line going through the center of force, for $\mathbf{L} = 0$ requires $\mathbf{r}$ to be parallel to $\dot{\mathbf{r}}$, which can be satisfied only in straight-line motion.* Thus, central force motion is always motion in a plane.

Expressed now in plane polar coordinates, the Lagrangian is

$$\begin{aligned} L &= T - V \\ &= \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r). \end{aligned} \tag{3.6}$$

As was forseen, θ is a cyclic coordinate, whose corresponding canonical momentum is the angular momentum of the system:

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

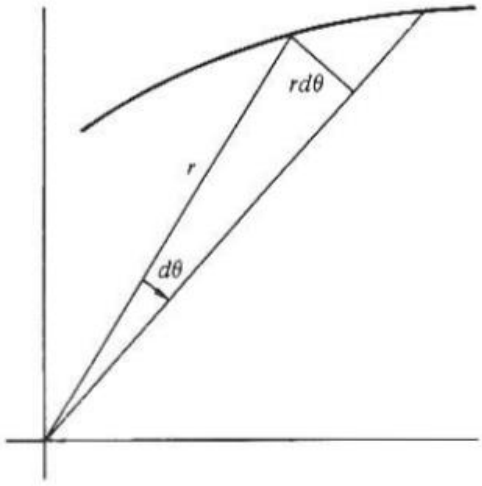


FIGURE 3.2 The area swept out by the radius vector in a time dt.

One of the two equations of motion is then simply

$$\dot{p}_\theta = \frac{d}{dt}\left(mr^2\dot{\theta}\right) = 0. \tag{3.7}$$

with the immediate integral

$$mr^2\dot{\theta} = l, \tag{3.8}$$

where l is the constant magnitude of the angular momentum. From (3.7) is also follows that

$$\frac{d}{dt}\left(\frac{1}{2}r^2\dot{\theta}\right) = 0. \tag{3.9}$$

The factor $\frac{1}{2}$ is inserted because $\frac{1}{2}r^2\dot{\theta}$ is just the *areal velocity*—the area swept out by the radius vector per unit time. This interpretation follows from Fig. 3.2, the differential area swept out in time dt being

$$dA = \tfrac{1}{2}r(r\,d\theta),$$

and hence

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt}.$$

The conservation of angular momentum is thus equivalent to saying the areal velocity is constant. Here we have the proof of the well-known Kepler's second law of planetary motion: The radius vector sweeps out equal areas in equal times. It should be emphasized however that the conservation of the areal velocity is a general property of central force motion and is not restricted to an inverse-square law of force.

The remaining Lagrange equation, for the coordinate r, is

$$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r} = 0. \tag{3.10}$$

Designating the value of the force along $\mathbf{r}$, $-\partial V/\partial r$, by $f(r)$ the equation can be rewritten as

$$m\ddot{r} - mr\dot{\theta}^2 = f(r). \tag{3.11}$$

By making use of the first integral, Eq. (3.8), $\dot{\theta}$ can be eliminated from the equation of motion, yielding a second-order differential equation involving r only:

$$m\ddot{r} - \frac{l^2}{mr^3} = f(r). \tag{3.12}$$

There is another first integral of motion available, namely the total energy, since the forces are conservative. On the basis of the general energy conservation theorem, we can immediately state that a constant of the motion is

$$E = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + V(r), \tag{3.13}$$

where E is the energy of the system. Alternatively, this first integral could be derived again directly from the equations of motion (3.7) and (3.12). The latter can be written as

$$m\ddot{r} = -\frac{d}{dr}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right). \tag{3.14}$$

If both sides of Eq. (3.14) are multiplied by $\dot{r}$ the left side becomes

$$m\ddot{r}\dot{r} = \frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2\right).$$

The right side similarly can be written as a total time derivative, for if $g(r)$ is any function of r, then the total time derivative of g has the form

$$\frac{d}{dt}g(r) = \frac{dg}{dr}\frac{dr}{dt}.$$

Hence, Eq. (3.14) is equivalent to

$$\frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2\right) = -\frac{d}{dt}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right)$$

or

$$\frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2 + \frac{1}{2}\frac{l^2}{mr^2} + V\right) = 0,$$

and therefore

$$\frac{1}{2}m\dot{r}^2 + \frac{1}{2}\frac{l^2}{mr^2} + V = \text{constant}. \tag{3.15}$$

Equation (3.15) is the statement of the conservation of total energy, for by using (3.8) for l, the middle term can be written

$$\frac{1}{2}\frac{l^2}{mr^2} = \frac{1}{2mr^2}m^2r^4\dot{\theta}^2 = \frac{mr^2\dot{\theta}^2}{2},$$

and (3.15) reduces to (3.13).

$$\dot{r} = \sqrt{\frac{2}{m}\left(E - V - \frac{l^2}{2mr^2}\right)}, \tag{3.16}$$

or

$$dt = \frac{dr}{\sqrt{\frac{2}{m}\left(E - V - \frac{l^2}{2mr^2}\right)}}. \tag{3.17}$$

At time $t = 0$, let r have the initial value r_0. Then the integral of both sides of the equation from the initial state to the state at time t takes the form

$$t = \int_{r_0}^{r} \frac{dr}{\sqrt{\frac{2}{m}\left(E - V - \frac{l^2}{2mr^2}\right)}}. \tag{3.18}$$

As it stands, Eq. (3.18) gives t as a function of r and the constants of integration E, l, and r_0. However, it may be inverted, at least formally, to give r as a function of t and the constants. Once the solution for r is found, the solution θ follows immediately from Eq. (3.8), which can be written as

$$d\theta = \frac{l\,dt}{mr^2}. \tag{3.19}$$

If the initial value of θ is θ_0, then the integral of (3.19) is simply

$$\theta = l\int_0^t \frac{dt}{mr^2(t)} + \theta_0. \tag{3.20}$$

THE EQUIVALENT ONE-DIMENSIONAL PROBLEM, AND CLASSIFICATION OF ORBITS

For example, with a system of known energy and angular momentum, the magnitude and direction of the velocity of the particle can be immediately determined in terms of the distance r. The magnitude v follows at once from the conservation of energy in the form

$$E = \tfrac{1}{2}mv^2 + V(r)$$

or

$$v = \sqrt{\frac{2}{m}(E - V(r))}. \tag{3.21}$$

The radial velocity—the component of $\dot{\mathbf{r}}$ along the radius vector—has been given in Eq. (3.16). Combined with the magnitude v, this is sufficient information to furnish the direction of the velocity.* These results, and much more, can also be obtained from consideration of an equivalent one-dimensional problem.

The equation of motion in r, with $\dot{\theta}$ expressed in terms of l, Eq. (3.12), involves only r and its derivatives. It is the same equation as would be obtained for a

fictitious one-dimensional problem in which a particle of mass m is subject to a force

$$f' = f + \frac{l^2}{mr^3}. \tag{3.22}$$

The significance of the additional term is clear if it is written as $mr\dot{\theta}^2 = mv_\theta^2/r$, which is the familiar centrifugal force. An equivalent statement can be obtained from the conservation theorem for energy. By Eq. (3.15) the motion of the particle in r is that of a one-dimensional problem with a fictitious potential energy:

$$V' = V + \frac{1}{2}\frac{l^2}{mr^2}. \tag{3.22'}$$

As a check, note that

$$f' = -\frac{\partial V'}{\partial r} = f(r) + \frac{l^2}{mr^3},$$

which agrees with Eq. (3.22). The energy conservation theorem (3.15) can thus also be written as

$$E = V' + \tfrac{1}{2}m\dot{r}^2. \tag{3.15'}$$

As an illustration of this method of examining the motion, consider a plot of V' against r for the specific case of an attractive inverse-square law of force:

$$f = -\frac{k}{r^2}.$$

(For positive k, the minus sign ensures that the force is *toward* the center of force.) The potential energy for this force is

$$V = -\frac{k}{r},$$

and the corresponding fictitious potential is

$$V' = -\frac{k}{r} + \frac{l^2}{2mr^2}.$$

Such a plot is shown in Fig. 3.3; the two dashed lines represent the separate components

$$-\frac{k}{r} \quad \text{and} \quad \frac{l^2}{2mr^2},$$

and the solid line is the sum V'.

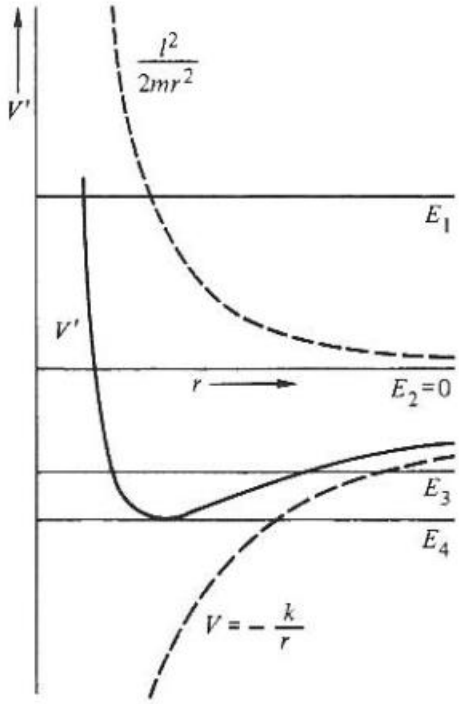


FIGURE 3.3 The equivalent one-dimensional potential for attractive inverse-square law of force.

THE VIRIAL THEOREM

Another property of central force motion can be derived as a special case of a general theorem valid for a large variety of systems—the *virial theorem*. It differs in character from the theorems previously discussed in being *statistical* in nature; i.e., it is concerned with the time averages of various mechanical quantities.

Consider a general system of mass points with position vectors $\mathbf{r}_i$ and applied forces $\mathbf{F}_i$ (including any forces of constraint). The fundamental equations of motion are then

$$\dot{\mathbf{p}}_i = \mathbf{F}_i. \tag{1.3}$$

We are interested in the quantity

$$G = \sum_i \mathbf{p}_i \cdot \mathbf{r}_i,$$

where the summation is over all particles in the system. The total time derivative of this quantity is

$$\frac{dG}{dt} = \sum_i \dot{\mathbf{r}}_i \cdot \mathbf{p}_i + \sum_i \dot{\mathbf{p}}_i \cdot \mathbf{r}_i. \tag{3.23}$$

The first term can be transformed to

$$\sum_i \dot{\mathbf{r}}_i \cdot \mathbf{p}_i = \sum_i m_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i = \sum_i m_i v_i^2 = 2T,$$

while the second term by (1.3) is

$$\sum_i \dot{\mathbf{p}}_i \cdot \mathbf{r}_i = \sum_i \mathbf{F}_i \cdot \mathbf{r}_i.$$

Equation (3.23) therefore reduces to

$$\frac{d}{dt} \sum_i \mathbf{p}_i \cdot \mathbf{r}_i = 2T + \sum_i \mathbf{F}_i \cdot \mathbf{r}_i. \tag{3.24}$$

The time average of Eq. (3.24) over a time interval τ is obtained by integrating both sides with respect to t from 0 to τ, and dividing by τ:

$$\frac{1}{\tau} \int_0^\tau \frac{dG}{dt} dt \equiv \overline{\frac{dG}{dt}} = \overline{2T} + \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i}$$

or

$$\overline{2T} + \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i} = \frac{1}{\tau} [G(\tau) - G(0)]. \tag{3.25}$$

If the motion is periodic, i.e., all coordinates repeat after a certain time, and if τ is chosen to be the period, then the right-hand side of (3.25) vanishes. A similar conclusion can be reached even if the motion is not periodic, provided that the coordinates and velocities for all particles remain finite so that there is an upper bound to G. By choosing τ sufficiently long, the right-hand side of Eq. (3.25) can be made as small as desired. In both cases, it then follows that

$$\overline{T} = -\frac{1}{2} \overline{\sum_i \mathbf{F}_i \cdot \mathbf{r}_i}. \tag{3.26}$$

Equation (3.26) is known as the *virial theorem*, and the right-hand side is called the *virial of Clausius*. In this form the theorem is imporant in the kinetic theory

of gases since it can be used to derive ideal gas law for perfect gases by means of the following brief argument.

$$d\mathbf{F}_i = -P\mathbf{n}\,dA,$$

or

$$\frac{1}{2}\sum_i \mathbf{F}_i \cdot \mathbf{r}_i = -\frac{P}{2}\int \mathbf{n}\cdot\mathbf{r}\,dA.$$

But, by Gauss's theorem,

$$\int \mathbf{n}\cdot\mathbf{r}\,dA = \int \nabla\cdot\mathbf{r}\,dV = 3V.$$

The virial theorem, Eq. (3.26), for the system representing a perfect gas can therefore be written

$$\tfrac{3}{2}Nk_BT = \tfrac{3}{2}PV,$$

which, cancelling the common factor of $\frac{3}{2}$ on both sides, is the familiar ideal gas law. Where the interparticle forces contribute to the virial, the perfect gas law of course no longer holds. The virial theorem is then the principal tool, in classical kinetic theory, for calculating the equation of state corresponding to such imperfect gases.

If the forces are derivable from a potential, then the theorem becomes

$$\overline{T} = \frac{1}{2}\overline{\sum_i \nabla V \cdot \mathbf{r}_i}, \tag{3.27}$$

and for a single particle moving under a central force it reduces to

$$\overline{T} = \frac{1}{2}\overline{\frac{\partial V}{\partial r} r}. \tag{3.28}$$

If V is a power-law function of r,

$$V = ar^{n+1},$$

where the exponent is chosen so that the force law goes as r^n, then

$$\frac{\partial V}{\partial r} r = (n+1)V,$$

and Eq. (3.28) becomes

$$\overline{T} = \frac{n+1}{2}\overline{V}. \tag{3.29}$$

By an application of Euler's theorem for homogeneous functions (cf. p. 62), it is clear that Eq. (3.29) also holds whenever V is a homogeneous function in r of degree $n + 1$. For the further special case of inverse-square law forces, n is -2, and the virial theorem takes on a well-known form:

$$\overline{T} = -\tfrac{1}{2}\overline{V}. \tag{3.30}$$

THE DIFFERENTIAL EQUATION FOR THE ORBIT, AND INTEGRABLE POWER-LAW POTENTIALS

In treating specific details of actual central force problems, a change in the orientation of our discussion is desirable. Hitherto solving a problem has meant finding r and θ as functions of time with E, l, etc., as constants of integration. But most often what we really seek is the equation of the orbit, i.e., the dependence of r upon θ, eliminating the parameter t. For central force problems, the elimination is particularly simple, since t occurs in the equations of motion only as a variable of differentiation. Indeed, one equation of motion, (3.8), simply provides a definite

relation between a differential change dt and the corresponding change $d\theta$:

$$l\,dt = mr^2\,d\theta. \tag{3.31}$$

The corresponding relation between derivatives with respect to t and θ is

$$\frac{d}{dt} = \frac{l}{mr^2}\frac{d}{d\theta}. \tag{3.32}$$

These relations may be used to convert the equation of motion (3.12) or (3.16) to a differential equation for the orbit. A substitution into Eq. (3.12) gives a second-order differential equation, while a substitution into Eq. (3.17) gives a simpler first-order differential equation.

The substitution into Eq. (3.12) yields

$$\frac{1}{r^2}\frac{d}{d\theta}\left(\frac{1}{mr^2}\frac{dr}{d\theta}\right) - \frac{l^2}{mr^3} = f(r), \tag{3.33}$$

which upon substituting $u = 1/r$ and expressing the results in terms of the potential gives

$$\frac{d^2u}{d\theta^2} + u = -\frac{m}{l^2}\frac{d}{du}V\left(\frac{1}{u}\right). \tag{3.34}$$

$$u = u(0), \qquad \left(\frac{du}{d\theta}\right)_0 = 0, \qquad \text{for } \theta = 0,$$

will likewise be unaffected. Hence, the orbit equation must be the same whether expressed in terms of θ or $-\theta$, which is the desired conclusion. *The orbit is therefore invariant under reflection about the apsidal vectors.* In effect, this means that

THE KEPLER PROBLEM: INVERSE-SQUARE LAW OF FORCE

The inverse-square law is the most important of all the central force laws, and it deserves detailed treatment. For this case, the force and potential can be written

as

$$f = -\frac{k}{r^2} \qquad V = -\frac{k}{r}. \tag{3.49}$$

There are several ways to integrate the equation for the orbit, the simplest being to substitute (3.49) in the differential equation for the orbit (3.33). Another approach is to start with Eq. (3.39) with n set equal to -2 for the gravitational force

$$\theta = \theta' - \int \frac{du}{\sqrt{\frac{2mE}{l^2} + \frac{2mku}{l^2} - u^2}}, \tag{3.50}$$

where the integral is now taken as indefinite. The quantity θ' appearing in (3.50) is a constant of integration determined by the initial conditions and will not necessarily be the same as the initial angle θ_0 at time $t = 0$. The indefinite integral is of the standard form,

$$\int \frac{dx}{\sqrt{\alpha + \beta x + \gamma x^2}} = \frac{1}{\sqrt{-\gamma}} \arccos -\frac{\beta + 2\gamma x}{\sqrt{q}}, \tag{3.51}$$

where

$$q = \beta^2 - 4\alpha\gamma.$$

To apply this to (3.50), we must set

$$\alpha = \frac{2mE}{l^2}, \qquad \beta = \frac{2mk}{l^2} \qquad \gamma = -1, \tag{3.52}$$

and the discriminant q is therefore

$$q = \left(\frac{2mk}{l^2}\right)^2 \left(1 + \frac{2El^2}{mk^2}\right). \tag{3.53}$$

With these substitutes, Eq. (3.50) becomes

$$\theta = \theta' - \arccos \frac{\frac{l^2 u}{mk} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}}. \tag{3.54}$$

Finally, by solving for $u, \equiv 1/r$, the equation of the orbit is found to be

$$\frac{1}{r} = \frac{mk}{l^2}\left(1 + \sqrt{1 + \frac{2El^2}{mk^2}} \cos(\theta - \theta')\right). \tag{3.55}$$

$$mr^2\, d\theta = l\, dt,$$

by means of (3.55), we must additionally specify the initial angle θ_0.

Now, the general equation of a conic with one focus at the origin is

$$\frac{1}{r} = C[1 + e\cos(\theta - \theta')], \tag{3.56}$$

where e is the eccentricity of the conic section. By comparison with Eq. (3.55), it follows that the orbit is always a conic section, with the eccentricity

$$e = \sqrt{1 + \frac{2El^2}{mk^2}}. \tag{3.57}$$

The nature of the orbit depends upon the magnitude of e according to the following scheme:

$e > 1$,	$E > 0$:	hyperbola,
$e = 1$,	$E = 0$:	parabola,
$e < 1$,	$E < 0$:	ellipse,
$e = 0$,	$E = -\frac{mk^2}{2l^2}$:	circle.

This classification agrees with the qualitative discussion of the orbits on the energy diagram of the equivalent one-dimensional potential V'. The condition for circular motion appears here in a somewhat different form, but it can easily be derived as a consequence of the previous conditions for circularity. For a circular orbit, T and V are constant in time, and from the virial theorem

$$E \equiv T + V = -\frac{V}{2} + V = \frac{V}{2}.$$

Hence

$$E = -\frac{k}{2r_0}. \tag{3.58}$$

But from Eq. (3.41), the statement of equilibrium between the central force and the "effective force," we can write

$$\frac{k}{r_0^2} = \frac{l^2}{mr_0^3},$$

or

$$r_0 = \frac{l^2}{mk}. \tag{3.59}$$

With this formula for the orbital radius, Eq. (3.58) becomes

$$E = -\frac{mk^2}{2l^2},$$

the above condition for circular motion.

$$E - \frac{l^2}{2mr^2} + \frac{k}{r} = 0,$$

or

$$r^2 + \frac{k}{E}r - \frac{l^2}{2mE} = 0. \tag{3.60}$$

Now, the coefficient of the linear term in a quadratic equation is the negative of the sum of the roots. Hence, the semimajor axis is given by

$$a = \frac{r_1 + r_2}{2} = -\frac{k}{2E}. \tag{3.61}$$

Note that in the circular limit, Eq. (3.61) agrees with Eq. (3.58). In terms of the semimajor axis, the eccentricity of the ellipse can be written

$$e = \sqrt{1 - \frac{l^2}{mka}}, \tag{3.62}$$

(a relation we will have use for in a later chapter). Further, from Eq. (3.62) we have the expression

$$\frac{l^2}{mk} = a(1 - e^2), \tag{3.63}$$

in terms of which the elliptical orbit equation (3.55) can be written

$$r = \frac{a(1 - e^2)}{1 + e\cos(\theta - \theta')}. \tag{3.64}$$

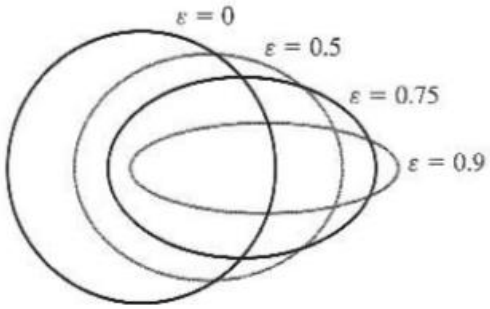


FIGURE 3.14 Ellipses with the same major axes and eccentricities from 0.0 to 0.9.

THE MOTION IN TIME IN THE KEPLER PROBLEM

The orbital equation for motion in a central inverse-square force law can thus be solved in a fairly straightforward manner with results that can be stated in simple closed expressions. Describing the motion of the particle in time as it traverses the orbit is however a much more involved matter. In principle, the relation between the radial distance of the particle r and the time (relative to some starting point) is given by Eq. (3.18), which here takes on the form

$$t = \sqrt{\frac{m}{2}} \int_{r_0}^{r} \frac{dr}{\sqrt{\frac{k}{r} - \frac{l^2}{2mr^2} + E}}. \tag{3.65}$$

Similarly, the polar angle θ and the time are connected through the conservation of angular momentum,

$$dt = \frac{mr^2}{l}\, d\theta,$$

which combined with the orbit equation (3.51) leads to

$$t = \frac{l^3}{mk^2} \int_{\theta_0}^{\theta} \frac{d\theta}{[1 + e\cos(\theta - \theta')]^2}. \tag{3.66}$$

passage. Using the trigonometric identity

$$1 + \cos\theta = 2\cos^2\frac{\theta}{2},$$

Eq. (3.66) then reduces for parabolic motion to the form

$$t = \frac{l^3}{4mk^2}\int_0^\theta \sec^4\frac{\theta}{2}\,d\theta.$$

The integration is easily performed by a change of variable to $x = \tan(\theta/2)$, leading to the integral

$$t = \frac{l^3}{2mk^2}\int_0^{\tan(\theta/2)} (1 + x^2)\,dx,$$

or

$$t = \frac{l^3}{2mk^2}\left(\tan\frac{\theta}{2} + \frac{1}{3}\tan^3\frac{\theta}{2}\right). \tag{3.67}$$

In this equation, $-\pi < \theta < \pi$, where for $t \to -\infty$ the particle starts approaching from infinitely far away located at $\theta = -\pi$. The time $t = 0$ corresponds to $\theta = 0$, where the particle is at perihelion. Finally $t \to +\infty$ corresponds to $\theta \to \pi$ as the particle moves infinitely far away. This is a straightforward relation for t as a function of θ; inversion to obtain θ at a given time requires solving a cubic equation for $\tan(\theta/2)$, then finding the corresponding arctan. The radial distance at a given time is given through the orbital equation.

For elliptical motion, Eq. (3.65) is most conveniently integrated through an auxiliary variable ψ, denoted as the *eccentric anomaly*,* and defined by the relation

$$r = a(1 - e\cos\psi). \tag{3.68}$$

By comparison with the orbit equation, (3.64), it is clear that ψ also covers the interval 0 to 2π as θ goes through a complete revolution, and that the perihelion occurs at $\psi = 0$ (where $\theta = 0$ by convention) and the aphelion at $\psi = \pi = \theta$.

Expressing E and ℓ in terms of a, e, and k, Eq. (3.65) can be rewritten for elliptic motion as

$$t = -\sqrt{\frac{m}{2k}} \int_{r_0}^{r} \frac{r\,dr}{\sqrt{r - \frac{r^2}{2a} - \frac{a(1-e^2)}{2}}}, \tag{3.69}$$

where, by the convention on the starting time, r_0 is the perihelion distance. Substitution of r in terms of ψ from Eq. (3.68) reduces this integral, after some algebra, to the simple form

$$t = \sqrt{\frac{ma^3}{k}} \int_0^{\psi} (1 - e\cos\psi)\,d\psi. \tag{3.70}$$

First, we may note that Eq. (3.70) provides an expression for the period, τ, of elliptical motion, if the integral is carried over the full range in ψ of 2π:

$$\tau = 2\pi a^{3/2}\sqrt{\frac{m}{k}}. \tag{3.71}$$

This important result can also be obtained directly from the properties of an ellipse. From the conservation of angular momentum, the areal velocity is constant and is given by

$$\frac{dA}{dt} = \frac{1}{2}r^2\theta = \frac{l}{2m}. \tag{3.72}$$

The area of the orbit, A, is to be found by integrating (3.72) over a complete period τ:

$$\int_0^{\tau} \frac{dA}{dt}\,dt = A = \frac{l\tau}{2m}.$$

Now, the area of an ellipse is

$$A = \pi ab,$$

where, by the definition of eccentricity, the semiminor axis b is related to a according to the formula

$$b = a\sqrt{1 - e^2}.$$

By (3.62), the semiminor axis can also be written as

$$b = a^{1/2}\sqrt{\frac{l^2}{mk}},$$

and the period is therefore

$$\tau = \frac{2m}{l}\pi a^{3/2}\sqrt{\frac{l^2}{mk}} = 2\pi a^{3/2}\sqrt{\frac{m}{k}},$$

as was found previously. Equation (3.71) states that, other things being equal, the square of the period is proportional to the cube of the major axis, and this conclusion is often referred to as the third of Kepler's laws.* Actually, Kepler was concerned with the specific problem of planetary motion in the gravitational field of the Sun. A more precise statement of this third law would therefore be: *The square of the periods of the various planets are proportional to the cube of their major axes.* In this form, the law is only approximately true. Recall that the motion of a planet about the Sun is a two-body problem and m in (3.71) must be replaced by the reduced mass: (cf. Eq. (3.4))

$$\mu = \frac{m_1 m_2}{m_1 + m_2},$$

where m_1 may be taken as referring to the planet and m_2 to the Sun. Further, the gravitational law of attraction is

$$f = -G\frac{m_1 m_2}{r^2},$$

so that the constant k is

$$k = Gm_1 m_2. \tag{3.73}$$

Under these conditions, (3.71) becomes

$$\tau = \frac{2\pi a^{3/2}}{\sqrt{G(m_1 + m_2)}} \approx \frac{2\pi a^{3/2}}{\sqrt{Gm_2}}, \tag{3.74}$$

$$\omega = \frac{2\pi}{\tau} = \sqrt{\frac{k}{ma^3}}. \tag{3.75}$$

The integration in Eq. (3.70) is of course easily performed, resulting in the relation

$$\omega t = \psi - e \sin \psi, \tag{3.76}$$

known as *Kepler's equation*. The quantity ωt goes through the range 0 to 2π, along with ψ and θ, in the course of a complete orbital revolution and is therefore also denoted as an anomaly, specifically the *mean anomaly*.

To find the position in orbit at a given time t, Kepler's equation, (3.76), would first be inverted to obtain the corresponding eccentric anomaly ψ. Equation (3.68) then yields the radial distance, while the polar angle θ can be expressed in terms of ψ by comparing the defining equation (3.68) with the orbit equation (3.64):

$$1 + e \cos \theta = \frac{1 - e^2}{1 - e \cos \psi}.$$

With a little algebraic manipulation, this can be simplified, to

$$\cos \theta = \frac{\cos \psi - e}{1 - e \cos \psi}. \tag{3.77}$$

By successively adding and subtracting both sides of Eq. (3.77) from unity and taking the ratio of the resulting two equations, we are led to the alternative form

$$\tan \frac{\theta}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{\psi}{2}. \tag{3.78}$$

SCATTERING IN A CENTRAL FORCE FIELD

In its one-body formulation, the scattering problem is concerned with the scattering of particles by a *center of force*. We consider a uniform beam of particles—whether electrons, or α-particles, or planets is irrelevant—all of the same mass and energy incident upon a center of force. It will be assumed that the force falls off to zero for very large distances. The incident beam is characterized by specifying its *intensity I* (also called flux density), which gives the number of particles crossing unit area normal to the beam in unit time. As a particle approaches the center of force, it will be either attracted or repelled, and its orbit will deviate from the incident straight-line trajectory. After passing the center of force, the force acting on the particle will eventually diminish so that the orbit once again approaches a straight line. In general, the final direction of motion is not the same as the incident direction, and the particle is said to be scattered. The *cross section for scattering in a given direction*, $\sigma(\mathbf{\Omega})$, is defined by

$$\sigma(\mathbf{\Omega})\, d\Omega = \frac{\text{number of particles scattered into solid angle } d\Omega \text{ per unit time}}{\text{incident intensity}}, \tag{3.88}$$

where $d\Omega$ is an element of solid angle in the direction $\mathbf{\Omega}$. Often $\sigma(\mathbf{\Omega})$ is also designated as the *differential scattering cross section*. With central forces there must be complete symmetry around the axis of the incident beam; hence the element of solid angle can be written

$$d\Omega = 2\pi \sin\Theta\, d\Theta, \tag{3.89}$$

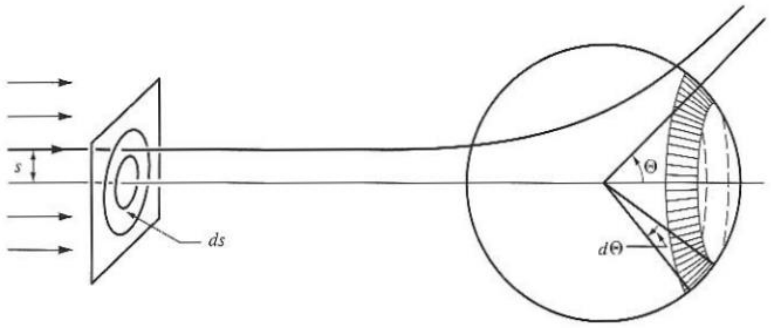


FIGURE 3.19 Scattering of an incident beam of particles by a center of force.

where Θ is the angle between the scattered and incident directions, known as the *scattering angle* (cf. Fig. 3.19, where repulsive scattering is illustrated). Note that the name "cross section" is deserved in that $\sigma(\Omega)$ has the dimensions of an area.

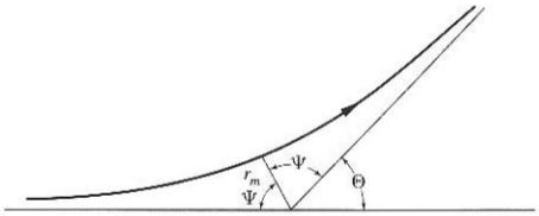


FIGURE 3.20 Relation of orbit parameters and scattering angle in an example of repulsive scattering.

then the dependence of the differential cross section on Θ is given by

$$\sigma(\Theta) = \frac{s}{\sin\Theta}\left|\frac{ds}{d\Theta}\right|. \tag{3.93}$$

A formal expression for the scattering angle Θ as a function of s can be directly obtained from the orbit equation, Eq. (3.36). Again, for simplicity, we will consider the case of purely repulsive scattering (cf. Fig. 3.20). As the orbit must be symmetric about the direction of the periapsis, the scattering angle is given by

$$\Theta = \pi - 2\Psi, \tag{3.94}$$

where Ψ is the angle between the direction of the incoming asymptote and the periapsis (closest approach) direction. In turn, Ψ can be obtained from Eq. (3.36) by setting $r_0 = \infty$ when $\theta_0 = \pi$ (the incoming direction), whence $\theta = \pi - \Psi$ when $r = r_m$, the distance of closest approach. A trivial rearrangement then leads to

$$\Psi = \int_{r_m}^{\infty} \frac{dr}{r^2\sqrt{\frac{2mE}{l^2} - \frac{2mV}{l^2} - \frac{1}{r^2}}}. \tag{3.95}$$

Expressing l in terms of the impact parameter s (Eq. (3.90)), the resultant expression for $\Theta(s)$ is

$$\Theta(s) = \pi - 2\int_{r_m}^{\infty} \frac{s\,dr}{r\sqrt{r^2\left(1 - \frac{V(r)}{E}\right) - s^2}}, \tag{3.96}$$

or, changing r to $1/u$

$$\Theta(s) = \pi - 2\int_0^{u_m} \frac{s\,du}{\sqrt{1 - \frac{V(u)}{E} - s^2u^2}}. \tag{3.97}$$

Equations (3.96) and (3.97) are rarely used except for direct numerical computation of the scattering angle. However, when an analytic expression is available for the orbits, the relation between Θ and s can often be obtained almost by inspection. An historically important illustration of such a procedure is the repulsive scattering of charged particles by a Coulomb field. The scattering force field is that produced by a fixed charge $-Ze$ acting on the incident particles having a charge $-Z'e$ so that the force can be written as

$$f = \frac{ZZ'e^2}{r^2},$$

i.e., a repulsive inverse-square law. The results of Section 3.7 can be taken over here with no more change that writing the force constant as

$$k = -ZZ'e^2. \tag{3.98}$$

The energy E is greater than zero, and the orbit is a hyperbola with the eccentricity given by*

$$\epsilon = \sqrt{1 + \frac{2El^2}{m(ZZ'e^2)^2}} = \sqrt{1 + \left(\frac{2Es}{ZZ'e}\right)^2}, \tag{3.99}$$

where use has been made of Eq. (3.90). If θ' in Eq. (3.55) is chosen to be π, periapsis corresponds to $\theta = 0$ and the orbit equation becomes

$$\frac{1}{r} = \frac{mZZ'e}{l^2}(\epsilon \cos\theta - 1). \tag{3.100}$$

This hyperbolic orbit equation has the same form as the elliptic orbit equation (3.56) except for a change in sign. The direction of the incoming asymptote, Ψ, is then determined by the condition $r \to \infty$:

$$\cos\Psi = \frac{1}{\epsilon}$$

or, by Eq. (3.94),

$$\sin\frac{\Theta}{2} = \frac{1}{\epsilon}.$$

Hence,

$$\cot^2\frac{\Theta}{2} = \epsilon^2 - 1,$$

and using Eq. (3.99)

$$\cot\frac{\Theta}{2} = \frac{2Es}{ZZ'e}.$$

The desired functional relationship between the impact parameter and the scattering angle is therefore

$$s = \frac{ZZ'e^2}{2E}\cot\frac{\Theta}{2}, \tag{3.101}$$

so that on carrying through the manipulation required by Eq. (3.93), we find that $\sigma(\Theta)$ is given by

$$\sigma(\Theta) = \frac{1}{4}\left(\frac{ZZ'e^2}{2E}\right)^2 \csc^4\frac{\Theta}{2}. \tag{3.102}$$

Equation (3.102) gives the famous Rutherford scattering cross section, originally derived by Rutherford for the scattering of α particles by atomic nuclei. Quantum mechanics in the nonrelativistic limit yields a cross section identical with this classical result.

In atomic physics, the concept of a *total scattering cross section* σ_T, defined as

$$\sigma_T = \int_{4\pi} \sigma(\mathbf{\Omega})\, d\Omega = 2\pi \int_0^{\pi} \sigma(\Theta)\sin\Theta\, d\Theta,$$

is of considerable importance. However, if we attempt to calculate the total cross section for Coulomb scattering by substituting Eq. (3.102) in this definition, we obtain an infinite result! The physical reason behind this behavior is not difficult to discern. From its definition the total cross section is the number of particles scattered in all directions per unit time for unit incident intensity. Now, the Coulomb field is an example of a "long-range" force; its effects extend to infinity.

TRANSFORMATION OF THE SCATTERING PROBLEM TO LABORATORY COORDINATES

In the previous section we were concerned with the one-body problem of the scattering of a particle by a fixed center of force. In practice, the scattering always involved two bodies; e.g., in Rutherford scattering we have the α particle and the atomic nucleus. The second particle, m_2, is not fixed but recoils from its initial position as a result of the scattering. Since it has been shown that any two-body

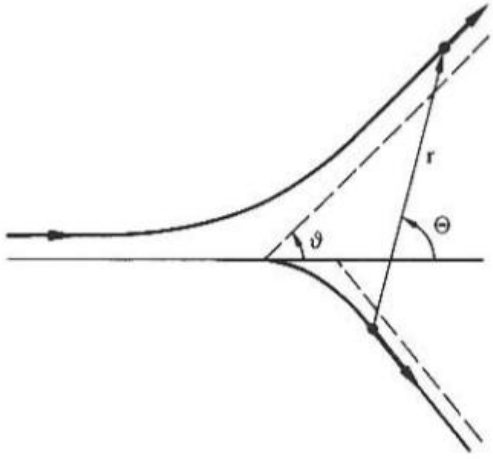


FIGURE 3.24 Scattering of two particles as viewed in the laboratory system.

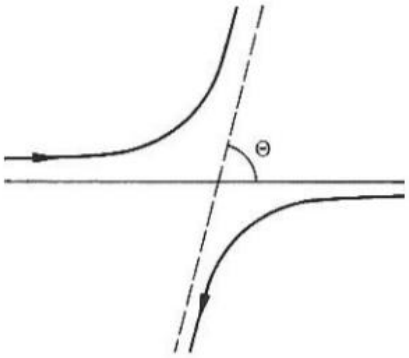


FIGURE 3.25 Scattering of two particles as viewed in the center of mass system.

It is convenient here to use the terminology of Section 3.1, with slight modifications:

$\mathbf{r}_1$ and $\mathbf{v}_1$ are the position and velocity, after scattering, of the incident particle, m_1, in the laboratory system,

$\mathbf{r}_1'$ and $\mathbf{v}_1'$ are the position and velocity, after scattering, of particle m_1 in the center of mass system, and

$\mathbf{R}$ and $\mathbf{V}$ are the position and (constant) velocity in the center of mass in the laboratory system.

At any instant, by definition

$$\mathbf{r}_1 = \mathbf{R} + \mathbf{r}_1',$$

and consequently

$$\mathbf{v}_1 = \mathbf{V} + \mathbf{v}_1'. \tag{3.104}$$

Figure 3.26 graphically portrays this vector relation evaluated *after* the scattering has taken place; at which time $\mathbf{v}_1$ and $\mathbf{v}_1'$ make the angles ϑ and Θ, respectively,

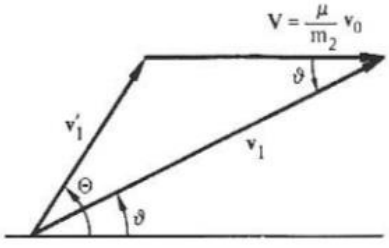


FIGURE 3.26 The relations between the velocities in the center of mass and laboratory coordinates.

with the vector $\mathbf{V}$ lying along the initial direction. Since the target is initially stationary in the laboratory system, the incident velocity of particle 1 in that system, $\mathbf{v}_0$, is the same as the initial relative velocity of the particles. By conservation of total linear momentum, the constant velocity of the center of mass is therefore given by

$$(m_1 + m_2)\mathbf{V} = m_1\mathbf{v}_0,$$

or

$$\mathbf{V} = \frac{\mu}{m_2}\mathbf{v}_0, \tag{3.105}$$

where $\mu = m_1m_2/(m_1 + m_2)$. From Fig. 3.26, it is readily seen that

$$v_1 \sin\vartheta = v_1' \sin\Theta$$

and

$$v_1 \cos\vartheta = v_1' \cos\Theta + V. \tag{3.106}$$

The ratio of these two equations gives a relation between ϑ and Θ:

$$\tan\vartheta = \frac{\sin\Theta}{\cos\Theta + \rho}, \tag{3.107}$$

where ρ is defined as

$$\rho \equiv \frac{\mu}{m_2}\frac{v_0}{v_1'}. \tag{3.108}$$

An alternative relation can be obtained by expressing v_1 in terms of the other speeds through the cosine law as applied to the triangle of Fig. 3.26:

$$v_1^2 = v_1'^2 + V^2 + 2v_1'V\cos\Theta. \tag{3.109}$$

When this is used to eliminate v_1 from Eq. (3.106) and V is expressed in terms of v_0 by Eq. (3.105), we find

$$\cos\vartheta = \frac{\cos\Theta + \rho}{\sqrt{1 + 2\rho\cos\Theta + \rho^2}}. \tag{3.110}$$

Both these relations still involve a ratio of speeds through ρ. By the definition of center of mass, the speed of particle 1 in the center-of-mass system, v_1', is connected with the relative speed v after the collision, by the equation (cf. Eq. (3.2)), where $v = |\dot{r}|$:

$$v_1' = \frac{\mu}{m_1}v.$$

Hence, ρ can also be written as

$$\rho = \frac{m_1}{m_2}\frac{v_0}{v}, \tag{3.108'}$$

where v, it should be emphasized, is the relative speed *after* the collision. When the collision is *elastic*, the total kinetic energy of the two particles remains unaltered and v must equal v_0 so that ρ is simply

$$\rho = \frac{m_1}{m_2}, \qquad \text{(elastic collision)} \tag{3.111}$$

independent of energies or speeds. If the collision is *inelastic*, the total kinetic energy of the two particles is altered (e.g., some of the kinetic energy goes into the form of internal excitation energy of the target). Since the total energy is conserved and momentum is conserved, the energy change resulting from the collision can be expressed as

$$\frac{\mu v^2}{2} = \frac{\mu v_0^2}{2} + Q. \tag{3.112}$$

The so-called Q value of the inelastic collision is clearly negative in magnitude, but the sign convention is chosen to conform to that used in general for atomic and nuclear reactions. From Eq. (3.112) the ratio of relative speeds before and after collision can be written

$$\frac{v}{v_0} = \sqrt{1 + \frac{m_1 + m_2}{m_2}\frac{Q}{E}}, \tag{3.113}$$

where $E = \frac{1}{2}mv_0^2$ is the energy of the incoming particle (in the laboratory system). Thus, for inelastic scattering ρ becomes

$$\rho = \frac{m_1}{m_2\sqrt{1 + \frac{m_1+m_2}{m_2}\frac{Q}{E}}}. \qquad \text{(inelastic scattering)} \tag{3.114}$$

Not only are the scattering angles ϑ and Θ in general different in magnitude, but the values of the differential scattering cross section depend upon which of the two angles is used as the argument of σ. The connection between the two functional forms is obtained from the observation that in a particular experiment the number of particles scattered into a given element of solid angle must be the same whether we measure the event in terms of ϑ or Θ. As an equation, this statement can be written

$$2\pi I\sigma(\Theta)\sin\Theta|\,d\Theta| = 2\pi I\sigma'(\vartheta)\sin\vartheta|\,d\vartheta|,$$

or

$$\sigma'(\vartheta) = \sigma(\Theta)\frac{\sin\Theta}{\sin\vartheta}\left|\frac{d\Theta}{d\vartheta}\right| = \sigma(\Theta)\left|\frac{d(\cos\Theta)}{d(\cos\vartheta)}\right|, \tag{3.115}$$

where $\sigma'(\vartheta)$ is the differential scattering cross section expressed in terms of the scattering angle in the laboratory system. The derivative can easily be evaluated from Eq. (3.110), leading to the result

$$\sigma'(\vartheta) = \sigma(\Theta)\frac{(1+2\rho\cos\Theta+\rho^2)^{3/2}}{1+\rho\cos\Theta}. \tag{3.116}$$

Note that $\sigma(\Theta)$ is *not* the cross section an observer would measure in the center-of-mass system. Both $\sigma(\Theta)$ and $\sigma'(\vartheta)$ are cross sections measured in the laboratory system; they are merely expressed in terms of different coordinates. An

The two scattering angles have a particularly simple relation for elastic scattering when the two masses of particles are equal. It then follows that $\rho = 1$, and from Eq. (3.110) we have

$$\cos\vartheta = \sqrt{\frac{1+\cos\Theta}{2}} = \cos\frac{\Theta}{2},$$

or

$$\vartheta = \frac{\Theta}{2}, \qquad (\rho = 1).$$

Thus, with equal masses, scattering angles greater than 90° cannot occur in the laboratory system; all the scattering is in the forward hemisphere. Correspondingly, the scattering cross section is given in terms of Θ from Eq. (3.116) as

$$\sigma'(\vartheta) = 4\cos\vartheta\cdot\sigma(\Theta), \qquad \vartheta \le \frac{\pi}{2}, \quad (\rho = 1).$$

particle. The degree of slowing down can be obtained from Eq. (3.109) if v_1' and V are expressed in terms of v_0 by Eqs. (3.108) and (3.105), respectively:

$$\frac{v_1^2}{v_0^2} = \left(\frac{\mu}{m_2\rho}\right)^2 (1 + 2\rho\cos\Theta + \rho^2). \tag{3.117}$$

For elastic collisions $\rho = m_1/m_2$, and Eq. (3.117) can be simplified to

$$\frac{E_1}{E_0} = \frac{1 + 2\rho\cos\Theta + \rho^2}{(1+\rho)^2}, \qquad \text{(elastic collision)} \tag{3.117'}$$

where E_0 is the initial kinetic energy of the incident particle in the laboratory system and E_1 the corresponding energy after scattering. When the particles are of equal mass, this relation becomes

$$\frac{E_1}{E_0} = \frac{1 + \cos\Theta}{2} = \cos\vartheta.$$

Thus, at the maximum scattering angle ($\Theta = \pi$, $\vartheta = \pi/2$), the incident particle loses all its energy and is completely stopped in the laboratory system.

THE THREE-BODY PROBLEM

The Newtonian three-body problem involves three masses m_1, m_2, and m_3 at the respective positions $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$, interacting with each other via gravitational forces. We assume that the position vectors $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$ are expressed in the center of mass system. It is easy to write the equation of motion of the first mass since by Newton's second law $m_1\ddot{\mathbf{r}}_1$ equals the gravitational forces that the other two masses exert on m_1:

$$\ddot{\mathbf{r}}_1 = -Gm_2\frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} - Gm_3\frac{\mathbf{r}_1 - \mathbf{r}_3}{|\mathbf{r}_1 - \mathbf{r}_3|^3} \tag{3.118}$$

and analogously for the other two masses. If we make use of the relative-position vectors defined by

$$\mathbf{s}_i = \mathbf{r}_j - \mathbf{r}_k \tag{3.119}$$

in Fig. 3.27, then clearly

$$\mathbf{s}_1 + \mathbf{s}_2 + \mathbf{s}_3 = 0. \tag{3.120}$$

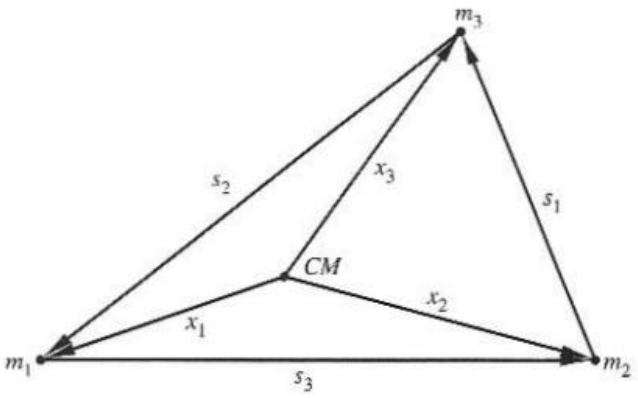


FIGURE 3.27 Position vectors $\mathbf{s}_i = \mathbf{r}_j - \mathbf{r}_k$ for the three-body problem. Adapted from Hestenes, *New Foundations for Classical Mechanics*, 1999, Fig. 5.1.

After a little algebra, the equations of motion assume the symmetrical form

$$\ddot{\mathbf{s}}_i = -mG\frac{\mathbf{s}_i}{s_i^3} + m_i\mathbf{G} \tag{3.121}$$

where $i = 1, 2, 3$, the quantity m is the sum of the three masses

$$m = m_1 + m_2 + m_3 \tag{3.122}$$

and the vector $\mathbf{G}$ is given by

$$\mathbf{G} = G\left(\frac{\mathbf{s}_1}{s_1^3} + \frac{\mathbf{s}_2}{s_2^3} + \frac{\mathbf{s}_3}{s_3^3}\right). \tag{3.123}$$

The three coupled equations in the symmetrical form, (3.121), cannot be solved in general, but they do provide solutions to the three-body problem for some simple cases.

There is a solution due to Euler in which mass m_2 always lies on the straight line between the other two masses so that $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$, and $\mathbf{G}$ are all collinear. Figure 3.28 shows Euler's negative-energy (i.e., bound-state) solution for the mass ratio $m_1 < m_2 < m_3$ in which the masses move along confocal ellipses with the same period τ. During each period, the masses pass through both a perihelion configuration, in which they lie close together along the axis of the ellipses, and an aphelion configuration, in which they lie along this same axis but far apart. The aphelion positions in the orbits are indicated in Figure 3.28.

If the vector $\mathbf{G} = 0$, the equations of motion decouple, and Eq. (3.121) reduces to the two-body form of the Kepler problem,

$$\ddot{\mathbf{s}}_i = -mG\frac{\mathbf{s}_i}{s_i^3}, \tag{3.124}$$

with each mass moving along an elliptical orbit lying in the same plane with the same focal point and the same period. This decoupling occurs when the three

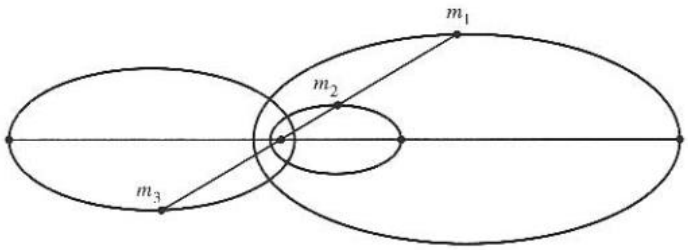


FIGURE 3.28 Euler's collinear solution to the three-body problem for the mass ratio $m_1 < m_2 < m_3$. Three of the dots show aphelion positions. Adapted from Hestenes, *New Foundations for Classical Mechanics*, 1999, Fig. 5.2.

Unit III: Rigid body Dynamics and Small Oscillations

The Independent coordinates of a Rigid Body - Orthogonal Transformations - The Euler Angles - Angular Momentum and Kinetic Energy of Motion about a Point – Tensors - the Inertia Tensor and the Moment of Inertia - The Eigen values of the Inertia Tensor and the Principal Axis Transformation - Solving Rigid Body Problems and the Euler Equations of Motion. Small Oscillations - Frequencies of Free Vibration, and Normal Coordinates, Linear Tri atomic Molecule.

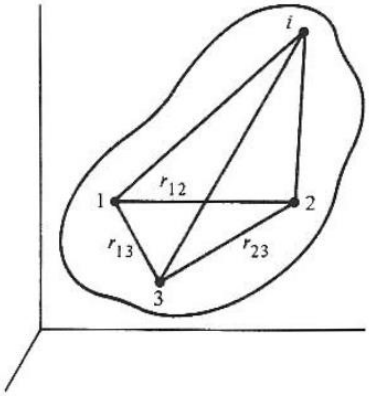
THE INDEPENDENT COORDINATES OF A RIGID BODY

Before discussing the motion of a rigid body, we must first establish how many independent coordinates are necessary to specify its configuration. From experience, we expect that there should be six independent coordinates. Three external coordinates are needed to specify the position of some reference point in the body and three more to specify how the body is oriented with respect to the external coordinates. In this section we show that these intuitive expectations are correct.

A rigid body with N particles can at most have $3N$ degrees of freedom, but these are greatly reduced by the constraints, which can be expressed as equations of the form

$$r_{ij} = c_{ij}. \tag{4.1}$$

Here r_{ij} is the distance between the ith and jth particles and the c's are constants. The actual number of degrees of freedom cannot be obtained simply by subtracting the number of constraint equations from $3N$, for there are $\frac{1}{2}N(N-1)$ possible equations of the form of Eq. (4.1), which is far greater than $3N$ for large N. In truth, the Eqs. (4.1) are not all independent.



The location of a point in a rigid body by its distances from three reference

To fix a point in the rigid body, it is not necessary to specify its distances to *all* other points in the body; we need only state the distances to any three other noncollinear points (cf. Fig. 4.1). Thus, once the positions of three of the particles of the rigid body are determined, the constraints fix the positions of all remaining particles. The number of degrees of freedom therefore cannot be more than nine. But the three reference points are themselves not independent; there are in fact three equations of rigid constraint imposed on them,

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

that reduce the number of degrees of freedom to *six*. That only six coordinates are needed can also be seen from the following considerations. To establish the position of one of the reference points, three coordinates must be supplied. But once point 1 is fixed, point 2 can be specified by only two coordinates, since it is constrained to move on the surface of a sphere centered at point 1. With these two points determined, point 3 has only one degree of freedom, for it can only rotate about the axis joining the other two points. Hence, a total of six coordinates is sufficient.

A rigid body in space thus needs six independent generalized coordinates to specify its configuration, no matter how many particles it may contain—even in the limit of a continuous body. Of course, there may be additional constraints on the body besides the constraint of rigidity. For example, the body may be constrained to move on a surface, or with one point fixed. In such case, the additional constraints will further reduce the number of degrees of freedom, and hence the number of independent coordinates.

How shall these coordinates be assigned? Note that the set of configuration of a rigid body is completely specified by locating a Cartesian set of coordinates

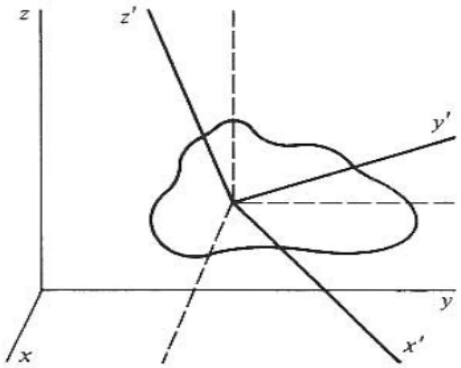


FIGURE 4.2 Unprimed axes represent an external reference set of axes; the primed axes are fixed in the rigid body.

fixed in the rigid body (the primed axes shown in Fig. 4.2) relative to the coordinate axes of the external space. Clearly three of the coordinates are needed to specify the coordinates of the origin of this "body" set of axes. The remaining three coordinates must then specify the orientation of the primed axes relative to a coordinate system parallel to the external axes, but with the same origin as the primed axes.

There are many ways of specifying the orientation of a Cartesian set of axes relative to another set with common origin. One fruitful procedure is to state the direction cosines of the primed axes relative to the unprimed. Thus, the x' axis could be specified by its three direction cosines α_1, α_2, α_3, with respect to the x, y, z axes. If, as customary, **i**, **j**, **k** are three unit vectors along x, y, z, and **i**′, **j**′, **k**′ perform the same function in the primed system (cf. Fig. 4.3), then these direction cosines are defined as

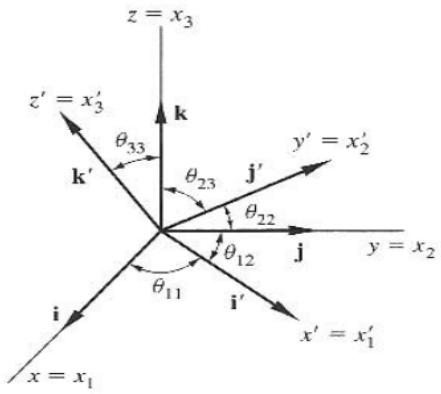


FIGURE 4.3 Direction cosines of the body set of axes relative to an external set of axes.

$$\cos\theta_{11} = \cos(\mathbf{i}' \cdot \mathbf{i}) = \mathbf{i}' \cdot \mathbf{i} = \mathbf{i} \cdot \mathbf{i}'$$
$$\cos\theta_{12} = \cos(\mathbf{i}' \cdot \mathbf{j}) = \mathbf{i}' \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i}'$$
$$\cos\theta_{21} = \cos(\mathbf{j}' \cdot \mathbf{i}) = \mathbf{j}' \cdot \mathbf{i} = \mathbf{i} \cdot \mathbf{j}'$$
$$\cos\theta_{22} = \cos(\mathbf{j}' \cdot \mathbf{j}) = \mathbf{j}' \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{j}' \tag{4.2}$$

and similarly for $\cos\theta_{13}$, $\cos\theta_{31}$, etc. Note that the angle θ_{ij} is defined so that the first index refers to the primed system and the second index to the unprimed system. These direction cosines can also be used to express the unit vector in the primed system in terms of the unit vectors of the unprimed system giving

$$\mathbf{i}' = \cos\theta_{11}\mathbf{i} + \cos\theta_{12}\mathbf{j} + \cos\theta_{13}\mathbf{k}$$
$$\mathbf{j}' = \cos\theta_{21}\mathbf{i} + \cos\theta_{22}\mathbf{j} + \cos\theta_{23}\mathbf{k}$$
$$\mathbf{k}' = \cos\theta_{31}\mathbf{i} + \cos\theta_{32}\mathbf{j} + \cos\theta_{33}\mathbf{k}. \tag{4.3}$$

These sets of nine directions cosines then completely specify the orientation of the x', y', z' axes relative to the x, y, z set. We can equally well invert the process, and use the direction cosines to express the **i**, **j**, **k** unit vectors in terms of their components along the primed axes. Thus, we can write

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}' \tag{4.4}$$

by

$$x' = (\mathbf{r} \cdot \mathbf{i}') = \cos\theta_{11}x + \cos\theta_{12}y + \cos\theta_{13}z$$
$$y' = (\mathbf{r} \cdot \mathbf{j}') = \cos\theta_{21}x + \cos\theta_{22}y + \cos\theta_{23}z$$
$$z' = (\mathbf{r} \cdot \mathbf{k}') = \cos\theta_{31}x + \cos\theta_{32}y + \cos\theta_{33}z \tag{4.5}$$

with analogous equations for **i**, **j** and **k**.

The direction cosines also furnish directly the relations between the coordinates of a given point in one system and the coordinates in the other system. Thus, the coordinates of a point in a given reference frame are the components of the position vector, **r**, along the primed and unprimed axes of the system, respectively. The primed coordinates are then given in terms of x, y, and z, as shown in Eq. (4.5). What has been done here for the components of the **r** vector can obviously be done for any arbitrary vector. If **G** is some vector, then the component of **G** along the x' axis will be related to its x-, y-, z-components by

$$G_{x'} = \mathbf{G} \cdot \mathbf{i}' = \cos\theta_{11}G_x + \cos\theta_{12}G_y + \cos\theta_{13}G_z, \tag{4.6}$$

and so on. The set of nine direction cosines thus completely spells out the transformation between the two coordinate systems.

If the primed axes are taken as fixed in the body, then the nine direction cosines will be functions of time as the body changes its orientation in the course of the

motion. In this sense, the direction cosines can be considered as coordinates describing the instantaneous orientation of the body, relative to a coordinate system fixed in space but with origin in common with the body system. But, clearly, they are not independent coordinates, for there are nine of them and it has been shown that only three coordinates are needed to specify an orientation.

The connections between the direction cosines arise from the fact that the basis vectors in both coordinate systems are orthogonal to each other and have unit magnitude; in symbols,

$$\begin{aligned} &\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = 0, \\ \text{and} \qquad &\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1, \end{aligned} \tag{4.7}$$

with similar relations for $\mathbf{i}'$, $\mathbf{j}'$, and $\mathbf{k}'$. We can obtain the conditions satisfied by the nine coefficients by forming all possible dot products among the three equations for $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in terms of $\mathbf{i}'$, $\mathbf{j}'$, and $\mathbf{k}'$ (as in Eq. (4.4)), making use of the Eqs. (4.7):

$$\begin{aligned} &\sum_{l=1}^{3} \cos\theta_{lm'} \cos\theta_{lm} = 0 \qquad m \neq m' \\ &\sum_{l=1}^{3} \cos^2\theta_{lm} = 1. \end{aligned} \tag{4.8}$$

These two sets of three equations each are exactly sufficient to reduce the number of independent quantities from nine to three. Formally, the six equations can be combined into one by using the Kronecker δ-symbol δ_{lm}, defined by

$$\begin{aligned} \delta_{lm} &= 1 \qquad l = m \\ &= 0 \qquad l \neq m. \end{aligned}$$

Equations (4.8) can then be written as

$$\sum_{l=1}^{3} \cos\theta_{lm'} \cos\theta_{lm} = \delta_{m'm} \tag{4.9}$$

It is therefore not possible to set up a Lagrangian and subsequent equations of motion with the nine direction cosines as generalized coordinates. For this purpose, we must use some set of three independent functions of the direction cosines. A number of such sets of independent variables will be described later, the most important being the Euler angles. The use of direction cosines to de-

■ ORTHOGONAL TRANSFORMATIONS

To study the properties of the nine direction cosines with greater ease, it is convenient to change the notation and denote all coordinates by x, distinguishing the axes by subscripts:

$$\begin{aligned} x &\rightarrow x_1 \\ y &\rightarrow x_2 \\ z &\rightarrow x_3 \end{aligned} \tag{4.10}$$

as shown in Fig. 4.3. We also change the notation for the direction cosines to

$$a_{ij} = \cos\theta_{ij} \tag{4.11}$$

Equations (4.5) and (4.6) constitute a group of transformation equations from a set of coordinates x_1, x_2, x_3 to a new set x_1', x_2', x_3'. In particular, they form an example of a *linear* or *vector* transformation, defined by transformation equations of the form

$$\begin{aligned} x_1' &= a_{1\,1}x_1 + a_{1\,2}x_2 + a_{1\,3}x_3 \\ x_2' &= a_{2\,1}x_1 + a_{2\,2}x_2 + a_{2\,3}x_3 \\ x_3' &= a_{3\,1}x_1 + a_{3\,2}x_2 + a_{3\,3}x_3, \end{aligned} \tag{4.12}$$

where the $a_{1\,1}, a_{1\,2}, \ldots$, are any set of constant (independent of x, x') coefficients.* To simplify the appearance of many of the expressions, we will also make use of the summation convention first introduced by Einstein: Whenever an index occurs two or more times in a term, it is implied, without any further symbols, that the terms are to be summed over all possible values of the index. Thus, Eqs. (4.12) can be written most compactly in accordance with this convention as

$$x_i' = a_{ij}x_j, \qquad i = 1, 2, 3. \tag{4.12'}$$

The repeated appearance of the index j indicates that the left-hand side of Eq. (4.12′) is a sum over the dummy index j for all possible values (here, $j = 1, 2, 3$). Some ambiguity is possible where powers of an indexed quantity occur, and for that reason, an expression such as

$$\sum_i x_i^2$$

appears under the summation convention as

$$x_i x_i .$$

For the rest of the book the summation convention should be automatically assumed in reading the equations unless otherwise explicitly indicated. Where convenient, or to remove ambiguity, the summation sign may be occasionally displayed explicitly, e.g., when certain values of the index are to be excluded from the summation.

The transformation represented by Eqs. (4.11) is only a special case of the general linear transformation, Eqs. (4.12), since the direction cosines are not all independent. The connections between the coefficients, Eqs. (4.8) are rederived here in terms of the newer notation. Since both coordinate systems are Cartesian, the magnitude of a vector is given in terms of the sum of squares of the components. Further, since the actual vector remains unchanged no matter which coordinate system is used, the magnitude of the vector must be the same in both systems. In symbols, we can state the invariance of the magnitude as

$$x_i' x_i' = x_i x_i. \tag{4.13}$$

The left-hand side of Eq. (4.13) is therefore

$$a_{ij} a_{ik} x_j x_k,$$

and it will reduce to the right-hand side of Eq. (4.13), if, and only if

$$\begin{aligned} a_{ij} a_{ik} &= 1 \qquad j = k \\ &= 0 \qquad j \neq k, \end{aligned} \tag{4.14}$$

or, in a more compact form, if

$$a_{ij} a_{ik} = \delta_{jk}, \qquad j, k = 1, 2, 3. \tag{4.15}$$

When the a_{ij} coefficients are expressed in terms of the direction cosines, the six equations contained in Eq. (4.15) become identical with the Eqs. (4.9).

Any linear transformation, Eq. (4.12), that has the properties required by Eq. (4.15) is called an *orthogonal* transformation, and Eq. (4.15) itself is known as the *orthogonality condition*. Thus, the transition from coordinates fixed in space to coordinates fixed in the rigid body (with common origin) is accomplished by means of an orthogonal transformation. The array of transformation quantities (the direction cosines), written as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \tag{4.16}$$

is called the *matrix of transformation*, and will be denoted by a capital letter **A**. The quantities a_{ij} are correspondingly known as the *matrix elements* of the transformation.

To make these formal considerations more meaningful, consider the simple example of motion in a plane, so that we are restricted to two-dimensional rotations,

and the transformation matrix reduces to the form

$$\begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The four matrix elements, a_{ij}, are connected by three orthogonality conditions:

$$a_{ij}a_{ik} = \delta_{jk}, \qquad j, k = 1, 2,$$

and therefore only one independent parameter is needed to specify the transformation. But this conclusion is not surprising. A two-dimensional transformation from one Cartesian coordinate system to another corresponds to a rotation of the axes in the plane (cf. Fig. 4.4), and such a rotation can be specified completely by only one quantity, the rotation angle ϕ. Expressed in terms of this single parameter, the transformation equations become

$$\begin{aligned} x_1' &= x_1 \cos\phi + x_2 \sin\phi \\ x_2' &= -x_1 \sin\phi + x_2 \cos\phi \\ x_3' &= x_3. \end{aligned}$$

The matrix elements are therefore

$$\begin{array}{lll} a_{11} = \cos\phi & a_{12} = \sin\phi & a_{13} = 0 \\ a_{21} = -\sin\phi & a_{22} = \cos\phi & a_{23} = 0 \\ a_{31} = 0 & a_{32} = 0 & a_{33} = 1, \end{array} \qquad (4.17)$$

so that the matrix **A** can be written

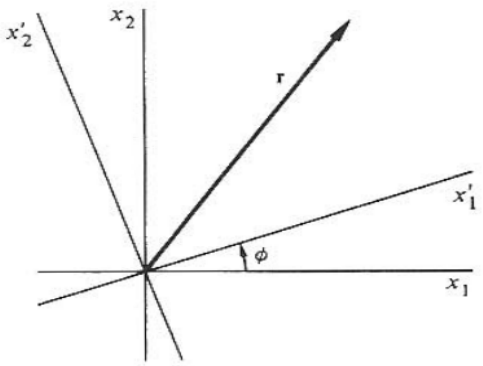


FIGURE 4.4 Rotation of the coordinate axes, as equivalent to two-dimensional orthogonal transformation.

$$\mathbf{A} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.17'}$$

The three nontrivial orthogonality conditions expand into the equations

$$a_{11}a_{11} + a_{21}a_{21} = 1$$
$$a_{12}a_{12} + a_{22}a_{22} = 1$$
$$a_{11}a_{12} + a_{21}a_{22} = 0.$$

These conditions are obviously satisfied by the matrix (4-17′), for in terms of the matrix elements (4.17) they reduce to the identities

$$\cos^2\phi + \sin^2\phi = 1$$
$$\sin^2\phi + \cos^2\phi = 1$$
$$\cos\phi\sin\phi - \sin\phi\cos\phi = 0.$$

The transformation matrix **A** can be thought of as an *operator* that, acting on the unprimed system, transforms it into the primed system. Symbolically, the process might be written

$$(\mathbf{r})' = \mathbf{A}\mathbf{r}, \tag{4.18}$$

which is to be read: The matrix **A** operating on the components of a vector in the unprimed system yields the components of the vector in the primed system. Note that in the development of the subject so far, **A** acts on the coordinate system only, the vector is unchanged, and we ask merely for its components in two different coordinate frames. Parentheses have therefore been placed around **r** on the left in Eq. (4.18) to make clear that the same vector is involved on both sides on the equation. Only the components have changed. In three dimensions, the transformation of coordinates, as shown earlier, is simply a rotation, and **A** is then identical with the *rotation* operator in a plane.

Despite this, note that without changing the formal mathematics, **A** can also be thought of as an operator acting on the *vector* **r**, changing it to a different vector **r**′:

$$\mathbf{r}' = \mathbf{A}\mathbf{r}, \tag{4.19}$$

with both vectors expressed in the same coordinate system. Thus, in two dimensions, instead of rotating the coordinate system counterclockwise, we can rotate the vector **r** *clockwise* by an angle ϕ to a new vector **r**′, as shown in Fig. 4.5. The

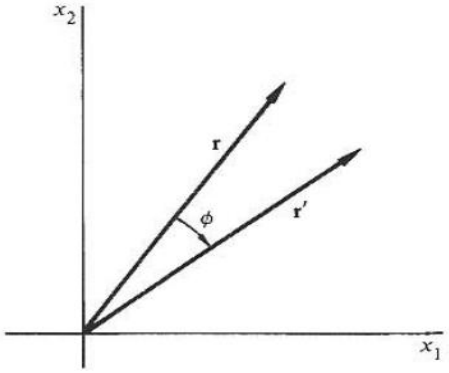


FIGURE 4.5 Interpretation of an orthogonal transformation as a rotation of the vector, leaving the coordinate system unchanged.

which of these two points of view is followed. The interpretation as an operator acting on the coordinates is the more pertinent one when using the orthogonal transformation to specify the orientation of a rigid body. On the other hand, the notion of an operator changing one vector into another has the more widespread application. In the mathematical discussion either interpretation will be freely used, as suits the convenience of the situation. Of course, note that the nature of the operation represented by **A** will change according to which interpretation is selected. Thus, if **A** corresponds to a *counterclockwise* rotation by an angle ϕ when applied to the coordinate system, it will correspond to a *clockwise* rotation when applied to the vector.

■ THE EULER ANGLES

We have noted (cf. p. 137) that the nine elements a_{ij} are not suitable as generalized coordinates because they are not independent quantities. The six relations that express the orthogonality conditions, Eqs. (4.9) or Eqs. (4.15), of course reduce the number of independent elements to three. But in order to characterize the motion of a rigid body, there is an additional requirement the matrix elements must satisfy, beyond those implied by orthogonality. In the previous section we pointed out that the determinant of a real orthogonal matrix could have the value $+1$ or -1. The following argument shows however that an orthogonal matrix whose determinant is -1 cannot represent a physical displacement of a rigid body.

Consider the simplest 3×3 matrix with the determinant -1:

$$\mathbf{S} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -\mathbf{1}.$$

The transformation **S** has the effect of changing the sign of each of the components or coordinate axes (cf. Fig. 4.6). Such an operation transforms a right-handed coordinate system into a left-handed one and is known as an *inversion* of the coordinate axes.

One method of performing an inversion is to rotate about a coordinate axis by 180° and then reflect in that coordinate axis direction. For the z-direction, this gives

$$\begin{pmatrix} \text{rotate} \\ \text{by } 180° \\ \text{about } z \end{pmatrix} \begin{pmatrix} \text{reflect} \\ \text{in the} \\ xy \text{ plane} \end{pmatrix} = \text{inversion}.$$

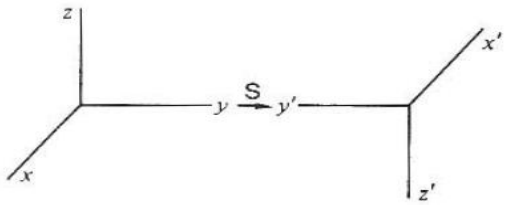


FIGURE 4.6 Inversion of the coordinate axes.

In matrix notation, this has the form

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

where the 180° rotation is obtained by setting $\phi = 180°$ in Eq. (4.17).

From the nature of this operation, it is clear that an inversion of a right-handed system into a left-handed one cannot be accomplished by any *rigid* change in the orientation of the coordinate axes. An inversion therefore never corresponds to a physical displacement of a rigid body. What is true for the inversion **S** is equally valid for any matrix whose determinant is -1, for any such matrix can be written as the product of **S** with a matrix whose determinant is $+1$, and thus includes the inversion operation. Consequently, it cannot describe a rigid change in orientation. Therefore, the transformations representing rigid body motion must be restricted to matrices having the determinant $+1$. Another method of reaching this conclusion starts from the fact that the matrix of transformation must evolve continuously from the unit matrix, which of course has the determinant $+1$. It would be incompatible with the continuity of the motion to have the matrix determinant suddenly change from its initial value $+1$ to -1 at some given time. Orthogonal transformations with determinant $+1$ are said to be *proper*, and those with the determinant -1 are called *improper*.

In order to describe the motion of rigid bodies in the Lagrangian formulation of mechanics, it will therefore be necessary to seek three independent parameters that specify the orientation of a rigid body in such a manner that the corresponding orthogonal matrix of transformation has the determinant $+1$. Only when such generalized coordinates have been found can we write a Lagrangian for the system and obtain the Lagrangian equations of motion. A number of such sets of parameters have been described in the literature, but the most common and useful are the *Euler* or *Eulerian angles*. We shall therefore define these angles at this point, and show how the elements of the orthogonal transformation matrix can be expressed in terms of them.

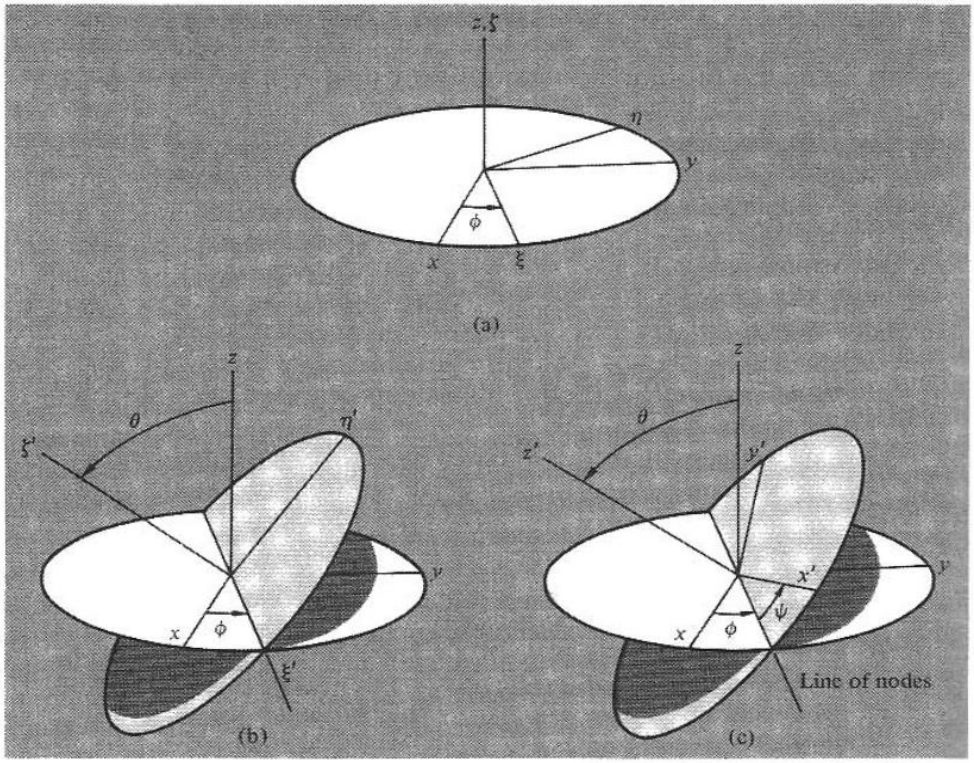


FIGURE 4.7 The rotations defining the Eulerian angles.

terclockwise by an angle ψ about the ζ' axis to produce the desired $x'y'z'$ system of axes. Figure 4.7 illustrates the various stages of the sequence. The Euler angles θ, ϕ, and ψ thus completely specify the orientation of the $x'y'z'$ system relative to the xyz and can therefore act as the three needed generalized coordinates.*

The elements of the complete transformation **A** can be obtained by writing the matrix as the triple product of the separate rotations, each of which has a relatively simple matrix form. Thus, the initial rotation about z can be described by a matrix **D**:

$$\boldsymbol{\xi} = \mathbf{D}\mathbf{x},$$

where $\boldsymbol{\xi}$ and $\mathbf{x}$ stand for column matrices. Similarly, the transformation from $\xi\eta\zeta$ to $\xi'\eta'\zeta'$ can be described by a matrix **C**,

$$\boldsymbol{\xi}' = \mathbf{C}\boldsymbol{\xi},$$

and the last rotation to $x'y'z'$ by a matrix $\mathbf{B}$,

$$\mathbf{x}' = \mathbf{B}\boldsymbol{\xi}'.$$

Hence, the matrix of the complete transformation,

$$\mathbf{x}' = \mathbf{A}\mathbf{x},$$

is the product of the successive matrices,

$$\mathbf{A} = \mathbf{BCD}.$$

Now the $\mathbf{D}$ transformation is a rotation about z, and hence has a matrix of the form (cf. Eq. (4.17))

$$\mathbf{D} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.43}$$

The $\mathbf{C}$ transformation corresponds to a rotation about ξ, with the matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \tag{4.44}$$

and finally $\mathbf{B}$ is a rotation about ζ' and therefore has the same form as $\mathbf{D}$:

$$\mathbf{B} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.45}$$

The product matrix $\mathbf{A} = \mathbf{BCD}$ then follows as

$$\mathbf{A} = \begin{bmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix}. \tag{4.46}$$

The inverse transformation from body coordinates to space axes

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{x}'$$

is then given immediately by the transposed matrix $\tilde{\mathbf{A}}$:

$$\mathbf{A}^{-1} = \tilde{\mathbf{A}} = \begin{bmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & \sin\theta\sin\phi \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{bmatrix}. \tag{4.47}$$

Verification of the multiplication, and demonstration that $\mathbf{A}$ represents a proper, orthogonal matrix will be left to the exercises.

ANGULAR MOMENTUM AND KINETIC ENERGY OF MOTION ABOUT A POINT

Chasles' theorem states that any general displacement of a rigid body can be represented by a translation plus a rotation. The theorem suggests that it ought to be possible to split the problem of rigid body motion into two separate phases, one concerned solely with the translational motion of the body, the other, with its rotational motion. Of course, if one point of the body is fixed, the separation is obvious, for then there is only a rotational motion about the fixed point, without any translation. But even for a general type of motion such a separation is often possible. The six coordinates needed to describe the motion have already been formed into two sets in accordance with such a division: the three Cartesian coordinates of a point fixed in the rigid body to describe the translational motion and, say, the three Euler angles for the motion about the point. If, further, the origin of the body system is chosen to be the center of mass, then by Eq. (1.28) the total angular momentum divides naturally into contributions from the translation of the center of mass and from the rotation about the center of mass. The former term will involve only the Cartesian coordinates of the center of mass, the latter only the angle coordinates. By Eq. (1.31), a similar division holds for the total kinetic energy T, which can be written in the form

$$T = \tfrac{1}{2}Mv^2 + T'(\phi, \theta, \psi),$$

ANGULAR MOMENTUM AND KINETIC ENERGY OF MOTION ABOUT A POINT

Chasles' theorem states that any general displacement of a rigid body can be represented by a translation plus a rotation. The theorem suggests that it ought to be possible to split the problem of rigid body motion into two separate phases, one concerned solely with the translational motion of the body, the other, with its rotational motion. Of course, if one point of the body is fixed, the separation is obvious, for then there is only a rotational motion about the fixed point, without any translation. But even for a general type of motion such a separation is often possible. The six coordinates needed to describe the motion have already been formed into two sets in accordance with such a division: the three Cartesian coordinates of a point fixed in the rigid body to describe the translational motion and, say, the three Euler angles for the motion about the point. If, further, the origin of the body system is chosen to be the center of mass, then by Eq. (1.28) the total angular momentum divides naturally into contributions from the translation of the center of mass and from the rotation about the center of mass. The former term will involve only the Cartesian coordinates of the center of mass, the latter only the angle coordinates. By Eq. (1.31), a similar division holds for the total kinetic energy T, which can be written in the form

$$T = \tfrac{1}{2}Mv^2 + T'(\phi, \theta, \psi),$$

Let $\mathbf{R}_1$ and $\mathbf{R}_2$ be the position vectors, relative to a fixed set of coordinates, of the origins of two sets of body coordinates (cf. Fig. 5.1). The difference vector is denoted by $\mathbf{R}$:

$$\mathbf{R}_2 = \mathbf{R}_1 + \mathbf{R}.$$

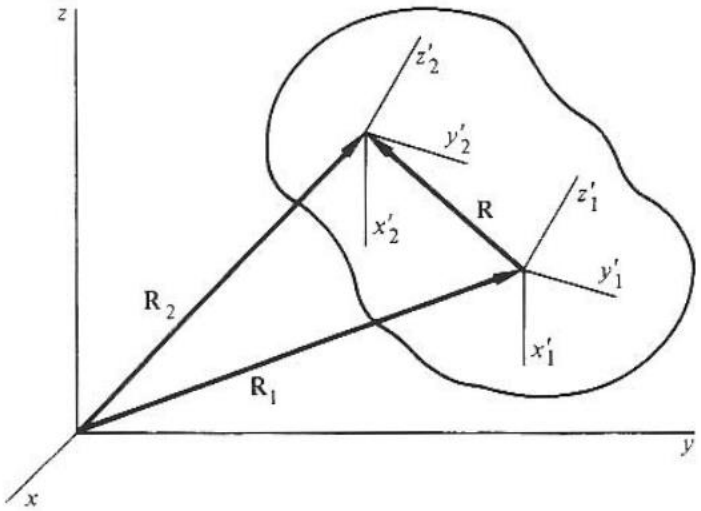


FIGURE 5.1 Vectorial relation between sets of rigid body coordinates with different origins.

If the origin of the second set of axes is considered as a point defined relative to the first, then the time derivative of $\mathbf{R}_2$ relative to the space axes is given by

$$\left(\frac{d\mathbf{R}_2}{dt}\right)_s = \left(\frac{d\mathbf{R}_1}{dt}\right)_s + \left(\frac{d\mathbf{R}}{dt}\right)_s = \left(\frac{d\mathbf{R}_1}{dt}\right)_s + \boldsymbol{\omega}_1 \times \mathbf{R}.$$

The last step follows from Eq. (4.86), recalling that the derivatives of $\mathbf{R}$ relative to any rigid body axes must vanish, and with $\boldsymbol{\omega}_1$ as being the angular velocity vector appropriate to the first coordinate system. Alternatively, the origin of the first coordinate system can be considered as fixed in the second system with the position vector $-\mathbf{R}$. In the same manner, then, the derivative of the position vector $\mathbf{R}_1$ to this origin relative to the fixed-space axes can be written as

$$\left(\frac{d\mathbf{R}_1}{dt}\right)_s = \left(\frac{d\mathbf{R}_2}{dt}\right)_s - \left(\frac{d\mathbf{R}}{dt}\right)_s = \left(\frac{d\mathbf{R}_2}{dt}\right)_s - \boldsymbol{\omega}_2 \times \mathbf{R}.$$

A comparison of these two expressions shows $(\boldsymbol{\omega}_1 - \boldsymbol{\omega}_2) \times \mathbf{R} = 0$. Any difference in the angular velocity vectors at two arbitrary points must lie along the line joining the two points. Assuming the $\boldsymbol{\omega}$ vector field is continuous, the only possible solution for all pairs of points is that the two angular velocity vectors must be equal:

$$\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2.^*$$

The angular velocity vector is the same for all coordinate systems fixed in the rigid body.

When a rigid body moves with one point stationary, the total angular momentum about that point is

$$\mathbf{L} = m_i(\mathbf{r}_i \times \mathbf{v}_i), \tag{5.1}$$

(employing the summation convention) where $\mathbf{r}_i$ and $\mathbf{v}_i$ are the radius vector and velocity, respectively, of the ith particle relative to the given point. Since $\mathbf{r}_i$ is a fixed vector relative to the body, the velocity $\mathbf{v}_i$ with respect to the space set of axes arises solely from the rotational motion of the rigid body about the fixed point. From Eq. (4.86), $\mathbf{v}_i$ is then

$$\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i. \tag{5.2}$$

Hence, Eq. (5.1) can be written as

$$\mathbf{L} = m_i\left[\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i)\right],$$

or, expanding the triple cross product,

$$\mathbf{L} = m_i\left[\boldsymbol{\omega} r_i^2 - \mathbf{r}_i(\mathbf{r}_i \cdot \boldsymbol{\omega})\right]. \tag{5.3}$$

Again expanding, the x-component of the angular momentum becomes

$$L_x = \omega_x m_i (r_i^2 - x_i^2) - \omega_y m_i x_i y_i - \omega_z m_i x_i z_i, \tag{5.4}$$

with similar equations for the other components of $\mathbf{L}$. Thus, each component of the angular momentum is a linear function of all the components of the angular velocity. *The angular momentum vector is related to the angular velocity by a linear transformation.* To emphasize the similarity of (5.4) with the equations of a linear transformation, (4.12), we may write L_x as

$$L_x = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z.$$

Analogously, for L_y and L_z we have

$$\begin{aligned} L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z, \\ L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z. \end{aligned} \tag{5.5}$$

The nine coefficients I_{xx}, I_{xy}, etc., are the nine elements of the transformation matrix. The diagonal elements are known as *moment of inertia coefficients*, and have the following form

$$I_{xx} = m_i (r_i^2 - x_i^2), \tag{5.6}$$

while the off-diagonal elements are designated as *products of inertia*, a typical one being

$$I_{xy} = -m_i x_i y_i. \tag{5.7}$$

In Eqs. (5.6) and (5.7), the matrix elements appear in the form suitable if the rigid body is composed of discrete particles. For continuous bodies the summation is replaced by a volume integration, with the particle mass becoming a mass density. Thus, the diagonal element I_{xx} appears as

$$I_{xx} = \int_V \rho(\mathbf{r})(r^2 - x^2)\, dV. \tag{5.6'}$$

With a slight change in notation, an expression for all matrix elements can be stated for continuous bodies. If the coordinate axes are denoted by x_j, $j = 1, 2, 3$, then the matrix element I_{jk} can be written

$$I_{jk} = \int_V \rho(\mathbf{r})(r^2 \delta_{jk} - x_j x_k)\, dV. \tag{5.8}$$

Thus far, the coordinate system used in resolving the components of $\mathbf{L}$ has not been specified. From now on, we will take it to be a system fixed in the body.* The various distances x_i, y_i, z_i are then constant in time, so that the matrix elements are likewise constants, peculiar to the body involved, and dependent on the origin and orientation of the particular body set of axes in which they are expressed.

Equations (5.5) relating the components of $\mathbf{L}$ and $\boldsymbol{\omega}$ can be summarized by a single operator equation,

$$\mathbf{L} = \mathbf{I}\boldsymbol{\omega}, \tag{5.9}$$

where the symbol $\mathbf{I}$ stands for the operator whose matrix elements are the inertia coefficients appearing in (5.5), and $\boldsymbol{\omega}$ and $\mathbf{L}$ are column matrices. Of the two interpretations that have been given to the operator of a linear transformation (cf. Section 4.2), it is clear that here $\mathbf{I}$ must be thought of as acting upon the vector $\boldsymbol{\omega}$, and not upon the coordinate system. The vectors $\mathbf{L}$ and $\boldsymbol{\omega}$ are two physically different vectors, having different dimensions, and are not merely the same vector expressed in two different coordinate systems. Unlike the operator of rotation, $\mathbf{I}$ will have dimensions—mass times length squared—and it is not restricted by any orthogonality conditions. Equation (5.9) is to be read as the operator $\mathbf{I}$ acting upon the vector $\boldsymbol{\omega}$ results in the physically new vector $\mathbf{L}$.

■ TENSORS

The quantity $\mathbf{I}$ may be considered as defining the quotient of $\mathbf{L}$ and $\boldsymbol{\omega}$ for the product of $\mathbf{I}$ and $\boldsymbol{\omega}$ gives $\mathbf{L}$. Now, the quotient of two quantities is often not a member of the same class as the dividing factors, but may belong to a more complicated class. Thus, the quotient of two integers is in general not an integer but rather a rational number. Similarly, the quotient of two vectors, as is well known, cannot be defined consistently within the class of vectors. It is not surprising, therefore to find that $\mathbf{I}$ is a new type of quantity, a *tensor of the second rank.*

In a Cartesian three-dimensional space, a tensor $\mathbf{T}$ of the Nth rank may be defined for our purposes as a quantity having 3^N components $T_{ijk\ldots}$ (with N indices) that transform under an orthogonal transformation of coordinates, $\mathbf{A}$, according to

the following scheme:*

$$T'_{ijk\ldots}(\mathbf{x}') = a_{il}a_{jm}a_{kn}\ldots T_{lmn\ldots}(\mathbf{x}). \tag{5.10}$$

By this definition, a tensor of the zero rank has one component, which is invariant under an orthogonal transformation. Hence, a *scalar is a tensor of zero rank*. A tensor of the first rank has three components transforming as

$$T'_i = a_{ij}T_j.$$

Comparison with the transformation equations for a vector, (4.12′), shows that a *tensor of the first rank is completely equivalent to a vector*.† Finally, the nine components of a tensor of the second rank transform as

$$T'_{ij} = a_{ik}a_{jl}T_{kl}. \tag{5.11}$$

■ THE INERTIA TENSOR AND THE MOMENT OF INERTIA

Considered as a linear operator that transforms $\boldsymbol{\omega}$ into L, the matrix $\mathbf{I}$ has elements that behave as the elements of a second-rank tensor. The quantity $\mathbf{I}$ is therefore identified as a second-rank tensor and is usually called the *moment of inertia tensor* or briefly the *inertia tensor*.

The kinetic energy of motion about a point is

$$T = \tfrac{1}{2}m_i v_i^2,$$

where $\mathbf{v}_i$ is the velocity of the ith particle relative to the fixed point as measured in the space axes. By Eq. (5.2), T may also be written as

$$T = \tfrac{1}{2}m_i\mathbf{v}_i \cdot (\boldsymbol{\omega} \times \mathbf{r}_i),$$

which, upon permuting the vectors in the triple dot product, becomes

$$T = \frac{\boldsymbol{\omega}}{2} \cdot m_i(\mathbf{r}_i \times \mathbf{v}_i).$$

The quantity summed over i will be recognized as the angular momentum of the body about the origin, and in consequence the kinetic energy can be written in the form

$$T = \frac{\boldsymbol{\omega} \cdot \mathbf{L}}{2} = \frac{\boldsymbol{\omega} \cdot \mathbf{I} \cdot \boldsymbol{\omega}}{2}. \tag{5.16}$$

Let $\mathbf{n}$ be a unit vector in the direction of $\boldsymbol{\omega}$ so that $\boldsymbol{\omega} = \omega\mathbf{n}$. Then an alternative form for the kinetic energy is

$$T = \frac{\omega^2}{2}\mathbf{n}\cdot\mathbf{I}\cdot\mathbf{n} = \frac{1}{2}I\omega^2, \tag{5.17}$$

where I is a scalar, defined by

$$I = \mathbf{n}\cdot\mathbf{I}\cdot\mathbf{n} = m_i\left[r_i^2 - (\mathbf{r}_i\cdot\mathbf{n})^2\right], \tag{5.18}$$

and known as the *moment of inertia about the axis of rotation.*

In the usual elementary discussions, the moment of inertia about an axis is defined as the sum, over the particles of the body, of the product of the particle mass and the square of the perpendicular distance from the axis. It must be shown that this definition is in accord with the expression given in Eq. (5.18). The perpendicular distance is equal to the magnitude of the vector $\mathbf{r}_i \times \mathbf{n}$ (cf. Fig. 5.2). Therefore, the customary definition of I may be written as

$$I = m_i(\mathbf{r}_i \times \mathbf{n})\cdot(\mathbf{r}_i \times \mathbf{n}). \tag{5.19}$$

Multiplying and dividing by ω^2, this definition of I may also be written as

$$I = \frac{m_i}{\omega^2}(\boldsymbol{\omega} \times \mathbf{r}_i)\cdot(\boldsymbol{\omega} \times \mathbf{r}_i).$$

But each vector in the dot product is exactly the relative velocity $\mathbf{v}_i$ as measured in the space system of axes. Hence, I so defined is related to the kinetic energy by

$$I = \frac{2T}{\omega^2},$$

which is the same as Eq. (5.17), and therefore I must be identical with the scalar defined by Eq. (5.19).

The value of the moment of inertia depends upon the direction of the axis of rotation. As $\boldsymbol{\omega}$ usually changes its direction with respect to the body in the course

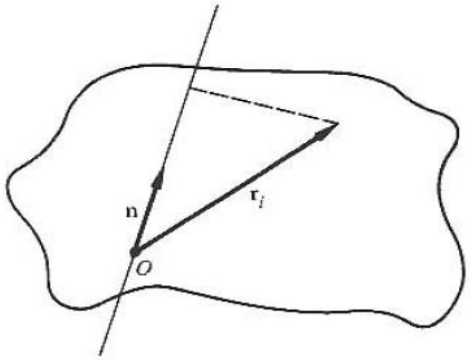


FIGURE 5.2 The definition of the moment of inertia.

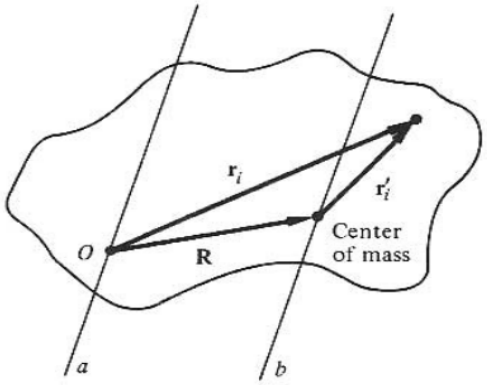


FIGURE 5.3 The vectors involved in the relation between moments of inertia about parallel axes.

of time, the moment of inertia must also be considered a function of time. When the body is constrained so as to rotate only about a fixed axis, then the moment of inertia is a constant. In such a case, the kinetic energy (5.16) is almost in the form required to fashion the Lagrangian and the equations of motion. The one further step needed is to express ω as the time derivative of some angle, which can usually be done without difficulty.

Along with the inertia tensor, the moment of inertia also depends upon the choice of origin of the body set of axes. However, the moment of inertia about some given axis is related simply to the moment about a parallel axis through the center of mass. Let the vector from the given origin O to the center of mass be $\mathbf{R}$, and let the radii vectors from O and the center of mass to the ith particle be $\mathbf{r}_i$ and $\mathbf{r}'_i$, respectively. The three vectors so defined are connected by the relation (cf. Fig. 5.3)

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}'_i. \tag{5.20}$$

The moment of inertia about the axis a is therefore

$$I_a = m_i(\mathbf{r}_i \times \mathbf{n})^2 = m_i[(\mathbf{r}'_i + \mathbf{R}) \times \mathbf{n}]^2$$

or

$$I_a = M(\mathbf{R} \times \mathbf{n})^2 + m_i(\mathbf{r}'_i \times \mathbf{n})^2 + 2m_i(\mathbf{R} \times \mathbf{n}) \cdot (\mathbf{r}'_i \times \mathbf{n}),$$

where M is the total mass of the body. The last term in this expression can be rearranged as

$$-2(\mathbf{R} \times \mathbf{n}) \cdot (\mathbf{n} \times m_i\mathbf{r}'_i).$$

By the definition of center of mass, the summation $m_i \mathbf{r}_i'$ vanishes. Hence, I_a can be expressed in terms of the moment about the parallel axis b as

$$\begin{aligned} I_a &= I_b + M(\mathbf{R} \times \mathbf{n})^2 \\ &= I_b + MR^2 \sin^2\theta. \end{aligned} \tag{5.21}$$

The magnitude of $\mathbf{R} \times \mathbf{n}$, which has the value $R \sin\theta$, where θ is the angle between $\mathbf{R}$ and $\mathbf{n}$, is the perpendicular distance of the center of mass from the axis passing through O. Consequently, the moment of inertia about a given axis is equal to the moment of inertia about a parallel axis through the center of mass plus the moment of inertia of the body, as if concentrated at the center of mass, with respect to the original axis.

The inertia tensor is defined in general from the kinetic energy of rotation about an axis, and is written as

$$T_{\text{rotation}} = \tfrac{1}{2} m_i (\boldsymbol{\omega} \times \mathbf{r}_i)^2 = \tfrac{1}{2} \omega_\alpha \omega_\beta m_i (\delta_{\alpha\beta} r_i^2 - r_{i\alpha} r_{i\beta}),$$

where Greek letters indicate the components of $\boldsymbol{\omega}$ and $\mathbf{r}_i$. In an inertial frame, the sum is over the particles in the body, and $r_{i\alpha}$ is the αth component of the position of the ith particle. Because T_{rotation} is a bilinear form in the components of $\boldsymbol{\omega}$, it can be written as

$$T_{\text{rotation}} = \tfrac{1}{2} I_{\alpha\beta} \omega_\alpha \omega_\beta,$$

where

$$I_{\alpha\beta} = m_i (\delta_{\alpha\beta} r_i^2 - r_{i\alpha} r_{i\beta}) \tag{5.22}$$

is the moment of inertia tensor. To get the moment of inertia about an axis through the center of mass, choose the rotation about this axis. For a body with a continuous distribution of density $\rho(r)$, the sums in the components of the moment of inertia tensor in Eq. (5.22) reduce to

$$I_{\alpha\beta} = \int_V \rho(r)(\delta_{\alpha\beta} r^2 - r_\alpha r_\beta)\, dV. \tag{5.23}$$

THE EIGENVALUES OF THE INERTIA TENSOR AND THE PRINCIPAL AXIS TRANSFORMATION

The preceding discussion emphasizes the important role the inertia tensor plays in the discussion of the motion of rigid bodies. An examination, at this point, of the properties of this tensor and its associated matrix will therefore prove of considerable interest. From the defining equation, (5.7), it is seen that the components of the tensor are symmetrical; that is

$$I_{xy} = I_{yx}. \tag{5.24}$$

This means that, while the inertia tensor will in general have nine components, only six of them will be independent—the three along the diagonal plus three of the off-diagonal elements.

The inertia coefficients depend both upon the location of the origin of the body set of axes and upon the orientation of these axes with respect to the body. This symmetry suggests that there exists a set of coordinates in which the tensor is diagonal with the three principal values I_1, I_2, and I_3. In this system, the components of $\mathbf{L}$ would involve only the corresponding component of $\boldsymbol{\omega}$, thus*

$$L_1 = I_1\omega_1, \qquad L_2 = I_2\omega_2, \qquad L_3 = I_3\omega_3. \tag{5.25}$$

A similar simplification would also occur in the form of the kinetic energy:

$$T = \frac{\boldsymbol{\omega}\cdot\mathbf{I}\cdot\boldsymbol{\omega}}{2} = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2. \tag{5.26}$$

We can show that it is always possible to find such axes, and the proof is based essentially on the symmetric nature of the inertia tensor.

There are several ways to understand vectors and tensors. For example, a vector is a quantity defined by its transformation properties. In any set of coordinates, a vector is specified by its three components, e.g.,

$$\mathbf{V} = V_x\mathbf{i} + V_y\mathbf{j} + V_z\mathbf{k}, \tag{5.27}$$

or by its magnitude and direction. In any frame, the magnitude is given by $\sqrt{V_x^2 + V_y^2 + V_z^2}$, and the direction is given by the polar angles θ and ϕ. An alternative is to use the first two Euler angles to specify a new z axis chosen such that the vector's direction is along that axis. Since the vector lies along that z axis, the third Euler angle is not needed.

An approach similar to this latter method can be used for the symmetric moment of inertia tensor. Consider the moment of inertia of a body about an axis passing through the center of mass of the body. A similarity transformation per-

formed by a rotation matrix **R** can be chosen such that

$$\mathbf{I}_D = \mathbf{R}\mathbf{I}\tilde{\mathbf{R}}. \tag{5.28}$$

This rotation can be expressed in terms of the Euler angles ϕ, θ, and ψ as shown in Eqs. (4.46) and (4.47). A proper choice of these angles will transform **I** into its diagonal form

$$\mathbf{I}_D = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{5.29}$$

where I_1, I_2, and I_3, which are the eigenvalues of **I**, are referred to as the components of the *principal moment of inertia* tensor. The directions of x', y', and z' defined by the rotation matrix in Eq. (5.28) are called the *principal axes*, or *eigenvectors* of the inertia tensor. These eigenvectors lie along the directions x', y', and z'.

Once the principal moments and their directions relative to the surface of a body are known, the inertia tensor relative to any other set of axis through the center of mass can be found by a similarity transformation defined by the Euler angles relating the two coordinate systems. If **S** is that transformation, then

$$\mathbf{I} = \mathbf{S}\mathbf{I}_D\tilde{\mathbf{S}}, \tag{5.30}$$

gives the moment of inertia in that frame. Equation (5.21) can then be used to transform the rotation center to any desired location. The principal values of **I** can be determined by the methods of matrix algebra.

The three principal values of the moment of inertia tensor in Eq. (5.29) can be found by solving the cubic equation for **I** that arises from the determinant

$$\begin{vmatrix} I_{xx} - I & I_{xy} & I_{zx} \\ I_{xy} & I_{yy} - I & I_{yz} \\ I_{zx} & I_{yz} & I_{zz} - I \end{vmatrix} = 0, \tag{5.31}$$

where the symmetry of **I** has been displayed explicitly. Equation (5.31) is the secular equation, whose three roots are the desired principal moments. For each of these roots, Eqs. (5.28) can be solved to obtain the direction of the corresponding principal axis. In most of the easily soluble problems in rigid dynamics, the principal axes can be determined by inspection. For example, we often have to deal with rigid bodies that are solids of revolution about some axis, with the origin of the body system on the symmetry axis. All directions perpendicular to the axis of symmetry are then alike, which is the mark of a double root to the secular equation. The principal axes are then the symmetry axis and any two perpendicular axes in the plane normal to the symmetry axis.

The principal moments of inertia cannot be negative, because as the diagonal elements in the principal axes system they have the form of sums of squares. Thus,

I_{xx} is given by (cf. Eq. (5.6))

$$I_{xx} = m_i(y_i^2 + z_i^2).$$

For one of the principal moments to vanish, all points of the body must be such that two coordinates of each particle are zero. Clearly this can happen only if all points of the body are collinear with the principal axis corresponding to the zero principal moment. Any two axes perpendicular to the line of the body will then be the other principal axes. Indeed, this is clearly a limiting case of a body with an axis of symmetry passing through the origin.

We can also understand the concept of principal axes through some geometrical considerations that historically formed the first approach to the subject. The moment of inertia about a given axis has been defined as $I = \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n}$. Let the direction cosines of the axis be α, β, and γ so that

$$\mathbf{n} = \alpha\mathbf{i} + \beta\mathbf{j} + \gamma\mathbf{k};$$

I then can be written as

$$I = I_{xx}\alpha^2 + I_{yy}\beta^2 + I_{zz}\gamma^2 + 2I_{xy}\alpha\beta + 2I_{yz}\beta\gamma + 2I_{zx}\gamma\alpha, \tag{5.32}$$

using the symmetry of $\mathbf{I}$ explicitly. It is convenient to define a vector $\boldsymbol{\rho}$ by the equation

$$\boldsymbol{\rho} = \frac{\mathbf{n}}{\sqrt{I}}. \tag{5.33}$$

The magnitude of $\boldsymbol{\rho}$ is thus related to the moment of inertia about the axis whose direction is given by $\mathbf{n}$. In terms of the components of this new vector, Eq. (5.32) takes on the form

$$1 = I_{xx}\rho_1^2 + I_{yy}\rho_2^2 + I_{zz}\rho_3^2 + 2I_{xy}\rho_1\rho_2 + 2I_{yz}\rho_2\rho_3 + 2I_{zx}\rho_3\rho_1. \tag{5.34}$$

Considered as a function of the three variables ρ_1, ρ_2, ρ_3, Eq. (5.34) is the equation of some surface in ρ space. In particular, Eq. (5.34) is the equation of an ellipsoid designated as the *inertial ellipsoid*. We can always transform to a set of Cartesian axes in which the equation of an ellipsoid takes on its normal form:

$$1 = I_1\rho'^2_1 + I_2\rho'^2_2 + I_3\rho'^2_3, \tag{5.35}$$

with the principal axes of the ellipsoid along the new coordinate axes. But (5.35) is simply the form Eq. (5.34) has in a system of coordinates in which the inertia tensor $\mathbf{I}$ is diagonal. Hence, the coordinate transformation that puts the equation of ellipsoid into its normal form is exactly the principal axis transformation previously discussed. The principal moments of inertia determine the lengths of the axes of the inertia ellipsoid. If two of the roots of the secular equation are equal, the inertia ellipsoid thus has two equal axes and is an ellipsoid of revolution. If all three principal moments are equal, the inertia ellipsoid is a sphere.

A quantity closely related to the moment of inertia is the *radius of gyration*, R_0, defined by the equation

$$I = MR_0^2. \tag{5.36}$$

In terms of the radius of gyration, the vector $\boldsymbol{\rho}$ can be written as

$$\boldsymbol{\rho} = \frac{\mathbf{n}}{R_0\sqrt{M}}.$$

The radius vector to a point on the inertia ellipsoid is thus inversely proportional to the radius of gyration about the direction of the vector.

■ SOLVING RIGID BODY PROBLEMS AND THE EULER EQUATIONS OF MOTION

For bodies without a fixed point, the most useful reference point is almost always the center of mass. We have already seen that the total kinetic energy and angular momentum then split neatly into one term relating to the translational

motion of the center of mass and another involving rotation *about* the center of mass. Thus, Eq. (1.31) can now be written

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

For many problems (certainly all those that will be considered here), a similar sort of division can be made for the potential energy. We can then solve individually for the translational motion of the center of mass and for the rotational motion about the center of mass. For example, the Newtonian equations of motion can be used directly: Eq. (1.22) for the motion of the center of mass and Eq. (1.26) for the motion about that point.

With holonomic conservative systems, the Lagrangian formulation is available, with the Lagrangian taking the form

$$L(q, \dot{q}) = L_c(q_c, \dot{q}_c) + L_b(q_b, \dot{q}_b).$$

Here L_c is that part of the Lagrangian involving the generalized coordinates q_c (and velocities $\dot{q}_c$) of the center of mass, and L_b the part relating to the orientation of the body about the center of mass, as described by q_b, $\dot{q}_b$. In effect then, there are two distinct problems, one with Lagrangian L_c and the other with Lagrangian L_b.

In both the Newtonian and Lagrangian formulations, it is convenient to work in terms of the principal axes system of the point of reference, so that the kinetic energy of rotation takes the simple form given in Eq. (5.26). So far, the only suitable generalized coordinates we have for the rotational motion of the rigid body are the Euler angles. Of course, the motion is often effectively confined to two dimensions, as in the motion of a rigid lamina in a plane. The axis of rotation is then fixed in the direction perpendicular to the plane; only one angle of rotation is necessary and we may dispense with the cumbersome machinery of the Euler angles.

For the rotational motion about a fixed point or the center of mass, the direct Newtonian approach leads to a set of equations known as Euler's equations of motion. We consider either an inertial frame whose origin is at the fixed point of the rigid body, or a system of space axes with origin at the center of mass. In these two situations, Eq. (1.26) holds, which here appears simply as

$$\left(\frac{d\mathbf{L}}{dt}\right)_s = \mathbf{N}.$$

The subscript s is used because the time derivative is with respect to axes that do not share the rotation of the body. However, Eq. (4.86) can be used to obtain the derivatives with respect to axes fixed in the body:

$$\left(\frac{d\mathbf{L}}{dt}\right)_s = \left(\frac{d\mathbf{L}}{dt}\right)_b + \boldsymbol{\omega} \times \mathbf{L},$$

or, by dropping the "body" subscript:

$$\frac{d\mathbf{L}}{dt} + \boldsymbol{\omega} \times \mathbf{L} = \mathbf{N}. \tag{5.37}$$

Equation (5.37) is thus the appropriate form of the Newtonian equation of motion relative to body axes. The ith component of Eq. (5.37) can be written

$$\frac{dL_i}{dt} + \epsilon_{ijk}\omega_j L_k = N_i. \tag{5.38}$$

If now the body axes are taken as the principal axes relative to the reference point, then the angular momentum components are $L_i = I_i\omega_i$. By Eq. (5.25), Eq. (5.38) takes the form (no summation on i*)

$$I_i\frac{d\omega_i}{dt} + \epsilon_{ijk}\omega_j\omega_k I_k = N_i, \tag{5.39}$$

since the principal moments of inertia are of course time independent. In expanded form, the three equations making up Eq. (5.39) look like

$$\begin{aligned} I_1\dot{\omega}_1 - \omega_2\omega_3(I_2 - I_3) &= N_1 \\ I_2\dot{\omega}_2 - \omega_3\omega_1(I_3 - I_1) &= N_2 \\ I_3\dot{\omega}_3 - \omega_1\omega_2(I_1 - I_2) &= N_3. \end{aligned} \tag{5.39'}$$

Equations (5.39) or (5.39′) are Euler's equations of motion for a rigid body with one point fixed. They can also be derived from Lagrange's equations in the form of Eq. (1.53) where the generalized forces Q_j are the torques, N_j, corresponding to the Euler angles of rotation. However, only one of the Euler angles has its associated torque along one of the body axes, and the remaining two Euler's equations must be obtained by cyclic permutation (cf. Derivation 4).

Consider the case where $I_1 = I_2 \neq I_3$. A torque with components N_1 or N_2 will cause both ω_1 and ω_2 to change without affecting ω_3. We shall return to a discussion of this in Section 5.7 when we consider the heavy symmetric top with one point fixed. Let us first consider the torque-free motion of a rigid body.

THE HEAVY SYMMETRICAL TOP WITH ONE POINT FIXED

As a further and more complicated example of the application of the methods of rigid dynamics, let us consider the motion of a symmetrical body in a uniform gravitational field when one point on the symmetry axis is fixed in space. A wide variety of physical systems, ranging from a child's top to complicated gyroscopic navigational instruments, are approximated by such a *heavy symmetrical top*. Both for its practical applications and as an illustration of many of the tech-

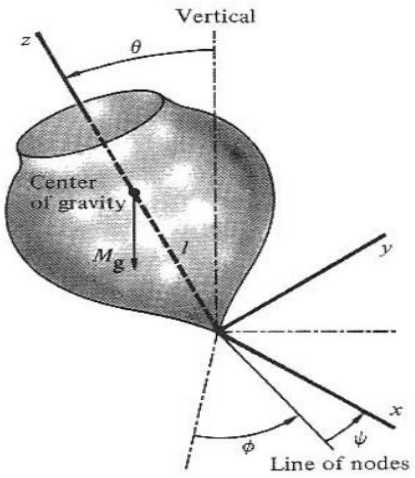


FIGURE 5.7 Euler's angles specifying the orientation of a symmetrical top.

niques previously developed, the motion of the heavy symmetrical top deserves a detailed exposition.

The symmetry axis is of course one of the principal axes and will be chosen as the z axis of the coordinate system fixed in the body.* Since one point is stationary, the configuration of the top is completely specified by the three Euler angles: θ gives the inclination of the z axis from the vertical, ϕ measures the azimuth of the top about the vertical, while ψ is the rotation angle of the top about its own z axis (cf. Fig. 5.7). The distance of the center of gravity (located on the symmetry axis) from the fixed point will be denoted by l.

The rate of change of these three angles give the characteristic motions of the top as

$\dot{\psi}$ = rotation of the top about its own figure axis, z

$\dot{\phi}$ = precession or rotation of the figure axis z about the vertical axis z'

$\dot{\theta}$ = nutation or bobbing up and down of the z figure axis relative to the vertical space axis z'.

For many cases of interest such as the top and the gyroscope, we have $\dot{\psi} \gg \dot{\theta} \gg \dot{\phi}$. Since $I_1 = I_2 \neq I_3$, Euler's equations (5.39′) become

$$I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) = N_1,$$
$$I_2\dot{\omega}_2 + \omega_1\omega_3(I_1 - I_3) = N_2,$$

and

$$I_3\dot{\omega}_3 = N_3.$$

Let us consider the case where initially $N_3 = 0 = N_2$, $N_1 \neq 0$, and $\omega_1 = \omega_2 = 0$, $\omega_3 \neq 0$, then ω_3 will be constant. The torque N_1 will cause ω_1 to change since $\dot{\omega}_1 \neq 0$. Since ω_1 is no longer zero, the second equation requires that ω_2 begin to change also. What this means in terms of an observation is not obvious. We observe the changes in the Euler angles $\dot{\psi}$, $\dot{\phi}$, $\dot{\theta}$ and their associated angles in the x', y', z' laboratory frame rather than the $\dot{\omega}_1$, $\dot{\omega}_2$, $\dot{\omega}_3$ and their associated angles in the principal axis system. This suggests that the Euler equations may not provide the most useful description of the motion.

The Lagrangian procedure, rather than Euler's equations, will be used to obtain a solution for the motion of the top. Since the body is symmetrical, the kinetic energy can be written as

$$T = \tfrac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \tfrac{1}{2}I_3\omega_3^2,$$

or, in terms of Euler's angles, and using Eqs. (4.87), as

$$T = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2, \tag{5.50}$$

where the $\dot{\phi}$, $\dot{\theta}$ cross terms in ω_1^2 and ω_2^2 cancel.

It is a well-known elementary theorem that in a constant gravitational field the potential energy is the same as if the body were concentrated at the center of mass. We will however give a brief formal proof here. The potential energy of the body is the sum over all the particles:

$$V = -m_i\mathbf{r}_i \cdot \mathbf{g},$$

where $\mathbf{g}$ is the constant vector for the acceleration of gravity. By Eq. (1.21), defining the center of mass, this is equivalent to

$$V = -M\mathbf{R} \cdot \mathbf{g}, \tag{5.51}$$

which proves the theorem. In terms of the Euler angles,

$$V = Mgl\cos\theta, \tag{5.51'}$$

so that the Lagrangian is

$$L = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 - Mgl\cos\theta. \tag{5.52}$$

We therefore have two immediate first integrals of the motion:

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = I_3\omega_3 = I_1 a \tag{5.53}$$

and

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = (I_1 \sin^2\theta + I_3\cos^2\theta)\dot{\phi} + I_3\dot{\psi}\cos\theta = I_1 b. \tag{5.54}$$

Here the two constants of the motion are expressed in terms of new constants a and b. There is one further first integral available; since the system is conservative, the total energy E is constant in time:

$$E = T + V = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}\omega_3^2 + Mgl\cos\theta. \tag{5.55}$$

Only three additional quadratures are needed to solve the problem, and they are easily obtained from these three first integrals without directly using the Lagrange equations. From Eq. (5.53), ψ is given in terms of ϕ by

$$I_3\dot{\psi} = I_1 a - I_3\dot{\phi}\cos\theta, \tag{5.56}$$

and this result can be substituted in (5.54) to eliminate ψ:

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

or

$$\dot{\phi} = \frac{b - a\cos\theta}{\sin^2\theta}. \tag{5.57}$$

Thus, if θ were known as a function of time, Eq. (5.57) could be integrated to furnish the dependence of ϕ on time. Substituting Eq. (5.57) back in Eq. (5.56) results in a corresponding expression for $\dot{\psi}$:

$$\dot{\psi} = \frac{I_1 a}{I_3} - \cos\theta \frac{b - a\cos\theta}{\sin^2\theta}. \tag{5.58}$$

which furnishes ψ if θ is known. Finally, Eqs. (5.57) and (5.58) can be used to eliminate $\dot{\phi}$ and $\dot{\psi}$ from the energy equation, resulting in a differential equation involving θ alone.

First notice that Eq. (5.53) says ω_3 is constant in time and equal to $(I_1/I_3)a$. Therefore, $E - I_3\omega_3^2/2$ is a constant of the motion, which we shall designate as E'. Making use of Eq. (5.57), the energy equation can thus be written as

$$E' = \frac{I_1\dot{\theta}^2}{2} + \frac{I_1}{2}\frac{(b - a\cos\theta)^2}{\sin^2\theta} + Mgl\cos\theta. \tag{5.59}$$

Equation (5.59) has the form of an equivalent one-dimensional problem in the variable θ, with the effective potential $V'(\theta)$ given by

$$V'(\theta) = Mgl\cos\theta + \frac{I_1}{2}\left(\frac{b - a\cos\theta}{\sin\theta}\right)^2. \tag{5.60}$$

Thus, we have four constants associated with the motion, the two angular momenta p_ψ and p_ϕ, the energy term $E - \frac{1}{2}I_3\omega_3^2$, and the potential energy term Mgl. It is common to define four normalized constants of the motion as

$$\begin{aligned} \alpha &= \frac{2E - I_3\omega_3^2}{I_1} \\ \beta &= \frac{2Mgl}{I_1} \\ a &= \frac{p_\psi}{I_1} \\ b &= \frac{p_\phi}{I_1}. \end{aligned} \tag{5.61}$$

In terms of these constants, the energy equation (5.55) can be written as

$$\alpha = \dot{\theta}^2 + \frac{(b - a\cos\theta)^2}{\sin^2\theta} + \beta\cos\theta. \tag{5.62}$$

We will use this one-dimensional problem to discuss the motion in θ, very similarly to what was done in Section 3.3 in describing the radial motion for the central force problem. It is more convenient to change variables as we did for the central force problem. Using the variable $u = \cos\theta$, rewrite Eq. (5.62) as

$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2, \tag{5.62'}$$

which can be reduced immediately to a quadrature:

$$t = \int_{u(0)}^{u(t)} \frac{du}{\sqrt{(1-u^2)(\alpha-\beta u)-(b-au)^2}}. \tag{5.63}$$

With this result, and Eqs. (5.57) and (5.58), ϕ and ψ can also be reduced to quadratures. However, the polynomial in the radical is a cubic so that we have to deal with elliptic integrals. These solutions can be generated on current desk-top computers. In the case of the force-free motion, the physics tends to be obscured in the profusion of mathematics. Fortunately, the general nature of the motion can be discovered without actually performing the integrations.

It is convenient to designate the right-hand side of Eq. (5.62′) as a function $f(u)$ and discuss the behavior of the cubic equation

$$f(u) = \beta u^3 - (\alpha + a^2)u^2 + (2ab - \beta)u + (\alpha - b^2).$$

For the gyroscope, $f(u)$ is only a quadratic equation since $\beta = 0$, while for the top the full cubic equation must be considered. Since many of the applications of the gyroscope use torque-free mountings, precession and nutations are suppressed so the gyroscope motions are trivial. To understand the general motions of a spinning body, we will consider only cases where $\beta > 0$.

The extent of the nutation under these given initial conditions is given by $u_1 - u_0$, where u_1 is the other physical root of $f(u)$. The initial conditions $E' = Mgl\cos\theta_0$ is equivalent to the equality

$$\alpha = \beta u_0.$$

With this relation, and the conditions of Eq. (5.65), $f(u)$ can be rewritten more simply as

$$f(u) = (u_0 - u)\left[\beta(1 - u^2) - a^2(u_0 - u)\right]. \tag{5.67}$$

The roots of $f(u)$ other than u_0 are given by the roots of the quadratic expression in the brackets, and the desired root u_1 therefore satisfies the equation

$$(1 - u_1^2) - \frac{a^2}{\beta}(u_0 - u_1) = 0. \tag{5.68}$$

Denoting $u_0 - u$ by x and $u_0 - u_1$ by x_1, Eq. (5.68) can be rewritten as

$$x_1^2 + px_1 - q = 0, \tag{5.69}$$

where

$$p = \frac{a^2}{\beta} - 2\cos\theta_0, \quad q = \sin^2\theta_0.$$

The condition for a "fast" top, Eq. (5.66), implies that p is much larger than q. This can be seen by writing the ratio a^2/β as

$$\frac{a^2}{\beta} = \left(\frac{I_3}{I_1}\right)\frac{I_3\omega_3^2}{2Mgl}.$$

Except in the case that $I_3 \ll I_1$ (which would correspond to a top in the unusual shape of a cigar), the ratio is much greater than unity, and $p \gg q$. To first order in the small quantity q/p, the only physically realizable root of Eq. (5.68) is then

$$x_1 = \frac{q}{p}.$$

Neglecting $2\cos\theta_0$ compared to a^2/β, this result can be written

$$x_1 = \frac{\beta\sin^2\theta_0}{a^2} = \frac{I_1}{I_3}\frac{2Mgl}{I_3\omega_3^2}\sin^2\theta_0. \tag{5.70}$$

Thus, the extent of the nutation, as measured by $x_1 = u_0 - u_1$, goes down as $1/\omega_3^2$. The faster the top is spun, the less is the nutation.

The *frequency* of nutation likewise can easily be found for the "fast" top. Since the amount of nutation is small, the term $(1-u^2)$ in Eq. (5.67) can be replaced by its initial value, $\sin^2\theta_0$. Equation (5.67) then reads, with the help of Eq. (5.70),

$$f(u) = \dot{x}^2 = a^2 x(x_1 - x).$$

If we shift the origin of x to the midpoint of its range, by changing variable to

$$y = x - \frac{x_1}{2},$$

then the differential equation becomes

$$\dot{y}^2 = a^2\left(\frac{x_1^2}{4} - y^2\right),$$

which on differentiation again reduces to the familiar equation for simple harmonic motion

$$\ddot{y} = -a^2 y.$$

In view of the initial condition $x = 0$ at $t = 0$, the complete solution is

$$x = \frac{x_1}{2}(1 - \cos at), \tag{5.71}$$

where x_1 is given by (5.70). The angular frequency of nutation of the figure axis between θ_0 and θ_1 is therefore

$$a = \frac{I_3}{I_1}\omega_3, \tag{5.72}$$

which *increases* the faster the top is spun initially.

Finally, the angular velocity of precession, from (5.57), is given by

$$\dot{\phi} = \frac{a(u_0 - u)}{\sin^2\theta} \approx \frac{ax}{\sin^2\theta_0},$$

or, substituting Eqs. (5.72) and (5.70),

$$\dot{\phi} = \frac{\beta}{2a}(1 - \cos at). \tag{5.73}$$

The rate of precession is therefore not uniform but varies harmonically with time, with the same frequency as the nutation. The *average* precession frequency however is

$$\overline{\dot{\phi}} = \frac{\beta}{2a} = \frac{Mgl}{I_3\omega_3}, \tag{5.74}$$

which indicates that the rate of precession decreases as the initial rotational velocity of the top is increased.

It is of interest to determine exactly what initial conditions will result in a tru regular precession. In such a case, the angle θ remains constant at its initial valu θ_0, which means that $\theta_1 = \theta_2 = \theta_0$. In other words, $f(u)$ must have a double roc at u_0 (cf. Fig. 5.10), or

$$f(u) = \dot{u}^2 = 0, \qquad \frac{df}{du} = 0; \qquad u = u_0.$$

The first of these conditions, from Eq. (5.62′) with $\dot{u} = 0$, implies

$$(\alpha - \beta u_0) = \frac{(b - au_0)^2}{1 - u_0^2}; \tag{5.75}$$

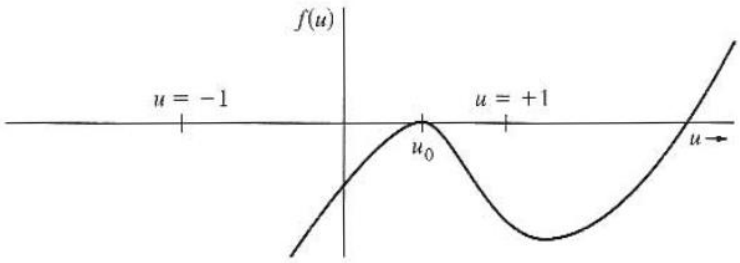


FIGURE 5.10 Appearance of $f(u)$ for a regular precession.

FORMULATION OF THE PROBLEM

We consider conservative systems in which the potential energy is a function of position only. It will be assumed that the transformation equations defining the generalized coordinates of the system, $q_1, \ldots, q_n$, do not involve the time explicitly. Thus, time-dependent constraints are to be excluded. The system is said to be in *equilibrium* when the generalized forces acting on the system vanish:

$$Q_i = -\left(\frac{\partial V}{\partial q_i}\right)_0 = 0. \tag{6.1}$$

The potential energy therefore has an extremum at the equilibrium configuration of the system, $q_{01}, q_{02}, \ldots, q_{0n}$. If the configuration is initially at the equilibrium position, with zero initial velocities $\dot{q}_n$, then the system will continue in equilibrium indefinitely. Examples of the equilibrium of mechanical systems are legion—a pendulum at rest, a suspension galvanometer at its zero position, an egg standing on end.

We shall be interested in the motion of the system within the immediate neighborhood of a configuration of stable equilibrium. Since the departures from equilibrium are too small, all functions may be expanded in a Taylor series about the equilibrium, retaining only the lowest-order terms. The deviations of the generalized coordinates from equilibrium will be denoted by η_i:

$$q_i = q_{0i} + \eta_i, \tag{6.2}$$

and these may be taken as the new generalized coordinates of the motion. Expanding the potential energy about q_{0i}, we obtain

$$V(q_1, \ldots, q_n) = V(q_{01}, \ldots, q_{0n}) + \left(\frac{\partial V}{\partial q_i}\right)_0 \eta_i + \frac{1}{2}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 \eta_i \eta_j + \cdots, \tag{6.3}$$

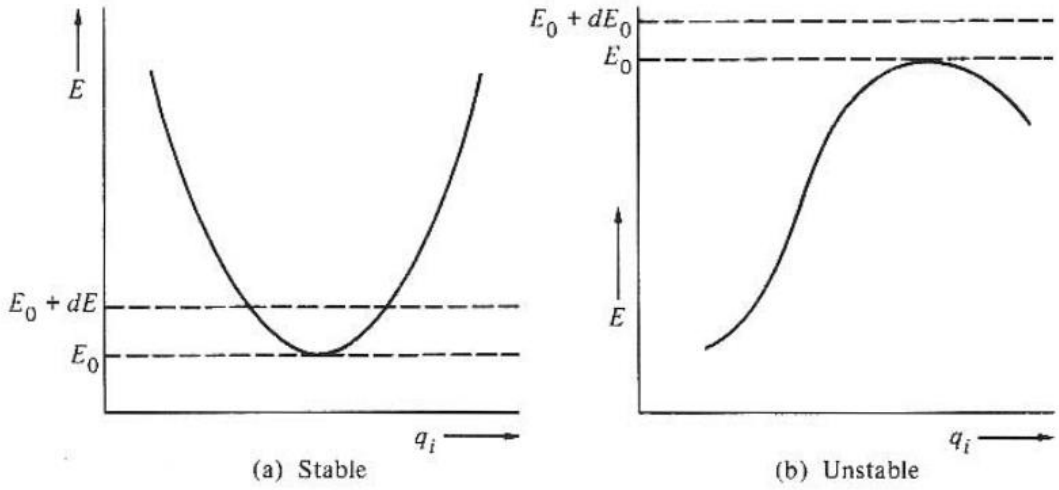


FIGURE 6.1 Shape of the potential energy curve at equilibrium.

where the summation convention has been invoked, as usual. The terms linear in η_i vanish automatically in consequence of the equilibrium conditions (6.1). The first term in the series is the potential energy of the equilibrium position, and by shifting the arbitrary zero of potential to coincide with the equilibrium potential, this term may also be made to vanish. We are therefore left with the quadratic terms as the first approximation to V:

$$V = \frac{1}{2}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 \eta_i \eta_j = \frac{1}{2} V_{ij} \eta_i \eta_j, \tag{6.4}$$

where the second derivatives of V have been designated by the constants V_{ij} depending only upon the equilibrium values of the q_i's. It is obvious from their definition that the V_{ij}'s are symmetrical, that is, that $V_{ij} = V_{ji}$. The V_{ij} coefficients can vanish under a variety of circumstances. Thus, the potential can simply be independent of a particular coordinate, so that equilibrium occurs at any arbitrary value of that coordinate. We speak of such cases as *neutral* or *indifferent* equilibrium. It may also happen, for example, that the potential behaves like a quadratic at that point, again causing one or more of the V_{ij}'s to vanish. Either situation calls for special treatment in the mathematical discussion that follows.

A similar series expansion can be obtained for the kinetic energy. Since the generalized coordinates do not involve the time explicitly, the kinetic energy is a homogeneous quadratic function of the velocities (cf. Eq. (1.71)):

$$T = \tfrac{1}{2} m_{ij} \dot{q}_i \dot{q}_j = \tfrac{1}{2} m_{ij} \dot{\eta}_i \dot{\eta}_j. \tag{6.5}$$

The coefficients m_{ij} are in general functions of the coordinates q_k, but they may be expanded in a Taylor series about the equilibrium configuration:

$$m_{ij}(q_1, \ldots, q_n) = m_{ij}(q_{01}, \ldots, q_{0n}) + \left(\frac{\partial m_{ij}}{\partial q_k}\right)_0 \eta_k + \cdots.$$

As Eq. (6.5) is already quadratic in the $\dot{\eta}_i$'s, the lowest nonvanishing approximation to T is obtained by dropping all but the first term in the expansions of m_{ij}. Denoting the constant values of the m_{ij} functions at equilibrium by T_{ij}, we can therefore write the kinetic energy as

$$T = \tfrac{1}{2} T_{ij} \dot{\eta}_i \dot{\eta}_j. \tag{6.6}$$

It is again obvious that the constants T_{ij} must be symmetric, since the individual terms in Eq. (6.6) are unaffected by an interchange of indices. From Eqs. (6.4) and (6.6), the Lagrangian is given by

$$L = \tfrac{1}{2}(T_{ij} \dot{\eta}_i \dot{\eta}_j - V_{ij} \eta_i \eta_j). \tag{6.7}$$

Taking the η's as the general coordinates, the Lagrangian of Eq. (6.7) leads to the following n equations of motion:

$$T_{ij} \ddot{\eta}_{ij} + V_{ij} \eta_j = 0, \tag{6.8}$$

where explicit use has been made of the symmetry property of the V_{ij} and T_{ij} coefficients. Each of Eqs. (6.8) will involve, in general, all of the coordinates η_i, and it is this set of simultaneous differential equations that must be solved to obtain the motion near the equilibrium.

In almost all cases of interest, the kinetic energy term can be easily written so as to have no cross terms.* This corresponds to the Lagrangian

$$L = \tfrac{1}{2}(T_i\dot{\eta}_i^2 - V_{ij}\eta_i\eta_j), \tag{6.9}$$

which generates the following equations of motion

$$T_i\ddot{\eta}_i + V_{ij}\eta_j = 0. \qquad \text{(no sum over } i\text{)} \tag{6.10}$$

FREQUENCIES OF FREE VIBRATION, AND NORMAL COORDINATES

The somewhat lengthy arguments of the preceding section demonstrate that the equations of motion will be satisfied by an oscillatory solution of the form (6.11), not merely for one frequency but in general for a set of n frequencies ω_k. A complete solution of the equations of motion therefore involves a superposition of oscillations with all the allowed frequencies. Thus, if the system is displaced slightly from equilibrium and then released, the system performs small oscillations about the equilibrium with the frequencies $\omega_1, \ldots, \omega_n$. The solutions of the secular equation are therefore often designated as the frequencies of *free vibration* or as the *resonant frequencies* of the system.

The general solution of the equations of motion may now be written as a summation over an index k:

$$\eta_i = C_k a_{ik} e^{-i\omega_k t}, \tag{6.35}$$

there being a complex scale factor C_k for each resonant frequency. It might be objected that for each solution λ_k of the secular equation there are two resonant frequencies $+\omega_k$ and $-\omega_k$. The eigenvector $\mathbf{a}_k$ would be the same for the two frequencies, but the scale factors C_k^+ and C_K^- could conceivably be different. On this basis, the general solution should appear as

$$\eta_i = a_{ik}(C_k^+ e^{+i\omega_k t} + C_k^- e^{-i\omega_k t}). \tag{6.35'}$$

Recall however that the actual motion is the real part of the complex solution, and the real part of either (6.35) or (6.35′) can be written in the form

$$\eta_i = f_k a_{ik} \cos(\omega_k t + \delta_k), \tag{6.36}$$

where the amplitude f_k and the phase δ_k are determined form the initial conditions. Either of the solutions ((6.35) and (6.36)) will therefore represent the actual motion, and the former of course is the more convenient.

The orthogonality properties of $\mathbf{A}$ greatly facilitate the determination of the scale factors C_k in terms of the initial conditions. At $t = 0$, the real part of Eq. (6.35) reduces to

$$\eta_i(0) = \text{Re}\, C_k a_{ik}, \tag{6.37}$$

where Re stands for "real part of." Similarly, the initial value of the velocities is obtained as

$$\dot{\eta}_i(0) = \text{Im}\, C_k a_{ik}\omega_k, \tag{6.38}$$

where Im C_k denotes the imaginary part of C_k. From these $2n$ equations, the real and imaginary parts of the n constants C_k may be evaluated. To solve Eq. (6.37), for example, let us first write it in terms of column matrices $\boldsymbol{\eta}(0)$ and $\mathbf{C}$:

$$\boldsymbol{\eta}(0) = \mathbf{A}\,\mathrm{Re}\,\mathbf{C}. \tag{6.37'}$$

If we multiply by $\tilde{\mathbf{A}}\mathbf{T}$ from the left and use Eq. (6.23), we immediately obtain a solution for Re $\mathbf{C}$:

$$\mathrm{Re}\,\mathbf{C} = \tilde{\mathbf{A}}\mathbf{T}\boldsymbol{\eta}(0),$$

or, taking the lth component,

$$\mathrm{Re}\,C_l = a_{jl}T_{jk}\eta_k(0). \tag{6.39}$$

A similar procedure leads to the imaginary part of the scale factors as*

$$\mathrm{Im}\,C_l = \frac{1}{\omega_l}\sum_{j,k} a_{jl}T_{jk}\dot{\eta}_k(0). \tag{6.40}$$

Equations (6.39) and (6.40) thus permit the direct computation of the complex factors C_l (and therefore the amplitudes and phases) in terms of the initial conditions and the matrices $\mathbf{T}$ and $\mathbf{A}$.

The solution for each coordinate, Eq. (6.35), is in general a sum of simple harmonic oscillations in all of the frequencies ω_k satisfying the secular equation. Unless it happens that all of the frequencies are commensurable, that is, rational fractions of each other, η_i never repeats its initial value and is therefore not itself a periodic function of time. However, it is possible to transform from the η_i to a new set of generalized coordinates that are all simple periodic functions of time—a set of variables known as the *normal coordinates*.

We define a new set of coordinates ζ_j

$$\eta_i = a_{ij}\zeta_j, \tag{6.41}$$

or, in terms of single-column matrices $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$,

$$\boldsymbol{\eta} = \mathbf{A}\boldsymbol{\zeta}. \tag{6.41'}$$

The potential energy, Eq. (6.4), is written in matrix notation as

$$V = \tfrac{1}{2}\tilde{\boldsymbol{\eta}}\mathbf{V}\boldsymbol{\eta}. \tag{6.42}$$

Now, the single-row transpose matrix $\tilde{\boldsymbol{\eta}}$ is related to $\tilde{\boldsymbol{\zeta}}$ by the equation

$$\tilde{\boldsymbol{\eta}} = \widetilde{\mathbf{A}\boldsymbol{\zeta}} = \tilde{\boldsymbol{\zeta}}\tilde{\mathbf{A}},$$

so that the potential energy can be written also as

$$V = \tfrac{1}{2}\tilde{\boldsymbol{\zeta}}\tilde{\mathbf{A}}\mathbf{V}\mathbf{A}\boldsymbol{\zeta}.$$

But **A** diagonalizes **V** by a congruence transformation (cf. Eq. (6.26)), and the potential energy therefore reduces simply to

$$V = \tfrac{1}{2}\tilde{\boldsymbol{\zeta}}\boldsymbol{\lambda}\boldsymbol{\zeta} = \tfrac{1}{2}\omega_k^2\zeta_k^2. \tag{6.43}$$

The kinetic energy has an even simpler form in the new coordinates. Since the velocities transform as the coordinates, T as given in Eq. (6.20) transforms to

$$T = \tfrac{1}{2}\dot{\tilde{\boldsymbol{\zeta}}}\tilde{\mathbf{A}}\mathbf{T}\mathbf{A}\dot{\boldsymbol{\zeta}}$$

which by virtue of Eq. (6.23) reduces to

$$T = \tfrac{1}{2}\dot{\tilde{\boldsymbol{\zeta}}}\dot{\boldsymbol{\zeta}} = \tfrac{1}{2}\dot{\zeta}_i\dot{\zeta}_i. \tag{6.44}$$

Equations (6.43) and (6.44) state that in the new coordinates both the potential and kinetic energies are sums of squares only, without any cross terms. Of course, this result is simply another way of saying that **A** produces a principal axis transformation. Recall that the principal axis transformation of the inertia tensor was specifically designed to reduce the moment of inertia to a sum of squares; the new axes being the principal axes of the inertia ellipsoid. Here the kinetic and potential energies are also quadratic forms (as was the moment of inertia) and both are diagonalized by **A**. For this reason, the principal axis transformation employed here is a particular example of the well-known algebraic process of the *simultaneous diagonalization of two quadratic forms*.

The equations of motion share in the simplification resulting from their use. The new Lagrangian is

$$L = \tfrac{1}{2}(\dot{\zeta}_k\dot{\zeta}_k - \omega_k^2\zeta_k^2) \tag{6.45}$$

so that the Lagrange equations for ζ_k are

$$\ddot{\zeta}_k + \omega_k^2\zeta_k = 0. \tag{6.46}$$

Equations (6.47) have the immediate solutions

$$\zeta_k = C_k e^{-i\omega_k t}, \tag{6.47}$$

which could have been seen of course directly from Eqs. (6.35) and (6.41). Each of the new coordinates is thus a simply periodic function involving only *one* of the resonant frequencies. As mentioned earlier, it is therefore customary to call the ζ's the *normal coordinates* of the system.

Each normal coordinate corresponds to a vibration of the system with only one frequency, and these component oscillations are spoken of as the *normal modes of vibration*. All of the particles in each mode vibrate with the same frequency and with the same phase;* the relative amplitudes being determined by the matrix

elements a_{jk}. The complete motion is then built up out of the sum of the normal modes weighted with appropriate amplitude and phase factors contained in the C_k's.

Harmonics of the fundamental frequencies are absent in the complete motion essentially because of the stipulation that the amplitude of oscillation be small. We are then allowed to represent the potential as a quadratic form, which is characteristic of simple harmonic motion. The normal coordinate transformation emphasizes this point, for the Lagrangian in the normal coordinates (6.45) is seen to be the sum of the Lagrangians for harmonic oscillators of frequencies ω_k. We can thus consider the complete motion for small oscillations as being obtained by exciting the various harmonic oscillators with different intensities and phases.*

■ FREE VIBRATIONS OF A LINEAR TRIATOMIC MOLECULE

To illustrate the technique for obtaining the resonant frequencies and normal modes, we shall consider in detail a model based on a linear symmetrical triatomic molecule. In the equilibrium configuration of the molecule, two atoms of mass m are symmetrically located on each side of an atom of mass M (cf. Fig. 6.3). All three atoms are on one straight line, the equilibrium distances apart being denoted by b. For simplicity, we shall first consider only vibrations along the line of the molecule, and the actual complicated interatomic potential will be approximated by two springs of force constant k joining the three atoms. There are three obvious coordinates marking the position of the three atoms on the line. In these coordinates, the potential energy is

$$V = \frac{k}{2}(x_2 - x_1 - b)^2 + \frac{k}{2}(x_3 - x_2 - b)^2. \tag{6.48}$$

We now introduce coordinates relative to the equilibrium positions:

$$\eta_i = x_i - x_{0i},$$

where

$$x_{02} - x_{01} = b = x_{03} - x_{02}.$$

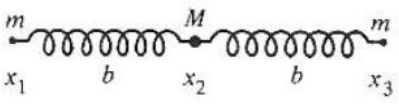


FIGURE 6.3 Model of a linear symmetrical triatomic molecule.

The potential energy then reduces to

$$V = \frac{k}{2}(\eta_2 - \eta_1)^2 + \frac{k}{2}(\eta_3 - \eta_2)^2,$$

or

$$V = \frac{k}{2}(\eta_1^2 + 2\eta_2^2 + \eta_3^2 - 2\eta_1\eta_2 - 2\eta_2\eta_3). \tag{6.49}$$

Hence, the **V** tensor has the form

$$\mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}. \tag{6.50}$$

The kinetic energy has an even simpler form:

$$T = \frac{m}{2}(\dot{\eta}_1^2 + \dot{\eta}_3^2) + \frac{M}{2}\dot{\eta}_2^2, \tag{6.51}$$

so that the **T** tensor is diagonal:

$$\mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}. \tag{6.52}$$

Combining these two tensors, the secular equation appears as

$$|\mathbf{V} - \omega^2\mathbf{T}| = \begin{vmatrix} k - \omega^2 m & -k & 0 \\ -k & 2k - \omega^2 M & -k \\ 0 & -k & k - \omega^2 m \end{vmatrix} = 0. \tag{6.53}$$

Direct evaluation of the determinant leads to the cubic equation in ω^2:

$$\omega^2(k - \omega^2 m)(k(M + 2m) - \omega^2 Mm) = 0, \tag{6.54}$$

with the obvious solutions

$$\omega_1 = 0, \quad \omega_2 = \sqrt{\frac{k}{m}}, \quad \omega^3 = \sqrt{\frac{k}{m}\left(1 + \frac{2m}{M}\right)}. \tag{6.55}$$

The first eigenvalue, $\omega_1 = 0$, may appear somewhat surprising and even alarming at first sight. Such a solution does not correspond to an oscillatory motion at all, for the equation of motion for the corresponding normal coordinate is

$$\ddot{\zeta}_1 = 0,$$

which produces a uniform translational motion. But this is precisely the key to the difficulty. The vanishing frequency arises from the fact that the molecule

Since the zero frequency found here is of no consequence for the vibration frequencies of interest, it is often desirable to phrase the problem so that the root is eliminated from the outset. We can do this here most simply by imposing the condition or constraint that the center of mass remain stationary at the origin:

$$m(x_1 + x_3) + Mx_2 = 0. \tag{6.56}$$

The components a_{ij} are determined for each frequency by the equations

$$\begin{aligned} (k - \omega_j^2 m)a_{1j} \qquad -ka_{2j} \qquad\qquad &= 0 \\ -ka_{1j} + (2k - \omega_j^2 M)a_{2j} \qquad -ka_{3j} \; &= 0 \\ -ka_{2j} + (k - \omega_j^2 m)a_{3j} \; &= 0, \end{aligned} \tag{6.57a}$$

along with the normalization condition:

$$m(a_{1j}^2 + a_{3j}^2) + Ma_{2j}^2 = 1. \tag{6.57b}$$

For $\omega_1 = 0$, it follows immediately from the first and third of Eqs. (6.57a) that all three coefficients are equal: $a_{11} = a_{21} = a_{31}$. This of course is exactly what was expected form the translational nature of the motion (cf. Fig. 6.4a). The normalization condition then fixes the value of a_{1j} so that

$$a_{11} = \frac{1}{\sqrt{2m + M}}, \quad a_{12} = \frac{1}{\sqrt{2m + M}}, \quad a_{13} = \frac{1}{\sqrt{2m + M}}. \tag{6.58a}$$

The factors $(k - \omega_2^2 m)$ vanish for the second mode, and Eqs. (6.57a) show immediately that $a_{22} = 0$ (as predicted) and $a_{12} = -a_{32}$. The numerical value of these quantities is then determined by Eq. (6.57b):

$$a_{12} = \frac{1}{\sqrt{2m}}, \quad a_{22} = 0, \quad a_{32} = -\frac{1}{\sqrt{2m}}. \tag{6.58b}$$

In this mode the center atom is at rest, while the two outer ones vibrate exactly out of phase (as they must in order to conserve linear momentum) (cf. Fig. 6.4b). Finally, when $\omega = \omega_3$, it can be seen from the first and third of Eqs. (6.57a) that a_{13} and a_{33} must be equal. The rest of the calculation for this mode is not quite as simple as for the others, and it will be sufficient to state the final result:

$$a_{13} = \frac{1}{\sqrt{2m\left(1 + \frac{2m}{M}\right)}}, \quad a_{23} = \frac{-2}{\sqrt{2M\left(2 + \frac{M}{m}\right)}}, \quad a_{33} = \frac{1}{\sqrt{2m\left(1 + \frac{2m}{M}\right)}}. \tag{6.58c}$$

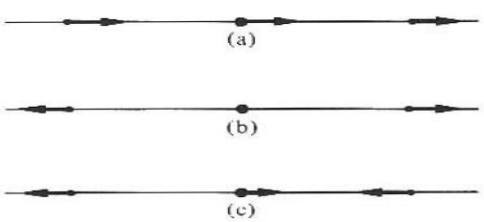


FIGURE 6.4 Longitudinal normal modes of the linear symmetric triatomic molecule.

Here the two outer atoms vibrate with the same amplitude, while the inner one oscillates out of phase with them and has a different amplitude, (cf. Fig. 6.4c.)

The normal coordinates may be found by inverting Eq. (6.41) as

$$\zeta_1 = \frac{1}{\sqrt{2m+M}}(\sqrt{m}\eta_1 + \sqrt{M}\eta_2 + \sqrt{m}\eta_3)$$

$$\zeta_2 = \sqrt{\frac{1}{2}}(\eta_1 - \eta_3) \tag{6.59}$$

$$\zeta_3 = \frac{1}{\sqrt{2m+M}}\left[\sqrt{\frac{M}{2}}(\eta_1 + \eta_3) - \sqrt{2m}\eta_2\right].$$

These normal modes describe each of the behaviors shown on Fig. 6.4. Any general longitudinal vibration of the molecule that does not involve a rigid translation will be some linear combination of the normal modes ω_2 and ω_3. The amplitudes of the normal modes, and their phases relative to each other, will of course be determined by the initial conditions (cf. Exercise 5).

normal to the molecular axis will depend upon the amplitudes and relative phases of the two degenerate modes. If both are excited, and they are exactly in phase, then the atoms will move on a straight line passing through the equilibrium configuration. But if they are out of phase, the composite motion is an elliptical Lissajous figure, exactly as in a two-dimensional isotropic oscillator. The two modes then represent a rotation, rather than a vibration.

It is obvious from the symmetry of the molecules that the amplitudes of the end atoms must be identical in magnitude. The complete calculation shows that the end atoms also travel in the same direction along the Lissajous figure. Hence, the center atom must revolve in the opposite direction, in order to conserve angular momentum. Figure 6.5 illustrates the motion for the two degenerate modes when they are 90° out of phase.

****

UNIT – IV Hamiltonian Dynamics
Legendre Transformations and the Hamilton Equations of Motion - Cyclic Coordinates and Conservation Theorems – Derivation of Hamilton's Equation from Variational principle – Principle of Least Action - Equations of Canonical Transformation - Examples of Canonical Transformations - The Harmonic Oscillator - The Symplectic Approach to Canonical Transformations - Poisson Brackets and Other Canonical Invariants - The Angular Momentum Poisson Bracket Relations.

LEGENDRE TRANSFORMATIONS AND THE HAMILTON EQUATIONS OF MOTION

In the Lagrangian formulation (nonrelativistic), a system with n degrees of freedom possesses n equations of motion of the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{8.1}$$

As the equations are of second order, the motion of the system is determined for all time only when $2n$ initial values are specified, for example, the n q_i's and n $\dot{q}_i$'s at a particular time t_1, or then n q_i's at *two* times, t_1 and t_2. We represent the state of the system by a point in an n-dimensional *configuration space* whose coordinates are the n generalized coordinates q_i and follow the motion of the system point in time as it traverses its path in configuration space. Physically, in the Lagrangian viewpoint a system with n independent degrees of freedom is a problem in n independent variables $q_i(t)$, and $\dot{q}_i$ appears only as a shorthand for the time derivative of q_i. All n coordinates must be independent. In the Hamiltonian formulation there can be no constraint equations among the coordinates. If the n coordinates are not independent, a reduced set of m coordinates, with $m < n$, must be used for the formulation of the problem before proceeding with the following steps.

The Hamiltonian formulation is based on a fundamentally different picture. We seek to describe the motion in terms of *first-order* equations of motion. Since the number of initial conditions determining the motion must of course still be $2n$, there must be $2n$ independent first-order equations expressed in terms of $2n$ *independent variables*. Hence, the $2n$ equations of the motion describe the behavior of the system point in a *phase space* whose coordinates are the $2n$ independent variables. In thus doubling our set of independent quantities, it is natural (though not inevitable) to choose half of them to be the n generalized coordinates q_i. As we shall see, the formulation is nearly symmetric if we choose the other half of the set to be the generalized or *conjugate momenta* p_i already introduced by the definition (cf. Eq. (2.44)):

$$p_i = \frac{\partial L(q_j, \dot{q}_j, t)}{\partial \dot{q}_i} \qquad \text{(no sum on } j\text{)} \tag{8.2}$$

where the j index shows the set of q's and $\dot{q}$'s. The quantities (q, p) are known as the *canonical variables.**

From the mathematical viewpoint, it can however be claimed that the q's and $\dot{q}$'s have been treated as distinct variables. In Lagrange's equations, Eq. (8.1), the partial derivative of L with respect to q_i means a derivative taken with all other q's and all $\dot{q}$'s constant. Similarly, in the partial derivatives with respect to $\dot{q}$, the q's are kept constant. Treated strictly as a mathematical problem, the transition from Lagrangian to Hamiltonian formulation corresponds to changing the variables in our mechanical functions from $(q, \dot{q}, t)$ to (q, p, t), where p is related to q and $\dot{q}$ by Eqs. (8.2). The procedure for switching variables in this manner is provided by the *Legendre transformation*, which is tailored for just this type of change of variable.

Consider a function of only two variables $f(x, y)$, so that a differential of f has the form

$$df = u\,dx + v\,dy, \tag{8.3}$$

where

$$u = \frac{\partial f}{\partial x}, \qquad v = \frac{\partial f}{\partial y}. \tag{8.4}$$

We wish now to change the basis of description from x, y to a new distinct set of variables u, y, so that differential quantities are expressed in terms of the differentials du and dy. Let g be a function of u and y defined by the equation

$$g = f - ux. \tag{8.5}$$

A differential of g is then given as

$$dg = df - u\,dx - x\,du,$$

or, by (8.3), as

$$dg = v\,dy - x\,du,$$

which is exactly in the form desired. The quantities x and v are now functions of the variables u and y given by the relations

$$x = -\frac{\partial g}{\partial u}, \qquad v = \frac{\partial g}{\partial y}, \tag{8.6}$$

which are the analogues of Eqs. (8.4).

The Legendre transformation so defined is used frequently in thermodynamics. The first law of thermodynamics relates the differential change in energy, dU, to the corresponding change in heat content, dQ, and the work done, dW:

$$dU = dQ - dW. \tag{8.7}$$

For a gas undergoing a reversible process, Eq. (8.7) can be written as

$$dU = T\,dS - P\,dV, \tag{8.8}$$

where $U(S, V)$ is written as a function of the entropy, S, and the volume, V, where the temperature, T, and the gas pressure, P, are given by

$$T = \frac{\partial U}{\partial S} \qquad \text{and} \qquad P = -\frac{\partial U}{\partial V}. \tag{8.9}$$

The enthalpy, $H(S, P)$ is generated by the Legendre transformation

$$H = U + PV, \tag{8.10}$$

which gives

$$dH = T\,dS + V\,dP, \tag{8.11}$$

where

$$T = \frac{\partial H}{\partial S} \qquad \text{and} \qquad V = \frac{\partial H}{\partial P}.$$

Additional Legendre transformations,

$$\begin{aligned} F &= U - TS \\ G &= H - TS, \end{aligned} \tag{8.12}$$

generate the Helmholtz free energy, $F(T, V)$, and the Gibbs free energy, $G(T, P)$.

The transformation from $(q, \dot{q}, t)$ to (q, p, t) differs from the type considered in Eqs. (8.3) to (8.12) only in that more than one variable is to be transformed. We begin by writing the differential of the Lagrangian, $L(q, \dot{q}, t)$, as

$$dL = \frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt. \tag{8.13}$$

The canonical momentum was defined in Eq. (2.44) as $p_i = \partial L / \partial \dot{q}_i$; substituting this into the Lagrange equation (8.1), we obtain

$$\dot{p}_i = \frac{\partial L}{\partial q_i}, \tag{8.14}$$

so Eq. (8.13) can be written as

$$dL = \dot{p}_i \, dq_i + p_i \, d\dot{q}_i + \frac{\partial L}{\partial t} dt. \tag{8.13'}$$

The Hamiltonian $H(q, p, t)$ is generated by the Legendre transformation

$$H(q, p, t) = \dot{q}_i p_i - L(q, \dot{q}, t), \tag{8.15}$$

which has the differential

$$dH = \dot{q}_i \, dp_i - \dot{p}_i \, dq_i - \frac{\partial L}{\partial t}, \tag{8.16}$$

where the term $p_i \, d\dot{q}_i$ is removed by the Legendre transformation. Since dH can also be written as

$$dH = \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt, \tag{8.17}$$

we obtain the $2n + 1$ relations

$$\left.\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ -\dot{p}_i &= \frac{\partial H}{\partial q_i} \end{aligned}\right\} \tag{8.18}$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}. \tag{8.19}$$

Equations (8.18) are known as the *canonical equations of Hamilton*; they constitute the desired set of $2n$ first-order equations of motion replacing the n second-order Lagrange equations.*

The first half of Hamilton's equations give the $\dot{q}_i$'s as functions of (q, p, t). They form therefore the inverse of the constitutive equations (8.2), which define the momenta p_i as functions of $(q, \dot{q}, t)$. It may therefore be said that they provide no new information. In terms of solving mechanical problems by means of the canonical equations, the statement is correct. But within the framework of the Hamiltonian picture, where $H(q, p, t)$ is some given function obtained no matter how, the two halves of the set of Hamiltonian equations are equally independent and meaningful. The first half says how $\dot{q}$ depends on q, p, and t; the second says the same thing for $\dot{p}$.

Of course, the Hamiltonian H is constructed in the same manner, and has identically the same value, as h, the energy function defined in Eq. (2.53). But they are functions of different variables: Like the Lagrangian, h is a function of q, $\dot{q}$ (and possibly t), while H must always be expressed as a function of q, p (and possibly t). It is to emphasize this difference in functional behavior that different symbols have been given to the quantities even though they have the same numerical values.

Nominally, the Hamiltonian for each problem must be constructed via the Lagrangian formulation. The formal procedure calls for a lengthy sequence of steps:

1. With a chosen set of generalized coordinates, q_i, the Lagrangian $L(q_i, \dot{q}_i, t) = T - V$ is constructed.
2. The conjugate momenta are defined as functions of q_i, $\dot{q}_i$, and t by Eqs. (8.2).
3. Equation (8.15) is used to form the Hamiltonian. At this stage we have some mixed function of q_i, $\dot{q}_i$, p_i, and t.
4. Equations (8.2) are then inverted to obtain $\dot{q}_i$ as functions of (q, p, t). Possible difficulties in the inversion will be discussed below.
5. The results of the previous step are then applied to eliminate $\dot{q}$ from H so as to express it solely as a function of (q, p, t).

Now we are ready to use the Hamiltonian in the canonical equations of motion.

ities of degree 0, 1, and 2, respectively. In that case, H by the prescription of Eq. (8.15) is given by (cf. Eqs. (2.53) and (2.55))

$$H = \dot{q}_i p_i - L = \dot{q}_i p_i - [L_0(q_i, t) + L_1(q_i, t)\dot{q}_k + L_2(q_i, t)\dot{q}_k\dot{q}_m] \quad (8.20)$$

(no sum on i in the square brackets) where L_0 is the part of the Lagrangian that is independent of the generalized velocities, L_1 represents the coefficients of the part of the Lagrangian that is homogeneous in $\dot{q}_i$ in the first degree, and L_2 is the part that is homogeneous in $\dot{q}_i$ in the second degree. Further, if the equations defining the generalized coordinates don't depend on time explicitly, then $L_2\dot{q}_k\dot{q}_m = T$ (the kinetic energy), and if the forces are derivable from a conservative potential V (that is, work is independent of the path), then $L_0 = -V$. When both these conditions are satisfied, the Hamiltonian is *automatically* the total energy:

$$H = T + V = E. \quad (8.21)$$

If either Eq. (8.20) or (8.21) holds, then much of the algebra in steps 3 and 4 above is eliminated.

We can at times go further. In large classes of problems, it happens that L_2 is a quadratic function of the generalized velocities and L_1 is a linear function of the same variables with the following specific functional dependencies:

$$L(q_i, \dot{q}_i, t) = L_0(q, t) + \dot{q}_i a_i(q, t) + \dot{q}_i^2 T_i(q, t), \quad (8.22)$$

where the a_i's and the T_i's are functions of the q's and t.

The algebraic manipulations required in steps 2–5 can then be carried out, at least formally, once and for all. To show this, let us form the $\dot{q}_i$'s into a single column matrix $\dot{\mathbf{q}}$. Under the given assumptions the Lagrangian can be written as

$$L(q, \dot{q}, t) = L_0(q, t) + \tilde{\dot{\mathbf{q}}}\mathbf{a} + \tfrac{1}{2}\tilde{\dot{\mathbf{q}}}\mathbf{T}\dot{\mathbf{q}}, \quad (8.23)$$

where the single row matrix $\tilde{\dot{\mathbf{q}}}$ has been written explicitly as the transpose of a single column matrix, $\dot{\mathbf{q}}$. Here $\mathbf{a}$ is a column matrix, and $\mathbf{T}$ is a square $n \times n$ matrix (much like the corresponding matrix introduced in Section 6.2). The elements of both are in general functions of q and t. To illustrate this formalism, let us consider the special case where $q_i = \{x, y, z\}$ and $\mathbf{T}$ is diagonal. We would then write

$$\frac{1}{2}\tilde{\dot{\mathbf{q}}}\mathbf{T}\dot{\mathbf{q}} = \frac{1}{2}(\dot{x}\dot{y}\dot{z})\begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix}\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad (8.24a)$$

and

$$\tilde{\dot{\mathbf{q}}}\mathbf{a} = (\dot{x}\dot{y}\dot{z})\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = a_x\dot{x} + a_y\dot{y} + a_z\dot{z} = \mathbf{a}\cdot\dot{\mathbf{r}}. \quad (8.24b)$$

In this notation the Hamiltonian, $H = \tilde{\dot{\mathbf{q}}}\mathbf{p} - L$, becomes

$$H = \tilde{\dot{\mathbf{q}}}(\mathbf{p} - \mathbf{a}) - \tfrac{1}{2}\tilde{\dot{\mathbf{q}}}\mathbf{T}\dot{\mathbf{q}} - L_0. \tag{8.24c}$$

The conjugate momenta, considered as a column matrix $\mathbf{p}$, is then, by Eq. (8.2), given as

$$\mathbf{p} = \mathbf{T}\dot{\mathbf{q}} + \mathbf{a}, \tag{8.25}$$

which can be inverted (step 4) to the column vector $\dot{\mathbf{q}}$

$$\dot{\mathbf{q}} = \mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}). \tag{8.26a}$$

This step presupposes that $\mathbf{T}^{-1}$ exists, which it normally does by virtue of the positive definite property of kinetic energy.

The corresponding equation for $\tilde{\dot{\mathbf{q}}}$ is

$$\tilde{\dot{\mathbf{q}}} = (\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}. \tag{8.26b}$$

To obtain the correct functional form for the Hamiltonian, Eqs. (8.26) must be used to replace $\dot{\mathbf{q}}$ and $\tilde{\dot{\mathbf{q}}}$, yielding the final form for the Hamiltonian:

$$H(q, p, t) = \tfrac{1}{2}(\tilde{\mathbf{p}} - \tilde{\mathbf{a}})\mathbf{T}^{-1}(\mathbf{p} - \mathbf{a}) - L_0(q, t). \tag{8.27}$$

If the Lagrangian can be written in the form of Eq. (8.23), then we can immediately skip the intervening steps and write the Hamiltonian as Eq. (8.27). The inverse matrix $\mathbf{T}^{-1}$ can usually most easily be obtained straightforwardly as

$$\mathbf{T}^{-1} = \frac{\tilde{\mathbf{T}}_c}{|\mathbf{T}|}, \tag{8.28}$$

where $\mathbf{T}_c$ is the cofactor matrix whose elements $(\mathbf{T}_c)_{jk}$ are $(-1)^{j+k}$ times the determinant of the matrix obtained by striking out the jth row and the kth column of $\mathbf{T}$.

In the example Eq. (8.24a), these three matrices are given explicitly by

$$\mathbf{T} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix}, \qquad \mathbf{T}^{-1} = \begin{bmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{m} & 0 \\ 0 & 0 & \frac{1}{m} \end{bmatrix}, \qquad \text{and}$$

$$\tilde{\mathbf{T}}_c = \begin{bmatrix} m^2 & 0 & 0 \\ 0 & m^2 & 0 \\ 0 & 0 & m^2 \end{bmatrix},$$

and the determinant $|\mathbf{T}| = m^3$. It is easy to see that for the usual case when $\mathbf{T}$ is diagonal, then $\mathbf{T}^{-1}$ is also diagonal with elements that are just the reciprocals of the corresponding elements of $\mathbf{T}$.

■ CYCLIC COORDINATES AND CONSERVATION THEOREMS

According to the definition given in Section 2.6, a cyclic coordinate q_j is one that does not appear explicitly in the Lagrangian; by virtue of Lagrange's equations

its conjugate momentum p_j is then a constant. But comparison of Eq. (8.14) with Eq. (8.16) has already told us that

$$\dot{p}_j = \frac{\partial L}{\partial q_j} = -\frac{\partial H}{\partial q_j}.$$

The considerations concerning h in Section 2.7 have already shown that if L (and in consequence of Eq. (8.15), also H) is not an explicit function of t, then H is a constant of motion. This can also be seen directly from the equations of motion (8.18) by writing the total time derivative of the Hamiltonian as

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t}.$$

In consequence of the equations of motion (8.18), the first two sums on the right cancel each other, and it therefore follows that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{8.41}$$

Thus if t doesn't appear explicitly in L, it will also not be present in H, and H will be constant in time.

Further, it was proved in Section 2.7 that if the equations of transformation that define the generalized coordinates (1.38),

$$\mathbf{r}_m = \mathbf{r}_m(q_1, \ldots, q_n; t),$$

do not depend explicitly upon the time, and if the potential is velocity independent, then H is the total energy, $T+V$. The identification of H as a constant of the motion and as the total energy are two separate matters, and the conditions sufficient for the one are not enough for the other. It can happen that the Eqs. (1.38) do involve time explicitly but that H does not. In this case, H is a constant of the motion but it is *not* the total energy. As was also emphasized in Section (2.6), the Hamiltonian is dependent both in magnitude and in functional form upon the initial choice of generalized coordinates. For the Lagrangian, we have a specific prescription, $L = T - V$, and a change of generalized coordinates within that prescription may change the functional appearance of L but cannot alter its magnitude. On the other hand, use of a different set of generalized coordinates in the definition for the Hamiltonian, Eq. (8.15), may lead to an entirely different quantity for the Hamiltonian. It may be that for one set of generalized coordinates H is conserved, but that for another it varies in time.

To illustrate some of these points in a simple example, we may consider a somewhat artificial one-dimensional system. Suppose a point mass m is attached to a spring, of force constant k, the other end of which is fixed on a massless cart that is being moved uniformly by an external device with speed v_0 (cf. Fig. 8.1). If we take as generalized coordinate the position x of the mass particle in the stationary system, then the Lagrangian of the system is obviously

$$L(x, \dot{x}, t) = T - V = \frac{m\dot{x}^2}{2} - \frac{k}{2}(x - v_0 t)^2. \tag{8.42}$$

(For simplicity, the origin has been chosen so that the cart passes through it at $t = 0$.) The corresponding equation of motion is clearly

$$m\ddot{x} = -k(x - v_0 t).$$

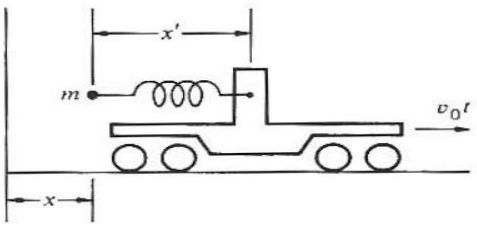


FIGURE 8.1 A harmonic oscillator fixed to a uniformly moving cart.

An obvious way of solving this equation is to change the unknown to $x'(t)$ defined as

$$x' = x - v_0 t, \tag{8.43}$$

and noting that $\ddot{x}' = \ddot{x}$, the equation of motion becomes

$$m\ddot{x}' = -kx'. \tag{8.44}$$

From Eq. (8.43), x' is the displacement of the particle relative to the cart; Eq. (8.44) says that to an observer on the cart the particle exhibits simple harmonic motion, as would be expected on the principle of equivalence in Galilean relativity.

Having looked at the nature of the motion, let us consider the Hamiltonian formulation. Since x is the Cartesian coordinate of the particle, and the potential does not involve generalized velocities, the Hamiltonian relative to x is the sum of the kinetic and potential energies, that is, the total energy. In functional form the Hamiltonian is given by

$$H(x, p, t) = T + V = \frac{p^2}{2m} + \frac{k}{2}(x - v_0 t)^2. \tag{8.45}$$

The Hamiltonian *is* the total energy of the system, but since it is explicitly a function of t, it is *not* conserved. Physically this is understandable; energy must flow into and out of the "external physical device" to keep the cart moving uniformly against the reaction of the oscillating particle.*

Suppose now we formulated the Lagrangian from the start in terms of the relative coordinate x'. The same prescription gives the Lagrangian as

$$L(x', \dot{x}') = \frac{m\dot{x}'^2}{2} + m\dot{x}'v_0 + \frac{mv_0^2}{2} - \frac{kx'^2}{2}. \tag{8.46}$$

In setting up the corresponding Hamiltonian, we note there is now a term linear in $\dot{x}'$, with the single component of **a** being mv_0. The new Hamiltonian is now

$$H'(x', p') = \frac{(p' - mv_0)^2}{2m} + \frac{kx'^2}{2} - \frac{mv_0^2}{2}. \tag{8.47}$$

Note that the last term is a constant involving neither x' nor p'; it could, if we wished, be dropped from H' without affecting the resultant equations of motion. Now H' is *not* the total energy of the system, but it *is* conserved. Except for the last term, it can be easily identified as the total energy of motion of the particle relative to the moving cart. The two Hamiltonian's are different in magnitude,

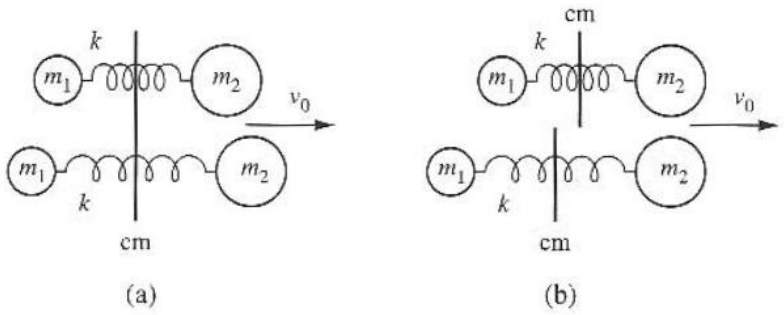


FIGURE 8.2 Vibrating dumbbell under two conditions: (a) freely oscillating, and (b) oscillating with mass m_2 kept at a constant velocity.

time dependence, and functional behavior. But the reader can easily verify that both lead to the same motion for the particle.

Additional insight into the problem of the mass cart previously discussed can be gained by considering a dumbbell of two masses connected by a spring of constant k. We shall consider the case where the center of mass of the dumbbell is in constant motion at a speed v_0 along the direction determined by the spring and allow oscillations of the masses only along this direction. This is shown in Fig. 8.2, where C-O-M denotes the center of mass.

The dumbbell is made to vibrate while its center of mass has an initial velocity v_0. It will continue with this velocity with uniform translational motion. This translational motion will have no effect on the oscillations. The motion of the center of mass and the motion relative to the center of mass separate as they do in the Kepler problem. Once the motion is started, energy is conserved and the Hamiltonian is the total conserved energy. The situation is different if the mass m_2 moves at the constant speed v_0 since a periodic force is applied. The center of mass and the mass m_1 then oscillate relative to m_2. Since a changing external force must be applied to the system to keep m_2 at the constant velocity v_0, the Hamiltonian is no longer conserved, nor is the Hamiltonian the total energy.

■ DERIVATION OF HAMILTON'S EQUATIONS FROM A VARIATIONAL PRINCIPLE

Lagrange's equations have been shown to be the consequence of a variational principle, namely, the Hamilton's principle of Section 2.1. Indeed, the variational method is often the preferable one for deriving Lagrange's equations, for it is applicable to types of systems not usually included within the scope of mechanics. It would be similarly advantageous if a variational principle could be found that

leads directly to the Hamilton's equations of motion. Hamilton's principle,

$$\delta I \equiv \delta \int_{t_1}^{t_2} L\, dt = 0, \tag{8.64}$$

$$\delta I = \delta \int_{t_1}^{t_2} \left(p_i \dot{q}_i - H(q, p, t)\right) dt = 0. \tag{8.65}$$

As a variational principle in phase space, Eq. (8.65) is sometimes referred to as the *modified Hamilton's principle*. Although it will be used most frequently in connection with transformation theory (see Chapter 9), the main interest in it here is to show that the principle leads to Hamilton's canonical equations of motion.

The modified Hamilton's principle is exactly of the form of the variational problem in a space of $2n$ dimensions considered in Section 2.3 (cf. Eq. (2.14)):

$$\delta I = \delta \int_{t_1}^{t_2} f(q, \dot{q}, p, \dot{p}, t)\, dt = 0, \tag{8.66}$$

for which the $2n$ Euler–Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}_j}\right) - \frac{\partial f}{\partial q_j} = 0 \qquad j = 1, \ldots, n \tag{8.67}$$

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{p}_j}\right) - \frac{\partial f}{\partial p_j} = 0 \qquad j = 1, \ldots, n. \tag{8.68}$$

The integrand f as given in Eq. (8.65) contains $\dot{q}_j$ only through the $p_i \dot{q}_i$ term, and q_j only in H. Hence, Eqs. (8.67) lead to

$$\dot{p}_j + \frac{\partial H}{\partial q_j} = 0. \tag{8.69}$$

On the other hand, there is no explicit dependence of the integrand in Eq. (8.65) on $\dot{p}_j$. Equations (8.68) therefore reduce simply to

$$\dot{q}_j - \frac{\partial H}{\partial p_j} = 0. \tag{8.70}$$

Equations (8.69) and (8.70) are exactly Hamilton's equations of motion, Eqs. (8.18). The Euler–Lagrange equations of the modified Hamilton's principle are thus the desired canonical equations of motion.

■ THE PRINCIPLE OF LEAST ACTION

Another variational principle associated with the Hamiltonian formulation is known as *the principle of least action*. It involves a new type of variation, which we shall call the Δ-variation, requiring detailed explanation. In the δ-variation process used in the discussion of Hamilton's principle in Chapter 2, the varied path in configuration space always terminated at end points representing the system configuration at the same time t_1 and t_2 as the correct path. To obtain Lagrange's equations of motion, we also required that the varied path return to the same end points in configuration space, that is, $\delta q_i(t_1) = \delta q_i(t_2) = 0$. The Δ-variation is less constrained; in general, the varied path over which an integral is evaluated may end at different times than the correct path, and there

may be a variation in the coordinates at the end points. We can however use the same parameterization of the varied path as in the δ-variation. In the notation of Section 2.3, a family of possible varied paths is defined by functions (cf. Eq. (2.15))

$$q_i(t, \alpha) = q_i(t, 0) + \alpha \eta_i(t), \tag{8.72}$$

where α is an infinitesimal parameter that goes to zero for the correct path. Here the functions η_i do not necessarily have to vanish at the end points, either the original or the varied. All that is required is that they be continuous and differentiable. Figure 8.3 illustrates the correct and varied path for a Δ-variation in configuration space.

Let us evaluate the Δ-variation of the action integral:

$$\Delta \int_{t_1}^{t_2} L\, dt \equiv \int_{t_1+\Delta t_1}^{t_2+\Delta t_2} L(\alpha)\, dt - \int_{t_1}^{t_2} L(0)\, dt, \tag{8.73}$$

where $L(\alpha)$ means the integral is evaluated along the varied path and $L(0)$ correspondingly refers to the actual path of motion. The variation is clearly composed of two parts. One arises from the change in the limits of the integral; to first-order infinitesimals, this part is simply the integrand on the actual path times the difference in the limits in time. The second part is caused by the change in the integrand on the varied path, but now between the same time limits as the original integral. We may therefore write the Δ-variation of the action integral as

$$\Delta \int_{t_1}^{t_2} L\, dt = L(t_2)\Delta t_2 - L(t_1)\Delta t_1 + \int_{t_1}^{t_2} \delta L\, dt. \tag{8.74}$$

Here the variation in the second integral can be carried out through a parameterization of the varied path, exactly as for Hamilton's principle except that the

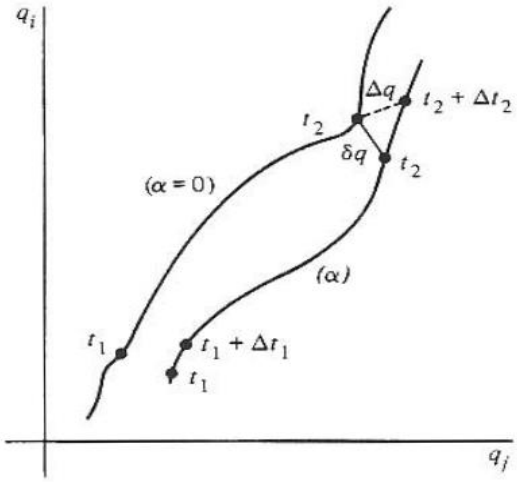


FIGURE 8.3 The Δ-variation in configuration space.

variation in q_i does not vanish at the end points. The end point terms arising in the integration by parts must be retained, and the integral term on the right appears as

$$\int_{t_1}^{t_2} \delta L\, dt = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right] \delta q_i\, dt + \frac{\partial L}{\partial \dot{q}_i}\, \delta q_i \Big|_1^2.$$

By Lagrange's equations the quantities in the square brackets vanish, and the Δ-variation therefore takes the form

$$\Delta \int_{t_1}^{t_2} L\, dt = (L\Delta t + p_i \delta q_i)\Big|_1^2. \tag{8.75}$$

In Eq. (8.75), δq_i refers to the variation in q_i at the original end point times t_1 and t_2. We would like to express the Δ-variation in terms of the change Δq_i between q_i at the end points of the actual path and q_i at the end points of the varied path, including the change in end point times. It is clear from Fig. 8.3 that these two variations are connected by the relation*

$$\Delta q_i = \delta q_i + \dot{q}_i \Delta t. \tag{8.76}$$

Hence, Eq. (8.75) can be rewritten as

$$\Delta \int_{t_1}^{t_2} L\, dt = (L\Delta t - p_i \dot{q}_i \Delta t + p_i \Delta q_i)\Big|_1^2$$

or

$$\Delta \int_{t_1}^{t_2} L\, dt = (p_i \Delta q_i - H\, \Delta t)\Big|_1^2. \tag{8.77}$$

To obtain the principle of least action, we restrict our further considerations by three important qualifications:

1. Only systems are considered for which L, and therefore H, are not explicit functions of time, and in consequence H is conserved.
2. The variation is such that H is conserved on the varied path as well as on the actual path.
3. The varied paths are further limited by requiring that Δq_i vanish at the end points (but not Δt).

The nature of the resultant variation may be illustrated by noting that the varied path satisfying these conditions might very well describe the same curve in configuration space as the actual path. The difference will be the speed with which the system point traverses this curve; that is, the functions $q_i(t)$ will be altered in the varied path. In order then to preserve the same value of the Hamiltonian at all points on the varied path, the times of the end points must be changed. With these three qualifications satisfied, the Δ-variation of the action integral, Eq. (8.77), reduces to

$$\Delta \int_{t_1}^{t_2} L\,dt = -H(\Delta t_2 - \Delta t_1). \tag{8.78}$$

But under the same conditions, the action integral itself becomes

$$\int_{t_1}^{t_2} L\,dt = \int_{t_1}^{t_2} p_i \dot{q}_i\,dt - H(t_2 - t_1),$$

the Δ-variation of which is

$$\Delta \int_{t_1}^{t_2} L\,dt = \Delta \int_{t_1}^{t_2} p_i \dot{q}_i\,dt - H(\Delta t_2 - \Delta t_1). \tag{8.79}$$

Comparison of Eqs. (8.78) and (8.79) finally gives the *principle of least action*:*

$$\Delta \int_{t_1}^{t_2} p_i \dot{q}_i\,dt = 0. \tag{8.80}$$

By way of caution, note that the modified Hamilton's principle can be written in a form with a superficial resemblance to Eq. (8.80). If the trajectory of the system point is described by a parameter θ, as in Sections 7.10 and 8.4, the modified Hamilton's principle appears as

$$\delta \int_{\theta_1}^{\theta_2} (p_i \dot{q}_i - H)t'\,d\theta = 0. \tag{8.81}$$

It will be recalled (cf. footnote on p. 351) that the momenta p_i do not change under the shift from t to θ, and that $\dot{q}_i t' = q_i'$. Further, the momentum conjugate to t is $-H$. Hence, Eq. (8.81) can be rewritten as

$$\delta \int_{\theta_1}^{\theta_2} \sum_{i=1}^{n+1} p_i q_i'\,d\theta = 0, \tag{8.82}$$

where t has been denoted by q_{n+1}. There should however be no confusion between Eq. (8.82) and the principle of least action. Equations (8.82) involve phase

space of $(2n + 2)$ dimensions, as is indicated by the explicit summation to $i = n + 1$, whereas Eq. (8.80) is in the usual configuration space. But most important, the principle of least action is in terms of a Δ-variation for constant H, while Eq. (8.82) employs the δ-variation, and H in principle could be a function of time. Equation (8.82) is nothing more than the modified Hamilton's principle, and the absence of a Hamiltonian merely reflects the phenomenon that the Hamiltonian vanishes identically for the "homogeneous problem."

The least action principle itself can be exhibited in a variety of forms. In non-relativistic mechanics, if the defining equations for the generalized coordinates do not involve the time explicitly, then the kinetic energy is a quadratic function of the $\dot{q}_i$'s (cf. Eq. (1.71)):

$$T = \tfrac{1}{2} M_{jk}(q)\dot{q}_j\dot{q}_k. \tag{8.83}$$

When in addition the potential is not velocity dependent, the canonical momenta are derived from T only, and in consequence

$$p_i\dot{q}_i = 2T.$$

The principle of least action for such systems can therefore be written as

$$\Delta \int_{t_1}^{t_2} T\, dt = 0. \tag{8.84}$$

If, further, there are no external forces on the system, as, for example, a rigid body with no net applied forces, then T is conserved along with the total energy H. The least action principle then takes the special form

$$\Delta(t_2 - t_1) = 0. \tag{8.85}$$

Equation (8.85) states that of all paths possible between two points, consistent with conservation of energy, the system moves along that particular path for which the time of transit is the least (more strictly, an extremum). In this form the principle of least action recalls Fermat's principle in geometrical optics that a light ray travels between two points along such a path that the time taken is the least. We discussed these considerations in Section 10–8 of the Second Edition when we considered the connection between the Hamiltonian formulation and geometrical optics.

space will be curvilinear and nonorthogonal. The element of path length in the space is then defined by (cf. Eq. (7.33′))

$$(d\rho)^2 = M_{jk}\, dq_j\, dq_k \tag{8.86}$$

so that the kinetic energy has the form

$$T = \frac{1}{2}\left(\frac{d\rho}{dt}\right)^2, \tag{8.87}$$

or equivalently

$$dt = \frac{d\rho}{\sqrt{2T}}. \tag{8.88}$$

Equation (8.88) enables us to change the variable in the abbreviated action integral from t to ρ, and the principle of least action becomes

$$\Delta \int_{t_1}^{t_2} T\, dt = 0 = \Delta \int_{\rho_1}^{\rho_2} \sqrt{T/2}\, d\rho,$$

or, finally

$$\Delta \int_{\rho_1}^{\rho_2} \sqrt{H - V(q)}\, d\rho = 0. \tag{8.89}$$

Equation (8.89) is often called *Jacobi's form of the least action principle*. It now

THE EQUATIONS OF CANONICAL TRANSFORMATION

There is one type of problem for which the solution of the Hamilton's equations is trivial. Consider a situation in which the Hamiltonian is a constant of the motion, and where *all* coordinates q_i are cyclic. Under these conditions, the conjugate momenta p_i are all constant:

$$p_i = \alpha_i,$$

and since the Hamiltonian cannot be an explicit function of either the time or the cyclic coordinates, it may be written as

$$H = H(\alpha_1, \ldots, \alpha_n).$$

Consequently, the Hamilton's equations for $\dot{q}_i$ are simply

$$\dot{q}_i = \frac{\partial H}{\partial \alpha_i} = \omega_i, \tag{9.1}$$

where the ω_i's are functions of the α_i's only and therefore are also constant in time. Equations (9.1) have the immediate solutions

$$q_i = \omega_i t + \beta_i, \tag{9.2}$$

where the β_i's are constants of integration, determined by the initial conditions.

It would seem that the solution to this type of problem, easy as it is, can only be of academic interest, for it rarely happens that all the generalized coordinates are cyclic. But a given system can be described by more than one set of generalized coordinates. Thus, to discuss motion of a particle in a plane, we may use as generalized coordinates either the Cartesian coordinates

$$q_1 = x, \qquad q_2 = y,$$

or the plane polar coordinates

$$q_1 = r, \qquad q_2 = \theta.$$

Both choices are equally valid, but one of the other set may be more convenient for the problem under consideration. Note that for central forces neither x nor y is cyclic, while the second set does contain a cyclic coordinate in the angle θ. The number of cyclic coordinates can thus depend upon the choice of generalized coordinates, and for each problem there may be one particular choice for which all coordinates are cyclic. If we can find this set, the remainder of the job is trivial. Since the obvious generalized coordinates suggested by the problem will not normally be cyclic, we must first derive a specific procedure for *transforming* from one set of variables to some other set that may be more suitable.

The transformations considered in the previous chapters have involved going from one set of coordinates q_i to a new set Q_i by transformation equations of the form

$$Q_i = Q_i(q, t). \tag{9.3}$$

For example, the equations of an orthogonal transformation, or of the change from Cartesian to plane polar coordinates, have the general form of Eqs. (9.3). As has been previously noted in Derivation 10 of Chapter 1, such transformations are known as *point transformations*. But in the Hamiltonian formulation the momenta are also independent variables on the same level as the generalized coordinates. The concept of transformation of coordinates must therefore be widened to include the simultaneous transformation of the independent *coordinates* and *momenta*, q_i, p_i, to a new set Q_i, P_i, with (invertible) equations of transformation:

$$\begin{aligned} Q_i &= Q_i(q, p, t), \\ P_i &= P_i(q, p, t). \end{aligned} \tag{9.4}$$

Thus, the new coordinates will be defined not only in terms of the old coordinates but also in terms of the old momenta. Equations (9.3) may be said to define

a *point transformation of configuration space*; correspondingly Eqs. (9.4) define a *point transformation of phase space*.

In developing Hamiltonian mechanics, only those transformations can be of interest for which the new Q, P are canonical coordinates. This requirement will be satisfied provided there exists some function $K(Q, P, t)$ such that the equations of motion in the new set are in the Hamiltonian form

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{9.5}$$

The function K plays the role of the Hamiltonian in the new coordinate set.* It is important for future considerations that the transformations considered be problem-independent. That is to say, (Q, P) must be canonical coordinates not only for some specific mechanical systems, but for all systems of the same number of degrees of freedom. Equations (9.5) must be the form of the equations of motion in the new coordinates and momenta no matter what the particular initial form of H. We may indeed be incited to develop a particular transformation from (q, p) to (Q, P) to handle, say, a plane harmonic oscillator. But the same transformation must then also lead to Hamilton's equations of motion when applied, for example, to the two-dimensional Kepler problem.

As was seen in Section 8.5, if Q_i and P_i are to be canonical coordinates, they must satisfy a modified Hamilton's principle that can be put in the form

$$\delta \int_{t_1}^{t_2} (P_i \dot{Q}_i - K(Q, P, t))\, dt = 0, \tag{9.6}$$

(where summation over the repeated index i is implied). At the same time the old canonical coordinates of course satisfy a similar principle:

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t))\, dt = 0. \tag{9.7}$$

The simultaneous validity of Eqs. (9.6) and (9.7) does not mean of course that the integrands in both expressions are equal. Since the general form of the modified Hamilton's principle has zero variation at the end points, both statements will be satisfied if the integrands are connected by a relation of the form

$$\lambda(p_i \dot{q}_i - H) = P_i \dot{Q}_i - K + \frac{dF}{dt}. \tag{9.8}$$

Here F is any function of the phase space coordinates with continuous second derivatives, and λ is a constant independent of the canonical coordinates and the time. The multiplicative constant λ is related to a particularly simple type of transformation of canonical coordinates known as a *scale transformation*.

a *point transformation of configuration space*; correspondingly Eqs. (9.4) define a *point transformation of phase space*.

In developing Hamiltonian mechanics, only those transformations can be of interest for which the new Q, P are canonical coordinates. This requirement will be satisfied provided there exists some function $K(Q, P, t)$ such that the equations of motion in the new set are in the Hamiltonian form

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}. \tag{9.5}$$

The function K plays the role of the Hamiltonian in the new coordinate set.* It is important for future considerations that the transformations considered be problem-independent. That is to say, (Q, P) must be canonical coordinates not only for some specific mechanical systems, but for all systems of the same number of degrees of freedom. Equations (9.5) must be the form of the equations of motion in the new coordinates and momenta no matter what the particular initial form of H. We may indeed be incited to develop a particular transformation from (q, p) to (Q, P) to handle, say, a plane harmonic oscillator. But the same transformation must then also lead to Hamilton's equations of motion when applied, for example, to the two-dimensional Kepler problem.

As was seen in Section 8.5, if Q_i and P_i are to be canonical coordinates, they must satisfy a modified Hamilton's principle that can be put in the form

$$\delta \int_{t_1}^{t_2} (P_i \dot{Q}_i - K(Q, P, t))\, dt = 0, \tag{9.6}$$

(where summation over the repeated index i is implied). At the same time the old canonical coordinates of course satisfy a similar principle:

$$\delta \int_{t_1}^{t_2} (p_i \dot{q}_i - H(q, p, t))\, dt = 0. \tag{9.7}$$

The simultaneous validity of Eqs. (9.6) and (9.7) does not mean of course that the integrands in both expressions are equal. Since the general form of the modified Hamilton's principle has zero variation at the end points, both statements will be satisfied if the integrands are connected by a relation of the form

$$\lambda(p_i \dot{q}_i - H) = P_i \dot{Q}_i - K + \frac{dF}{dt}. \tag{9.8}$$

Here F is any function of the phase space coordinates with continuous second derivatives, and λ is a constant independent of the canonical coordinates and the time. The multiplicative constant λ is related to a particularly simple type of transformation of canonical coordinates known as a *scale transformation*.

Suppose we change the size of the units used to measure the coordinates and momenta so that in effect we transform them to a set (Q', P') defined by

$$Q_i' = \mu q_i, \qquad P_i' = \nu p_i. \tag{9.9}$$

Then it is clear Hamilton's equations in the form of Eqs. (9.5) will be satisfied for a transformed Hamiltonian $K'(Q', P') = \mu\nu H(q, p)$. The integrands of the corresponding modified Hamilton's principles are, also obviously, related as

$$\mu\nu(p_i\dot{q}_i - H) = P_i'\dot{Q}_i' - K', \tag{9.10}$$

which is of the form of Eq. (9.8) with $\lambda = \mu\nu$. With the aid of suitable scale transformation, it will always be possible to confine our attention to transformations of canonical coordinates for which $\lambda = 1$. Thus, if we have a transformation of canonical coordinates $(q, p) \rightarrow (Q', P')$ for some $\lambda \neq 1$, then we can always find an intermediate set of canonical coordinates (Q, P) related to (Q', P') by a simple scale transformation of the form (9.9) such that $\mu\nu$ also has the same value λ. The transformation between the two sets of canonical coordinates (q, p) and (Q, P) will satisfy Eq. (9.8), but now with $\lambda = 1$:

$$p_i\dot{q}_i - H = P_i\dot{Q}_i - K + \frac{dF}{dt}. \tag{9.11}$$

Since the scale transformation is basically trivial, the significant transformations to be examined are those for which Eq. (9.11) holds.

A transformation of canonical coordinates for which $\lambda \neq 1$ will be called an *extended canonical transformation*. Where $\lambda = 1$, and Eq. (9.11) holds, we will speak simply of a *canonical transformation*. The conclusion of the previous paragraph may then be stated as saying that any extended canonical transformation can be made up of a canonical transformation followed by a scale transformation. Except where otherwise stated, all future considerations of transformations between canonical coordinates will involve only canonical transformations. It is also convenient to give a specific name to canonical transformations for which the equations of transformation Eqs. (9.4) do not contain the time explicitly; they will be called *restricted canonical transformations*.

The last term on the right in Eq. (9.11) contributes to the variation of the action integral only at the end points and will therefore vanish if F is a function of (q, p, t) or (Q, P, t) or any mixture of the phase space coordinates since these have zero variation at the end points. Further, through the equations of transformation, Eqs. (9.4) and their inverses F can be expressed partly in terms of the old set of variables and partly of the new. Indeed, F is useful for specifying the exact form of the canonical transformation only when half of the variables (beside the time) are from the old set and half are from the new. It then acts, as it were, as a bridge between the two sets of canonical variables and is called the *generating function* of the transformation.

To show how the generating function specifies the equations of transformation, suppose F were given as a function of the old and new generalized space

coordinates:

$$F = F_1(q, Q, t). \tag{9.12}$$

Equation (9.11) then takes the form

$$\begin{aligned} p_i\dot{q}_i - H &= P_i\dot{Q}_i - K + \frac{dF_1}{dt} \\ &= P_i\dot{Q}_i - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q_i}\dot{q}_i + \frac{\partial F_1}{\partial Q_i}\dot{Q}_i. \end{aligned} \tag{9.13}$$

Since the old and the new coordinates, q_i and Q_i, are separately independent, Eq. (9.13) can hold identically only if the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ each vanish:

$$p_i = \frac{\partial F_1}{\partial q_i}, \tag{9.14a}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i}, \tag{9.14b}$$

leaving finally

$$K = H + \frac{\partial F_1}{\partial t}. \tag{9.14c}$$

Equations (9.14a) are n relations defining the p_i as functions of q_j, Q_j, and t. Assuming they can be inverted, they could then be solved for the n Q_i's in terms of q_j, p_j, and t, thus yielding the first half of the transformation equations (9.4). Once the relations between the Q_i's and the old canonical variables (q, p) have been established, they can be substituted into Eqs. (9.14b) so that they give the n P_i's as functions of q_j, p_j, and t, that is, the second half of the transformation equations (9.4). To complete the story, Eq. (9.14c) provides the connection between the new Hamiltonian, K, and the old one, H. We must be careful to read Eq. (9.14c) properly. First q and p in H are expressed as functions of Q and P through the inverses of Eqs. (9.4). Then the q_i in $\partial F_1/\partial t$ are expressed in terms of Q, P in a similar manner and the two functions are added to yield $K(Q, P, t)$.

It sometimes happens that it is not suitable to describe the canonical transformation by a generating function of the type $F_1(q, Q, t)$. For example, the transformation may be such that p_i cannot be written as functions of q, Q, and t, but rather will be functions of q, P, and t. We would then seek a generating function that is a function of the old coordinates q and the new momenta P. Clearly Eq. (9.13) must then be replaced by an equivalent relation involving $\dot{P}_i$ rather than $\dot{Q}_i$. This can be accomplished by writing F in Eq. (9.11) as

$$F = F_2(q, P, t) - Q_i P_i. \tag{9.15}$$

Substituting this F in Eq. (9.11) leads to

$$p_i \dot{q}_i - H = -Q_i \dot{P}_i - K + \frac{d}{dt} F_2(q, P, t). \tag{9.16}$$

Again, the total derivative of F_2 is expanded and the coefficients of $\dot{q}_i$ and P_i collected, leading to the equations

$$p_i = \frac{\partial F_2}{\partial q_i}, \tag{9.17a}$$

$$Q_i = \frac{\partial F_2}{\partial P_i}, \tag{9.17b}$$

with

$$K = H + \frac{\partial F_2}{\partial t}. \tag{9.17c}$$

As before, Eqs. (9.17a) are to be solved for P_i as functions of q_j, p_j, and t to correspond to the second half of the transformation equations (9.4). The remaining half of the transformation equations is then provided by Eqs. (9.17b).

The corresponding procedures for the remaining two basic types of generating functions are obvious, and the general results are displayed in Table 9.1.

It is tempting to look upon the four basic types of generating functions as being related to each other through Legendre transformations. For example, the

TABLE 9.1 Properties of the Four Basic Canonical Transformations

Generating Function	Generating Function Derivatives	Trivial Special Case
$F = F_1(q, Q, t)$	$p_i = \frac{\partial F_1}{\partial q_i} \quad P_i = -\frac{\partial F_1}{\partial Q_i}$	$F_1 = q_i Q_i, \quad Q_i = p_i, \quad P_i = -q_i$
$F = F_2(q, P, t) - Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i} \quad Q_i = \frac{\partial F_2}{\partial P_i}$	$F_2 = q_i P_i, \quad Q_i = q_i, \quad P_i = p_i$
$F = F_3(p, Q, t) + q_i p_i$	$q_i = -\frac{\partial F_3}{\partial p_i} \quad P_i = -\frac{\partial F_3}{\partial Q_i}$	$F_3 = p_i Q_i, \quad Q_i = -q_i, \quad P_i = -p_i$
$F = F_4(p, P, t) + q_i p_i - Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i} \quad Q_i = \frac{\partial F_4}{\partial P_i}$	$F_4 = p_i P_i, \quad Q_i = p_i, \quad P_i = -q_i$

transition from F_1 to F_2 is equivalent to going from the variables q, Q to q, P with the relation

$$-P_i = \frac{\partial F_1}{\partial Q_i}. \tag{9.18}$$

This is just the form required for a Legendre transformation of the basis variables, as described in Section 8.1, and in analogy to Eq. (8.5) we would set

$$F_2(q, P, t) = F_1(q, Q, t) + P_i Q_i, \tag{9.19}$$

which is equivalent to Eq. (9.15) combined with Eq. (9.12). All the other defining equations for the generating functions can similarly be looked on, in combination with Eq. (9.12) as Legendre transformations from F_1, with the last entry in Table 9.1 describing a double Legendre transformation. The only drawback to this picture is that it might erroneously lead us to believe that any given canonical transformation can be expressed in terms of the four basic types of Legendre transformations listed in Table 9.1. This is not always possible. Some transformations are just not suitable for description in terms of these or other elementary forms of generating functions, as has been noted above and as will be illustrated in the next section with specific examples. If we try to apply the Legendre transformation process, we are then led to generating functions that are identically zero or are indeterminate. For this reason, we have preferred to define each type of generating function relative to F, which is some unspecified function of $2n$ independent coordinates and momenta.

Finally, note that a suitable generating function doesn't have to conform to one of the four basic types for *all* the degrees of freedom of the system. It is possible, and for some canonical transformations necessary, to use a generating function that is a mixture of the four types. To take a simple example, it may be desirable for a particular canonical transformation with two degrees of freedom to be defined by a generating function of the form

$$F'(q_1, p_2, P_1, Q_2, t). \tag{9.20}$$

This generating function would be related to F in Eq. (9.11) by the equation

$$F = F'(q_1, p_2, P_1, Q_2, t) - Q_1 P_1 + q_2 p_2, \tag{9.21}$$

and the equations of transformation would be obtained from the relations

$$\begin{aligned} p_1 &= \frac{\partial F'}{\partial q_1}, \qquad Q_1 = \frac{\partial F'}{\partial P_1}, \\ q_2 &= -\frac{\partial F'}{\partial p_2}, \qquad P_2 = -\frac{\partial F'}{\partial Q_2}, \end{aligned} \tag{9.22}$$

with

$$K = H + \frac{\partial F'}{\partial t}. \tag{9.23}$$

Specific illustrations are given in the next section and in the exercises.

■ EXAMPLES OF CANONICAL TRANSFORMATIONS

The nature of canonical transformations and the role played by the generating function can best be illustrated by some simple yet important examples. Let us consider, first, a generating function of the second type with the particular form

$$F_2 = q_i P_i \tag{9.24}$$

found in column 3 of Table 9.1. From Eqs. (9.17), the transformation equations are

$$\begin{aligned} p_i &= \frac{\partial F_2}{\partial q_i} = P_i, \\ Q_i &= \frac{\partial F_2}{\partial P_i} = q_i, \\ K &= H. \end{aligned} \tag{9.25}$$

The new and old coordinates are the same; hence F_2 merely generates the *identity transformation* (cf. Table 9.1). We also note, referring to Table 9.1, that the particular generating function $F_3 = p_i Q_i$ generates an identity transformation with negative signs; that is, $Q_i = -q_i$, $P_i = -p_i$.

A more general type of transformation is described by the generating function

$$F_2 = f_i(q_1, \ldots, q_n;\ t) P_i, \tag{9.26}$$

where the f_i may be any desired set of independent functions. By Eqs. (9.17b), the new coordinates Q_i are given by

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i(q_i, \ldots, q_n;\ t) \tag{9.27}$$

Thus, with this generating function the new coordinates depend only upon the old coordinates and the time and do not involve the old momenta. Such a transformation is therefore an example of the class of point transformations defined by Eqs. (9.3). In order to define a point transformation, the functions f_i must be independent and invertible, so that the q_j can be expressed in terms of the Q_i. Since the f_i are otherwise completely arbitrary, we may conclude that *all point transformations are canonical*. Equation (9.17c) furnishes the new Hamiltonian in terms of the old and of the time derivatives of the f_i functions.

Note that F_2 as given by Eq. (9.26) is not the only generating function leading to the point transformation specified by the f_i. Clearly the same point transformation is implicit in the more general form

$$F_2 = f_i(q_1, \ldots, q_n;\ t)P_i + g(q_1, \ldots, q_n;\ t), \tag{9.28}$$

where $g(q, t)$ is any (differentiable) function of the old coordinates and the time. Equations (9.27), the transformation equations for the coordinates, remain unaltered for this generating function. But the transformation equations of the momenta differ for the two forms. From Eqs. (9.17a), we have

$$p_j = \frac{\partial F_2}{\partial q_j} = \frac{\partial f_i}{\partial q_j}P_i + \frac{\partial g}{\partial q_j}, \tag{9.29}$$

using the form of F_2 given by Eq. (9.28). These equations may be inverted to give P as a function of (q, p), most easily by writing them in matrix notation:

$$\mathbf{p} = \frac{\partial \mathbf{f}}{\partial \mathbf{q}}\mathbf{P} + \frac{\partial g}{\partial \mathbf{q}}. \tag{9.29'}$$

Here $\mathbf{p}$, $\mathbf{P}$, and $\partial g/\partial \mathbf{q}$ are n-elements of single-column matrices, and $\partial \mathbf{f}/\partial \mathbf{q}$ is a square matrix whose ijth element is $\partial f_i/\partial q_j$. In two dimensions, Eq. (9.29') can be written as

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial q_1} & \dfrac{\partial f_1}{\partial q_2} \\ \dfrac{\partial f_2}{\partial q_1} & \dfrac{\partial f_2}{\partial q_2} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} + \begin{bmatrix} \dfrac{\partial g}{\partial q_1} \\ \dfrac{\partial g}{\partial q_2} \end{bmatrix}.$$

It follows that $\mathbf{P}$ is a linear function of $\mathbf{p}$ given by

$$\mathbf{P} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{q}}\right]^{-1}\left[\mathbf{p} - \frac{\partial g}{\partial \mathbf{q}}\right]. \tag{9.30}$$

In two dimensions, (9.30) becomes

$$\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial q_1} & \dfrac{\partial f_1}{\partial q_2} \\ \dfrac{\partial f_2}{\partial q_1} & \dfrac{\partial f_2}{\partial q_2} \end{bmatrix}^{-1} \left[\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} - \begin{bmatrix} \dfrac{\partial g}{\partial q_1} \\ \dfrac{\partial g}{\partial q_2} \end{bmatrix}\right]. \tag{9.31}$$

Thus, the transformation equations (9.27) for Q are independent of g and depend only upon the $f_i(q, t)$, but the transformation equations (9.29) for P do depend upon the form of g and are in general functions of both the old coordinates and momenta. The generating function given by Eq. (9.26) is only a special case of Eq. (9.28) for which $g = 0$, with correspondingly specialized transformation equations for P.

An instructive transformation is provided by the generating function of the first kind, $F_1(q, Q, t)$, of the form

$$F_1 = q_k Q_k.$$

The corresponding transformation equations, from (9.14a, b) are

$$p_i = \frac{\partial F_i}{\partial q_i} = Q_i, \tag{9.32a}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i} = -q_i. \tag{9.32b}$$

In effect, the transformation interchanges the momenta and the coordinates; the new coordinates are the old momenta and the new momenta are essentially the old coordinates. Table 9.1 shows that the particular generating function of type $F_4 = p_i P_i$ produces the same transformation. These simple examples should emphasize the independent status of generalized coordinates and momenta. They are both needed to describe the motion of the system in the Hamiltonian formulation. The distinction between them is basically one of nomenclature. We can shift the names around with at most no more than a change in sign. There is no longer present in the theory any lingering remnant of the concept of q_i as a spatial coordinate and p_i as a mass times a velocity. Incidentally, we may see directly from Hamilton's equations,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad \dot{q}_i = \frac{\partial H}{\partial p_i},$$

that this exchange transformation is canonical. If q_i is substituted for p_i, the equations remain in the canonical form only if $-p_i$ is substituted for q_i.

A transformation that leaves some of the (q, p) pairs unchanged, and interchanges the rest (with a sign change), is obviously a canonical transformation of a "mixed" form. Thus, in a system of two degrees of freedom, the transformation

$$Q_1 = q_1, \qquad P_1 = p_1,$$
$$Q_2 = p_2, \qquad P_2 = -q_2,$$

is generated by the function

$$F = q_1 P_1 + q_2 Q_2, \tag{9.33}$$

which is a mixture of the F_1 and F_2 types.

I THE HARMONIC OSCILLATOR

As a final example, let us consider a canonical transformation that can be used to solve the problem of the simple harmonic oscillator in one dimension. If the force

constant is k, the Hamiltonian for this problem in terms of the usual coordinates is

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}. \tag{9.34a}$$

Designating the ratio k/m by ω^2, H can also be written as

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 q^2). \tag{9.34b}$$

This form of the Hamiltonian, as the sum of two squares, suggests a transformation in which H is cyclic in the new coordinate. If we could find a canonical transformation of the form

$$p = f(P)\cos Q, \tag{9.35a}$$

$$q = \frac{f(P)}{m\omega}\sin Q, \tag{9.35b}$$

then the Hamiltonian as a function of Q and P would be simply

$$K = H = \frac{f^2(P)}{2m}(\cos^2 Q + \sin^2 Q) = \frac{f^2(P)}{2m}, \tag{9.36}$$

so that Q is cyclic. The problem is to find the form of the yet unspecified function $f(P)$ that makes the transformation canonical. If we use a generating function of the first kind given by

$$F_1 = \frac{m\omega q^2}{2}\cot Q, \tag{9.37}$$

Eqs. (9.14) then provide the equations of transformation,

$$p = \frac{\partial F_1}{\partial q} = m\omega q \cot Q, \tag{9.38a}$$

$$P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q}. \tag{9.38b}$$

Solving for q and p, we have*

$$q = \sqrt{\frac{2P}{m\omega}}\sin Q, \tag{9.39a}$$

$$p = \sqrt{2pm\omega}\cos Q, \tag{9.39b}$$

and comparison with Eq. (9.35a) evaluates $f(P)$:

$$f(P) = \sqrt{2m\omega P}. \tag{9.40}$$

It follows then that the Hamiltonian in the transformed variables is

$$H = \omega P. \tag{9.41}$$

Since the Hamiltonian is cyclic in Q, the conjugate momentum P is a constant. It is seen from Eq. (9.41) that P is in fact equal to the constant energy divided by ω:

$$P = \frac{E}{\omega}.$$

The equation of motion for Q reduces to the simple form

$$\dot{Q} = \frac{\partial H}{\partial P} = \omega,$$

with the immediate solution

$$Q = \omega t + \alpha, \tag{9.42}$$

where α is a constant of integration fixed by the initial conditions. From Eqs. (9.39), the solutions for q and p are

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \alpha), \tag{9.43a}$$

$$p = \sqrt{2mE}\cos(\omega t + \alpha). \tag{9.43b}$$

It is instructive to plot the time dependence of the old and new variables as is shown in Fig. 9.1. We see that q and p oscillate (Fig. 9.1a, b) whereas Q and P are linear plots (Fig. 9.1d, e). The figure also shows the phase space plots for p versus q (Fig. 9.1c) and for P versus Q (Fig. 9.1f). Fig. 9.1c is an ellipse with the following semimajor axes (for the q and p directions, respectively):

$$a = \sqrt{\frac{2E}{m\omega^2}} \qquad \text{and} \qquad b = \sqrt{2mE},$$

where m is the mass of the oscillator, ω its frequency, and E the oscillator's energy. The area, A, of this ellipse in phase space is

$$A = \pi ab = \frac{2\pi E}{\omega}.$$

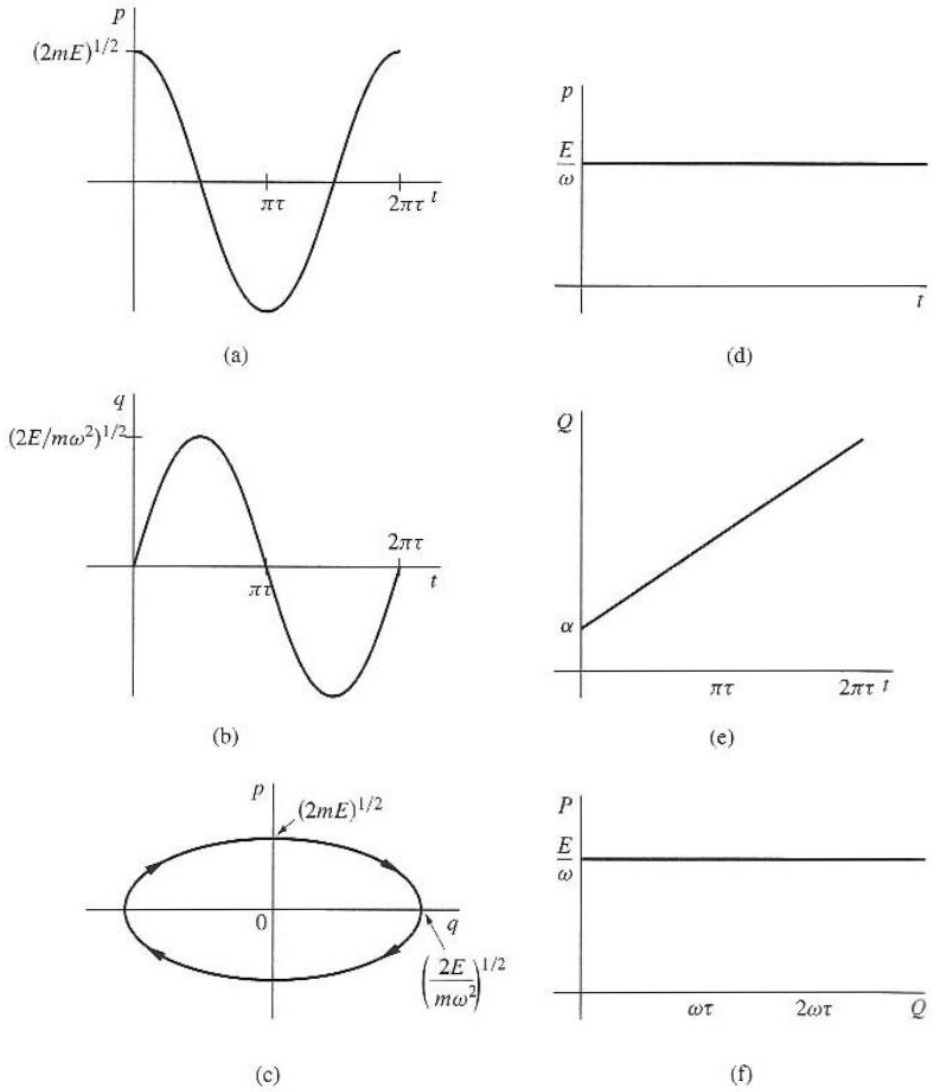


FIGURE 9.1 The harmonic oscillator in two canonical coordinate systems. Drawings (a)–(c) show the q, p system and (d)–(f) show the P, Q system.

When we invoke quantum mechanics, we write $E = \hbar\omega$, where $\hbar = h/2\pi$, and h is Planck's constant. The coordinate and momentum q and p can be normalized as

$$q' = \sqrt{\frac{m\omega^2}{2E}}q \qquad \text{and} \qquad p' = \frac{p}{\sqrt{2mE}},$$

to make the phase space plot of p' versus q' a circle of area π. This normalized

THE SYMPLECTIC APPROACH TO CANONICAL TRANSFORMATIONS

Another method of treating canonical transformations, seemingly unrelated to the generator formalism, can be expressed in terms of the matrix or symplectic formulation of Hamilton's equations. By way of introduction to this approach, let us consider a restricted canonical transformation, that is, one in which time does not appear in the equations of transformation:

$$\begin{aligned} Q_i &= Q_i(q, p), \\ P_i &= P_i(q, p). \end{aligned} \tag{9.44}$$

We know that the Hamiltonian function does not change in such a transformation. The time derivative of Q_i, on the basis of Eqs. (9.44), is to be found as

$$\dot{Q}_i = \frac{\partial Q_i}{\partial q_j}\dot{q}_j + \frac{\partial Q_i}{\partial p_j}\dot{p}_j = \frac{\partial Q_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial H}{\partial q_j}. \tag{9.45}$$

On the other hand, the inverses of Eqs. (9.44),

$$\begin{aligned} q_j &= q_j(Q, P), \\ p_j &= p_j(Q, P), \end{aligned} \tag{9.46}$$

enables us to consider $H(q, p, t)$ as a function of Q and P and to form the partial derivative

$$\frac{\partial H}{\partial P_i} = \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial P_i} + \frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial P_i}. \tag{9.47}$$

Comparing Eqs. (9.45) and (9.47), it can be concluded that

$$\dot{Q}_i = \frac{\partial H}{\partial P_i};$$

that is, the transformation is canonical, only if

$$\left(\frac{\partial Q_i}{\partial q_j}\right)_{q,p} = \left(\frac{\partial p_j}{\partial P_i}\right)_{Q,P}, \qquad \left(\frac{\partial Q_i}{\partial p_j}\right)_{q,p} = -\left(\frac{\partial q_j}{\partial P_i}\right)_{Q,P}. \tag{9.48a}$$

The subscripts on the derivatives are to remind us that on the left-hand side of these equations Q_i is considered as a function of (q, p) (cf. Eqs. (9.44)), while on the right-hand side the derivatives are for q_j and p_j as functions of (Q, P) (cf. Eqs. (9.46)). A similar comparison of $\dot{P}_i$ with the partial of H with respect to Q_j leads to the conditions

$$\left(\frac{\partial P_i}{\partial q_j}\right)_{q,p} = -\left(\frac{\partial p_j}{\partial Q_i}\right)_{Q,P}, \qquad \left(\frac{\partial P_i}{\partial p_j}\right)_{q,p} = \left(\frac{\partial q_j}{\partial Q_i}\right)_{Q,P}. \tag{9.48b}$$

The sets of Eqs. (9.48) together are sometimes known as the "direct conditions" for a (restricted) canonical transformation.

The algebraic manipulation that leads to Eqs. (9.48) can be performed in a compact and elegant manner if we make use of the symplectic notation for the Hamiltonian formulation introduced above at the end of Section 8.1. If $\boldsymbol{\eta}$ is a column matrix with the $2n$ elements q_i, p_i, then Hamilton's equations can be written, it will be remembered, as Eq. (8.39)

$$\dot{\boldsymbol{\eta}} = \mathbf{J}\frac{\partial H}{\partial \boldsymbol{\eta}},$$

where $\mathbf{J}$ is the antisymmetric matrix defined in Eq. (8.38a). Similarly the new set

canonical transformation the equations of transformation (9.44)

$$\boldsymbol{\zeta} = \boldsymbol{\zeta}(\boldsymbol{\eta}).$$

Analogously to Eq. (9.45) we can seek the equations of motion for the
ables by looking at the time derivative of a typical element of $\boldsymbol{\zeta}$:

$$\dot{\zeta}_i = \frac{\partial \zeta_i}{\partial \eta_j}\dot{\eta}_j, \qquad i, j = 1, \ldots, 2n.$$

In matrix notation, this time derivative can be written as

$$\dot{\boldsymbol{\zeta}} = \mathbf{M}\dot{\boldsymbol{\eta}},$$

where $\mathbf{M}$ is the Jacobian matrix of the transformation with elements

$$M_{ij} = \frac{\partial \zeta_i}{\partial \eta_j}.$$

Making use of the equations of motion for $\boldsymbol{\eta}$, Eq. (9.50) becomes

$$\dot{\zeta} = \mathbf{M}\mathbf{J}\frac{\partial H}{\partial \boldsymbol{\eta}}. \tag{9.52}$$

Now, by the inverse transformation H can be considered as a function of $\boldsymbol{\zeta}$, and the derivative with respect to η_i evaluated as

$$\frac{\partial H}{\partial \eta_i} = \frac{\partial H}{\partial \zeta_j}\frac{\partial \zeta_j}{\partial \eta_i},$$

or, in matrix notation*

$$\frac{\partial H}{\partial \boldsymbol{\eta}} = \widetilde{\mathbf{M}} \frac{\partial H}{\partial \boldsymbol{\zeta}}. \tag{9.53}$$

The combination of Eqs. (9.52) and (9.53) leads to the form of the equations of motion for any set of variables $\boldsymbol{\zeta}$ transforming, independently of time, from the canonical set $\boldsymbol{\eta}$:

$$\dot{\boldsymbol{\zeta}} = \mathbf{M}\mathbf{J}\widetilde{\mathbf{M}} \frac{\partial H}{\partial \boldsymbol{\zeta}}. \tag{9.54}$$

We have the advantage of knowing from the generator formalism that for a *restricted* canonical transformation the old Hamiltonian expressed in terms of the new variables serves as the new Hamiltonian:

$$\dot{\boldsymbol{\zeta}} = \mathbf{J} \frac{\partial H}{\partial \boldsymbol{\zeta}}. \tag{9.54'}$$

The transformation, Eq. (9.49), will therefore be canonical if $\mathbf{M}$ satisfies the condition

$$\mathbf{M}\mathbf{J}\widetilde{\mathbf{M}} = \mathbf{J}. \tag{9.55}$$

That Eq. (9.55) is also a necessary condition for a restricted canonical transformation is easily shown directly by reversing the order of the steps of the proof. Note that for an extended time-independent canonical transformation, where $K = \lambda H$, the condition of Eq. (9.55) would be replaced by

$$\mathbf{M}\mathbf{J}\widetilde{\mathbf{M}} = \lambda \mathbf{J}. \tag{9.56}$$

Equation (9.55) may be expressed in various forms. Multiplying from the right by the matrix inverse to $\widetilde{\mathbf{M}}$ leads to

$$\mathbf{M}\mathbf{J} = \mathbf{J}\widetilde{\mathbf{M}}^{-1}, \tag{9.57}$$

(since the transpose of the inverse is the inverse of the transpose). The elements of the matrix equation (9.57) will be found to be identical with Eqs. (9.48a) and (9.48b). If Eq. (9.57) is multiplied by $\mathbf{J}$ from the left and $-\mathbf{J}$ from the right, then by virtue of Eq. (8.38e) we have

$$\mathbf{J}\mathbf{M} = \widetilde{\mathbf{M}}^{-1}\mathbf{J},$$

or

$$\tilde{\mathbf{M}}\mathbf{J}\mathbf{M} = \mathbf{J}. \tag{9.58}$$

Equation (9.55), or its equivalent version, Eq. (9.58), is spoken of as the *symplectic condition* for a canonical transformation, and the matrix $\mathbf{M}$ satisfying the condition is said to be a *symplectic matrix*.

These concepts may become more obvious if we display the details of the $\mathbf{J}$ and $\mathbf{M}$ matrices corresponding to the mixed generating function $F = F_2(q_1, P_1) + F_1(q_2, Q_2)$ of Eq. (9.33). The variables $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$ are column vectors given by

$$\boldsymbol{\eta} = \begin{bmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\zeta} = \begin{bmatrix} Q_1 \\ Q_2 \\ P_1 \\ P_2 \end{bmatrix}.$$

The transformation $\dot{\boldsymbol{\zeta}} = \mathbf{M}\dot{\boldsymbol{\eta}}$ (cf. Eq. (9.50)) is made by the following $\mathbf{M}$ matrix:

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} \dot{q}_1 \\ \dot{p}_2 \\ \dot{p}_1 \\ -\dot{q}_2 \end{bmatrix},$$

in agreement with the expressions obtained by differentiating the results of the generating function with respect to time (cf. Column 3, Table 9.1). Hamilton's equations for the transformed variables $\dot{\zeta} = \mathbf{J}\frac{\partial H}{\partial \zeta}$ (Eq. (9.54′)) are expressed as follows independent of the generating function F

$$\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{P}_1 \\ \dot{P}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\dot{P}_1 \\ -\dot{P}_2 \\ \dot{Q}_1 \\ \dot{Q}_2 \end{bmatrix}$$

where $-\dot{P}_i = \partial H/\partial \zeta_i$ for ζ_1 and ζ_2 and $\dot{Q}_i = \partial H/\partial \zeta_i$ for ζ_3 and ζ_4. Note that $\mathbf{M}$ depends on F whereas J does not (cf. Eq. (8.38a)). This formalism is not applicable to all cases. For example, a simple $\mathbf{M}$ matrix cannot be written for the harmonic oscillator example discussed in Section 9.3.

transformations. For example, either the symplectic or the generator formalisms can be used to prove that canonical transformations have the four properties that characterize a group (cf. Appendix B).

1. The identity transformation is canonical.
2. If a transformation is canonical, so is its inverse.
3. Two successive canonical transformations (the group "product" operation) define a transformation that is also canonical.
4. The product operation is associative.

■ POISSON BRACKETS AND OTHER CANONICAL INVARIANTS

The *Poisson bracket* of two functions u, v with respect to the canonical variables (q, p) is defined as

$$[u, v]_{q,p} = \frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial v}{\partial q_i}. \tag{9.67}$$

In this bilinear expression we have a typical symplectic structure, as in Hamilton's equations, where q is coupled with p, and p with $-q$. The Poisson bracket thus lends itself readily to being written in matrix form, where it appears as

$$[u, v]_{\boldsymbol{\eta}} = \frac{\widetilde{\partial u}}{\partial \boldsymbol{\eta}}\mathbf{J}\frac{\partial v}{\partial \boldsymbol{\eta}}. \tag{9.68}$$

The transpose sign is used on the first matrix on the right-hand side to indicate explicitly that this matrix must be treated as a single-row matrix in the multiplication. On most occasions this specific reminder will not be needed and the transpose sign may be omitted.

Suppose we choose the functions u, v out of the set of canonical variables (q, p) themselves. Then it follows trivially from the definition, either as Eq. (9.67) or (9.68), that these Poisson brackets have the values

$$[q_j, q_k]_{q,p} = 0 = [p_j, q_k]_{q,p},$$

and

$$[q_j, p_k]_{q,p} = \delta_{jk} = -[p_j, q_k]_{q,p}. \tag{9.69}$$

We can summarize the relations of Eqs. (9.69) in one equation by introducing a *square matrix Poisson bracket*, $[\boldsymbol{\eta}, \boldsymbol{\eta}]$, whose lm element is $[\eta_l, \eta_m]$. Equations (9.69) can then be written as

$$[\boldsymbol{\eta}, \boldsymbol{\eta}]_{\eta} = \mathbf{J}. \tag{9.70}$$

Now let us take for u, v the members of the transformed variables (Q, P), or $\boldsymbol{\zeta}$, defined in terms of (q, p) by the transformation equations (9.59). The set of all the Poisson brackets that can be formed out of (Q, P) comprise the matrix Poisson bracket defined as

$$[\zeta, \zeta]_{\boldsymbol{\eta}} = \frac{\widetilde{\partial \boldsymbol{\zeta}}}{\partial \boldsymbol{\eta}}\mathbf{J}\frac{\partial \boldsymbol{\zeta}}{\partial \boldsymbol{\eta}}.$$

But we recognize the partial derivatives as defining the square Jacobian matrix of the transformation, so that the Poisson bracket relation is equivalent to

$$[\boldsymbol{\zeta}, \boldsymbol{\zeta}]_{\boldsymbol{\eta}} = \tilde{\mathbf{M}}\mathbf{J}\mathbf{M}. \tag{9.71}$$

If the transformation $\boldsymbol{\eta} \rightarrow \boldsymbol{\zeta}$ is canonical, then the symplectic condition holds and Eq. (9.71) reduces to (cf. Eq. (9.58))

$$[\boldsymbol{\zeta}, \boldsymbol{\zeta}]_{\boldsymbol{\eta}} = \mathbf{J}, \tag{9.72}$$

and conversely, if Eq. (9.72) is valid, then the transformation is canonical.

Poisson brackets of the canonical variables themselves, such as Eqs. (9.70) or (9.72), are referred to as the *fundamental Poisson brackets*. Since we have from Eq. (9.70) that

$$[\boldsymbol{\zeta}, \boldsymbol{\zeta}]_{\boldsymbol{\zeta}} = \mathbf{J}, \tag{9.73}$$

Eq. (9.72) states that the fundamental Poisson brackets of the ζ variables have the same value when evaluated with respect to *any* canonical coordinate set. In other words, the *fundamental Poisson brackets are invariant under canonical transformation*. We have seen from Eq. (9.71) that the invariance is a necessary and sufficient condition for the transformation matrix to be symplectic. The invariance of the fundamental Poisson brackets is thus in all ways equivalent to the symplectic condition for a canonical transformation.

It does not take many more steps to show that *all* Poisson brackets are invariant under canonical transformation. Consider the Poisson bracket of two functions u, v with respect to the η set of coordinates, Eq. (9.68). In analogy to Eq. (9.53), the partial derivative of v with respect to $\boldsymbol{\eta}$ can be expressed in terms of partial derivatives with respect to $\boldsymbol{\zeta}$ as

$$\frac{\partial v}{\partial \boldsymbol{\eta}} = \tilde{\mathbf{M}} \frac{\partial v}{\partial \boldsymbol{\zeta}}$$

(that is, the partial derivative transforms as a 1-form). In a similar fashion,

$$\widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}} = \widetilde{\tilde{\mathbf{M}} \frac{\partial u}{\partial \boldsymbol{\zeta}}} = \widetilde{\frac{\partial u}{\partial \boldsymbol{\zeta}}}\mathbf{M}.$$

Hence the Poisson bracket Eq. (9.68) can be written

$$[u, v]_{\boldsymbol{\eta}} = \widetilde{\frac{\partial u}{\partial \boldsymbol{\eta}}}\mathbf{J}\frac{\partial v}{\partial \boldsymbol{\eta}} = \widetilde{\frac{\partial u}{\partial \boldsymbol{\zeta}}}\mathbf{M}\mathbf{J}\tilde{\mathbf{M}}\frac{\partial u}{\partial \boldsymbol{\zeta}}.$$

If the transformation is canonical, the symplectic condition in the form of Eq. (9.55) holds, and we then have

$$[u, v]_{\boldsymbol{\eta}} = \widetilde{\frac{\partial u}{\partial \boldsymbol{\zeta}}}\mathbf{J}\frac{\partial v}{\partial \boldsymbol{\zeta}} \equiv [u, v]_{\boldsymbol{\zeta}}. \tag{9.74}$$

Thus, the Poisson bracket has the same value when evaluated with respect to any canonical set of variables—*all Poisson brackets are canonical invariants*. In writing the symbol for the Poisson bracket, we have so far been careful to indicate by the subscript the set of variables in terms of which the brackets are defined. So long as we use only canonical variables that practice is now seen to be unnecessary, and we shall in general drop the subscript.*

The algebraic properties of the Poisson bracket are therefore of considerable interest. We have already used the obvious properties

$$[u, u] = 0, \tag{9.75a}$$

$$[u, v] = -[v, u]. \qquad \text{(antisymmetry)} \tag{9.75b}$$

Almost equally obvious are the characteristics

$$[au + bv, w] = a[u, w] + b[v, w], \qquad \text{(linearity)} \tag{9.75c}$$

where a and b are constants, and

$$[uv, w] = [u, w]v + u[v, w]. \tag{9.75d}$$

One other property is far from obvious, but is very important in defining the nature of the Poisson bracket. It is usually given in the form of *Jacobi's identity*, which states that if u, v, and w are three functions with continuous second derivatives, then

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0; \tag{9.75e}$$

that is, the sum of the cyclic permutations of the double Poisson bracket of three functions is zero. There seems to be no simple way of proving Jacobi's identity for the Poisson bracket without lengthy algebra. However, it is possible to mitigate the complexity of the manipulations by introducing a special nomenclature. We

shall use subscripts on u, v, w (or functions of them) to denote partial derivatives by the corresponding canonical variable. Thus,

$$u_i \equiv \frac{\partial u}{\partial \eta_i}, \qquad \text{and} \qquad v_{ij} \equiv \frac{\partial v}{\partial \eta_i \, \partial \eta_j}.$$

In this notation the Poisson bracket of u and v can be expressed as

$$[u, v] = u_i J_{ij} v_j.$$

Here J_{ij}, as usual, is simply the ijth element of $\mathbf{J}$. In the proof, the only property of $\mathbf{J}$ that we shall need is its antisymmetry.

Now let us consider the first double Poisson bracket in Eq. (9.75e):

$$[u, [v, w]] = u_i J_{ij} [v, w]_j = u_i J_{ij} (v_k J_{kl} w_l)_j.$$

Because the elements J_{kl} are constants, the derivative with resect to η doesn't act on them, and we have

$$[u, [v, w]] = u_i J_{ij} (v_k J_{kl} w_{lj} + v_{kj} J_{kl} w_l). \tag{9.76}$$

The other double Poisson brackets can be obtained from Eq. (9.76) by cyclic permutation of u, v, w. There are thus six terms in all, each being a fourfold sum over dummy indices i, j, k, and l. Consider the term in Eq. (9.76) involving a second derivative of w:

$$J_{ij} J_{kl} u_i v_k w_{lj}.$$

The only other second derivative of w will appear in evaluating the second double Poisson bracket in (Eq. 9.75e):

$$[v, [w, u]] = v_k J_{kl} (w_j J_{ji} u_i)_l.$$

Here the term in the second derivative in w is

$$J_{ji} J_{kl} u_i v_k w_{jl}.$$

Since the order of differentiation is immaterial, $w_{lj} = w_{jl}$, and the sum of the two terms is given by

$$(J_{ij} + J_{ji}) J_{kl} u_i v_k w_{lj} = 0,$$

by virtue of the antisymmetry of J. The remaining four terms are cyclic permutations and can similarly be divided in two pairs, one involving second derivatives of u and the other of v. By the same reasoning, each of these pairs sums to zero, and Jacobi's identity is thus verified.

If the Poisson bracket of u, v is looked on as defining a "product" operation of the two functions, then Jacobi's identity is the replacement for the associa-

If two constants of the motion are known, the Jacobi identity provides a possible way for obtaining further constants. Suppose u and v are two constants of the motion not explicitly functions of time. Then if w in Eq. (9.75e) is taken to be H, the Jacobi identity says

$$[H, [u, v]] = 0;$$

that is, the Poisson bracket of u and v is also a constant in time. Even when the conserved quantities depend upon time explicitly, it can be shown with a bit more algebra (cf. Exercise 30) that *the Poisson bracket of any two constants of the motion is also a constant of the motion* (Poisson's theorem). Repeated application of the Jacobi identity in this manner can in principle lead to a complete sequence of constants of the motion. Quite often, however, the process is disappointing. The Poisson bracket of u and v frequently turns out to be a trivial function of u and v themselves, or even identically zero. Still, the possibility of generating new independent constants of motion by Poisson's theorem should be kept in mind.

THE ANGULAR MOMENTUM POISSON BRACKET RELATIONS

The identification of the canonical angular momentum as the generator of a rigid rotation of the system leads to a number of interesting and important Poisson bracket relations. Equations (9.103) for the change of a function u under an infinitesimal canonical transformation (on the "active" view) is also valid if u is taken as the component of a vector along a *fixed* axis in ordinary space. Thus, if $\mathbf{F}$ is a vector function of the system configuration, then (cf. Eq. (9.116))

$$\partial F_i = d\alpha[F_i, G].$$

Note that the direction along which the component is taken must be fixed, that is, not affected by the canonical transformation. If the direction itself is determined in terms of the system variables, then the transformation changes not only the value of the function but the nature of the function, just as with the Hamiltonian. With this understanding the change in a vector $\mathbf{F}$ under a rotation of the system about a fixed axis $\mathbf{n}$, generated by $\mathbf{L} \cdot \mathbf{n}$, can be written in vector notation (cf. Eq. (9.115))

$$\partial \mathbf{F} = d\theta[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}]. \tag{9.121}$$

To put it in other words, Eq. (9.121) implies that the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ that form the basis set for $\mathbf{F}$ are not themselves rotated by $\mathbf{L} \cdot \mathbf{n}$.

The words describing what is meant by Eq. (9.121) must be chosen carefully for another reason. What is spoken of is the rotation of the system under the I.C.T., not necessarily the rotation of the vector **F**. The generator $\mathbf{L} \cdot \mathbf{n}$ induces a spatial rotation of the system variables, not for example of some external vector such as a magnetic field or the vector of the acceleration of gravity. Under what conditions then does $\mathbf{L} \cdot \mathbf{n}$ generate a spatial rotation of **F**? The answer is clear—when **F** is a function only of the system variables (q, p) and does not involve any external quantities or vectors not affected by the I.C.T. Only under these conditions does a spatial rotation imply a corresponding rotation of **F**. We shall designate such vectors as *system vectors*. The change in a vector under infinitesimal rotation about an axis **n** has been given several times before (cf. Eq. (2.50) and Eq. (4.75)):

$$d\mathbf{F} = \mathbf{n}\, d\theta \times \mathbf{F}.$$

For a system vector **F**, the change induced under an I.C.T. generated by $\mathbf{L} \cdot \mathbf{n}$ can therefore be written as

$$\partial\mathbf{F} = d\theta[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n}\, d\theta \times \mathbf{F}. \tag{9.122}$$

Equation (9.122) implies an important Poisson bracket identity obeyed by all system vectors:

$$[\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{F}. \tag{9.123}$$

Note that in Eq. (9.123) there is no longer any reference to a canonical transformation or even to a spatial rotation. It is simply a statement about the value of certain Poisson brackets for a specific class of vectors and, as such, can be verified by direct evaluation in any given case. Suppose, for example, we had a system of an unconstrained particle and used the Cartesian coordinates as the canonical space coordinates. Then the Cartesian vector **p** is certainly a suitable system vector. If **n** is taken as a unit vector in the z direction, then by direct evaluation we have

$$[p_x, xp_y - yp_x] = -p_y,$$
$$[p_y, xp_y - yp_x] = p_x,$$
$$[p_z, xp_y - yp_x] = 0.$$

The right-hand sides of these identities is clearly the same as the components of $\mathbf{n} \times \mathbf{p}$, as predicted by Eq. (9.123).

On the other hand, suppose that in the same problem we tried to use for **F** the vector $\mathbf{A} = \frac{1}{2}(\mathbf{r} \times \mathbf{B})$ where $\mathbf{B} = B\mathbf{i}$ is a fixed vector along the x axis. The vector **A** will be recognized as the vector potential corresponding to a uniform magnetic field **B** in the x-direction. As **A** depends upon a vector external to the system, we would expect it not to fit the characteristics of a system vector and Eq. (9.123) should not hold for it. Indeed, we see that the Poisson brackets involved are here

$$[0, xp_y - yp_x] = 0,$$

$$\left[\tfrac{1}{2}zB, xp_y - yp_x\right] = 0,$$

$$\left[-\tfrac{1}{2}yB, xp_y - yp_x\right] = -\tfrac{1}{2}Bx,$$

whereas the vector $\mathbf{n} \times \mathbf{A}$ has instead the components $(-\frac{1}{2}Bz, 0, 0)$.

The relation (9.123) may be expressed in various notations. Perhaps the most advantageous is a form using the Levi–Civita density to express the cross product (cf. Eq. (4.77′)). The ith component of Eq. (9.123) for arbitrary $\mathbf{n}$ then can be written

$$[F_i, L_j n_j] = \epsilon_{ijk} n_j F_k, \tag{9.124}$$

which implies the simple result

$$[F_i, L_j] = \epsilon_{ijk} F_k. \tag{9.125}$$

An alternative statement of Eq. (9.125) is to note that if l, m, n are three indices in cyclic order, then

$$[F_l, L_m] = F_n, \quad l, m, n \text{ in cyclic order.} \tag{9.125′}$$

Another consequence of Eq. (9.123) relates to the dot product of two system vectors: $\mathbf{F} \cdot \mathbf{G}$. Being a scalar, such a dot product should be invariant under rotation, and indeed the Poisson bracket of the dot product with $\mathbf{L} \cdot \mathbf{n}$ is easily shown to vanish:

$$\begin{aligned}[\mathbf{F} \cdot \mathbf{G}, \mathbf{L} \cdot \mathbf{n}] &= \mathbf{F} \cdot [\mathbf{G}, \mathbf{L} \cdot \mathbf{n}] + \mathbf{G} \cdot [\mathbf{F}, \mathbf{L} \cdot \mathbf{n}] \\ &= \mathbf{F} \cdot \mathbf{n} \times \mathbf{G} + \mathbf{G} \cdot \mathbf{n} \times \mathbf{F} \\ &= \mathbf{F} \cdot \mathbf{n} \times \mathbf{G} + \mathbf{F} \cdot \mathbf{G} \times \mathbf{n} \\ &= 0. \end{aligned} \tag{9.126}$$

The magnitude of any system vector therefore has a vanishing Poisson bracket with any component of $\mathbf{L}$.

Perhaps the most frequent application of these results arises from taking F to be the vector $\mathbf{L}$ itself. We then have

$$[\mathbf{L}, \mathbf{L} \cdot \mathbf{n}] = \mathbf{n} \times \mathbf{L}, \tag{9.127}$$

$$[L_i, L_j] = \epsilon_{ijk} L_k, \tag{9.128}$$

and

$$[L^2, \mathbf{L} \cdot \mathbf{n}] = 0. \tag{9.129}$$

A number of interesting consequences follow from Eqs. (9.127)

$$[p, \mathbf{L}\cdot\mathbf{n}] = \mathbf{n}\times\mathbf{p}$$

$$[p_i, L_j] = \epsilon_{ijk}p_k.$$

If L_x and L_y are constants of the motion, Poisson's theorem then states that $[L_x, L_y] = L_z$ is also a constant of the motion. Thus, if any two components of the angular momentum are constant, the total angular momentum vector is conserved. As a further instance, let us assume that in addition to L_x and L_y being conserved there is a Cartesian vector of canonical momentum $\mathbf{p}$ with p_z a constant of the motion. Not only is L_z conserved but we have two further constants of the motion:

$$[p_z, L_x] = p_y$$

and

$$[p_z, L_y] = -p_x,$$

that is, both $\mathbf{L}$ and $\mathbf{p}$ are conserved. We have here an instance in which Poisson's theorem does yield new constants of the motion. Note, however, that if p_x, p_y, and L_z were the given constants of the motion, then their Poisson brackets are

$$[p_x, p_y] = 0,$$

$$[p_x, L_z] = -p_y,$$

$$[p_y, L_z] = p_x.$$

Here no new constants can be obtained from Poisson's theorem.

Recall from the fundamental Poisson brackets, Eqs. (9.69), that the Poisson bracket of any two canonical momenta must always be zero. But, from Eq. (9.128), L_i does not have a vanishing Poisson bracket with any of the other components of $\mathbf{L}$. Thus, while we have described $\mathbf{L}$ as the total canonical angular momentum by virtue of its definition as $\mathbf{r}_i \times \mathbf{p}_i$ (summed over all particles), no two components of $\mathbf{L}$ can simultaneously be canonical variables. However, Eq. (9.129) shows that any one of the components of $\mathbf{L}$, and its magnitude L, can be chosen to be canonical variables at the same time.*

****

Unit V: The Special Theory of Relativity
Basic Postulates of the Special Theory – Newton's Law of Motion – Non-Variance Maxwell's Equation - Michelson Morley Experiment – Null results - Lorentz Transformations – Concept of Inertial frame – Velocity Addition and Thomas Precession – Length Contraction – Vectors and Metric Tensor – Relativistic Kinematics – Relativistic Angular Momentum – Introduction to the General theory of Relativity – Gravitation and acceleration and their relation to non-inertial frames of reference – Minkowski space and Lorenz transformation.

■ BASIC POSTULATES OF THE SPECIAL THEORY

Einstein used two postulates to develop what became known as the special theory:

1. The laws of physics are the same to all inertial observers.
2. The speed of light is the same to all inertial observers.

A formulation of physics that explicitly incorporates these two postulates is said to be *covariant*. Since the speed of light, c, is the same in all coordinate systems, it is reasonable to consider the numerical value of c as a conversion factor between the units used in measuring space and the units used in measuring time. So, $c\,dt$ is the time interval measured in the same units used to measure space units. In the SI system of units, $c\,dt$ has dimensions of meters. Many books and articles on relativity set $c = 1$ and measure time and space in meters. In the material that follows, we shall show the explicit dependence upon c.

To satisfy the two postulates, the space and time of the special theory consist of a single entity that we refer to as *spacetime*. This spacetime is the geometric framework within which we perform physics. We cannot assume that all observers make the same division into time and space in the same way. The separation is unique to each inertial frame. The square of the distance in that spacetime, Δs^2, between two points $\mathcal{A}$ and $\mathcal{B}$ is given by

$$(\Delta s)^2 = c^2(\text{time interval})^2 - (\text{space interval})^2, \qquad (7.4)$$

where the interval is between the two points $\mathcal{A}$ and $\mathcal{B}$. If the separation of the interval is assumed to be infinitesimal, the Δ is replaced by the differential symbol d. Since a point in spacetime consists of a specification of three spatial coordinate values and one time value, the usual convention is to refer to a point in spacetime as an **event**. The term *event* is used because such a point has a definite location and a definite time in any frame.

The choice of opposite signs for the time and space intervals is intrinsic to the theory; however, the choice of a positive sign for $(c\,dt)^2$ is arbitrary. Some authors define a $(ds)^2$, which is the negative of the choice given in Eq. (7.4). All sign choices makes $(ds)^2 = 0$ according to the definition in Eq. (7.4) for light, since the space interval is $\pm(c \times$ time interval). The choice made here for the relative signs used for space and time is such that real bodies moving at a velocity less than light have $(ds)^2 > 0$. This makes ds real for bodies moving slower than light speed. If $(ds)^2 > 0$, the interval is called *timelike*. If $(ds)^2 < 0$, the interval is called *spacelike*. Intervals for which $(ds)^2 = 0$ are called *lightlike* or *null*.

Since, to all inertial observers, objects that travel on timelike paths move less than the speed of light, they are called *tardyons*. Hypothetical bodies that always move faster than light are called *tachyons*, but such bodies will not concern us here. Objects moving at the speed of light are called *null* or *lightlike*.

In the limit of small displacements (differential displacements), Eq. (7.4) becomes, in a Cartesian coordinate system,

$$(ds)^2 = (c\,dt)^2 - (dx^2 + dy^2 + dz^2). \tag{7.4'}$$

The four-dimensional space with an interval defined by Eqs. (7.4) or (7.4′), is often called *Minkowski* space to distinguish it from a four-dimensional Euclidean space for which there would be no minus sign in Eqs. (7.4) or (7.4′). The idea of using *ict* for the time coordinate to make the space Euclidean is no longer useful since it obscures the non-Euclidean nature of spacetime and makes the generalization to noninertial frames more difficult.

Since the interval between two events of spacetime is a geometric quantity, all inertial observers measure coordinates that preserve the value of the interval squared, $(ds)^2$. If S and S' are two different inertial frames, then

$$ds'^2 = ds^2. \tag{7.5}$$

As a special case of Eq. (7.4), consider the relation between the proper time, τ, measured by an observer at rest with respect to an object in frame S' with coordinates (τ, x', y', z'), which is moving at a velocity, $\mathbf{v}$, with respect to a laboratory frame S with coordinates (t, x, y, z). In the rest frame of the object, there is no motion, so Eqs. (7.4′) and (7.5) give

$$c^2(d\tau)^2 = c^2(dt)^2 - v^2(dt)^2 = c^2(dt)^2\left(\frac{1 - v^2}{c^2}\right)$$

or

$$dt = \frac{d\tau}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{7.6}$$

Since Eq. (7.6) makes $d\tau < dt$, this effect on dt is called "time dilation": moving clocks appear to run slower.

■ LORENTZ TRANSFORMATIONS

The simplest set of transformations that preserve the invariance of the interval, ds^2, are called the *Lorentz transformations*. These transformations are simplest in the sense that they are linear in the coordinates and as the relative velocity goes to zero, the transformations become identity transformations. If we consider parallel Cartesian coordinate systems, S and S', whose origins coincide at $t = t' = 0$, and whose relative velocity is v along the x axis as measured by S, and define

$$\beta = \frac{v}{c}, \qquad \text{and} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}, \tag{7.7}$$

then the following four equations relate the two sets of coordinates

$$ct' = \frac{ct - \beta x}{\sqrt{1-\beta^2}} = \gamma(ct - \beta x) \tag{7.8a}$$

$$x' = \frac{x - \beta ct}{\sqrt{1-\beta^2}} = \gamma(x - \beta ct) \tag{7.8b}$$

$$y' = y \tag{7.8c}$$

$$z' = z. \tag{7.8d}$$

Here we are only interested in transformations for which $t' \to t$ and $x' \to x$ as $\beta \to 0$. As matrices, these transformations appear as

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}. \tag{7.8'}$$

In the limit of $\beta \ll 1$, Eqs. (7.8) reduce to the Galilean transformations as expected.

The generalization to arbitrary orientation of the velocity relative to the axes is straightforward. Since we are considering spacetime a four-dimensional entity, we would expect to deal with four-dimensional vectors. Using the notation $(ct, x, y, z) = (ct, r)$ allows the writing of the generalization of Eqs. (7.8′) to the case where $\mathbf{v}$ is not parallel to an axis, as

$$\begin{aligned} ct' &= \gamma(ct - \boldsymbol{\beta} \cdot \mathbf{r}) \\ \mathbf{r}' &= \mathbf{r} + \frac{(\boldsymbol{\beta} \cdot \mathbf{r})\boldsymbol{\beta}(\gamma - 1)}{\beta^2} - \boldsymbol{\beta}\gamma ct, \end{aligned} \tag{7.9}$$

provided the two sets of axes are aligned. Another way to express this arbitrary velocity is to consider the Lorentz transformation between two inertial coordinate systems with aligned axes, as a matrix transformation relating the two 4-quantities, $\mathbf{x} = (ct, r)$ and $\mathbf{x}' = (ct', r')$, where

$$\mathbf{x}' = \mathbf{L}\mathbf{x} \tag{7.10}$$

We treat $\mathbf{x}'$ and $\mathbf{x}$ as column matrices and $\mathbf{L}$ as the symmetric matrix

$$\mathbf{L} = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + (\gamma - 1)\frac{\beta_x^2}{\beta^2} & (\gamma - 1)\frac{\beta_x\beta_y}{\beta^2} & (\gamma - 1)\frac{\beta_x\beta_z}{\beta^2} \\ -\gamma\beta_y & (\gamma - 1)\frac{\beta_x\beta_y}{\beta^2} & 1 + (\gamma - 1)\frac{\beta_y^2}{\beta^2} & (\gamma - 1)\frac{\beta_y\beta_z}{\beta^2} \\ -\gamma\beta_z & (\gamma - 1)\frac{\beta_x\beta_z}{\beta^2} & (\gamma - 1)\frac{\beta_z\beta_y}{\beta^2} & 1 + (\gamma - 1)\frac{\beta_z^2}{\beta^2} \end{bmatrix}. \tag{7.11}$$

This reduces to the results given in Eqs. (7.8′) when $\beta_x = \beta$, $\beta_y = \beta_z = 0$.

These transformations map the origin of S and the origin of S' to $(0, 0, 0, 0)$. Hence the coordinates of both origins correspond to the same location in spacetime. If this is not desired, there is a more general transformation of the form

$$\mathbf{x}' = \mathbf{L}\mathbf{x} + \mathbf{a} \tag{7.12}$$

where **L** is a spacetime rotation (boost) and **a** is a spacetime translation. This is the *Poincaré transformation* or the *inhomogeneous Lorentz transformation*. We shall consider only homogeneous transformations for which **a** of Eq. (7.12) is zero.

12.7. THE MICHELSON-MORLEY EXPERIMENTS

Michelson-Morley experiment was designed to determine the motion of the earth relative to a frame of reference in which the speed of light is c in all the directions. This frame is the previledge or absolute frame of reference. The earth itself cannot be the absolute frame of reference because it is moving around the sun in its orbit with a speed $v = 3 \times 10^4$ m/sec and therefore at any time if the earth is identified with the motion of the preferred frame, it will have a speed of 6×10^4 m/sec after 6 months relative to the absolute (inertial) frame of reference. In any case, the earth has a speed at least $v = 3 \times 10^4$ m/sec relative to an inertial frame at some instant or the other during 6 months of the year. Assuming that the absolute frame is situated at the centre of the sun and the earth is moving with a speed 3×10^4 m/sec relative to it. Therefore, an observer on earth will measure a speed $(c - v)$ for light moving along the direction of its motion and $(c + v)$ for light moving in opposite direction. We shall see how this assumption (*the change of velocity of light due to the motion of the observer*) was contradicted by the famous experiments of Michelson and Morley, conducted in the year 1880.

In principle, in the experiment of Michelson and Morley, a ray of light starts from a monochromatic source of light S and falls on a half-silvered glass plate P, where a part of light is reflected towards the mirror M_1 and part of it is transmitted towards the mirror M_2 [Fig 12.3(a)]. PM_1 and PM_2 are mutually perpendicular directions and they are nearly equal. Let $PM_1 = PM_2 = l$. The two rays, reflected from the mirrors M_1 and M_2 unite again at P and interference fringes are obtained. These fringes may be seen through the telescope T. The entire apparatus is floated on mercury, contained in a large vessel so that the interferometer may be rotated in any desired direction.

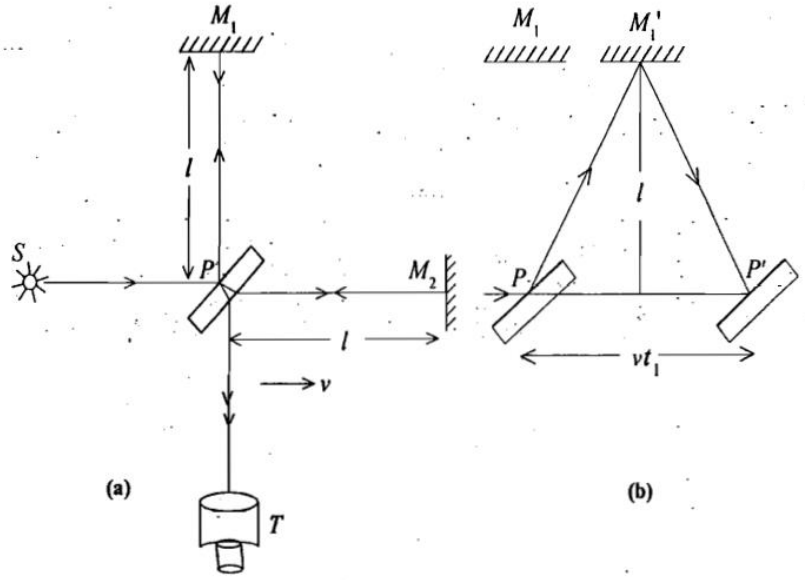


Fig. 12.3 : Michelson-Morley Experiment

In the experiment the instrument is first set up in such a way that arm PM_2 is parallel to the motion of the earth in space. Therefore, the apparatus moves in the direction PM_2 with velocity v relative to the absolute frame. After reflection at P, the ray moves towards the mirror M_1. In the time, the ray reaches the mirror, it moves to the position M_1' so that the reflection occurs at this position M_1'. If t_1 is the time taken by the ray in traversing the path $PM'_1\ P'$, in this duration plate P will move a distance $PP' = vt_1$. Hence from fig 12.3 (b), we have

$$PM_1' = \sqrt{PN^2 + M_1'N^2} = \sqrt{(vt_1/2)^2 + l^2}$$

Hence, distance $PM_1'P' = 2PM_1' = 2\sqrt{l^2 + v^2t_1^2/4}$.

This distance $PM_1'P'$ has been traversed by the ray in the absolute frame, in which the velocity of light is c in all the directions. Hence

$$ct_1 = 2\sqrt{l^2 + v^2t_1^2/4} \quad \text{or} \quad c^2t_1^2 = 4l^2 + v^2t_1^2 \quad \text{or} \quad t_1^2 = \frac{4l^2}{c^2 - v^2}$$

Thus
$$t_1 = \frac{2l}{\sqrt{c^2 - v^2}} = \frac{2l}{c(1 - v^2/c^2)^{1/2}} = \frac{2l}{c}\left(1 - \frac{v^2}{c^2}\right)^{-1/2}$$

$$= \frac{2l}{c}\left(1 + \frac{v^2}{2c^2}\right) \qquad \left[\text{by Binomial theorem for } \frac{v^2}{c^2} << 1\right]$$

The transmitted light ray at P is travelling with velocity $(c - v)$ relative to the mirror M_2, because the mirror is moving with the velocity of the earth (v). This ray after reflection at M_2 will travel with a velocity $(c + v)$ relative to the plate P, because now the plate is moving opposite to the ray. If t_2 be the time taken by the transmitted ray to travel the distance l to the mirror M_2 and back, then

$$t_2 = \frac{l}{c - v} + \frac{l}{c + v} \quad (\text{where } PM_2 = l)$$

$$= \frac{2lc}{c^2 - v^2} = \frac{2l}{c}\left(1 - \frac{v^2}{c^2}\right)^{-1} = \frac{2l}{c}\left(1 + \frac{v^2}{c^2}\right) \quad [\text{using Binomial series for } v^2/c^2 << 1]$$

Hence the difference in times taken to traverse by the two paths is

$$\Delta t = t_2 - t_1 = \frac{2l}{c}\left(1 + \frac{v^2}{c^2}\right) - \frac{2l}{c}\left(1 + \frac{v^2}{2c^2}\right) = \frac{2l}{c} \cdot \frac{v^2}{2c^2} = \frac{lv^2}{c^3}$$

Therefore, the difference in the distance travelled by the two rays of light, *i.e.*, the path difference = $c\Delta t = lv^2/c^2$.

Finally, the whole apparatus is turned through 90° so that the other arm PM_1 becomes coincident with the earth's velocity (v) in space [Fig. 12.4]. This causes the difference of path in the opposite direction and hence the displacement of fringes should correspond to the path difference $2lv^2/c^2$. In experiment, distance l was taken nearly 11 metres. Hence the displacement of the fringes should correspond to the path difference

$$\frac{2 \times 11 \times (3 \times 10^4)^2}{(3 \times 10^8)^2} = 2200 \times 10^{-1} \text{ m.}$$

The shift for yellow colour (λ = 5000 Å) is expected to be 0.4 (= 2200×10^{-10}/ 5800×10^{-10}) of a fringe-width. The expected shift of 0.4 of a fringe could be measured easily in the experiment. However, no shift of fringes was observed. The experiment was repeated several times, but such displacement was never seen.

One may argue that accidently the absolute frame has the same velocity 3×10^4 m /sec with respect to sun as does the earth. Therefore, at this instant, the earth is at rest in the absolute reference system. However, Michelson and Mor-ley repeated their experiment six months later and a four times magnified effect could be expected ; but again nothing was observed [Fig. 12.5]. The experiments since then have been repeated several times under different circumstances and always the same negative result was obtained.

The negative results of Michelson-Morley suggest that the value of v relative to the absolute frame should be zero. In other words, ***the speed of light in vacuum must be the same*** (c) ***in all inertial frames***. It does not depend upon the motion of the observer or source expected on the basis of Galilean transformations. The negative results of Mechelson-Morley experiments show the validity of the principle of relativity in the fields of elctrodynamics and optics. In fact, *the principle of relativity is a fundamental truth applicable to all areas of physics*.

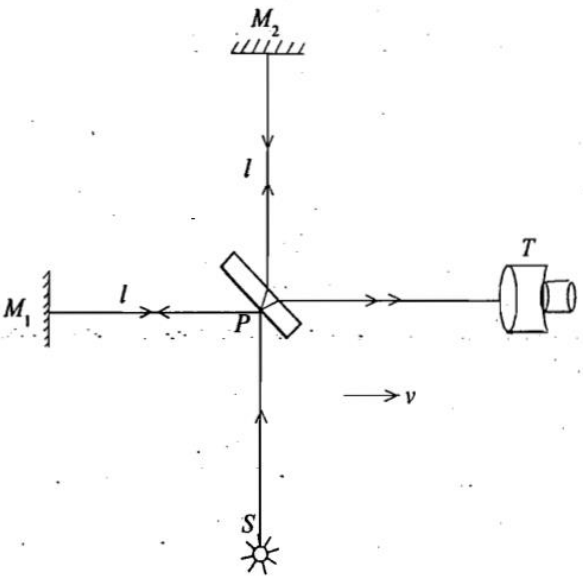


Fig. 12.4 : Michelson-Morley apparatus rotated by $\pi/2$.

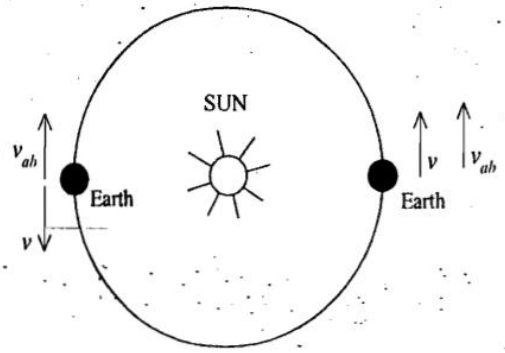


Fig. 12.5

12.8. ETHER HYPOTHESIS

In the 19th century, physicists made a false analogy between light waves and sound waves or other purely mechanical disturbances. In order to propagate the sound waves a material medium (*e.g.*, air) is necessary. If we say that the speed of sound in air is 332 m/sec, it means that this is the speed which is measured with respect to reference frame fixed in the air. Therefore, these physicists postulate the existence of a hypothetical medium for transmission of light and called it *ether*. It was supposed to fill all the space. To explain the very high speed of light, the density of the ether was supposed to be vanishingly small while its elastic moduli were assumed to be quite large. These workers considered that there is a fixed frame of reference of ether in which light travels with velocity c ($= 3 \times 10^8$ m/sec) in all directions. Since the earth is moving at a speed of $v = 3 \times 10^4$ m/sec. around the sun in its orbit, the supporters of ether theory reasoned that there must be times of the year when the earth has a velocity of at least 3×10^4 m/sec with respect to the ether. The negative results of Michelson-Morley experiments suggest that the effects of ether are undetectable and therefore, ether theory must be discarded. In fact, what Einstein said that there is no necessity of any material medium for the propagation of light waves and analogy between the electromagnatic waves and mechanical waves is not correct.

In conclusion, *the Michelson-Morley experiments discard the idea of a priviledged (absolute) frame of reference or ether and suggest that* ***the velo-city of light c is constant in vaccum in all inertial frames***. The later fact is the root of the relativistic discussion of physical laws.

12.11. CONSEQUENCES OF LORENTZ TRANSFORMATIONS

Now, let us discuss the consequences of Lorentz transformations regarding the lengths of the bodies and the time intervals between given events.

(1) Length contraction : In order to measure the length of an object in motion, relative to an observer, the positions of the two end points must be recorded simultaneously. Consider a frame S relative to which a rod is moving with velocity v along the X-axis. Let us associate a frame S' with the rod so that the rod is at rest in S'. If in this frame, the x-coordinates of the ends of the rod are x_1' and x_2', then

$$l_0 = x_2' - x_1'$$

This length l_0 has been measured by a stationary observer relative to the rod and is called the ***proper length*** of the rod. It is not necessary that the observer O' should measure the positions of the end points of the rod simultaneously, because the rod is at rest relative to him.

If the X-coordinates of the end points of that rod in frame S are measured to be x_1 and x_2 at the same time t, then in this frame the observed length of the rod is

$$l = x_2 - x_1$$

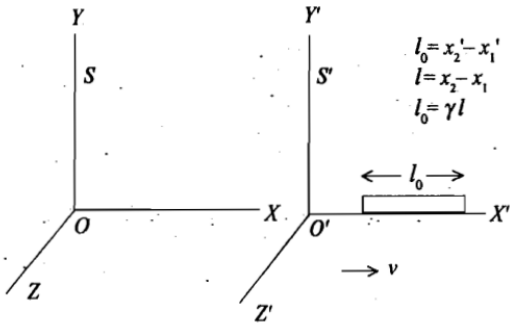


Fig. 12.7 : Contraction of moving rod

According to Lorentz transformations

$$x_1' = \gamma(x_1 - vt) \text{ and } x_2' = \gamma(x_2 - vt)$$

$$\therefore \quad x_2' - x_1' = \gamma(x_2 - x_1) \text{ or } l_0 = \gamma l$$

Thus $$l = l_0\sqrt{1 - v^2/c^2} \qquad ...(17)$$

As the factor $\sqrt{1 - v^2/c^2}$ is smaller than unity, we have $l < l_0$. This means that the length of the rod (l), as measured by an observer relative to which the rod is in motion, is smaller than its proper length. Such a contradiction of length in the direction of motion relative to an observer is called ***Lorentz-Fitzgerald contradiction***. However, there will be no change in length in the perpendicular direction of motion.

If the rod is at rest in the frame S, then its proper length is

$$l_0 = x_2 - x_1$$

Now the observer of S' at time t' measures the end coordinates of the rod as x_2' and x_1', then according to Lorentz transformations, we have

$$x_1 = \gamma(x_1' + vt') \text{ and } x_2 = \gamma(x_2' + vt')$$

Hence, $$x_2 - x_1 = \gamma(x_2' - x_1') \text{ or } l_0 = \gamma l \text{ or } l = l_0\sqrt{1 - v^2/c^2}.$$

This means that a rod at rest in S appears to be contracted to the observer O'. Thus, *a length is contracted, if there is relative motion between it and the observer.*

(2) Simultaneity : If two events occur at the same time in a frame, they are said to be *simultaneous*. Suppose that S' frame is moving relative to S along positive direction of X-axis with velocity v. Let two events occur simultaneously in frame S at the points P_1 and P_2 with the coordinates (x_1, y_1, z_1, t_1) and $(x_2,$

y_2, z_2, t_2) respectively as measured by the observer O of the S-frame. As the two events are simultaneous in frame S, we have $t_2 = t_1$. If t_1' and t_2' are the corresponding times of the same two events are measured by O' of frame S', then the use of Lorentz transformation equations gives

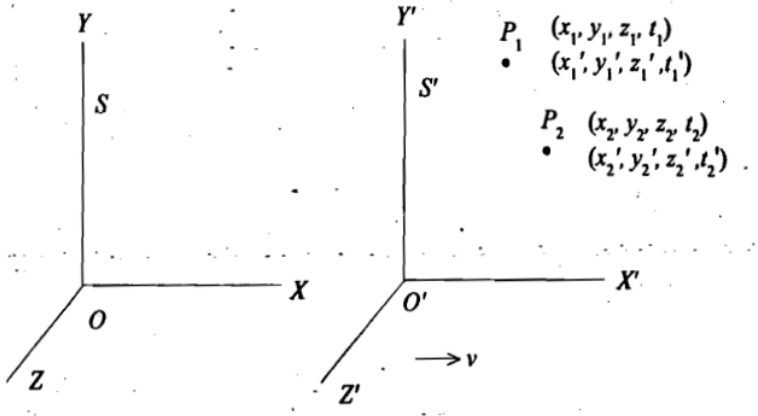


Fig. 12.8 : Representation of two events in two inertial frames

$$t_1' = \frac{t_1 - vx_1/c^2}{\sqrt{1 - v^2/c^2}} \quad \text{and} \quad t_2' = \frac{t_2 - vx_2/c^2}{\sqrt{1 - v^2/c^2}}$$

Therefore, $$t_2' - t_1' = \frac{t_2 - t_1}{\sqrt{1 - v^2/c^2}} - \frac{(v/c^2)(x_2 - x_1)}{\sqrt{1 - v^2/c^2}}$$

As $t_2 = t_1$ or $t_2 - t_1 = 0$, we have

$$t_2' - t_1' = -\frac{(v/c^2)(x_2 - x_1)}{\sqrt{1 - v^2/c^2}} \qquad ...(18)$$

We observe that $t_2' - t_1' \neq 0$. This means that the two events at two different points P_1 and P_2 which are simultaneous for O in frame S are not simultaneous for the observer O' of the frame S', moving with speed v along X-axis relative to S. Thus the simultaneity is not absolute, but relative.

(3) Time Dilation : Let a frame S' be moving along X-axis with velocity v relative to S. Now, if a clock being at rest in the frame S', mea- sures the time t_1' and t_2' of two events occurring at a fixed position x' in this frame, then the interval of time between these events is

$$\Delta t' = t_2' - t_1' = \Delta t_0 \text{ (say)}$$

Now, according to Lorentz transformations, we have

$$t_1 = \gamma(t_1' + vx'/c^2) \quad \text{and} \quad t_2 = \gamma(t_2' + vx'/c^2)$$

Therefore $$t_2 - t_1 = \gamma(t_2' - t_1') \text{ or } \Delta t = \gamma \Delta t_0$$

Thus $$\Delta t = \frac{\Delta t_0}{\sqrt{1 - v^2/c^2}} \qquad ...(19)$$

As $$1/\sqrt{1 - v^2/c^2} > 1, \ \Delta t > \Delta t_0 .$$

(4) Addition of Velocities : Let the coordinates of a particle in frame S be (x, y, z, t) and in frame S' (x', y', z', t'), then the components of its velocity in two frames can be written as

$$u_x = dx/dt,\ u_y = dy/dt,\ u_z = dz/dt \qquad \text{in } S$$

and $$u'_x = dx'/dt',\ u_y' = dy'/dt',\ u_z' = dz'/dt' \qquad \text{in } S'.$$

According to the inverse Lorentz transformations

$$x = \gamma(x' + vt'),\ y = y',\ z = z',\ t = \gamma(t' + vx'/c^2)$$

Therefore, $$dx = \gamma(dx' + vdt'),\ dy = dy',\ dz = dz' \text{ and } dt = \gamma(dt' + vdx'/c^2)$$

$$\therefore \quad u_x = \frac{dx}{dt} = \frac{\gamma(dx' + vdt')}{\gamma\left(dt' + \frac{vdx'}{c^2}\right)} = \frac{\frac{dx'}{dt'} + v}{1 + \frac{v\,dx'}{c^2\,dt'}}$$

or $$u_x = \frac{u'_x + v}{1 + vu'_x/c^2} \qquad \text{...(22 a)}$$

Also, $$u_y = \frac{dy}{dt} = \frac{dy'}{\gamma\left(dt' + \frac{vdx'}{c^2}\right)} = \frac{dy'/dt'}{\gamma\left(1 + \frac{v\,dx'}{c^2\,dt'}\right)}$$

or $$u_y = \frac{u'_y}{\gamma(1 + vu'_x/c^2)} \qquad \text{...(22 b)}$$

Similarly, $$u_z = \frac{u_z}{\gamma(1 + vu'_x/c^2)} \qquad \text{...(22 c)}$$

This is the ***relativistic law of addition of velocities*** while in classical mechanics $u_x = u_x' + v$, $u_y' = u_y$ and $u_z' = u_z$. We get the later (Galilean) equations, when v is much less than the speed of light c.

If we take the Lorentz transformations, we can prove that

$$u_x' = \frac{u_x - v}{1 - vu_x/c^2};\quad u_y' = \frac{u_y}{\gamma(1 - vu_x/c^2)};\quad u_z' = \frac{u_z}{\gamma(1 - vu_x/c^2)} \qquad \text{...(23)}$$

In case a particle (as photon) is moving with a velocity c in the frame S' and S' is moving with velocity c relative to S along positive X-axis direction, then from eq. (22 *a*), we have

$$u_x = \frac{c + c}{1 + cc/c^2} = c \qquad [\text{because } u_x' = c,\ v = c]$$

Thus the speed of photon in the second frame is also c, *i.e.,* the velocity of light is the same for all inertial frames whatever their relative speeds may be. This result is in accordance with the Michelson-Morley experiment. In fact, the Lorentz transformations and hence the law of addition of velocities have been deduced by assuming the velocity of light constant for all inertial observers.

■ VELOCITY ADDITION AND THOMAS PRECESSION

The most general homogeneous Lorentz transformation will involve both a velocity change and a rotation of the coordinates. The velocity transformation is termed a boost and has the form of Eq. (7.11). Any homogeneous Lorentz transformation, $\mathbf{L}$, can be written as

$$\mathbf{L} = \mathbf{R}\mathbf{L}_0 = \mathbf{L}_0'\mathbf{R}' \tag{7.13}$$

where $\mathbf{R}$ is a rotation matrix as discussed in Chapter 4, and $\mathbf{L}_0$, which is called a restricted or proper Lorentz transformation, corresponds to a pure boost. The restricted Lorentz transformations form a representation of the Lorentz group.* Since $\mathbf{R}$ is not symmetric and $\mathbf{L}_0$ is symmetric, $\mathbf{L}$ will, in general, have no symmetry. Also, since $\mathbf{L}_0$ and $\mathbf{R}$ are matrices, $\mathbf{R}\mathbf{L}_0 \neq \mathbf{L}_0\mathbf{R}$. There will exist two other transformations $\mathbf{L}_0'$ and $\mathbf{R}'$ such that $\mathbf{R}\mathbf{L}_0 = \mathbf{L}_0'\mathbf{R}'$.

For any Lorentz transformation, $\mathbf{L}$, there is an inverse transformation, $\mathbf{L}^{-1}$, such that

$$\mathbf{L}\mathbf{L}^{-1} = \mathbf{L}^{-1}\mathbf{L} = \mathbf{1}, \tag{7.14}$$

where $\mathbf{1}$ is the diagonal unit 4×4 matrix with elements $\delta_{\alpha\beta}$. The existence of an inverse places four constraints on the diagonal element and six on the off-diagonal elements for a total of ten constraints on the Lorentz transformation. There are then only six independent components. Three of these correspond to the components of the relative velocity vector and three correspond to the Euler angles of the rotation (see Section 4.4).

Consider three inertial systems, S_1, S_2, and S_3, with x axes aligned. Let S_2 be moving at a velocity v along the common x-direction with respect to S_1 and let S_3 be moving at velocity v' along the common x-direction with respect to S_2. The Lorentz transformation from S_1 to S_3 is given by

$$\mathbf{L}_{1-3} = \begin{bmatrix} \gamma' & -\gamma'\beta' & 0 & 0 \\ -\gamma'\beta' & \gamma' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \gamma\gamma'(1+\beta\beta') & -\gamma\gamma'(\beta+\beta') & 0 & 0 \\ -\gamma\gamma'(\beta+\beta') & \gamma\gamma'(1+\beta\beta') & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where Eq. (7.7) defines β and γ for v and β' and γ' for v'. Let β'' be the speed of S_3 relative to S_1 and γ'' the associated factor, then since $\mathbf{L}_{1-3}$ can be written as a single Lorentz transformation with a velocity β'' with its associated γ'' as

$$\mathbf{L}_{1-3} = \begin{bmatrix} \gamma'' & -\gamma''\beta'' & 0 & 0 \\ -\gamma''\beta'' & \gamma'' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and, since these two forms of $\mathbf{L}_{1-3}$ must be the same, we have

$$\beta'' = \frac{\beta + \beta'}{1 + \beta\beta'} \tag{7.15}$$

This is the relativistic addition of velocity formula for parallel velocities.

The product of any two transformations, $\mathbf{L}_1$ and $\mathbf{L}_2$ is itself a Lorentz transformation, $\mathbf{L}_3$. Such a Lorentz transformation will, in general, involve not only a boost, but may also include a rotation of coordinate axes. If both $\mathbf{L}_1$ and $\mathbf{L}_2$ are pure boosts but their two velocities are not parallel, $\mathbf{L}_3$ will involve a rotation in addition to a boost. This rotation is called the *Thomas precession* rotation. The usual form for the Thomas precession assumes the second boost, $\mathbf{L}_2$ has a velocity small compared to the first boost, $\mathbf{L}_1$ and also that it is small compared to the speed of light. For example, the Thomas precession can be observed for a gyroscope orbiting the Earth or for electrons in atoms.

Consider three inertial frames S_1, S_2, and S_3, with S_2 moving at a velocity $\boldsymbol{\beta}$ with respect to S_1 and S_3 moving at a velocity of $\boldsymbol{\beta}'$ with respect to S_2. Without loss of generality, we can arrange the axes of S_1 so that $\boldsymbol{\beta}$ is along the x axis of S_1 and $\boldsymbol{\beta}'$ lies in the $x'y'$ plane of S_2; that is, β, β' define the $x'y'$ plane of S_2. Let $\mathbf{L}$ represent the transformation from S_1 to S_2 and $\mathbf{L}'$ the transformation from S_2 to S_3 with γ and γ' associated with $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$. Then from Eq. (7.11),

$$\mathbf{L} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.16}$$

and

$$\mathbf{L}' = \begin{bmatrix} \gamma' & -\gamma'\beta'_{x'} & -\gamma'\beta'_{y'} & 0 \\ -\gamma'\beta'_{x'} & 1 + (\gamma' - 1)\frac{\beta'^2_{x'}}{\beta'^2} & (\gamma' - 1)\frac{\beta'_{x'}\beta'_{y'}}{\beta'^2} & 0 \\ -\gamma'\beta'_{y'} & (\gamma' - 1)\frac{\beta'_{x'}\beta'_{y'}}{\beta'^2} & 1 + (\gamma' - 1)\frac{\beta'^2_{y'}}{\beta'^2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.17}$$

We assume that the components of $\boldsymbol{\beta}'$ are small and only need be retained to first order giving via matrix multiplications of Eq. (7.16) and Eq. (7.17)

$$\mathbf{L}'' = \mathbf{L}'\mathbf{L} = \begin{bmatrix} \gamma\gamma' & -\gamma\gamma'\beta & -\gamma'\beta'_{y'} & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ -\gamma\gamma'\beta'_{y'} & \gamma\beta\gamma'\beta'_{y'} & \gamma' & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.18}$$

Since $\mathbf{L}''$ is not symmetric, it must correspond to a rotation and a boost. We shall write the velocity of S_3 as observed by S_1 as $\boldsymbol{\beta}''$.

Since the off-diagonal elements corresponding to the z axis are zero, this rotation is about an axis perpendicular to the xy plane. The boost from S_1 to S_3 is denoted by $\boldsymbol{\beta}''$, and we assume that $\boldsymbol{\beta}'$ is small compared to $\boldsymbol{\beta}$ and also small compared to the speed of light ($\gamma' \approx 1$). Then, to first order, the nonvanishing components of β'' are (Since the velocity perpendicular to x is small we can ignore to first order the distinction among y, y', and y'')

$$\beta''_x = \beta, \qquad \beta''_y = \frac{\beta'_y}{\gamma}, \quad \beta''^2 = \beta^2, \qquad \text{and} \qquad \gamma'' = \gamma, \tag{7.19}$$

and Eq. (7.18) becomes

$$\mathbf{L}'' \approx \begin{bmatrix} \gamma'' & -\gamma''\beta''_x & -\gamma''\beta''_y & 0 \\ -\gamma''\beta''_x & \gamma'' & 0 & 0 \\ -\gamma''\beta''_y & \gamma''\beta''_x\beta''_y & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.18'}$$

In this approximation, a pure Lorentz transformation from S_3 to S_1 (the inverse transformation) would correspond to a large boost in the x'' axis of $-\beta''_x$ and a small boost in the y'' axis of $-\beta''_y$. The Lorentz boost for that transformation

$$\mathbf{L}_{3-1} = \begin{bmatrix} \gamma'' & \gamma''\beta''_x & \gamma''\beta''_y & 0 \\ \gamma''\beta''_x & \gamma'' & (\gamma''-1)\frac{\beta''_y}{\beta''_x} & 0 \\ \gamma''\beta''_y & (\gamma''-1)\frac{\beta''_y}{\beta''_x} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.20}$$

Finally, the rotation matrix induced by the rotation from S_1 to S_3, after some algebraic simplification and the dropping of higher-order terms in $\boldsymbol{\beta}''$, is found to be

$$\mathbf{R} = \mathbf{L}''\mathbf{L}_{3-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & (\gamma-1)\frac{\beta''_y}{\beta} & 0 \\ 0 & -(\gamma-1)\frac{\beta''_y}{\beta} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.21}$$

Comparison with Eq. (4.44) shows that $\mathbf{R}$ implies S_3 is rotated with respect to S_1 about the z axis through an infinitesimal angle:

$$\Delta\Omega = (\gamma - 1)\frac{\beta_y''}{\beta} = \beta_y''\beta\left(\frac{\gamma - 1}{\beta^2}\right). \tag{7.22}$$

The spatial rotation resulting from the successive application of two nonparallel Lorentz transformations has been declared every bit as paradoxical as the more frequently discussed apparent violations of common sense, such as the so-called "twin paradox." But the present apparent paradox has important applications, especially in atomic physics, and therefore has been abundantly verified experimentally.

Suppose now that S_1 is the laboratory system, while S_2 and S_3 are two of the instantaneous rest systems a time Δt apart in the particle's motion. By Eq. (7.22), the laboratory observer will see a change in the particle's velocity in this time, $\Delta\mathbf{v}$, which has only a y-component $\beta_y'' c = \Delta v$. Since the initial x axis has been chosen along the direction of $\mathbf{v} = \beta c$, the vector of the infinitesimal rotation in this time can be written as

$$\Delta\boldsymbol{\Omega} = -(\gamma - 1)\frac{\mathbf{v} \times \Delta\mathbf{v}}{v^2} \tag{7.23}$$

Hence, if the particle has some specific direction attached to it (such as a spin vector), it will be observed from the laboratory system that this direction precesses with an angular velocity

$$\boldsymbol{\omega} = \frac{d\boldsymbol{\Omega}}{dt} = -(\gamma - 1)\frac{\mathbf{v} \times \mathbf{a}}{v^2} \tag{7.24}$$

where $\mathbf{a}$ is the particle's acceleration as seen from S_1. Equation (7.24) is frequency encountered in the form it takes when v is small enough that γ can be approximated (using $\gamma \approx 1 + \frac{1}{2}\beta^2$) as

$$\boldsymbol{\omega} = \frac{1}{2c^2}(\mathbf{a} \times \mathbf{v}). \tag{7.25}$$

In either form, $\boldsymbol{\omega}$ is known as the *Thomas precession frequency*.

■ VECTORS AND THE METRIC TENSOR

We will use the notation that the coordinates, which need not be Cartesian, are written as x^μ where $x^0 = ct$ is the time coordinate, and x^1, x^2, x^3 are the space coordinates. This change in notation is needed to be consistent with the developments in the following sections.

Consider an arbitrary one-dimensional curve in 4-dimensional spacetime, $\mathcal{P}$, described by a parameter λ, where for a given λ the coordinates of a point of the curve can be written as $x^0(\lambda)$, $x^1(\lambda)$, $x^2(\lambda)$, $x^3(\lambda)$. In introductory texts a 4-vector, v, is defined by this curve as an arrow whose tail is located at an event $\mathcal{A}$ on the curve and whose head is at an event $\mathcal{B}$ on the curve where $v_{\mathcal{AB}} = \mathcal{P}_\mathcal{B} - \mathcal{P}_\mathcal{A}$. However, instead of defining the vector at two points, we can use the parameter λ, which is a measure of the length along the curve from $\mathcal{A}$ to $\mathcal{B}$, by writing

$$v_{\mathcal{AB}} = \left(\frac{d\mathcal{P}}{d\lambda}\right)_{\lambda=0}. \tag{7.26}$$

Such a 4-vector is a *tangent vector* to the curve. We adopt the notation that the components of vectors are written with superscripts such as v^0, v^1, v^2, v^3. In spite of the way we draw tangent vectors, they do not have any extension in spacetime. The arrows we draw simply help us visualize the vector. At each point along the curve, the tangent vector has a direction and a magnitude. For curves that are timelike, the proper time, τ, is usually chosen as the parameter λ. The laboratory coordinates are then $x^0 = ct(\tau)$, $x^1 = x(\tau)$, $x^2 = y(\tau)$, $x^3 = z(\tau)$, and the tangent to the curve is the *four-velocity*, u, of a particle traveling along the curve $\mathcal{P}$. Equation (7.26) becomes

$$u^0 = \frac{dct}{d\tau} = \gamma c, \qquad u^i = \frac{dx^i}{d\tau} = \gamma v^i \tag{7.27}$$

where $v^i = dx^i/dt$ is the normal three-velocity with $\mathbf{v}^2 = (v^x)^2 + (v^y)^2 + (v^z)^2$. We shall assume that Greek letters can take on the values 0–3 and Latin letters the values 1–3. Repeated indices are summed. Since the 4-velocity of a particle is defined over a range of the parameter λ, there is an infinite set of 4-velocities for the particle, one for each value of λ. Such a set of vectors is termed a *vector field*. Some common examples of vector fields are given in Table 7.1.

We assume that the components of any 4-vector can be expressed by the values of the vector's projections along a set of basis vectors, $\boldsymbol{e}_0$, $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, and that the coordinates are measured along the direction given by the basis vectors. Such a system is called a *coordinates basis*.* Cartesian, spherical, and cylindrical coordinate systems, among many possible systems, can have such a basis set. The position of a point on the curve $\mathcal{P}(\tau)$ can be written as

$$\mathcal{P}(\tau) = x^\mu(\tau)\boldsymbol{e}_\mu, \tag{7.28}$$

TABLE 7.1 Examples of Vector Fields

Name	Time Portion	Space Portion	(Magnitude)^2	Type
Coordinate	ct	$\mathbf{r}$	$c^2t^2 - r^2$	spacelike, null, or timelike
Velocity	γc	$\gamma \mathbf{v}$	c^2	timelike
Momentum	$\frac{E}{c}$	$\mathbf{p}$	m^2c^2	timelike
Force	$\frac{\gamma}{c}\frac{dE}{dt}$	$\gamma\frac{d\mathbf{p}}{dt} = \gamma\mathbf{F}$	$-(\mathbf{F}_{\text{Newtonian}})^2$	spacelike
Current density	$\gamma\rho c$	$\gamma\mathbf{J}$	ρ^2c^2	timelike

where repeated Greek indices, one raised and one lowered, are summed from 0 to 3. In particular, the 4-velocity given in Eq. (7.27) becomes

$$u = \frac{d\mathcal{P}}{d\tau} = \frac{dx^\mu}{d\tau}\mathbf{e}_\mu = u^\mu \mathbf{e}_\mu. \tag{7.29}$$

The magnitude of the 4-velocity is a scalar whose values can vary as we change λ. This set of magnitudes is an example of a *scalar field*. To convert a 4-vector field to a scalar field, we need what is called a functional,* which can convert a pair of vectors into a scalar function at each point in spacetime. In other words, we wish to define the scalar product of two vectors or vector fields. This conversion of a 4-vector field (or two different vector fields) to a scalar field is an example of a *mapping*. If both the vectors are the same, then this scalar would be the square of the length of the vector, and when the vectors are different, it is called the scalar product of the vectors. Such a functional is called the *metric tensor*, g.† The metric tensor functional can be considered as a machine with two slots into which you can insert two vectors to produce a scalar (real-valued function). That is,

$$g(u, v) = g(v, u) = u \cdot v, \tag{7.30}$$

is the scalar product. In particular if the basis vectors are inserted into the metric,

$$g_{\alpha\beta} = \mathbf{g}(\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta) = \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta. \tag{7.31}$$

The $g_{\alpha\beta}$ are the components of the metric tensor associated with the basis vectors $\boldsymbol{e}_\alpha$. For example, consider a two-dimensional Minkowski space with coordinates ct and x and a vector $v = (a, b)$. Then $g(v, v) = a^2 - b^2$ and $g_{00} = 1$, $g_{11} = -1$.

The form of the $g_{\alpha\beta}$ is defined by the form for the interval. This suggests that we consider small displacements. If the relative displacement vector between two

points is small, it can be written as

$$d\zeta = \Delta x^{\alpha} \boldsymbol{e}_{\alpha}. \tag{7.32}$$

Recasting Eq. (7.32) in the language of Eq. (7.4′), we see for Minkowski coordinates

$$\begin{aligned}(\Delta s)^2 = d\zeta \cdot d\zeta &= \Delta x^{\alpha} \Delta x^{\beta} \boldsymbol{e}_{\alpha} \cdot \boldsymbol{e}_{\beta} = g_{\alpha\beta} \Delta x^{\alpha} \Delta x^{\beta} \\ &= (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2.\end{aligned}$$

In the limit of infinitesimal displacements this can be written as

$$ds^2 = g_{\alpha\beta} dx^{\alpha} dx^{\beta}, \tag{7.32′}$$

which holds for any metric tensor. The metric tensor for a Minkowski coordinate system, using the $+---$ sign convention, has the following tensor representation*

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{7.33}$$

The scalar product of two vectors in this coordinate system is

$$u \cdot v = u^{\alpha} v^{\beta} g_{\alpha\beta} = u^0 v^0 - u^1 v^1 - u^2 v^2 - u^3 v^3. \tag{7.34}$$

It is straightforward to show that in any coordinate system, the square of the magnitude of the four-velocity is

$$u \cdot u = c^2. \tag{7.35}$$

The 4-momentum can be defined from Eq. (7.27)

$$p = mu, \tag{7.36}$$

where the mass, m, is a scalar. So the length squared of the four-momentum is

$$p \cdot p = m^2 c^2, \tag{7.37}$$

or from Eqs. (7.27) and (7.34),

$$p \cdot p = m^2 c^2 = m^2 c^2 \gamma^2 - m^2 v^2 \gamma^2 = \frac{E^2}{c^2} - \mathbf{p}^2 \tag{7.38}$$

where $\mathbf{p}$ is the length of the 3-momentum. This last form of Eq. (7.38) is often written as

$$E^2 = m^2c^4 + \mathbf{p}^2c^2. \tag{7.38'}$$

The relativistic kinetic energy, T, is defined as

$$T = E - mc^2 = mc^2(\gamma - 1) \tag{7.39}$$

$$= \sqrt{(mc^2)^2 + \mathbf{p}^2c^2} - mc^2. \tag{7.39'}$$

For $\beta \ll 1$, a power series expansion gives

$$T = \tfrac{1}{2}mv^2 + \mathrm{O}(\beta^4). \tag{7.40}$$

Since $\mathbf{p} = m\gamma\mathbf{v}$, Eq. (7.39) shows that the kinetic energy of a body with finite rest mass tends to infinity as the speed approaches that of light (as $\beta \to 1$, $\gamma \to \infty$). In other words, it takes an infinite amount of energy to increase the speed of a mass particle (or a space ship) from any velocity less than c to c itself. This is another proof that it is impossible to attain or exceed the speed of light starting from any finite speed less than c.

■ FORCES IN THE SPECIAL THEORY; ELECTROMAGNETISM

The preceding material has been concerned with the kinematics of the special theory. The dynamics of the theory follows from the assumption that Newton's laws are correct for objects at rest in the rest frame of the observer, nearly correct for objects moving slowly relative to the speed of light, and require generalizations to covariant equations. The correct generalization of the three-velocity to the four-velocity was given in Eq. (7.27). So we must generalize the force law,

$$\mathbf{F}^i = \frac{d(m\mathbf{v}^i)}{dt}, \tag{7.65}$$

to a covariant form.

Since Maxwell's equations are assumed to be a correct description, we shall briefly consider a covariant reformulation of electromagnetic theory as a guide for the correct form of the force laws of mechanics. The vector and scalar electromagnetic potentials form a four-vector $A^\mu = (\phi/c, \boldsymbol{A})$. If the potentials satisfy the Lorentz condition (in SI units), which is the vanishing of the four-divergence of the electromagnetic potential 4-vector,

$$\Box \cdot A = \boldsymbol{\nabla} \cdot A = \frac{\partial A^\mu}{\partial x^\mu} = \boldsymbol{\nabla} \cdot \mathbf{A} + \mu_0\varepsilon_0 \frac{\partial \phi}{\partial t} = 0, \tag{7.66}$$

they separately satisfy the wave equations of the form (where $\mu_0\varepsilon_0 = 1/c^2$)

$$\Box^2\mathbf{A} = \boldsymbol{\nabla}^2\mathbf{A} = \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = \mu_0\mathbf{j} \tag{7.67a}$$

for the space components and for the time component

$$\Box^2\phi = \nabla^2\phi = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} - \nabla^2\phi = \frac{\rho}{\varepsilon_0}. \tag{7.67b}$$

In terms of ϕ and $\mathbf{A}$, the Lorentz force is $\mathbf{F} = e\{-\nabla\phi + \frac{1}{c^2}\frac{\partial \mathbf{A}}{\partial t} + \frac{1}{c}[\mathbf{v} \times (\nabla \times \mathbf{A})]\}$. This suggests that we should generalize the Lorentz force law to

$$\frac{dp_\mu}{d\tau} = e\left(\frac{\partial(u^\nu A_\nu)}{\partial x^\mu} - \frac{dA_\mu}{d\tau}\right). \tag{7.68}$$

For the three-momentum, $\boldsymbol{p}_3$, and three-velocity, $\boldsymbol{v}$, Eq. (7.68) becomes

$$\frac{d\boldsymbol{p}_3}{dt} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{7.68′}$$

with $\boldsymbol{E}$ the electric field, $\boldsymbol{B}$ the magnetic field, and e the electric charge. The geometric approach is to define a tensor F, named **Faraday**, whose components will be the electromagnetic field tensor and write, with u the 4-velocity,

$$\frac{dp}{d\tau} = eF(u). \tag{7.69}$$

In component notation, this becomes

$$\frac{dp^\mu}{d\tau} = eF^\mu{}_\beta u^\beta. \tag{7.70}$$

This produces Maxwell's equations, provided (according to Eq. (7.68)) $F^\alpha{}_\beta$ is given by

$$F^\alpha{}_\beta = \begin{pmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & cB_z & -cB_y \\ E_y & -cB_z & 0 & cB_x \\ E_z & cB_y & -cB_x & 0 \end{pmatrix}. \tag{7.71}$$

In Minkowski space, the indices are raised and lowered by the metric tensor (Eq. (7.33)), so

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix}, \tag{7.71′}$$

and

$$F_{\alpha\beta} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix}. \tag{7.71″}$$

The Faraday tensor can be written in at least two different ways using either the tensor product, Eq. (7.58), or the wedge product, Eq. (7.64), as

$$\mathbf{F} = F_{\alpha\beta}\, \boldsymbol{dx}^{\alpha} \otimes \boldsymbol{dx}^{\beta} = \tfrac{1}{2} F_{\alpha\beta}\, \boldsymbol{dx}^{\alpha} \wedge \boldsymbol{dx}^{\beta}.$$

The latter expression explicitly shows the antisymmetry.

We can write Maxwell's equation in their normal component form using geometric notation:

$$\boldsymbol{\nabla} F = 0 \qquad \text{and} \qquad \boldsymbol{\nabla} \cdot F = J, \tag{7.72}$$

where J is the 4-current density with components $(\rho c, \mathbf{j})$, where ρ is the charge density and $\mathbf{j}$ is the three-current density. The first of these equations produces (using three-dimensional notation) $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ and $\partial \boldsymbol{B}/\partial t + \boldsymbol{\nabla} \times \boldsymbol{E} = 0$, while the second gives $\boldsymbol{\nabla} \cdot \boldsymbol{E} = \rho/\varepsilon_0$ and $(1/c^2)\, \partial \boldsymbol{E}/\partial t - \boldsymbol{\nabla} \times \boldsymbol{B} = -\mu_0 \boldsymbol{j}$.

Following the guide provided by the covariant formulation of electromagnetic theory, the proper generalization of Newton's second law, Eq. (7.65), is

$$\frac{dp^{\mu}}{d\tau} = K^{\mu}, \tag{7.73}$$

where K^{μ} is a 4-vector force, known as the *Minkowski force*. The spatial components of K^{μ} are not the components of the force in Eq. (7.65), but rather they are quantities that reduce to the F^{i} as $\beta \to 0$. The exact form clearly results from the Lorentz transformation properties of the forces present. Some aspects of the 4-force are listed in Table 7.1.

The general question (which cannot be uniquely resolved) is, How do we find the proper relativistic expression for force? Electromagnetism is used to justify the special theory, so we should expect no problem with it. As we saw in the previous paragraphs, this is trivial for electromagnetic forces because the special theory and the Lorentz transformations are constructed to make Maxwell's electromagnetic theory covariant. For example, the electromagnetic force is given by Eq. (7.68) as

$$K_{\mu} = -q \left(\frac{\partial u_{v} A^{v}}{\partial x^{\mu}} - \frac{dA_{\mu}}{d\tau} \right), \tag{7.74}$$

with q the charge on the particles and A_{μ} the components of the four-potential given by $(\phi/c, \boldsymbol{A})$. Note that ϕ is the scalar potential and $\boldsymbol{A}$ is the three-dimensional electromagnetic vector potential. So the ordinary force, F_i, and the spatial component of the Minkowski electromagnetic force, K_i, are related by

$$F_i = K_i \sqrt{1 - \beta^2}. \tag{7.75}$$

What about other forces? Two methods are commonly used to deduce acceptable transformation properties of forces and hence the correct relativistic form of the forces.

The first method is to argue that there are only four fundamental forces in nature—gravitational, weak nuclear, electromagnetic, and strong nuclear. A cor-

The second approach of determining the correct relativistic force is to simply define force as being the time rate of change of the momentum. Then we write

$$\frac{dp_i}{dt} = F_i \tag{7.76}$$

where the p_i in Eq. (7.76) is some relativistic generalization of the Newtonian momentum that reduces to mv_i in the limit of small β. The simplest generalization is the one given in Eq. (7.36). This second approach has thus far failed to produce any results other than those predicted by the first approach.

■ RELATIVISTIC ANGULAR MOMENTUM

In Chapter 1, it was proven that the nonrelativistic angular momentum obeys an equation of motion much like that for the linear momentum, but with torques

replacing forces. It was shown that for an isolated system obeying the law of action and reaction the total angular momentum is conserved, and that in the C-O-M system it is independent of the point of reference. All of these statements have their relativistic counterparts, at times involving some additional restrictions.

For a single particle, let us define an antisymmetric tensor of rank $\binom{2}{0}$ in Minkowski space using the formalism of Eq. (7.64)

$$m = x \wedge p \tag{7.119}$$

whose elements would be

$$m^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu. \tag{7.120}$$

The 3×3 subtensor m^{ij} clearly corresponds, as was seen in Section 5.1, with the spatial angular momentum of the particle. An equation of motion for $m^{\mu\nu}$ can be found by taking its derivative with respect to the particle's proper time and making use of Eq. (7.73) giving

$$\frac{dm}{d\tau} = u \wedge p + x \wedge K = x \wedge K, \tag{7.121}$$

where the first term vanishes by the antisymmetry of the wedge product and K is the Minkowski force. In component notation, Eq. (7.121) becomes

$$\frac{dm^{\mu\nu}}{d\tau} = x^\mu K^\nu - x^\nu K^\mu. \tag{7.122}$$

This suggests we define the relativistic generalization of the torque by

$$N = x \wedge K, \tag{7.123}$$

whose components are

$$N^{\mu\nu} = x^\mu K^\nu - x^\nu K^\mu. \tag{7.124}$$

Thus, m obeys the equations of motion

$$\frac{dm}{d\tau} = N, \tag{7.125}$$

whose component form is

$$\frac{dm^{\mu\nu}}{d\tau} = N^{\mu\nu}, \tag{7.126}$$

with Eq. (1.11) as the nonrelativistic limiting form.

For a system involving a collection of particles, a total angular momentum 4-tensor can be defined (analogously to the total linear momentum 4-vector) as

$$M = \sum_s m_s \tag{7.127}$$

or in component form

$$M^{\mu\nu} = \sum_s m_s^{\mu\nu}, \tag{7.128}$$

The relativistic angular momentum obeys the same kind of theorem regarding translation of the reference point as does its nonrelativistic counterpart. In the definition, Eq. (7.120) or Eq. (7.128), the reference point (really reference "event") is the arbitrary origin of the Lorentz system. With respect to some other reference event $\mathbf{a}_\lambda$, the total angular momentum is

$$M(a_\lambda) = \sum_s (x_s - a_\lambda) \wedge p_s \tag{7.129}$$

$$= M(0) - a_\lambda \wedge P \tag{7.130}$$

As in the nonrelativistic case, the change in the angular momentum components is equal to the angular momentum, relative to the origin, that the whole system would have if it were located at $\mathbf{a}_\lambda$.

In Chapter 1, one particular reference point played an important role—the center of mass. We can find something similar here, at least in one Lorentz frame, by examining the nature of the mixed time and space components of $M^{\mu\nu}$, namely, $M^{0j} = -M^{j0}$. By definition, in some particular Lorentz frame, these components are given by

$$M^{0j} = \sum_s (x_s^0 p_s^j - x_s^j p_s^0) \tag{7.131}$$

$$= c \sum_s \left(t p_s^j - \frac{x_s^j E_s}{c^2} \right). \tag{7.132}$$

In the C-O-M frame, the total linear momentum $p = \sum p_s$ vanishes, and M^{0j} in this frame has the form

$$M^{0j} = -c \sum_s \frac{x_s E_s}{c^2}. \tag{7.133}$$

If the system is such that the total angular momentum is conserved, as described above, then along with other components M is conserved and hence

$$\sum_s x_s^j E_s = \text{constant}.$$

Conservation of total linear momentum means that $E = \sum E_s$ is also conserved. It is therefore possible to define a *spatial* point R_j,

$$R_j = \frac{\sum_s x_s^j E_s}{\sum_s E_s}, \tag{7.134}$$

associated with the system, which is stationary in the C-O-M coordinate frame. In the nonrelativistic limit, where to first approximation $E_s = m_s c^2$, Eq. (7.134) reduces to the usual definition, Eq. (1.21). Thus, a meaningful center of mass (sometimes called *center of energy*) can be defined in special relativity only in terms of the angular-momentum tensor, and only for a particular frame of reference. Finally, it should be noted that by Eq. (7.130) the spatial part of the angular momentum tensor, M, is independent of reference point in the C-O-M system, exactly as in the nonrelativistic case.

Except for the special case of point collisions, we have so far carefully skirted the problem of finding the motion of a relativistic particle given the Minkowski forces. To this more general problem we address ourselves in the next section, within the nominal framework of the Lagrangian formulation.

I INTRODUCTION TO THE GENERAL THEORY OF RELATIVITY

Thus far we have been careful to use the term "special theory of relativity" and not to introduce the term "special relativity," by which we endeavored to make clear that it is the theory that is special, not the relativity. The special theory uses ideal inertial frames that are assumed to exist over all of spacetime. The general theory not only removes that requirement, but also has a spacetime whose nature is part of the solution to the question of motion. To paraphrase John A. Wheeler: "*Matter tells space how to bend, and space returns the compliment by telling matter how to move*." The general theory is often interpreted in terms of non-Euclidean geometry, so terms like geodesic (paths of shortest distance) and curvature of spacetime are often used. In this brief section we can only outline the formalism of the general theory to show how the full tensor notation is used.

Five principles guided Einstein in the development of the general theory:

1. *Mach's principle*—the special theory used inertial frames. E. Mach observed that Newtonian inertial frames were not rotating with respect to the fixed stars. This suggests Mach's principle, whereby inertial properties are determined by the presence of other bodies in the universe.
2. *Principle of equivalence*—whereby the gravitational mass for each body in the universe can be consistently and universally chosen to equal its inertial mass. To the best accuracy of all experiments performed to date, the ratio of the gravitational mass (the mass that appears in Newton's force law for gravity) to the inertial mass (the mass that appears in the second law) of any object is independent of both the total mass and of the composition of the object. This means that no local experiments can distinguish nonrotating free fall in a gravitational field from uniform motion in the absence of any gravitational fields. Likewise, local experiments cannot distinguish between being at rest in a uniform gravitational field and undergoing uniform acceleration in the absence of any gravitational field (that is, in a rocket).
3. *Principle of covariance*—in the special theory, all inertial observers are equivalent. The general theory extends this idea by postulating the principle of covariance. This principle is that all observers, inertial or not, observe the same laws of physics. That means the laws of physics can be expressed in terms of tensors, since tensors are geometric objects defined independent of any coordinate system.
4. *Correspondence principle*—in weak gravitational fields with velocities small compared to light, the general theory should make predictions that approximate the predictions of gravitational behavior in Newtonian mechanics. As gravitational fields go to zero, the correspondence principle states the predictions of the general theory should approach those of the special theory.
5. *Principle of minimal gravitational coupling*—this principle postulates that no terms explicitly containing the curvature should be added in making the transition from the special theory to the general theory.

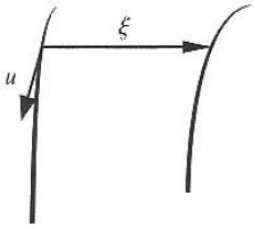


FIGURE 7.5 Tangent vector, u, and deviation vector, ξ.

We shall use the proper time at the tail of the deviation vector and have the head point to where the other test particle is at that time. In general, as the motion progresses, the proper time of the first test particle will not be the same proper time for the other test particle. A straightforward calculation, in the Newtonian limit, for the example of two falling balls, gives for the space components of ξ perpendicular to the direction toward Earth's center,

$$\frac{d^2\boldsymbol{\xi}^i}{dt^2} = R\xi^i, \tag{7.169}$$

where R depends upon the distance to Earth's center and other physical constants. Equation (7.169) says the acceleration in the separation of two geodesics is proportional to their separation. A two-dimensional example is the geodesics on the surface of a sphere. Consider two initially parallel geodesics on a sphere. These geodesics will meet after they have traveled one-quarter of the circumference of the sphere. For this case, Eq. (7.169) has $R = 1/a^2$, where a is the radius of the sphere.

Riemann produces

$$\nabla_{\mathbf{u}} \nabla_{\mathbf{u}} \boldsymbol{\xi} + Riemann(\ldots, u, \xi, u) = 0, \tag{7.170}$$

where $\nabla_{\mathbf{u}} \nabla_{\mathbf{u}} = \frac{D^2}{ds^2}$. In component notation, Eq. (7.170) is

$$\frac{d^2\xi^\alpha}{ds^2} + R^\alpha{}_{\beta\gamma\delta} \frac{dx^\beta}{d\tau} \xi^\gamma \frac{dx^\delta}{d\tau} = 0. \tag{7.171}$$

If we contract *Riemann* on slots 1 and 3, we produce a tensor called *Ricci*, defined as

$$Ricci(u, v) = Riemann(\boldsymbol{w}^\alpha, u, \boldsymbol{e}_\alpha, v), \tag{7.172}$$

whose components are

$$R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}. \tag{7.173}$$

Another critical contraction produces the *curvature scalar*, called R

$$R = Ricci(\boldsymbol{w}^\alpha, \boldsymbol{e}_\alpha) = R^\alpha{}_\alpha. \tag{7.174}$$

Of all these possible contractions of *Riemann*, only one tensor of rank $\binom{2}{0}$ retains all the differential symmetries of *Riemann*. That tensor is called *Einstein* (denoted by G) and is defined as

$$G = Ricci - \tfrac{1}{2} g R, \tag{7.175}$$

with components

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R. \tag{7.176}$$

Using T to denote the stress-energy tensor, Einstein's field equations make *Einstein* proportional to T.

$$G = kT. \tag{7.177}$$

These equations for weak gravitational fields and for speeds much less than light approach Newtonian gravitational theory, and for no gravitational fields produce the results of the special theory. They also correctly predict all the measured first- and second-order corrections to the special theory of relativity in experiments thus far performed. In addition, the theory predicts the existence of gravitational waves from moving masses. Although these waves have not, at this writing, been directly observed, measured changes in the periods of several binary star systems are consistent with the existence of such radiation existing.

14.2. MINKOWSKI SPACE AND LORENTZ TRANSFORMATIONS

Minkowski considered a four dimensional cartesian space in which the position is specified by three coordinates x, y, z and the time is referred by a fourth coordinate ict. If we write $x_1 = x$, $x_2 = y$, $x_3 = z$ and $x_4 = ict$, then an event is represented by the position vector (x_1, x_2, x_3, x_4) in this four dimensional space. Of course the fourth dimension, referring to time, is imaginary. This four dimensional space is called ***Minkowski*** or ***world space***. It is also referred as ***space-time continuum*** and sometimes briefly as ***four-space***. The square of the magnitude of the position vector in such a four-space has the form

$$s^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = x^2 + y^2 + z^2 - c^2t^2 \qquad ...(1)$$

Lorentz transformations are designed so that the speed of light remains constant in S and S' inertial frames (S' is moving with constant velocity v relative to S) and this condition is equivalent to require that the position vector in the four-space is held invariant under the transformations, *i.e.*,

$$s'^2 = x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2$$

or $$s'^2 = x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

or $$s^2 = \sum_{\mu=1}^{4} x_\mu'^2 = \sum_{\mu=1}^{4} x_\mu^2 \qquad ...(2)$$

Deduction of Lorentz Tansformations : In order to prove the statement that the Lorentz transformations can be regarded as orthogonal transformations due to rotation of axes in the Minkowski space, we deduce these transformations in the four-space.

The frame S' is moving with constant velocity v along X-axis relative to the inertial frame S and hence we may have

$$y' = y \quad \text{and} \quad z' = z \quad \text{or} \quad x'_2 = x_2 \quad \text{and} \quad x'_3 = x_3 \qquad \ldots(3)$$

Thus from (2), the transformations should be such that

$$x'^2_1 + x'^2_4 = x_1^2 + x_4^2 \qquad \ldots(4)$$

In order to keep this requirement, we consider two orthogonal coordinate systems X_1X_4 and $X'_1X'_4$ in the same plane (plane of the paper) with the same origin O. The axes of $X'_1X'_4$ system correspond to rotation θ with respect to those of X_1X_4 system, *i.e.*, the axes of the former coordinate system are inclined with the later through an angle θ. We observe that

$$OP^2 = x_1^2 + x_4^2 = x'^2_1 + x'^2_4$$

where the coordinates in two coordinate systems are related as

$$x'_1 = x_1\cos\theta + x_4\sin\theta \qquad \ldots(5)$$

$$x'_4 = -x_1\sin\theta + x_4\cos\theta$$

In matrix notation

$$\begin{pmatrix} x'_1 \\ x'_4 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_4 \end{pmatrix} \qquad \ldots(6)$$

Also, $x_1 = x'_1\cos\theta - x'_4\sin\theta$

$$x_4 = x'_1\sin\theta + x'_4\cos\theta \qquad \ldots(7)$$

When $x'_1 = 0$, $x_1 = -x'_4\sin\theta$

and $x_4 = x'_4\cos\theta$

So that $$\tan\theta = -\frac{x_1}{x_4} = -\frac{x}{ict} = \frac{iv}{c} \qquad \ldots(8)$$

where $x'_1 = x' = 0$ corresponds to the coordinate of the point O' (S'-frame) relative to O (S-frame); *i.e.*, $x = vt$ or $\frac{x}{t} = v$.

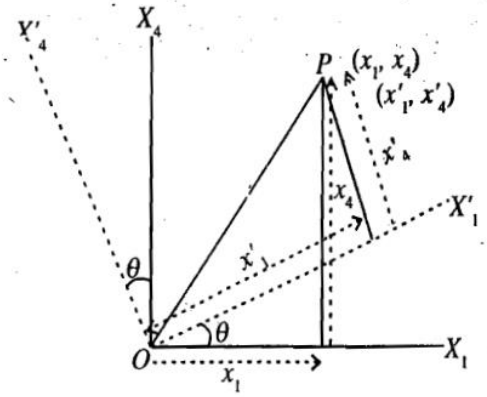


Fig. 14.1 : Rotation of orthogonal coordinates axes and invariance of $OP_2 = x_1^2 + x_4^2 - x'^2_1 + x'^2_4$

Therefore from (8),

$$\sin\theta = \frac{iv/c}{\sqrt{1-v^2/c^2}} = \frac{i\gamma v}{c} \text{ and } \cos\theta = \frac{1}{\sqrt{1-v^2/c^2}} = \gamma \quad \text{(say)}$$

Hence eqs. (5) can be expressed as

$$x'_1 = \gamma x_1 + i\gamma \frac{v}{c} x_4 = \gamma\left(-x_1 + i\frac{v}{c}x_4\right) \text{ and } x'_4 = -i\gamma\frac{v}{c}x_1 + \gamma x_4 = \gamma\left(-\frac{iv}{c}x_1 + x_4\right).$$

If we add $x'_2 = x_2$, and $x'_3 = x_3$, the transformation equations are

$$x'_1 = \gamma\left(x_1 + i\frac{v}{c}x_4\right), x'_2 = x_2, x'_3 = x_3, \text{ and } x'_4 = \gamma\left(-i\frac{v}{c}x_1 + x_4\right) \quad ...(9)$$

In fact, these are the Lorentz transformations. This may be seen by putting $x_1 = x, x_2 = y, x_3 = z$ and $x_4 = ict$ in eq . (9), *i.e.*,

$$x' = \gamma(x - vt),\ y' = y,\ z' = z \text{ and } t' = \gamma\left(t - vx/c^2\right) \quad ...(10a)$$

In matrix notation, the Lorentz transformations from S-frame to S'-frame can be represented as

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad ...(10b)$$

or $$x'_\mu = \sum_{\nu=1}^{4} a_{\mu\nu} x_\nu \quad ...(10c)$$

where $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. and $a_{\mu\nu}$ are the elements of the above square matrix .

The inverse Lorentz transformations are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ i\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} \quad ...(11a)$$

or $$x_\mu = \sum_{\nu=1}^{4} a_{\nu\mu} x'_\nu \quad ...(11b)$$

because $\sum_\nu a_{\nu\mu} x'_\nu = \sum_\nu a_{\nu\mu} \sum_\lambda a_{\nu\lambda} x_\lambda = \sum_\lambda \sum_\nu a_{\nu\mu} a_{\nu\lambda} x_\lambda = \sum_\lambda \delta_{\mu\lambda} x_\lambda = x_\mu$.

Remember that for orthogonal transformations

$$\sum_{\nu} a_{\mu\nu} a_{\lambda\nu} = \sum_{\nu} a_{\nu\mu} a_{\nu\lambda} = \delta_{\mu\lambda}$$

Here x_μ and x'_μ satisfy the condition (2), *i.e.*,

$$\sum_{\mu=1}^{4} x'^{2}_{\mu} = \sum_{\mu=1}^{4} x^2_{\mu}$$

The four coordinates x_1, x_2, x_3 and x_4 or x, y, z and ict, define the position vector in the four-space and may be termed as ***four-position vector***. We shall discuss more about four-vectors later.

15.6. COVARIANCE OF MAXWELL'S FIELD EQUATIONS IN TERMS OF FOUR VECTORS

The covariance (or invariance) of Maxwell's field equations means that these equations have the same form in any two inertial systems.

The Maxwell's field equations in terms of electromagnetic potentials **A** and ϕ are obtained as

$$\Box^2 \mathbf{A} = -\mu_0 \mathbf{j} \qquad ...(31)$$

and $$\Box^2 \phi = -\rho/\epsilon_0 \qquad ...(32)$$

with the Lorentz condition $\text{div}\,\mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$ ...(33)

The electromagnetic four vector potential A_μ and current four vector $\mathbf{j}_\mu$ are represented as

$$A_\mu = (\mathbf{A}, i\phi/c) \text{ or } A_\mu = (A_1, A_2, A_3, A_4,) \text{ with } A_4 = i\phi/c \qquad ...(34)$$

and $$j_\mu = (\mathbf{j}, ic\rho) \text{ or } j_\mu = (j_1, j_2, j_3, j_4) \text{ with } j_4 = ic\rho \qquad ...(35)$$

These four vectors transform like Lorentz transformations and the transformation equations are as follows :

$$A'_1 = \gamma(A_1 - v\phi/c^2) \qquad j_1' = \gamma(j_1 - v\rho)$$
$$A'_2 = A_2 \qquad j_2' = j_2$$
$$A'_3 = A_3 \qquad j_3' = j_3$$
$$\phi' = \gamma(\phi - vA_1) \qquad \rho' = \gamma(\rho - vj_1/c^2)$$

In terms of A_μ and j_μ, the Maxwell's field equations are represented in the form

$$\Box^2 A_\mu = -\mu_0 j_\mu \qquad ...(36a)$$

with the Lorentz condition

$$\sum_{\mu=1}^{4} \frac{\partial A_\mu}{\partial x_\mu} = 0 \qquad ...(36b)$$

Let the eq. (36) be in S-frame. The covariance of Maxwell's field equations requires that in any inertial frame S', eq. (36) must have the same form *i.e.*,

$$\Box^2 A_\mu' = -\mu_0 j'_\mu \qquad \ldots(37a)$$

with Lorentz conditon $\sum_{\mu=1}^{4} \frac{\partial A'_\mu}{\partial x'_\mu} = 0$ $\qquad \ldots(37b)$

In order that the statement (37) is true, let us consider

$$\Box'^2 A_1' = \Box^2 \gamma\left(A_1 - \frac{v\phi}{c^2}\right) \qquad (\text{because } \Box'^2 = \Box^2)$$

$$= \gamma\left[\Box^2 A_1 - \frac{v}{c^2}\Box^2\phi\right] = \gamma\left[-\mu_0 j_1 - \frac{v}{c^2}\left(-\frac{\rho}{\epsilon_0}\right)\right]$$

$$= -\gamma\left[\mu_0 j_1 - \mu_0 v\rho\right] = -\mu_0\gamma\left[j_1 - v\rho\right] = -\mu_0 j_1'$$

$$\Box'^2 A_2' = \Box^2 A_2 = -\mu_0 j_2 = -\mu_0 j_2'$$

$$\Box'^2 A_3' = \Box^2 A_3 = -\mu_0 j_3 = -\mu_0 j_3'$$

$$\Box'^2 A_4' = \Box^2 \frac{i\phi'}{c} = \Box^2 \frac{i}{c}\gamma(\phi - vA_1) = \frac{i\gamma}{\epsilon_0 c}\left[-\frac{\rho}{\epsilon_0} + v\mu_0 j_1\right]$$

$$= \frac{i\gamma}{\epsilon_0 c}\left[\rho - \frac{vj_1}{c^2}\right] = -\frac{i\rho'}{c} = -\mu_0 j_4'$$

Also $A_\mu' = \sum_\nu a_{\mu\nu} A_\nu$, $x_\mu' = \sum_\nu a_{\mu\nu} x_\nu$, hence $\frac{\partial x'_\mu}{\partial x_\nu} = a_{\mu\nu}$

Therefore, $\frac{\partial A'_\mu}{\partial x'_\mu} = \sum_{\mu\nu} a_{\mu\nu}\frac{\partial A_\nu}{\partial x'_\mu} = \sum_{\mu\nu}\frac{\partial x'_\mu}{\partial x_\nu}\frac{\partial A_\nu}{\partial x'_\mu}$

$$= \sum_\nu \frac{\partial A_\nu}{\partial x_\nu} = \sum_\mu \frac{\partial A_\mu}{\partial x_\mu}$$

But $\sum_\mu \frac{\partial A_\mu}{\partial x_\mu} = 0$, hence $\sum_\mu \frac{\partial A'_\mu}{\partial x'_\mu} = 0$

Thus eqs. (37) are same as eqs. (36). This means that the *Maxwell's equations are covariant under Lorentz transformations.*

15.7. THE ELECTROMAGNETIC FIELD TENSOR

The electrromagnetic field vectors **E** and **B** can be expressed in terms of electromagnetic potential **A** and ϕ as

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi \qquad \ldots(38)$$

The vector **B** has three components B_x, B_y, B_z (or B_1, B_2, B_3,) and **E** has E_x, E_y, E_z (or E_1, E_2, E_3) components. **A** also has three components A_x, A_y, A_z (or A_1, A_2, A_3).

Now $$\mathbf{B} = \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ A_1 & A_2 & A_3 \end{bmatrix}$$

Therefore, $$B_x = \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} = F_{23} \quad \text{...(39a)}$$

$$B_y = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} = F_{31} \quad \text{...(39b)}$$

$$B_z = \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} = F_{12} \quad \text{...(39c)}$$

Also from (38) $E_x = \frac{\partial A_1}{\partial t} - \frac{\partial \phi}{\partial x}$

or $$\frac{iE_x}{c} = -\frac{i}{c}\frac{\partial A_1}{\partial t} - \frac{i}{c}\frac{\partial \phi}{\partial x} = \frac{\partial A_1}{\partial (ict)} - \frac{\partial (i\phi/c)}{\partial x_1}$$

Thus $$\frac{iE_x}{c} = \frac{\partial A_1}{\partial x_4} - \frac{\partial A_4}{\partial x_1} = F_{41} \text{ (say)} \quad \text{...(40a)}$$

Similarly, $$\frac{iE_y}{c} = \frac{\partial A_2}{\partial x_4} - \frac{\partial A_4}{\partial x_2} = F_{42} \quad \text{...(40b)}$$

and $$\frac{iE_z}{c} = \frac{\partial A_3}{\partial x_4} - \frac{\partial A_4}{\partial x_3} = F_{43} \quad \text{...(40c)}$$

here in general $$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} \quad \text{...(41)}$$

We observe that $F_{\mu\nu} = -F_{\nu\mu}$ and $F_{\mu\mu} = 0$ ...(42)

Thus we can form an antisymmetric tensor whose components are given by

$$F_{\mu\nu} = \begin{bmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{bmatrix} = \begin{bmatrix} 0 & B_z & -B_y & -\frac{iE_x}{c} \\ -B_z & 0 & B_x & -\frac{iE_y}{c} \\ B_y & -B_x & 0 & -\frac{iE_z}{c} \\ \frac{iE_x}{c} & \frac{iE_y}{c} & \frac{iE_z}{c} & 0 \end{bmatrix} \quad \text{...(43)}$$

This is called the ***electromagnetic field tensor of rank two***.

15.9. COVARIANT FORM OF MAXWELL'S FIELD EQUATION IN TERMS OF ELECTROMAGNETIC FIELD TENSOR

Maxwell's equations in free space are given by

$$(i)\ \text{div}\ \mathbf{E} = \frac{\rho}{\epsilon_0}, \qquad (ii)\ \text{div}\ \mathbf{B} = 0$$

$$(iii)\ \text{curl}\ \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad (iv)\ \text{curl}\ \mathbf{B} = \mu_0\left(\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j}\right) \qquad ...(52)$$

Let us consider eqs. (*i*) and (*iv*)

$$\nabla \times \mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{j} \text{ and } \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

or $$\nabla \times \mathbf{B} = \frac{\partial (i\mathbf{E}/c)}{\partial (ict)} + \mu_0 \mathbf{j} \text{ and } \nabla \cdot \frac{i\mathbf{E}}{c} = \frac{i\rho}{\epsilon_0 c}$$

or $$\nabla \times \mathbf{B} - \frac{\partial (i\mathbf{E}/c)}{\partial x_4} = \mu_0 \mathbf{j} \text{ and } \nabla \cdot \frac{i\mathbf{E}}{c} = \mu_0 j_4$$

Writing in component form, the two equations are

$$\left.\begin{aligned}
0 + \frac{\partial B_z}{\partial x_2} - \frac{\partial B_y}{\partial x_3} + \frac{\partial (iE_x/c)}{\partial x_4} &= \mu_0 j_1 \\
-\frac{\partial B_z}{\partial x_1} + 0 + \frac{\partial B_x}{\partial x_3} + \frac{\partial (-iE_y/c)}{\partial x_4} &= \mu_0 j_2 \\
\frac{\partial B_y}{\partial x_1} - \frac{\partial B_x}{\partial x_2} + 0 + \frac{\partial (-iE_z/c)}{\partial x_4} &= \mu_0 j_3 \\
\frac{\partial (iE_x/c)}{\partial x_1} + \frac{\partial (iE_y/c)}{\partial x_2} + \frac{\partial (iE_z/c)}{\partial x_3} + 0 &= \mu_0 j_4
\end{aligned}\right\} \qquad ...(53)$$

Introducing the electromagnetic field tensor given by

$$F_{\mu\nu} = \begin{pmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{pmatrix} = \begin{pmatrix} 0 & B_z & -B_y & -iE_x/c \\ -B_z & 0 & B_x & -iE_y/c \\ B_y & -B_x & 0 & -iE_z/c \\ iE_x/c & iE_y/c & iE_z/c & 0 \end{pmatrix} \quad ...(54)$$

Hence eqs. (53) are

$$\frac{\partial F_{11}}{\partial x_1} + \frac{\partial F_{12}}{\partial x_2} + \frac{\partial F_{13}}{\partial x_3} + \frac{\partial F_{14}}{\partial x_4} = \mu_0 j_1$$

$$\frac{\partial F_{21}}{\partial x_1} + \frac{\partial F_{22}}{\partial x_2} + \frac{\partial F_{23}}{\partial x_3} + \frac{\partial F_{24}}{\partial x_4} = \mu_0 j_2$$

$$\frac{\partial F_{31}}{\partial x_1} + \frac{\partial F_{32}}{\partial x_2} + \frac{\partial F_{33}}{\partial x_3} + \frac{\partial F_{34}}{\partial x_4} = \mu_0 j_3$$

$$\frac{\partial F_{41}}{\partial x_1} + \frac{\partial F_{42}}{\partial x_2} + \frac{\partial F_{43}}{\partial x_3} + \frac{\partial F_{44}}{\partial x_4} = \mu_0 j_4$$

These equations and hence the Maxwell's field equations (*i*) and (*iv*) are obtained in the compact form

$$\sum_{\nu=1}^{4} \frac{\partial F_{\mu\nu}}{\partial x_\nu} = \mu_0 j_\mu \quad ...(55)$$

Maxwell's equations (ii) and (iii) are written as

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \text{ and } \nabla \cdot \mathbf{B} = 0$$

First of these equations can be written as

$$\nabla \times \mathbf{E} + ic\frac{\partial \mathbf{B}}{\partial x_4} = 0 \text{ or } \nabla \times \left(\frac{-i\mathbf{E}}{c}\right) + \frac{\partial \mathbf{B}}{\partial x_4} = 0$$

Thus the two equations can be written as

$$\frac{\partial(-iE_z/c)}{\partial x_2} + \frac{\partial(iE_y/c)}{\partial x_3} + \frac{\partial B_x}{\partial x_4} = 0$$

$$\frac{\partial(iE_z/c)}{\partial x_1} + \frac{\partial(iE_x/c)}{\partial x_3} + \frac{\partial B_y}{\partial x_4} = 0$$

$$\frac{\partial(-iE_y/c)}{\partial x_1} + \frac{\partial(iE_x/c)}{\partial x_2} + \frac{\partial B_z}{\partial x_4} = 0$$

and $$\frac{\partial B_x}{\partial x_1} + \frac{\partial B_y}{\partial x_2} + \frac{\partial B_z}{\partial x_3} = 0$$

which can be written as

$$0+\frac{\partial F_{34}}{\partial x_2}+\frac{\partial F_{42}}{\partial x_3}+\frac{\partial F_{23}}{\partial x_4}=0$$

$$\frac{\partial F_{43}}{\partial x_1}+0+\frac{\partial F_{14}}{\partial x_3}+\frac{\partial F_{31}}{\partial x_4}=0$$

$$\frac{\partial F_{24}}{\partial x_1}+\frac{\partial F_{41}}{\partial x_2}+0+\frac{\partial F_{12}}{\partial x_4}=0$$

and $$\frac{\partial F_{23}}{\partial x_1}+\frac{\partial F_{31}}{\partial x_2}+\frac{\partial F_{12}}{\partial x_3}+0=0$$

These equations and hence the Maxwell's equations (*ii*) and (*iii*) can be written in compact form as

$$\frac{\partial F_{\lambda\mu}}{\partial x_\nu}+\frac{\partial F_{\mu\nu}}{\partial x_\lambda}+\frac{\partial F_{\nu\lambda}}{\partial x_\mu}=0 \qquad \ldots(56)$$

Equations (55) and (56) represent the Maxwell's field equations in terms of electromagnetic field tensor $F_{\mu\nu}$ defined by (54). As tensor equations are invariant under coordinate transformations, eqs. (55) and (56) represent the ***covariant form of the Maxwell's field equations.***

****

KARPAGAM ACADEMY OF HIGHER EDUCATION,COIMBATORE-21
DEPARTMENT OF PHYSICS
I M.Sc., PHYSICS (2019-2021)
CLASSICAL MECHANICS AND RELATIVITY (19PHP103)

UNIT - I

QUESTIONS

Total energy of body is sum of

Energy can neither be created nor be destroyed, but it can be changed from one form

An artificial Satellite revolves round the Earth in circular orbit, which quantity remai

A man presses more weight on earth at :

The rotational effect of a force on a body about an axis of rotation is described in terr

If no external force acts on a system of bodies, the total linear momentum of the syst

Which law is also called the law of inertia ?

Energy possessed by a body in motion is called

Lagrangian L =

The path adopted by the system during its motion can be represented by a space of __

Co-ordinate transformation equations should not involve _____ explicitly

The frequency of Harmonic oscillator is given by

If the total energy of the particle is conserved then,

Constraint relations do not depend on time is

Constraint relations depend on time is

Constraint relations can be made independent of velocities

The Branchistochrone problem is to find

"If no external torque is applied on a body, then total angular momentum remains cor

Which one of the following choices is an example of a non-conservative force?

Which one of the following choices is an example of a conservative force?

A man of mass 50 kg jumps to a height of 1 m. His potential energy at the highest p

The type of energy possessed by a simple pendulum, when it is at the mean position

If air resistance is negligible, the sum total of potential and kinetic energies of a free

Name the physical quantity which is equal to the product of force and velocity.

The P.E. of a body at a certain height is 200 J. The kinetic energy possessed by it wh

The point, through which the whole weight of the body acts, irrespective of its positi

According to the law of moments, if a number of coplaner forces acting on a particle

The motion of a particle round a fixed axis is

The principle of transmissibility of forces states that, when a force acts upon a body,

The centre of gravity of a semi-circle lies at a distance of _____ from its base r

Concurrent forces are those forces whose lines of action

The velocity ratio in case of an inclined plane inclined at angle θ to the horizontal an
One complete round trip of a vibrating body about it's mean position is

Potential energy of mass attached to spring at mean position is

Velocity of bob in SHM becomes zero at
If potential energies and kinetic energies are equal then displacement of an object in S
Kinetic energy of mass attached to spring at extreme position is
Potential energy of mass attached to spring at extreme position is
Hamiltonian H =
Advantage of Action and Angle variable is that one can obtain the frequencies of
For non-interacting particle in a quantum state the energy E is given by
Co-ordinate transformation equations should not involve _____ explicity.
Generating function have _____ forms.
Hamilton's principal function is denoted by _____.
Hamilton's characteristic function W is identified as _____.
Hamilton's characteristic function is denoted by _____.
The number of independent ways in which a mechanical system can move without vi

UNIT-II

Canonical transformations are the transformations of
The Hamilton's principle function is a generating function, which give rise to canon
All function whose Poisson bracket with the Hamiltonian vanishes will be
Let L and P represent the matrices of Lagrange and Poisson brackets respectively, th

The given transformation is not canonical when
The function p =1/Q and q = PQ2 is
In point transformation one set of co-ordinates qj to a new set Qj can be expressed
The problem consists on finding the path of a charged particle under the action if a ce
Hamilton – Jacobi method is used to find the solution of problem in
Hamilton equation of motion is
Poisson and Lagrange brackets are _____ under Canonical Transformation
Equation of motion in Poisson bracket from depends on

In Kepler problem, the path of the particle is
In Poisson bracket
In Poisson bracket
In Poisson bracket
In Poisson bracket
In Lagrange bracket
In of Lagrange bracket
In of Lagrange bracket
Poisson bracket of two operator X and Y in quantum mechanics is given by
If the Lagrangian of the system does not contain a paricular co-ordinate q, then

Hamilton-Jacobi is a partial differential equation in _____ variables.
_____ is a partial differential equation in (n+1) variables.
Hamilton's characteristic function W is identified as _____.
Hamilton's characteristic function is denoted by _____.
The number of independent ways in which a mechanical system can move without vi

Path in phase space almost refers to actual _____path.
The one way of obtaining the solution of mechanical problem is to transform _____
If the operators X, Y commute, then [X, Y] = _____.
If [X, Y] = 0, then X and Y behave like _____ variables of classical mechanics
If Poisson bracket of two variables in classical mechanics is zero, then the operators
The Lagrange's bracket is _____ under canonical transformation.
Lagrange's equation of motion are second order equations with _____ degrees
The greatest advantage of action and angle variable is that we can obtain the _____
The generalized co-ordinate conjugate to Jj are called _____.
Jj has the dimension of _____.
If F does not involve time explicitly, then the Poisson bracket of F with H _____
If the Poisson bracket of F with H vanishes then F will be a _____.
If Poisson bracket of momentum with H vanishes, then _____ is conser
If Poisson bracket of momentum with H vanishes, then the co-ordinate momenta is _
Lagrange's bracket does not obey the _____law.
H = _____.
L = _____.
In case of either of the set of conjugate variables with (q, p) or with (Q, P), the value
In new set of co-ordinates all Qj are _____.
In new set of co-ordiantes all Pj are _____.
If H is conserved then the new Hamiltonian K is _____.
An assembly of particles with _____inter-particle distance is called as rigid bod
Degree of freedom to fix the configuration of a rigid body is
These are most useful set of generalised co-ordinates for a rigid body and are angles
Angular momentum of a rigid body is
A mathematical structure having nine components in 3-D is termed as tensor of rank
The rotation about space z-axis (angle f) is called _____
Rotation about intermediate X1 axis (angle q) or line of nodes is called
The rotation about z' axis (angle Y) is called
The variation of angle q is referred as _____of the symmetry axis of the top and i
Precession can be
_____is ordinarily observed with a rapidly spinning top.
In case of _____ top amplitude of nutation is small, nutation is sinusoidal,
The minimum spin angular velocity below which top cannot spin stably about vertica
When wz < wmin then the top begins to

Angular velocity of a rigid body is given by
Angular momentum of a rigid body is L =
The diagonal elements Ixx, Iyy, Izz of inertia _____I are moments of inertia
Tensor I is _____to principal axes

UNIT - III

Rotational kinetic energy of a rigid body is
In certain system of body axes with respect to which the off-diagonal elements
If wz = wz' > wmin, atop will spin with its axis vertical continuously , therefore it is
A rigid body with N particles have _____degrees of freedom.
The configuration of a rigid body with respect to some cartesian co-ordinate system i
The most useful set of generalised co-ordinates for a rigid body are _____ angles.
The transformation worked out through three _____ rotations performed only in
The distance between any two points of a rigid body is _____
A rigid body can possesses simultaneously the translational and _____ motion
A mathematical structure having nine components in three dimensions is termed as a
The products of inertia of all vanish when one of the axes of the body lies along the a
If the symmetry axis of the body is taken as axis of rotation and the origin of body ax
The motion of a rigid body with one point fixed will take place under the action of to
The assembly of particles with fixed inter-particle distance is called_____
The orientation of the body by locating a cartesian set of co-ordinates fixed in the rig
The fixed point in the body which registers its translation and coincident with the cer
The generation of body set of axes from the space set of axes through three successiv
The system of body axes in which off-diagonal elements disappear and the diagonal e
The system of body axes in which off-diagonal elements disappear, and the diagonal
The secular equation of inertia tensor and its solution is called _____
A rigid body can possesses simulataneous the _____ and _____ motion.
Rigid body possessing rotational and translational motion simulataneously will have
If we consider three non-collinear points in a rigid body, then each particle will have
Three non-collinear points in a rigid body will have the total of _____degrees of
All the space set of axis if rotated wbout the space z-axis, then the yz plane takes ___
The inverse transformation matrix from body set of axes to space set of axes is given
The position vector of any point p relative to the origin O of the body set of axes is
The configuration of a rigid body is completely specified by _____degrees of free
If a is the column matrix representing the co-ordinates having single frequency and a'
If a is the column matrix representing the co-ordinates having single frequency and
The generalised co-ordinate in which each one of them executing oscillations of one
In parallel pendula the two pendula oscillates in _____
In parallel pendular, if the two pendula are independent i.e., there is no _____
In paralle pendula _____ force due to spring will come into action.
If the system possesses two identical frequencies, then it is therefore said to be ____
A continuous string has infinite number of normal modes and _____
The use of nomal co-ordinate in the coupled system reduces it to one of a system of

A continuous string has a linear _____
If the system is in stable equilibrium, then the frequency wl2 should be a _____ qua
If wl2 are real and positive, then all co-ordinate always remain _____ for any time
If wl2 are not real and positive, then all the co-ordinate becomes _____ for any tim
The system is said to be unstable if the frequency wl2 are not _____

UNIT - IV

When the forces acting on the particle vanishes, then the particle is said to be in
Potential energy is minimum at stable equilibrium and _____at unstable
In case of stable equilibrium the system undergoes bounded motion and in case of
When a system at stable equilibrium is disturbed its potential energy increases and k
When a system at unstable equilibrium is disturbed its potential energy decreases an
The example for stable equilibrium.
If a slight displacement of a system from its equilibrium results only in small
If a slight displacement of a system from its equilibrium results only in unbounded
The example for unstable equilibrium.
The two modes of motion involving a single frequency are called _____ modes
The eigen frequency in case of oscillatory motion about the point of stable
The generalised co-ordinates each of them executing oscillations of one single
Two pendula in parallel pendula oscillate in phase with frequency
Two pendula in parallel pendula oscillate out of phase with frequency
Triple pendulum is a
Triple pendulum is a degenerate system, since the two normal modes frequency
Example for linear triatomic molecule is
In case of linear triatomic molecule when w1 = 0 , the system undergoes
In case of linear triatomic molecule when w2 = (K/M)1/2 and
In case of linear triatomic molecule when _____ the central atom does not
In linear triatomic molecule when _____ , the end atoms vibrate
The example for continuous system is
A continuous system has _____ number of normal modes of frequency.
If the linear triatomic molecule is stretched symmetrically, the absorption band
A system of mutually interacting particles is called _____
When the forces acting on a particle vanishes, the particle is said to be _____
The two modes of motion involving a single frequency are referred to as the _____
The system of two equal masses joined by identical springs to each other is called___
A system of particles is said to be in stable equilibrium if all the particles _____
The system consists of two identical simple pendula, each of mass m, length l and co
All the other co-ordinates except one co-ordinate are zero for all times, then it corres
If the motion for a given wl2 is completely oscillatory about the position of stable
If the eigenfunctions is imaginary, then the motion is said to be _____ equilibriun
If the solution of equation of motion has one single frequency, then in such a case the
If the parallel pendula move in a vertical plane in equilibrium position, then the two
In the two pendula it can vibrate as if they are independent i.e., there is no stretching

In triple pendulum, if the system possesses two identical frequencies, then it is theref
In linear triatomic molecule, the displacement of all the atoms are in the same directi
The continuous string has infinite number of normal modes and _____
A continuous string has a linear _____
The use of normal co-ordinates in the coupled system reduces it to one of a system of
The volume integral of the function of the Lagrangian functions within the braces
Lagrangian density is a function of _____ and _____ derivative of
The system consists of two equal masses joined by identical springs to each other and
In case of two-coupled oscillators, the potential energy V of the system is the sum of
The force tending to change any generalised co-ordinate depends on the _____ of
If two pendula oscillate in phase, then the frequency of motion is
In case of linear triatomic molecule there exists _____ bond between the central
The system consists of infinite chain of equal mass points spaced equally at a distanc
The continuous system is a function of the continuous variables ____ and _____ to
In discrete system, the continuous variables changes only by_____
The propagation velocity of the wave in continuous system is similar to that velocity
In linear triatomic molecule if the molecule is assymmetrically stretched, then _____
For small oscillation, the displacement of the particles are restricted to _____
The motion with imaginary frequency would give rise to an unbounded exponential r
If the particle oscillates about the equilibrium point performing bound motion, then t
In the conservative force-field, generalised forces acting on each particle must _____
The displacement of the generalised co-ordinates from their equilibrium value will be
Michelson-Morley experiment proves
Michelson-Morley experiment proves that
The special theory of relativity was proposed by
If we transform set into another form of n equations, then it involves only a_____
Michelson-Morley experiment proved that
Special theory of relativity deals with the events in the frames of reference which mo
Michelson-Morley experiment to detect the presence of either is based on the phenon
Michelson and Morley experiment showed that
Length contraction happens only

UNIT - V

The mass of 70 kg man moving in car at 66kmh is
Special theory of relativity treats problems involving
According to special theory of relativity which one is not an absolute quantity
Conversion of solar energy into carbohydrates and starch by leaf of a plant is an exan
A reference frame attached to the earth:
Two photons approach each other, their relative velocity will be
An inertial frame is
All the inertial frames are equivalent" this statement is called the principle of ---------
According to relativity, the length of a rod in motion:

If v = c, the length of a rod in motion is:
According to special theory of relativity:
James travels at high speed from the Earth to the star Alpha Centauri, four light years
Relativity mechanics is applicable for a particle which is moving with a velocityà
The relativistic measurement depends uponà
A frame which is moving with zero acceleration is called
When we specific the place of occurrence of a phenomenon as well as the time of oc
Newton's law's remain unchanged or invariant
The laws of mechanics in all initial frame of reference are
The acceleration of a particle under Galilean transformation is
The mass energy relation was proposed by
The Lorentz transformation will converted to Galilean transformation when the relati
the length of an object is maximum in a reference frame in which it is
the length of a rod in uniform motion relative to an observer
The time interval between two event in a reference in a reference frame which is in n
A moving clock
If the velocity of a moving particle is comparable to velocity of light then the mass o
Einistein's mass energy equation E=mc2 implies that
How fast a particle must travel so that its mass becomes twice its rest mass?
Relative velocity of two particles moving with velocity of light of light in opposite di
For a photon particle which is moving with a velocity of light, the rest mass is
The fictitious force, which acts on particle in motion relative to a rotating frame of re
If the particle is at rest relative to the rotating frame of reference the coriolis force is
When the particle is at a non-rotating of reference the Coriolis force
The Coriolis acceleration on a freely falling body under the action of gravitational fo
According to theory of relative mass of an object is
Radiation with energy that is easily detected as quanta
lf the kinetic energy of a body becomes four times its initial value, the new momentu
Lorentz transformation equations hold for
lf the kinetic energy of a body becomes four times its initial value, the new momentu

If the radius of the earth were to shrink, its mass remaining the same, the value of ac

What do we mean by the straightest possible path between two points on Earth's surf

Which of the following statements is not a prediction of the general theory of relativi

What does the equivalence principle say?

Each of the following is a prediction of the theory of relativity. Which one is crucial

According to general relativity, how is time affected by gravity?

According to general relativity, a black hole is

According to general relativity, why does Earth orbit the Sun?
If you draw a spacetime diagram, the worldline of an object that is accelerating away
If you draw a spacetime diagram, the worldline of an object that is traveling by you a
If you draw a spacetime diagram, the worldline of an object that is stationary in your

What do we mean by dimension in the context of relativity?

Suppose you claim that you are feeling the effects of a gravitational field. How can y

Einstein's Theory of General Relativity states that
Einstein said that gravity exists because
According to Einstein, what is considered the fourth dimension?

Einstein's famous equation E = mc2 states that
A person is riding a moped that is traveling at 20.0 m/s. What is the speed of a ball if
A beam of light travels at 3.00 × 108 m/s. If a moped moving at 20.0 m/s turns on its

Einstein's Second Postulate of Special Relativity states that the speed of light
A particular task requires 3.46 J of energy. Using E = mc2, how much mass is neede
Mass of 700 N man moving in car at 66 kmh^{-1} is

Special theory of relativity treats problems involving

A	B	C	D
kinetic energies	potential energies.	forces.	both a and b.
kinetic energy.	potential energies.	conservation of e	conservation principle.
Angular Momentum	Linear Momentum	Angular Displace	None of these
Sitting position	Standing Position	Lying Position	None of these
Centre of gravity	Centripetal force	Centrifugal force	Moment of force
Newton's first law	Newton's Second Law	Newton's Third I	Principle of conservation of linea
Newton's first law	Newton's Second Law	Newton's Third L	All of these
kinetic energy.	potential energies.	conservation of e	conservation principle.
T-V	T+V	(T-V)2	(T+V)1/2
3N	6N	9N	N
time	position	momentum	velocity
[1/2p(k/m)5/2]	[1/2p(k/m)3/2]	[1/2p(k/m)1/2]	[1/2p(k/m)]
T+V =constant	b. T-V=0	c. T-V =constant	None of these
scleronomic	b. rheonomic	c. unilateral	None of these
scleronomic	b. rheonomic	c. unilateral	None of these
scleronomic	b. rheonomic	c. unilateral	.d holonomic
shape of a curve	blength of a curve c.	elasticity of a cur	None of these
A. law of conservatio	A. law of conservation of	A. law of conserv	A. law of conservation of angula
elastic spring force	kinetic frictional force	torque	gravitational force
elastic spring force	kinetic frictional force	torque	gravitational force
50J	500J	12J	30J
KE	PE	KE+PE	KE-PE
increases	increases	becomes zero	remains the same
WORK	ENERGY	POWER	ACCELERATION
>PE	<PE	P.E	Not Known
centre of mass	centre of percussion	moment of inertia	centre of gravity
the algebraic sum of	the algebraic sum of their	their lines of actic	their algebraic sum is zero
translatory as well as	translatory	rotary	circular
different at different	maximum, if it acts at the	minimum, if it ac	same at every point on its line of
3r/4π	4r/ 3π	3r/ 8	8r/3
meet on the same pla	lie on the same line	meet at one point	none of these

cos θ	sin θ	tan θ	cot θ
frequency	time period	amplitude	vibration
maximum	moderate	zero	minimum
mean position	in air	extreme position	middle of mean and extreme pos
4	0	3	1
maximum	moderate	zero	minimum
maximum	moderate	zero	minimum
T-V	T+V	(T-V)2	(T+V)1/2
Vibratory motion	periodic motion	circular mation	all the above
p/2m	p2/m	p/m	p2/2m
position	momentum	time	force
four	two	three	five
H	K	P	S
kinetic energy	potential energy	work	action A
S	K	W	H
action-angle variable	generalized variables	degrees of freedo	co-ordinates
Phase space	Hillbert space	Minkowski space	Space phase
both constant momen	constant moments only	co-ordinates only	constant momenta and co-ordina
constant of motion	constant of momentum	constant of co-or	all the above
LP = 1	LP = -1	LP = -1/2	LP = 1/2
[Q,P] = 1	[Q,P] = -1	[Q,P] = 1/2	[Q,P] = 0
conjugate	canonical	identical	hyrebolic
Qj = Qj (qj, t)	Qj = -Qj (qj, t)	Qj = Pj (qj, t)	Qj = -Pj (qj, t)
Jacobi problem	cononical problem	Kepler problem	Poission problem
Vibratory motion	periodic motion	circular mation	all the above
convergent	divergent	variant	invariant
convergent	divergent	invariant	variant
position	momentum	time	all the three
circular	parabolic	elliptical	zig-zag
[X,Y] = [Y,X]	[X,Y] = - [Y,X]	[X,Y] = 2[Y,X]	[X,Y] = - 2[Y,X]
[X,X] =0	[X,X] =1	[X,X] =2	[X,X] = -2
[X,Y+Z] = [X,Y] - [X	[X,Y+Z] = [X,Y] * [X,Z]	[X,Y+Z] = [X,Y]	[X,Y+Z] = [X,Y] / [X,Z]
[X,YZ] = Y[X,Z] * [	[X,YZ] = Y[X,Z] - [X,Y]	[X,YZ] = Y[X,Z]	[X,YZ] = Y[X,Z] + [X,Y]Z
[X,qj]Q,P = - [qj,X]Q	[X,qj]Q,P = [qj,X]Q,P	[X,qj]Q,P = 2 [qj,	2[X,qj]Q,P = - [qj,X]Q,P
[X,Y]Q,P = -[X,Y]q,	[X,Y]Q,P = [X,Y]q,p	[X,Y]Q,P = 2[X,Y	[X,Y]Q,P = -2[X,Y]q,p
[X,X]q,p = [X,X]Q,I	[X,X]q,p = [X,X]Q,P = -	[X,X]q,p = [X,X	[X,X]q,p = [X,X]Q,P = 1/2
[X,Y] = - 2p/h[XY-Y	[X,Y] = - 2p/h[XY+YX]	[X,Y] = - p/h[XY	[X,Y] = 2p/h[XY-YX]
cyclic co-ordinates	cylindrical co-ordinates	polar co-ordinates	spherical polar co-ordinates

n	n+1	n-1	n+2
Hamilton-Jacobi equ	Lagrangian	Hamiltonian	Jacobian
kinetic energy	potential energy	work	action A
S	K	W	H
action-angle variable	generalized variables	degrees of freedo	co-ordinates

statistical	N	3N	dynamical
old to new	new to old	new to new	old to old
1	-1	0	-2
statistical	dynamical	proportional	inversely proportional
vanish	be multiplied twice	proportional	commute
invariant	variant	not applicable	exponentially variant
n+1	n	2n+1	3n
displacement	frequencies	total time	accelerations
action variable	dynamic variable	statistical variable	angle variable
angular momentum	angular velocity	linear momentum	linear velocity
is proportional with	is proportional with K	Vanishes	exist
positive value	constant of motion	negative value	same value
linear velocity	energy	angular momentu	linear momentum
cyclic	rotational	irrotational	spherical
associative	kepler’s	commutative	Hamilton's variational law
T- V	T + V	T	V
T + V	T	V	T-V
same	proportional	inversely proporti	exponentially proportional
rotational	irrotational	cyclic	variable
cyclic	constant	rotational	irrotational
same	variable	different	constant of motion
fixed	different	1 mm	2 mm
3	6	5	0
Lagrangian angle	azimuthal angle	Euler’s angle	Larmor's precession angle
L = Iw/2	L = 2Iw	L = Iw2	L = Iw
2	3	4	0
translation	precession	nutation	spin.
translation	precession	nutation	spin.
translation	precession	nutation	spin.
translation	precession	nutation	spin.
slow or fast	always slow	always fast	neither fast nor slow
fast precession	slow precession	slow nutation	fast nutation
slow	rotating	fast	both a & b
wmin = (4mglI1/I32)	wmin = (4mglI1/I32)3/2	wmin = (4mglI1/I	wmin = (4mglI1/I32)1/2
wobble	precesse	nutate	spin.

Vi = w2 x ri	Vi = (w x ri)1/2	Vi = w x ri	Vi = w3 x ri
S m2(ri x Vi)	S m(ri x Vi)2	S m2(ri x Vi)2	S m(ri x Vi)
tensor	vector	scalar	donar
symmetric	antisymmetric	parallel	perpendicular
½ w2 I2	w2 I	½ w2 I	2w2 I.
symmetric	antisymmetric	principal	perpendicular
sleeping top	spinning top	rotating top	symmetric top
2N	3N	N	4N
momentum	inertia	orientation	angular momentum
rotation	specified	auxillary	euler’s
successive	different	independent	dependent
varied	fixed	proportional	exponentially proportional
arbitrary	circular	rotational	orbital
tensor	matrix	covariant tensor	contravariant tensor
rotation	vibration	motion	symmetry
unsymmetry	rotational	symmetry	b and c
displacement	torque	time	rotational motion
fluid	vapor	colloidal	rigid body
body set of axes	space set of axes	both a and b	rotational set of axes
body set of axes	space or external set of ax	rotational set of a	vibrational set of axes
direction cosines	successive angles	rotational angles	Euler’s angles
principle axes	secondary axes	primary axes	catesian axes
principle moment of	secondary moment of ine	moments of inerti	inertia
constant of motion	tensor of rank two	covariant tensor	eigen values
translation and rotati	linear and harmonic	periodic and non-	symmetrical around
polar and cartesian	generalised and canonical	translation and ro	both a and b
four	three	six	nine
six	three	nine	twelve
same	alternate	orthogonal	new
AT	adj (A)	co-factor of A	determinant of A
Different	constant	proportional	both a and c
two	three	six	nine
0	1	a	1
0	1	1	a2
normal co-ordinate	cartesain co-ordinate	polar co-ordinate	rectangular co-ordinate
out or phase	phase	damped motion	undamped motion
unstretching	rarefying	transiting	stretching
impulsive	repulsive	restoring	attractive
degenerate	generate	distorted	in harmonic motion
velocities	frequencies	vibrations	motion
dependent	single	independent	double

velocity	acceleration	displacement	mass density
real	imaginary	complex	integer
infinite	same	different	finite
infinite	finite	equal	exponential
equal	finite	real	infinite

equilibrium	stable equlibrium	unstable	neutral equilibrium
maximum	minimum	zero	infinity
same	unbounded	harmonic	distorted
increases	decreases	zero	constant
increases	decreases	constant	neither increase nor decrease
Bar pendulum at rest	compound pendulum at re	simple pendulum	pendulum in motion
unstable	stable	neither stable nor	neutral
unstable	stable	neither stable nor	neutral
Rod standing on its o	rod stretched on two ends	rod in motion	rod in simple harmonic motion
abnormal	normal	transverse	longitudinal
imaginary	real	complex	whole number
normal co-ordinates	genaral co-ordinates	spherical co-ordir	polar co-ordinates
w = (g/l)1/2	w = (g/l)1/3	w = (g/l)1/4	w = (g/l)
w = (g/l +2k/m)	w = (g/l +2k/m)1/4	w = (g/l +2k/m)	w = (g/l +2k/m)1/2
generate system	stable system	degenerate systen	unstable system
w1 = w2 = (g/l +2k/	w1 = w2 = (g/l +2k/m)1	w1 = w2 = (g/l +	w1 = w2 = (g/l +2k/m)1/2
HPO3	H2SO4	HNO3	CO2
periodic motion	non-periodic motion	translatory motio	SHM motion
Oscillatory motion	translatory motion	periodic motion	SHM motion
w = (K/M)1/2	w = (K/M)	w = (K/M)1/3	w = (K/M)1/4
w = { K/M(1+2m/M)	w = { K/M(1+2m/M)}1/2	w = { K/M(1+2m	w = { K/M(1+2m/M)}4
Continuous string	string stretched at one end	String stretched a	String with load at one end
Finite	infinite	Constant	Same
Ultra-violet region	Infra-red region	Visible region	Microwave region
uncoupled system	Translatory system	Coupled system	harmonic system
Equilibrium	Stable equilibrium	unstable equilibri	Neutral equilibrium
abnormal	normal	Damped	undamped
Uncoupled	single coupled	Three-coupled	two-coupled
rest	periodic motion	damped motion	simple harmonic motion
series pendula	compound pendula	paralled pendula	complex pendula
abnormal	standard	variable	normal
imaginary	Real	complex	integer
unstable	Stable	neutral	neither stable nor neutral
Cartesian	canonical	polar	normal
different	identical	relative to each ot	Away from each other
rest	oscillate infinitely	action	neither action nor oscillate infini

periodic	non-periodic	degenerate	harmonic
unequal	equal	infinite	finite
vibrations	displacement	a & b together	frequencies
momentum	volume density	mass density	specific density
dependent	harmonic	periodic	independent
Hamiltonian	Lagrangian	linear	volume
space and time	angle and r	x and y co-ordian	y and z co-ordinates
damped	harmonic	periodic	undamped
kinetic	potential	rest energy	a & b
velocity	accelecration	displacement	momentum
wl =Ög/l	wl =g/l	wl =1/2pÖg/l	wl =2pÖg/l
Inelastic	covalent	Elastic	ionic
Discontinus	continuous	harmonic	linear
w and t	x,y and z	r and w	x and t
twice	thrice	unity	0
inelastic	elastic	damped	undamped
magnetic	quadrapole	oscillating dipole	both a & b
stable	periodic	non-periodic	small
Uj	Vj	pj	qj
unstable	stable	neutral	neither neutral nor stable
finite	infinite	vanish	a constant
Vj	wj	pj	Uj
The existence of ethe	The non-existence of ethe	None	Ether pervades
The speed of light in	The speed of light is char	None	variable light velocity
Einstein	newton	eigen	galileo
Single	double	triple	more than three
speed of light is relati	there is no preferred fram	earth is an inertial	earth is a non-inertial frame
speed	velocity	acceleration	momentum.
interference	diffraction	polarization	dispersion
Newtonian mechanic	There is an absolute ether	There is no absol	Velocity of light is relative in all
perpendicular to direc	along the direction of mot	parallel to directi	both a and b
70 kg	100 kg	infinite	zero
inertial frame of refer	non-inertial frame of refer	non-accelerated fi	accelerated frame of reference
time	mass	height	both a and b
energy into mass	mass in to energy	momentum into v	velocity into momentum
is an inertial frame by	is an inertial frame becaus	Cannot be an iner	Cannot be an inertial frame beca
c/2	Zero	c/8	c
Accelerated	decelarated	Moving with unif	May be accelerated, decelerated
relative motion	equivalence	inertia	Correspondence.
is same as its rest len	is more than its rest lengtl	is less than its res	may be more or less than or equa

zero	equal to proper length	less than proper le	more than proper length.
speed of light is relati	speed of light is same in a	time is relative	mass is relative
the trip takes more tir	James travels to Alpha Ce	clocks on Earth ar	Alpha Centauri travels to James
Greater than that of li	Less than that of light	Comparable to tha	equal to velocity of light
The state of motion o	The state of motion of the	The quantity that	absolute motion
Non-inertial frame	Inertial frame	rest frame	decelerated frame
a point	an event	an incident	an accident
Under Galilean transf	under lorentz transformati	cartesean transfor	new transformation
same	different	none	variable
invariant	non-variant	none	variable
Newton	Einstein	Kepler	Michelson
v>>c	v=c	v<<c	v=0
at rest	in motion	neither rest nor in	varying speed
appears to be shorte	appears to be lengthened	equal to aboslute	invariant length
Maximum	minimum	zero	varying speed
Runs slower than a st	Runs than a stationary ide	neither slow nor f	very fast
Greater than when it	Smaller than when it is at	Equal	very smaller
Energy disappears to	Mass disappears to reapp	All the above stat	nothing can be done
0.5 c	2 c	0.866 c	0.9c
0	2c	c	3c
0	1	2	3
Coriolis force	Newtonian force	Pseudo force	centripetal force
0	1	10	2
1	0	2	3
Directed towards the	Directed towards the west	directed towards r	directed towards south
depends on particles	speed of light.	volume of object	area of object.
1 eV.	1 keV.	1 MeV.	. 10-10 eV.
Three times the initia	Four times the initial valu	Two times the in	unchanged
Non-relativistic veloc	Relativistic velocities onl	. All velocities: re	Photons only
Three times the initia	Four times the initial valu	. Two times the ir	unchanged
Increase and decreas	Decrease and increase re	Increase at both p	Decrease at both places
a path that actually is	a path that follows a circl	a path that follow	the shortest path between the two
Time runs slightly slo	The Universe has no boun	The curvature of s	Different observers can disagree
Gravity is the same th	The effects of gravity are	All observers mus	The effects of relativity are exact
If you observe someo	Gravity is curvature of spa	E = mc2	Time runs slower on the surface

Time is not affected b	Time is stopped by any gr	Time runs slower	Time is stopped by any gravitatic
an object that cannot	a hole in the observable u	a place where the	a place where light travels faster
Earth is following the vertical.	Earth is following the stra curved.	The mysterious fo horizontal.	The mysterious force that we call slanted.
vertical.	curved.	horizontal.	slanted.
vertical.	curved.	horizontal.	slanted.
the size of an object	the number of independen	the letter used to r	the height of an object
She is weightless bec	She is weightless because	She is weightless	If you are in a gravitational field,
gravity and accelerati	the speed of light is const	physics for accele	physics for nonmoving and movi
massive objects warp	massive objects attract on	light moves randc	of the existence of black holes.
horizontal dimension	curled dimension	me dimension	space dimension
mass is always greate	energy and mass are equiv	energy and the sp	mass and the speed of light are e
20.0 m/s	3.00 × 108 m/s	24.0 m/s	3.00 × 108 m/s + 20.0 m/s
20.0 m/s	3.00 × 108 m/s	3.00 × 108 m/s +	3.00 × 108 m/s – 20.0 m/s
is constant regardless of the speed of the	can increase if the speed of the light source increases.	can decrease if the speed of the observer	randomly changes depending upon its original light source.
3.11 × 1017 kg	3.84 × 10–17 kg	3.46 × 10–8 kg	1.15 × 10–8 kg
70 kg.	100Kg	0	10Kg
inertial frame of reference.	non-inertial frame of reference.	non-accelerated	accelerated frame of reference.

ANSWER
both a and b.
conservation of energy
Angular Momentum
Standing Position
Moment of force

ar momentı Principle of conservation of linear momentum
Newton's first law
kinetic energy.
T-V
6N
time
[1/2p(k/m)1/2]
T+V =constant
scleronomic
rheonomic
unilateral
shape of a curve

r speed. A. law of conservation of angular momentum.
kinetic frictional force
elastic spring force
500J
KE
remains the same
POWER
>PE
centre of gravity

the algebraic sum of their moments about any point in their plane is zero

circular

action minimum, if it acts at the centre of gravity of the body
4r/ 3π
meet at one point

sin θ
vibration
zero

ition extreme position
0
zero
maximum
T+V
periodic motion
p2/2m
time
four
S
action A
W
degrees of freedom

Phase space
tes both constant moments and co-ordinates
constant of motion
LP = -1

[Q,P] = 0
canonical
Qj = Qj (qj, t)
Kepler problem
periodic motion
invariant
invariant
all the three

elliptical
[X,Y] = - [Y,X]
[X,X] =0
[X,Y+Z] = [X,Y] + [X,Z]
[X,YZ] = Y[X,Z] + [X,Y]Z
2[X,qj]Q,P = - [qj,X]Q,P
[X,Y]Q,P = [X,Y]q,p
[X,X]q,p = [X,X]Q,P = 0
[X,Y] = - 2p/h[XY-YX]
cyclic co-ordinates

n+1
Hamilton-Jacobi equation
action A
W
degrees of freedom

dynamical
old to new
0
dynamical
commute
invariant
2n+1
frequencies
angle variable
angular momentum
Vanishes
constant of motion
linear momentum
cyclic
commutative
T + V
T-V
same
cyclic
constant
constant of motion
fixed
6
Euler’s angle
L = Iw
2
precession
nutation
spin.
nutation
always slow
slow precession
fast
wmin = (4mglI1/I32)1/2
wobble

Vi = w x ri
S m(ri x Vi)
tensor
symmetric

½ w2 I
principal
sleeping top
3N
orientation
euler’s
successive
fixed
rotational
tensor
symmetry
symmetry
torque
rigid body
body set of axes
space or external set of axes
Euler’s angles
secondary axes
inertia
eigen values
translation and rotational
translation and rotational
three
nine
new
AT
constant
six
1
1
normal co-ordinate
phase
stretching
restoring
degenerate
frequencies
independent

mass density
real
finite
infinite
real

equilibrium
maximum
unbounded
decreases
increases
Bar pendulum at rest
stable
stable
Rod standing on its one end
normal
real
normal co-ordinates
w = (g/l)
w = (g/l +2k/m)1/3
unstable system
w1 = w2 = (g/l +2k/m)1/3
HNO3
periodic motion
Oscillatory motion
w = (K/M)
w = { K/M(1+2m/M)}
string stretched at one end
infinite
Visible region
uncoupled system
Stable equilibrium
undamped
Uncoupled
damped motion
complex pendula
standard
imaginary
neither stable nor neutral
canonical
relative to each other
tely neither action nor oscillate infinitely

non-periodic
unequal
frequencies
volume density
periodic
Hamiltonian
x and y co-ordiantes
harmonic
a & b
displacement
wl =Ög/l
Elastic
continuous
x and t
unity
elastic
oscillating dipole
small
Uj
stable
vanish

Uj
The non-existence of ether medium (i.e. absolute rest frame)
The speed of light in free space in invariant
Einstein
Single
there is no preferred frame like ether
velocity
interference
cases. There is no absolute ether frame, but all frames are relative
along the direction of motion

70 kg
inertial frame of reference
both a and b

energy into mass
use the ear is an inertial frame by definition
c
or moving Moving with uniform velocity or at rest.
relative motion
l to rest le is less than its rest length

zero
speed of light is same in all inertial frames
over a leng Alpha Centauri travels to James over a length that is shorter than four light years.
Comparable to that of light
The quantity that is being measured
Inertial frame
an event
Under Galilean transformation
same
invariant
Einstein
v<<c
at rest
appears to be shortened when it at rest w.r.t. to the observer
Maximum
Runs slower than a stationary identical clock
Greater than when it is rest
All the above statements are correct except d
0.866 c
c
0
Coriolis force
0
0
Directed towards the east
speed of light.
1 MeV.
Two times the initial value
All velocities: relativistic & non-relativistic
Two times the initial value

Increase at both places

› points the shortest path between the two points

about the Different observers can disagree about the fundamental structure of spacetime.

ly equival The effects of gravity are exactly equivalent to the effects of acceleration.

of the Sun E = mc2

ɔnal field. Time runs slower in stronger gravitational fields.

than the n a hole in the observable universe

l centripeta Earth is following the straightest path possible, but spacetime is curved in such a wa
curved.
slanted.
vertical.

the number of independent directions in which movement is possible

, then she c She is weightless because she is in free-fall.

ng frames gravity and acceleration are equivalent.
massive objects warp space
me dimension

quivalent. energy and mass are equivalent.
24.0 m/s
3.00 × 108 m/s

is constant regardless of the speed of the observer or the light source.
3.84×10^{-17} kg
70 kg.
inertial

y that this path goes around the Sun