Los Alamos Electronic Archives: physics/9909035

CLASSICAL MECHANICS

HARET C. ROSU
rosu@ifug3.ugto.mx

$$\vec{F}_a = M\vec{a}$$

$$\vec{F}_p = \frac{d\vec{p}}{dt} \qquad \vec{F}_v = -\vec{\nabla} V(\vec{r})$$

graduate course

León, Guanajuato, Mexico
v1: September 1999.

CONTENTS

1. THE "MINIMUM" PRINCIPLES

Forward: The history of "minimum" principles in physics is long and interesting. The study of such principles is based on the idea that the nature acts always in such a way that the important physical quantities are minimized whenever a real physical process takes place. The mathematical background for these principles is the variational calculus.

CONTENTS

1. Introduction

The empirical evidence has shown that the motion of a particle in an inertial system is correctly described by Newton's second law $\vec{F} = d\vec{p}/dt$, whenever possible to neglect the relativistic effects. When the particle happens not to be forced to a complicated motion, the Cartesian coordinates are sufficient to describe the movement. If none of these conditions are fulfilled, rather complicated equations of motion are to be expected.
In addition, when the particle moves on a given surface, certain forces called constraint forces must exist to maintain the particle in contact with the surface. Such forces are not so obvious from the phenomenological point of view; they require a separate postulate in Newtonian mechanics, the one of action and reaction. Moreover, other formalisms that may look more general have been developed. These formalisms are equivalent to Newton's laws when applied to simple practical problems, but they provide a general approach for more complicated problems. The Hamilton principle is one of these methods and its corresponding equations of motion are called the Euler-Lagrange equations.
If the Euler-Lagrange equations are to be a consistent and correct description of the dynamics of particles, they should be equivalent to Newton's equations. However, Hamilton's principle can also be applied to phenomena generally not related to Newton's equations. Thus, although HP does not give a new theory, it unifies many different theories which appear as consequences of a simple fundamental postulate.

The first "minimum" principle was developed in the field of optics by Heron of Alexandria about 2,000 years ago. He determined that the law of the reflection of light on a plane mirror was such that the path taken by a light ray to go from a given initial point to a given final point is always the shortest one. However, Heron's minimum path principle does not give the right law of reflection. In 1657, Fermat gave another formulation of the principle by stating that the light ray travels on paths that require the shortest time. Fermat's principle of minimal time led to the right laws of reflection and refraction. The investigations of the minimum principles went on, and in the last half of the XVII century, Newton, Leibniz and Bernoulli brothers initiated the development of the variational calculus. In the following years, Lagrange (1760) was able to give a solid mathematical base to this principle.

In 1828, Gauss developed a method of studying Mechanics by means of his principle of minimum constraint. Finally, in a sequence of works published during 1834-1835, Hamilton presented the dynamical principle of minimum action.This principle has always been the base of all Mechanics and also of a big part of Physics.
Action is a quantity of dimensions of length multiplied by the momentum or energy multiplied by time.

2. The action principle

The most general formulation of the law of motion of mechanical systems is the *action or Hamilton principle*. According to this principle every mechanical system is characterized by a function defined as:

$$L\left(q_1, q_2, ..., q_s, \dot{q_1}, \dot{q_2}, \dot{q_s}, t\right),$$

or shortly $L\left(q, \dot{q}, t\right)$, and the motion of the system satisfies the following condition: assume that at the moments t_1 and t_2 the system is in the positions given by the set of coordinates $q^{(1)}$ y $q^{(2)}$; the system moves between these positions in such a way that the integral

$$S = \int_{t_1}^{t_2} L\left(q, \dot{q}, t\right) dt \tag{1}$$

takes the minimum possible value. The function L is called the *Lagrangian* of the system, and the integral (1) is known as the *action* of the system. The Lagrange function contains only q and $\dot{q}$, and no other higher-order derivatives. This is because the mechanical state is completely defined by its coordinates and velocities.
Let us establish now the difererential equations that determine the minimum of the integral (1). For simplicity we begin by assuming that the system has only one degree of freedom, therefore we are looking for only one function $q\left(t\right)$. Let $q = q\left(t\right)$ be the function for which S is a minimum. This means that S grows when one $q\left(t\right)$ is replaced by an arbitrary function

$$q\left(t\right) + \delta q\left(t\right), \tag{2}$$

where $\delta q\left(t\right)$ is a small function through the interval from t_1 to t_2 [it is called the variation of the function $q\left(t\right)$]. Since at t_1 and t_2 all the functions (2) should take the same values $q^{(1)}$ and $q^{(2)}$, one gets:

$$\delta q\left(t_1\right)=\delta q\left(t_2\right)=0. \tag{3}$$

What makes S change when q is replaced by $q+\delta q$ is given by:

$$\int_{t_1}^{t_2} L\left(q+\delta q,\dot{q}+\delta\,\dot{q},t\right)dt-\int_{t_1}^{t_2} L\left(q,\dot{q},t\right)dt.$$

An expansion in series of this difference in powers of δq and $\delta\,\dot{q}$ begins by terms of first order. The necessary condition of minimum (or, in general, extremum) for S is that the sum of all terms turns to zero; Thus, the action principle can be written down as follows:

$$\delta S=\delta\int_{t_1}^{t_2} L\left(q,\dot{q},t\right)dt=0, \tag{4}$$

or by doing the variation:

$$\int_{t_2}^{t_1}\left(\frac{\partial L}{\partial q}\delta q+\frac{\partial L}{\partial\,\dot{q}}\delta\,\dot{q}\right)dt=0\ .$$

Taking into account that $\delta\,\dot{q}=d/dt\left(\delta q\right)$, we make an integration by parts to get:

$$\delta S=\left[\frac{\partial L}{\partial\,\dot{q}}\delta q\right]_{t_1}^{t_2}+\int_{t_2}^{t_1}\left(\frac{\partial L}{\partial q}-\frac{d}{dt}\frac{\partial L}{\partial\,\dot{q}}\right)\delta q dt=0\ . \tag{5}$$

Considering the conditions (3), the first term of this expresion disappears. Only the integral remains that should be zero for all values of δq. This is possible only if the integrand is zero, which leads to the equation:

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

For more degrees of freedom, the s different functions $q_i(t)$ should vary independently. Thus, it is obvious that one gets s equations of the form:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\,\dot{q}_i}\right)-\frac{\partial L}{\partial q_i}=0\qquad\qquad(i=1,2,...,s) \tag{6}$$

These are the equations we were looking for; in Mechanics they are called *Euler-Lagrange equations*. If the Lagrangian of a given mechanical system

is known, then the equations (6) form the relationship between the accelerations, the velocities and the coordinates; in other words, they are the equations of motion of the system. From the mathematical point of view, the equations (6) form a system of s differential equations of second order for s unknown functions $q_i(t)$. The general solution of the system contains $2s$ arbitrary constants. To determine them means to completely define the movement of the mechanical system. In order to achieve this, it is necessary to know the initial conditions that characterize the state of the system at a given moment (for example, the initial values of the coordinates and velocities.

3. D'Alembert principle

The virtual displacement of a system is the change in its configurational space under an arbitrary infinitesimal variation of the coordinates $\delta\mathbf{r}_i$,*which is compatible with the forces and constraints imposed on the system at the given instant t.* It is called virtual in order to distinguish it from the real one, which takes place in a time interval dt, during which the forces and the constraints can vary.

The constraints introduce two types of difficulties in solving mechanics problems:

(1) Not all the coordinates are independent.

(2) In general, the constraint forces are not known *a priori*; they are some unknowns of the problem and they should be obtained from the solution looked for.

In the case of holonomic constraints the difficulty (1) is avoided by introducing a set of independent coordinates ($q_{1,}q_{2,...,}q_m$, where m is the number of degrees of freedom involved). This means that if there are m constraint equations and $3N$ coordenates $(x_1, ..., x_{3N})$, we can eliminate these n equations by introducing the independent variables $(q_1, q_2, .., , q_n)$. A transformation of the following form is used

$$x_1 = f_1(q_1, ..., q_m, t)$$

$$\vdots$$

$$x_{3N} = f_{3N}(q_1, ..., q_n, t) \ ,$$

where $n = 3N - m$.

To avoid the difficulty (2) Mechanics needs to be formulated in such a way that the forces of constraint *do not occur* in the solution of the problem. This is the essence of the "*principle of virtual work*".

Virtual work: We assume that a system of N particles is described by $3N$ coordenates $(x_1, x_2, ..., x_{3N})$ and let $F_{1,}F_{2,...,}F_{3N}$ be the components of the forces acting on each particle. If the particles of the system display infinitesimal and instantaneous displacements $\delta x_1, \delta x_2, ..., \delta x_{3N}$ under the action of the $3N$ forces, then the performed work is:

$$\delta W = \sum_{j=1}^{3N} F_j \delta x_j \ . \tag{7}$$

Such displacements are known as *virtual displacements* and δW is called *virtual work*; (7*)* can be also written as:

$$\delta W = \sum_{\alpha=1}^{N} \mathbf{F}_\alpha \cdot \delta \mathbf{r} \ . \tag{8}$$

Forces of constraint: besides the applied forces $\mathbf{F}_\alpha^{(e)}$, the particles can be acted on by forces of constraint $\mathbf{F}_\alpha$.
The principle of virtual work: Let $\mathbf{F}_\alpha$ be the force acting on the particle α of the system. If we separate $\mathbf{F}_\alpha$ in a contribution from the outside $\mathbf{F}_\alpha^{(e)}$ and the constraint $\mathbf{R}_\alpha$

$$\mathbf{F}_\alpha = \mathbf{F}_\alpha^{(e)} + \mathbf{R}_\alpha \ . \tag{9}$$

and if the system is in equilibrium, then

$$\mathbf{F}_\alpha = \mathbf{F}_\alpha^{(e)} + \mathbf{R}_\alpha = 0 \ . \tag{10}$$

Thus, the virtual work due to all possible forces $\mathbf{F}_\alpha$ is:

$$W = \sum_{\alpha=1}^{N} \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha = \sum_{\alpha=1}^{N} \left(\mathbf{F}_\alpha^{(e)} + \mathbf{R}_\alpha \right) \cdot \delta \mathbf{r}_\alpha = 0 \ . \tag{11}$$

If the system is such that the constraint forces do not make virtual work, then from (11) we obtain:

$$\sum_{\alpha=1}^{N} \mathbf{F}_\alpha^{(e)} \cdot \delta \mathbf{r}_\alpha = 0 \ . \tag{12}$$

Taking into account the previous definition, we are now ready to introduce the D'Alembert principle. According to Newton, the equation of motion is:

$$\mathbf{F}_\alpha = \dot{\mathbf{p}}_\alpha$$

and can be written in the form

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

which tells that the particles of the system would be in equilibrium under the action of a force equal to the real one plus an inverted force $-\dot{\mathbf{p}}_i$. Instead of (12) we can write

$$\sum_{\alpha=1}^{N} \left(\mathbf{F}_\alpha - \dot{\mathbf{p}}_\alpha \right) \cdot \delta \mathbf{r}_\alpha = 0 \tag{13}$$

and by doing the same decomposition in applied and constraint forces ($\mathbf{f}_\alpha$), we obtain:

$$\sum_{\alpha=1}^{N} \left(\mathbf{F}_\alpha^{(e)} - \dot{\mathbf{p}}_\alpha \right) \cdot \delta \mathbf{r}_\alpha + \sum_{\alpha=1}^{N} \mathbf{f}_\alpha \cdot \delta \mathbf{r}_\alpha = 0 \ .$$

Again, let us limit ourselves to systems for which the virtual work due to the forces of constraint is zero leading to

$$\sum_{\alpha=1}^{N} \left(\mathbf{F}_\alpha^{(e)} - \dot{\mathbf{p}}_\alpha \right) \cdot \delta \mathbf{r}_\alpha = 0 \ , \tag{14}$$

which is the *D'Alembert's principle*. However, this equation does not have a useful form yet for getting the equations of motion of the system. Therefore, we should change the principle to an expression entailing the virtual displacements of the generalized coordinates, which being independent from each other, imply zero coefficients for $\delta \mathbf{q}_\alpha$. Thus, the velocity in terms of the generalized coordinates reads:

$$\mathbf{v}_\alpha = \frac{d\mathbf{r}_\alpha}{dt} = \sum_k \frac{\partial \mathbf{r}_\alpha}{\partial q_k} \dot{q_k} + \frac{\partial \mathbf{r}_\alpha}{\partial t} \qquad \text{where} \qquad \mathbf{r}_\alpha = \mathbf{r}_\alpha \left(q_1, q_2, ..., q_n, t \right) \ .$$

Similarly, the arbitrary virtual displacement $\delta \mathbf{r}_\alpha$ can be related to the virtual displacements $\delta \mathbf{q}_j$ through

$$\delta \mathbf{r}_\alpha = \sum_j \frac{\partial \mathbf{r}_\alpha}{\partial q_j} \delta q_j \ .$$

Then, the virtual work $\mathbf{F}_\alpha$ expressed in terms of the generalized coordinates will be:

$$\sum_{\alpha=1}^{N} \mathbf{F}_\alpha \cdot \delta \mathbf{r}_\alpha = \sum_{j,\alpha} \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_j} \delta q_j = \sum_j Q_j \delta q_j \ , \tag{15}$$

where the Q_j are the so-called components of the generalized force, defined in the form

$$Q_j = \sum_\alpha \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_j} \ .$$

Now if we see eq. (14) as:

$$\sum_\alpha \dot{\mathbf{p}} \cdot \delta \mathbf{r}_\alpha = \sum_\alpha m_\alpha \ddot{\mathbf{r}}_\alpha \cdot \delta \mathbf{r}_\alpha \tag{16}$$

and by substituting in the previous results we can see that (16) can be written:

$$\sum_\alpha \left\{ \frac{d}{dt} \left(m_\alpha \mathbf{v}_\alpha \cdot \frac{\partial \mathbf{v}_\alpha}{\partial \dot{q}_j} \right) - m_\alpha \mathbf{v}_\alpha \cdot \frac{\partial \mathbf{v}_\alpha}{\partial q_j} \right\} = \sum_j \left[\left\{ \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right\} - Q_j \right] \delta q_j = 0 \, . \tag{17}$$

The variables q_j can be an arbitrary system of coordinates describing the motion of the system. However, if the constraints are holonomic, it is possible to find systems of independent coordinates q_j containing implicitly the constraint conditions already in the equations of transformation $x_i = f_i$ if one nullifies the coefficients by separate:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}} \right) - \frac{\partial T}{\partial q_\alpha} = Q_j \ . \tag{18}$$

There are m equations. The equations (18) are sometimes called the Lagrange equations, although this terminology is usually applied to the form they take when the forces are conservative (derived from a scalar potential V)

$$\mathbf{F}_\alpha = -\nabla_i V.$$

Then Q_j can be written as:

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

The equations (18) can also be written in the form:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q_j}}\right) - \frac{\partial (T-V)}{\partial q_j} = 0 \tag{19}$$

and defining the *Lagrangian* L in the form $L = T - V$ one gets

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_j}}\right) - \frac{\partial L}{\partial q_j} = 0 \ . \tag{20}$$

These are the **Lagrange equations of motion**.

4. - Phase space

In the geometrical interpretation of mechanical phenomena the concept of *phase space* is very much in use. It is a space of $2s$ dimensions whose axes of coordinates are the s generalized coordinates and the s momenta of the given system. Each point of this space corresponds to a definite mechanical state of the system. When the system is in motion, the representative point in phase space performs a curve called *phase trajectory.*

5. - Space of configurations

The state of a system composed of n particles under the action of m constraints connecting some of the $3n$ cartesian coordinates is completely determined by $s = 3n - m$ generalized coordinates. Thus, it is possible to descibe the state of such a system by a point in the s dimensional space usually called the *configuration space*, for which each of its dimensions corresponds to one q_j. The time evolution of the system will be represented by a curve in the configuration space made of points describing the instantaneous configuration of the system.

6. - Constraints

One should take into account the *constraints* that act on the motion of the system. The constraints can be classified in various ways. In the general case in which the constraint equations can be written in the form:

$$\sum_i c_{\alpha i}\ \dot{q_i} = 0 \ ,$$

where the $c_{\alpha i}$ are functions of only the coordinates (the index α counts the constraint equations). If the first members of these equations are not total derivatives with respect to the time they cannot be integrated. In other words, they cannot be reduced to relationships between only the coordinates, that might be used to express the position by less coordinates, corresponding to the real number of degrees of freedom. Such constraints are called *non holonomic* (in contrast to the previous ones which are *holonomic* and which connect only the coordinates of the system*)*.

7. Hamilton's equations of motion

The formulation of the laws of Mechanics by means of the Lagrangian assumes that the mechanical state of the system is determined by its generalized coordinates and velocities. However, this is not the unique possible method; the equivalent description in terms of its generalized coordinates and momenta has a number of advantages.

Turning from one set of independent variables to another one can be achieved by what in mathematics is called *Legendre transformation.* In this case the transformation takes the following form where the total differential of the Lagrangian as a function of coordinates and velocities is:

$$dL = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} d\,\dot{q}_i \quad ,$$

that can be written as:

$$dL = \sum_i \dot{p}_i \; dq_i + \sum_i p_i d\,\dot{q}_i \quad , \tag{21}$$

where we already know that the derivatives $\partial L/\partial\, \dot{q}_i$, are by definition the generalized momenta and moreover $\partial L \;/\; \partial q_i = \dot{p}_i$ by Lagrange equations. The second term in eq. (21)can be written as follows

$$\sum_i p_i d\,\dot{q}_i = d\left(\sum p_i\, \dot{q}_i\right) - \sum \dot{q}_i\; dq_i \;.$$

By attaching the total diferential $d\left(\sum p_i\, \dot{q}_i\right)$ to the first term and changing the signs one gets from (21):

$$d\left(\sum p_i\, \dot{q}_i - L\right) = -\sum \dot{p}_i\; dq_i + \sum p_i\, \dot{q}_i \quad . \tag{22}$$

The quantity under the diferential is the energy of the system as a function of the coordinates and momenta and is called *Hamiltonian function or Hamiltonian* of the system:

$$H\left(p,q,t\right)=\sum_{i}p_{i}\,\dot{q}_{i}\,-L\ . \tag{23}$$

Then from ec. (22)

$$dH=-\sum\dot{p}_{i}\,dq_{i}+\sum p_{i}\,\dot{q}_{i}$$

where the independent variables are the coordinates and the momenta, one gets the equations

$$\dot{q}_{i}=\frac{\partial H}{\partial p_{i}}\qquad\qquad\dot{p}_{i}=-\frac{\partial H}{\partial q_{i}}\ . \tag{24}$$

These are the equations of motion in the variables q y p and they are called *Hamilton's equations.*

8. Conservation laws

8.1 Energy

Consider first the conservation theorem resulting from the *homogeneity of time*. Because of this homogeneity, the Lagrangian of a closed system does not depend explicitly on time. Then, the total time diferential of the Lagrangian (not depending explicitly on time) can be written:

$$\frac{dL}{dt}=\sum_{i}\frac{\partial L}{\partial q_{i}}\,\dot{q}_{i}+\sum_{i}\frac{\partial L}{\partial\,\dot{q}_{i}}\,\ddot{q}_{i}$$

and according to the Lagrange equations we can rewrite the previous equation as follows:

$$\frac{dL}{dt}=\sum_{i}\dot{q}_{i}\,\frac{d}{dt}\left(\frac{\partial L}{\partial\,\dot{q}_{i}}\right)+\sum_{i}\frac{\partial L}{\partial\,\dot{q}_{i}}\,\ddot{q}_{i}=\sum_{i}\frac{d}{dt}\left(\dot{q}_{i}\,\frac{\partial L}{\partial\,\dot{q}_{i}}\right)\ ,$$

or

$$\sum_{i}\frac{d}{dt}\left(\dot{q}_{i}\,\frac{\partial L}{\partial\,\dot{q}_{i}}-L\right)=0\ .$$

From this one concludes that the quantity

$$E\equiv\sum_{i}\dot{q}_{i}\,\frac{\partial L}{\partial\,\dot{q}_{i}}-L \tag{25}$$

remains constant during the movement of the closed system, that is it is an integral of motion. This constant quantity is called the *energy* E of the system.

8.2 Momentum

The *homogeneity of space* implies another conservation theorem. Because of this homogeneity, the mechanical properties of a cosed system do not vary under a parallel displacement of the system as a whole through space. We consider an infinitesimal displacement ϵ (i.e., the position vectors $\mathbf{r}_\alpha$ turned into $\mathbf{r}_a + \epsilon$) and look for the condition for which the Lagrangian does not change. The variation of the function L resulting from the infinitesimal change of the coordinates (maintaining constant the velocities of the particles) is given by:

$$\delta L = \sum_a \frac{\partial L}{\partial \mathbf{r}_a} \cdot \delta \mathbf{r}_a = \epsilon \cdot \sum_a \frac{\partial L}{\partial \mathbf{r}_a} \ ,$$

extending the sum over all the particles of the system. Since ϵ is arbitrary, the condition $\delta L = 0$ is equivalent to

$$\sum_a \frac{\partial L}{\partial \mathbf{r}_a} = 0 \tag{26}$$

and taking into account the already mentioned equations of Lagrange

$$\sum_a \frac{d}{dt}\left(\frac{\partial L}{\partial \mathbf{v}_a}\right) = \frac{d}{dt}\sum_a \frac{\partial L}{\partial \mathbf{v}_a} = 0 \ .$$

Thus, we reach the conclusion that for a closed mechanical system the vectorial quantity called *impetus/momentum*

$$\mathbf{P} \equiv \sum_a \frac{\partial L}{\partial \mathbf{v}_a}$$

remains constant during the motion.

8.3 Angular momentum

Let us study now the conservation theorem coming out from *the isotropy of space*. For this we consider an infinitesimal rotation of the system and look for the condition under which the Lagrangian does not change.

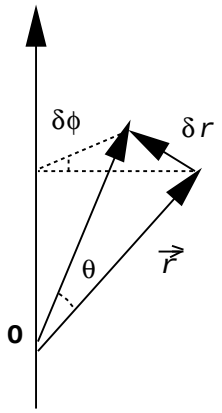
We shall call an infinitesimal rotation vector $\delta\phi$ a vector of modulus equal to the angle of rotation $\delta\phi$ and whoose direction coincide with that of the rotation axis. We shall look first to the increment of the position vector

of a particle in the system, by taking the origin of coordinates on the axis of rotation. The lineal displacement of the position vector as a function of angle is

$$|\delta\mathbf{r}| = r\sin\theta\delta\phi \ ,$$

(see the figure). The direction of the vector $\delta\mathbf{r}$ is perpendicular to the plane defined by $\mathbf{r}$ and $\delta\phi$, and therefore,

$$\delta\mathbf{r} = \delta\phi \times \mathbf{r} \ . \tag{27}$$



The rotation of the system changes not only the directions of the position vectors but also the velocities of the particles that are modified by the same rule for all the vectors. The velocity increment with respect to a fixed frame system will be:

$$\delta\mathbf{v} = \delta\phi \times \mathbf{v} \ .$$

We apply now to these expressions the condition that the Lagrangian does not vary under rotation:

$$\delta L = \sum_a \left(\frac{\partial L}{\partial \mathbf{r}_a} \cdot \delta\mathbf{r}_a + \frac{\partial L}{\partial \mathbf{v}_a} \cdot \delta\mathbf{v}_a \right) = 0$$

and substituting the definitions of the derivatives $\partial L/\partial \mathbf{v}_a$ por $\mathbf{p}_a$ and $\partial L/\partial \mathbf{r}_a$ from the Lagrange equations by $\dot{\mathbf{p}}_a$; we get

$$\sum_a \left(\dot{\mathbf{p}}_a \cdot \delta\phi \times \mathbf{r}_a + \mathbf{p}_a \cdot \delta\phi \times \mathbf{v}_a \right) = 0 \ ,$$

or by circular permutation of the factors and getting $\delta\phi$ out of the sum:

$$\delta\phi \sum_a \left(\mathbf{r}_a \times \dot{\mathbf{p}}_a + \mathbf{v}_a \times \mathbf{p}_a \right) = \delta\phi \cdot \frac{d}{dt} \sum_a \mathbf{r}_a \times \mathbf{p}_a = 0 \ ,$$

because $\delta\phi$ is arbitrary, one gets

$$\frac{d}{dt} \sum_a \mathbf{r}_a \times \mathbf{p}_a = 0$$

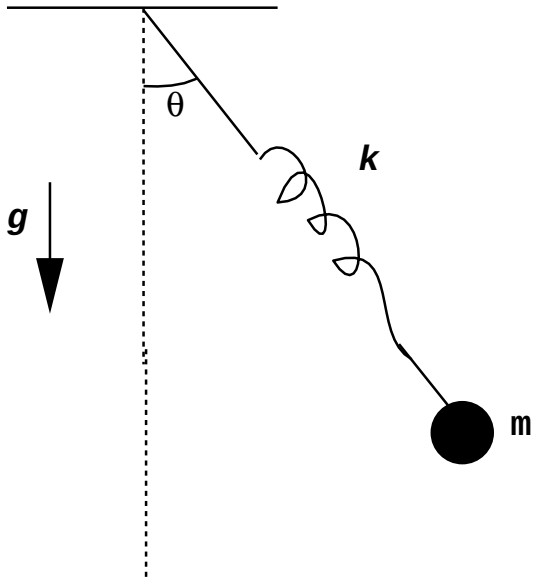
Thus, the conclusion is that during the motion of a closed system the vectorial quantity called the *angular (or kinetic) momentum* is conserved.

$$M \equiv \sum_a \mathbf{r}_a \times \mathbf{p}_a \ .$$

9.- Applications of the action principle

a) Equations of motion

Find the eqs of motion for a pendular mass sustained by a resort, by directly applying Hamilton's principle.



For the pendulum in the figure the Lagrangian function is

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\,\dot{\theta}^2) + mgr\cos\theta - \frac{1}{2}k(r - r_o)^2 \ ,$$

therefore

$$\int_{t_1}^{t_2} \delta L dt = \int_{t_1}^{t_2} \left[m\left(\dot{r}\ \delta\ \dot{r} + r\ \dot{\theta}^2 + r^2\ \dot{\theta}\ \delta\ \dot{\theta}\right) + mg\delta r\cos\theta - mgr\delta\theta\sin\theta - k(r - r_o)\delta r \right] dt$$

$$m\ \dot{r}\ \delta\ \dot{r}\ dt = m\ \dot{r}\ d(\delta r) = d\left(m\ \dot{r}\ \delta r\right) - m\delta r\ \ddot{r}\ dt \ .$$

In the same way

$$mr^2\theta^2\delta\ \dot{\theta}\ dt = d\left(mr^2\ \dot{\theta}\ \delta\ \dot{\theta}\right) - \delta\theta\frac{d\left(mr^2\ \dot{\theta}\right)}{dt}dt$$

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

Therefore, the previous integral can be written

$$\int_{t_1}^{t_2} \left[\left\{ m\ \ddot{r} - mr\ \dot{\theta}^2 - mg\cos\theta + k\left(r - r_o\right) \right\} + \left\{ mr^2\ \ddot{\theta} + 2mr\ \dot{r}\dot{\theta} + mgr\sin\theta \right\} \delta\theta \right] dt$$

$$- \int_{t_1}^{t_2} \left[d\left(m\ \dot{r}\ \delta r\right) + d\left(mr^2\theta^2\ \dot{\theta}\ \delta\theta\right) \right] = 0 \ .$$

Assuming that both δr and $\delta\theta$ are equal zero at t_1 and t_2, the second integral is obviously nought. Since δr and $\delta\theta$ are completely independent of each other, the first integral can be zero only if

$$m\ \ddot{r} - mr\ \dot{\theta}^2 - mg\cos\theta + k(r - r_o) = 0$$

and

$$mr^2\ \ddot{\theta} + 2mr\ \dot{r}\dot{\theta} + mgr\sin\theta = 0 \ ,$$

These are the equations of motion of the system.

b) Exemple of calculating a minimum value

Prove that the shortest line between two given points p_1 and p_2 on a cilinder is a helix.

The length S of an arbitrary line on the cilinder between p_1 and p_2 is given by

$$S = \int_{p_1}^{p_2} \left[1 + r^2 \left(\frac{d\theta}{dz} \right)^2 \right]^{1/2} dz \ ,$$

where r, θ and z are the usual cilindrical coordinates for $r = const.$ A relationship between θ and z can be determined for which the last integral has an extremal value by means of

$$\frac{d}{dz} \left(\frac{\partial \phi}{\partial \theta'} \right) - \frac{\partial \phi}{\partial \theta} = 0 \ ,$$

where $\phi = \left[1 + r^2 \theta'^2 \right]^{1/2}$ y $\theta' = \frac{d\theta}{dz}$, but since $\partial \phi / \partial \theta = 0$ we have

$$\frac{\partial \phi}{\partial \theta'} = \left(1 + r^2 \theta'^2 \right)^{-1/2} r^2 \theta' = c_1 = const. \ ,$$

therefore $r\theta' = c_2$. Thus, $r\theta = c_2 z + c_3$, which is the parametric equation of a helix. Assuming that in p_1 we have $\theta = 0$ and $z = 0$, then $c_3 = 0$. In p_2, make $\theta = \theta_2$ and $z = z_2$, therefore $c_2 = r\theta_2 / z_2$, and $r\theta = (r\theta_2 / z_2)\, z$ is the final equation.

References

L. D. Landau and E. M Lifshitz, *Mechanics*, Theoretical Physics, vol I, (Pergammon, 1976)

H. Goldstein, *Classical Mechanics*, (Addison-Wesley, 1992)

2. MOTION IN CENTRAL FORCES

Forward: Because of astronomical reasons, the motion under the action of central forces has been the physical problem on which pioneer researchers focused more, either from the observational standpoint or by trying to disentangle the governing laws of motion. This movement is a basic example for many mathematical formalisms. In its relativistic version, Kepler's problem is yet an area of much interest.

CONTENTS:

2.1 Two-body problem: Reduction to the one-body problem

Consider a system of two material points of masses m_1 and m_2, in which there are forces due only to an interaction potential V. We suppose that V is a function of any position vector between m_1 and m_2, $\mathbf{r}_2 - \mathbf{r}_1$, or of their relative velocities $\dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1$, or of the higher-order derivatives of $\mathbf{r}_2 - \mathbf{r}_1$. Such a system has 6 degrees of freedom and therefore 6 independent generalized coordinates.

We suppose that these are the vector coordinates of the center-of-mass $\mathbf{R}$, plus the three components of the relative difference vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. The Lagrangian of the system can be written in these coordinates as follows:

$$L = T(\dot{\mathbf{R}}, \dot{\mathbf{r}}) - V(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}}, ...). \tag{1}$$

The kinetic energy T is the sum of the kinetic energy of the center-of-mass plus the kinetic energy of the motion around it, T':

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

being

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

Here, $\mathbf{r}_1'$ and $\mathbf{r}_2'$ are the position vectors of the two particles with respect to the center-of-mass, and they are related to $\mathbf{r}$ by means of

$$\mathbf{r}_1' = -\frac{m_2}{m_1 + m_2}\mathbf{r}, \ \mathbf{r}_2' = \frac{m_1}{m_1 + m_2}\mathbf{r} \ . \tag{2}$$

Then, T' takes the form

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

and the total Lagrangian as given by equation (1) is:

$$L = \frac{1}{2}(m_1 + m_2)\dot{\mathbf{R}}^2 + \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\dot{\mathbf{r}}^2 - V(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}}, ...) \ , \tag{3}$$

where from the reduced mass is defined as

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad o' \quad \frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \ .$$

Then, the equation (3) can be written as follows

$$L = \frac{1}{2}(m_1 + m_2)\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 - V(\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}}, ...).$$

From this equation we see that the coordinates $\dot{\mathbf{R}}$ are cyclic implying that the center-of-mass is either fixed or in uniform motion.
Now, none of the equations of motion for $\mathbf{r}$ will contain a term where $\mathbf{R}$ or $\dot{\mathbf{R}}$ will occur. This term is exactly what we will have if a center of force would have been located in the center of mass with an additional particle at a distance $\mathbf{r}$ away of mass μ (the reduced mass).
Thus, the motion of two particles around their center of mass, which is due to a central force can be always reduced to an equivalent problem of a single body.

2.2 Equations of motion

Now we limit ourselves to conservative central forces for which the potential is a function of only r, $V(r)$, so that the force is directed along $\mathbf{r}$. Since in order to solve the problem we only need to tackle a particle of mass m moving around the fixed center of force, we can put the origin of the reference frame there. As the potential depends only on r, the problem has spherical symmetry, that is any arbitrary rotation around a fixed axis has no effect on the solution. Therefore, an angular coordinate representing that rotation should be cyclic providing another considerable simplification to the problem. Due to the spherical symmetry, the total angular momentum

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

is conserved. Thus, it can be inferred that $\mathbf{r}$ is perpendicular to the fixed axis of $\mathbf{L}$. Now, if $\mathbf{L} = 0$ the motion should be along a line passing through the center of force, since for $\mathbf{L} = 0$ $\mathbf{r}$ and $\dot{\mathbf{r}}$ are parallel. This happens only in the rectilinear motion, and therefore central force motions proceed in one plane.
By taking the z axis as the direction of $\mathbf{L}$, the motion will take place in the (x, y) plane. The spherical angular coordinate ϕ will have the constant value $\pi/2$ and we can go on as foollows. The conservation of the angular momentum provides three independent constants of motion. As a matter of fact, two of them, expressing the constant direction of the angular momentum, are used to reduce the problem of three degrees of freedom to only two. The third coordinate corresponds to the conservation of the modulus of $\mathbf{L}$.
In polar coordinates the Lagrangian is

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r) \ . \tag{4}$$

As we have seen, θ is a cyclic coordinate whose canonically conjugate momentum is the angular momentum

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

then, one of the equations of motion will be

$$\dot{p}_\theta = \frac{d}{dt}(mr^2\dot{\theta}) = 0 \ . \tag{5}$$

This leads us to

$$mr^2\dot{\theta} = l = cte \ , \tag{6}$$

where l is constant modulus of the angular momentum. From equation (5) one also gets

$$\frac{d}{dt}\left(\frac{r^2\dot{\theta}}{2}\right) = 0. \tag{7}$$

The factor $1/2$ is introduced because $(r^2\dot{\theta})/2$ is the *areolar velocity* (the area covered by the position vector per unit of time).
The conservation of the angular momentum is equivalent to saying that the areolar velocity is constant. This is nothing else than a proof of Kepler's second law of planetary motion: *the position vector of a planet covers equal areas in equal time intervals*. However, we stress that the constancy of the areolar velocity is a property valid for any central force not only for inverse square ones.
The other Lagrange equation for the r coordinates reads

$$\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r} = 0 \ . \tag{8}$$

Denoting the force by $f(r)$, we can write this equation as follows

$$m\ddot{r} - mr\dot{\theta}^2 = f(r) \ . \tag{9}$$

Using the equation (6), the last equation can be rewritten as

$$m\ddot{r} - \frac{l^2}{mr^3} = f(r). \tag{10}$$

Recalling now the conservation of the total energy

$$E = T + V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) \ . \tag{11}$$

we say that E is a constant of motion. This can be derived from the equations of motion. The equation (10) can be written as follows

$$m\ddot{r} = -\frac{d}{dr}\left[V(r) + \frac{1}{2}\frac{l^2}{mr^2}\right] \ , \tag{12}$$

and by multiplying by $\dot{r}$ both sides, we get

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

or

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

Thus

$$\frac{1}{2}m\dot{r}^2 + V(r) + \frac{1}{2}\frac{l^2}{mr^2} = cte \tag{13}$$

and since $(l^2/2mr^2) = (mr^2\dot{\theta}/2)$, the equation (13) is reduced to (11).
Now, let us solve the equations of motion for r and θ. Taking $\dot{r}$ from equation (13), we have

$$\dot{r} = \sqrt[2]{\frac{2}{m}(E - V - \frac{l^2}{2mr^2})} \ , \tag{14}$$

or

$$dt = \frac{dr}{\sqrt[2]{\frac{2}{m}(E - V - \frac{l^2}{2mr^2})}} \ . \tag{15}$$

Let r_0 be the value of r at $t = 0$. The integral of the two terms of the equation reads

$$t = \int_{r_0}^{r} \frac{dr}{\sqrt[2]{\frac{2}{m}(E - V - \frac{l^2}{2mr^2})}}. \tag{16}$$

This equation gives t as a function of r and of the constants of integration E, l and r_0. It can be inverted, at least in a formal way, to give r as a function of t and of the constants. Once we have r, there is no problem to get θ starting from equation (6), that can be written as follows

$$d\theta = \frac{l dt}{mr^2} \ . \tag{17}$$

If θ_0 is the initial value of θ, then (17) will be

$$\theta = l\int_0^t \frac{dt}{mr^2(t)} + \theta_0. \tag{18}$$

Thus, we have already get the equations of motion for the variables r and θ.

2.3 The differential equation of the orbit

A change of our standpoint regarding the approach of real central force problems prove to be convenient. Till now, solving the problem meant seeking r and θ as functions of time and some constants of integration such as E, l, etc. However, quite often, what we are really looking for is the equation of the orbit, that is the direct dependence between r and θ, by eliminating the time parameter t. In the case of central force problems, this elimination is particularly simple because t is to be found in the equations of motion only in the form of a variable with respect to which the derivatives are performed. Indeed, the equation of motion (6) gives us a definite relationship between dt and $d\theta$

$$ldt = mr^2 d\theta. \tag{19}$$

The corresponding relationship between its derivatives with respect to t and θ is

$$\frac{d}{dt} = \frac{l}{mr^2}\frac{d}{d\theta}. \tag{20}$$

This relationship can be used to convert (10) in a differential equation for the orbit. At the same time, one can solve for the equations of motion and go on to get the orbit equation. For the time being, we follow up the first route.

From equation (20) we can write the second derivative with respect to t

$$\frac{d^2}{dt^2} = \frac{d}{d\theta}\frac{l}{mr^2}\left(\frac{d}{d\theta}\frac{l}{mr^2}\right)$$

and the Lagrange equation for r, (10), will be

$$\frac{l}{r^2}\frac{d}{d\theta}\left(\frac{l}{mr^2}\frac{dr}{d\theta}\right) - \frac{l}{mr^3} = f(r)\ . \tag{21}$$

But

$$\frac{1}{r^2}\frac{dr}{d\theta} = -\frac{d(1/r)}{d\theta}\ .$$

Employing the change of variable $u = 1/r$, we have

$$\frac{l^2u^2}{m}\left(\frac{d^2u}{d\theta^2} + u\right) = -f\left(\frac{1}{u}\right)\ . \tag{22}$$

Since

$$\frac{d}{du} = \frac{dr}{d\theta}\frac{d}{dr} = -\frac{1}{u^2}\frac{d}{dr} \ ,$$

equation (22) can be written as follows

$$\frac{d^2u}{d\theta^2} + u = -\frac{m}{l^2}\frac{d}{du}V\left(\frac{1}{u}\right). \tag{23}$$

Any of the equations (22) or (23) is the differential equation of the orbit if we know the force f or the potential V. Vice versa, if we know the orbit equation we can get f or V.
For an arbitrary particular force law, the orbit equation can be obtained by integrating the equation (22). Since a great deal of work has been done when solving (10), we are left with the task of eliminating t in the solution (15) by means of (19),

$$d\theta = \frac{ldr}{mr^2 \cdot \sqrt[2]{\frac{2}{m}\left[E - V(r) - \frac{l^2}{2mr^2}\right]}} \ , \tag{24}$$

or

$$\theta = \int_{r_0}^{r} \frac{dr}{r^2 \cdot \sqrt[2]{\frac{2mE}{l^2} - \frac{2mU}{l^2} - \frac{1}{r^2}}} + \theta_0 \ . \tag{25}$$

By the change of variable $u = 1/r$,

$$\theta = \theta_0 - \int_{u_0}^{u} \frac{du}{\sqrt[2]{\frac{2mE}{l^2} - \frac{2mU}{l^2} - u^2}} \ , \tag{26}$$

which is the formal solution for the orbit equation.

2.4 Kepler's problem: the case of inverse square force
The inverse square central force law is the most important of all and therefore we shall pay more attention to this case. The force and the potential are:

$$f = -\frac{k}{r^2} \quad \text{y} \quad V = -\frac{k}{r} \ . \tag{27}$$

To integrate the orbit equation we put (23) in (22),

$$\frac{d^2u}{d\theta^2} + u = -\frac{mf(1/u)}{l^2u^2} = \frac{mk}{l^2} \ . \tag{28}$$

Now, we perform the change of variable $y = u - \frac{mk}{l^2}$, in order that the differential equation be written as follows

$$\frac{d^2y}{d\theta^2} + y = 0 \ ,$$

possessing the solution

$$y = B\cos(\theta - \theta') \ ,$$

where B and θ' are the corresponding integration constants. The solution in terms of r is

$$\frac{1}{r} = \frac{mk}{l^2}\left[1 + e\cos(\theta - \theta')\right], \tag{29}$$

where

$$e = B\frac{l^2}{mk} \ .$$

We can get the orbit equation from the formal solution (26). Although the procedure is longer than solving the equation (28), it is nevertheless to do it since the integration constant e is directly obtained as a function of E and l.

We write equation (26) as follows

$$\theta = \theta' - \int \frac{du}{\sqrt[2]{\frac{2mE}{l^2} - \frac{2mU}{l^2} - u^2}} \ , \tag{30}$$

where now one deals with a definite integral. Then θ' of (30) is an integration constant determined through the initial conditions and is not necessarily the initial angle θ_0 at $t = 0$. The solution for this type of integrals is

$$\int \frac{dx}{\sqrt[2]{\alpha + \beta x + \gamma x^2}} = \frac{1}{\sqrt[2]{-\gamma}} \arccos\left[-\frac{\beta + 2\gamma x}{\sqrt[2]{q}}\right] \ , \tag{31}$$

where

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

In order to apply this type of solutions to the equation (30) we should make

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

and the discriminant q will be

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

With these substitutions, (30) is

$$\theta = \theta' - \arccos\left[\frac{\frac{l^2 u}{mk} - 1}{\sqrt[2]{1 + \frac{2El^2}{mk^2}}}\right] .$$

For $u \equiv 1/r$, the resulting orbit equation is

$$\frac{1}{r} = \frac{mk}{l^2}\left[1 + \sqrt[2]{1 + \frac{2El^2}{mk^2}}\cos(\theta - \theta')\right] . \tag{32}$$

Comparing (32) with the equation (29) we notice that the value of e is:

$$e = \sqrt[2]{1 + \frac{2El^2}{mk^2}} . \tag{33}$$

The type of orbit depends on the value of e according to the following table:

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

2.5 Dispersion by a center of force

From a historical point of view, the interest on central forces was related to the astronomical problem of planetary motions. However, there is no reason to consider them only under these circumstances. Another important issue that one can study within Classical Mechanics is the dispersion of particles by central forces. Of course, if the particles are of atomic size, we should keep in mind that the classical formalism may not give the right result because of quantum effects that begin to be important at those scales. Despite this, there are classical predictions that continue to be correct. Moreover, the main concepts of the dispersion phenomena are the same in both Classical Mechanics and Quantum Mechanics; thus, one can learn this scientific idiom in the classical picture, usually considered more convenient.

In its one-body formulation, the dispersion problem refers to the action of the center of force on the trajectories of the coming particles. Let us consider a uniform beam of particles, (say electrons, protons, or planets and comets), but all of the same mass and energy impinging on a center of force.

We can assume that the force diminishes to zero at large distances. The incident beam is characterized by its *intensity* I (also called flux density), which is the number of particles that pass through per units of time and normal surface. When one particle comes closer and closer to the center of force will be attracted or repelled, and its orbit will deviate from the initial rectilinear path. Once it passed the center of force, the perturbative effects will diminish such that the orbit tends again to a streight line. In general, the final direction of the motion does not coincide with the incident one. One says that the particle has been dispersed. By definition, the *differential cross section* $\sigma(\Omega)$ is

$$\sigma(\Omega)d\Omega = \frac{dN}{I}, \tag{34}$$

where dN is the number of particles dispersed per unit of time in the element of solid angle $d\Omega$ around the Ω direction. In the case of central forces there is a high degree of symmetry around the incident beam axis. Therefore, the element of solid angle can be written

$$d\Omega = 2\pi \sin\Theta d\Theta, \tag{35}$$

where Θ is the angle between two incident dispersed directions, and is called *the dispersion angle*.

For a given arbitrary particle the constants of the orbit and therefore the degree of dispersion are determined by its energy and angular momentum. It is convenient to express the latter in terms of a function of energy and the so-called impact parameter s, which by definition is the distance from the center of force to the straight suport line of the incident velocity. If u_0 is the incident velocity of the particle, we have

$$l = mu_0 s = s \cdot \sqrt[2]{2mE}. \tag{36}$$

Once E and s are fixed, the angle of dispersion Θ is uniquely determined. For the time being, we suppose that different values of s cannot lead to the same dispersion angle. Therefore, the number of dispersed particles in the element of solid angle $d\Omega$ between Θ and $\Theta + d\Theta$ should be equal to the number of incident particles whose impact parameter ranges within the corresponding s and $s + ds$:

$$2\pi I s \left|ds\right| = 2\pi\sigma(\Theta) I \sin\Theta \left|d\Theta\right|. \tag{37}$$

In the equation (37) we have introduced absolute values because while the number of particles is always positive, s and Θ can vary in opposite directions. If we consider s as a function of the energy and the corresponding

dispersion angle,

$$s = s(\Theta, E),$$

the dependence of the cross section of Θ will be given by

$$\sigma(\Theta) = \frac{s}{\sin\Theta} \left| \frac{ds}{d\Theta} \right| . \tag{38}$$

From the orbit equation (25), one can obtain directly a formal expression for the dispersion angle. In addition, for the sake of simplicity, we tackle the case of a pure repulsive dispersion. Since the orbit should be symmetric with respect to the direction of the periapsis, the dispersion angle is

$$\Theta = \pi - 2\Psi \ , \tag{39}$$

where Ψ is the angle between the direction of the incident asymptote and the direction of the periapsis. In turn, Ψ can be obtained from the equation (25) by making $r_0 = \infty$ when $\theta_0 = \pi$ (incident direction). Thus, $\theta = \pi - \Psi$ when $r = r_m$, the closest distance of the particle to the center of force. Then, one can easily obtain

$$\Psi = \int_{r_m}^{\infty} \frac{dr}{r^2 \cdot \sqrt[2]{\frac{2mE}{l^2} - \frac{2mV}{l^2} - \frac{1}{r^2}}} \ . \tag{40}$$

Expressing l as a function of the impact parameter s (eq. (36)), the result is

$$\Theta = \pi - 2\int_{r_m}^{\infty} \frac{sdr}{r \cdot \sqrt[2]{r^2\left[1 - \frac{V(r)}{E}\right] - s^2}} \ , \tag{41}$$

or

$$\Theta = \pi - 2\int_{0}^{u_m} \frac{sdu}{\sqrt[2]{1 - \frac{v(u)}{E} - s^2u^2}} \ . \tag{42}$$

The equations (41) and (42) are used rarely, as they do not enter in a direct way in the numerical calculation of the dispersion angle. However, when an analytic expression for the orbits is available, one can often get, merely by simple inspection, a relationship between Θ and s.

EXAMPLE:
This example is very important from the historical point of view. It refers to the repulsive dispersion of charged particles in a Coulomb field. The field

is produced by a fixed charge $-Ze$ and acts on incident particles of charge $-Z'e$; therefore, the force can be written as follows

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

that is, one deals with a repulsive inverse square force. The constant is

$$k = -ZZ'e^2. \tag{43}$$

The energy E is positive implying a hyperbolic orbit of eccentricity

$$\epsilon = \sqrt[2]{1 + \frac{2El^2}{m(ZZ'e^2)^2}} = \sqrt[2]{1 + \left(\frac{2Es}{ZZ'e^2}\right)^2}, \tag{44}$$

where we have taken into account the equation (36). If the angle θ' is taken to be π, then from the equation (29) we come to the conclusion that the periapse corresponds to $\theta = 0$ and the orbit equation reads

$$\frac{1}{r} = \frac{mZZ'e^2}{l^2} \left[\epsilon \cos\theta - 1\right]. \tag{45}$$

The direction Ψ of the incident asymptote is thus determined by the condition $r \to \infty$:

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

that is, according to equation (39),

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

Then,

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

and by means of equation (44)

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

The functional relationship between the impact parameter and the dispersion angle will be

$$s = \frac{ZZ'e^2}{2E} \cot\frac{\Theta}{2}, \tag{46}$$

and by effecting the transformation required by the equation (38) we find that $\sigma(\Theta)$ is given by

$$\sigma(\Theta) = \frac{1}{4}\left(\frac{ZZe^2}{2E}\right)^2 \csc^4\frac{\Theta}{2}. \tag{47}$$

The equation (47) gives the famous Rutherford scattering cross section derived by him for the dispersion of α particles on atomic nuclei. In the non-relativistic limit, the same result is provided by the quantum mechanical calculations.
The concept of *total cross section* σ_T is very important in atomic physics. Its definition is

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

However, if we calculate the total cross section for the Coulombian dispersion by substituting the equation (47) in the definition above we get an infinite result. The physical reason is easy to see. According to the definition, the total cross section is the number of particles per unit of incident intensity that are dispersed in all directions. The Coulombian field is an example of long-range force; its effects are still present in the infinite distance limit. The small deviation limit is valid only for particles of large impact parameter. Therefore, for an incident beam of infinite lateral extension all the particles will be dispersed and should be included in the total cross section. It is clear that the infinite value of σ_T is not a special property of the Coulombian field and occurs for any type of long-range field.

Further reading

L.S. Brown, *Forces giving no orbit precession*, Am. J. Phys. 46, 930 (1978)

H. Goldstein, *More on the prehistory of the Laplace-Runge-Lenz vector*, Am. J. Phys. 44, 1123 (1976)

3. THE RIGID BODY

Forward: Due to its particular features, the study of the motion of the rigid body has generated several interesting mathematical techniques and methods. In this chapter, we briefly present the basic rigid body concepts.

CONTENTS:

3.1 Definition

A rigid body (RB) is defined as a system of particles whose relative distances are forced to stay constant during the motion.

3.2 Degrees of freedom

In order to describe the general motion of a RB in the three-dimensional space one needs six variables, for example the three coordinates of the center of mass measured with respect to an inertial frame and three angles for labeling the orientation of the body in space (or of a fixed system within the body with the origin in the center of mass). In other words, in the three-dimensional space the RB can be described by at most six degrees of freedom.
The number of degrees of freedom may be less when the rigid body is subjected to various conditions as follows:

- If the RB rotates around a single axis there is only one degree of freedom (one angle).
- If the RB moves in a plane, its motion can be described by five degrees of freedom (two coordinates and three angles).

3.3 Tensor of inertia.

We consider a body made of N particles of masses m_α, $\alpha = 1, 2, 3..., N$. If the body rotates at angular velocity ω around a fixed point in the body and this point, in turn, moves at velocity $\mathbf{v}$ with respect to a fixed inertial system, then the velocity of the αth particle w.r.t. the inertial system is given by

$$\mathbf{v}_\alpha = \mathbf{v} + \omega \times \mathbf{r}_\alpha. \tag{1}$$

The kinetic energy of the αth particle is

$$T_\alpha = \frac{1}{2} m_\alpha \mathbf{v}_\alpha^2 \ , \tag{2}$$

where

$$\mathbf{v}_\alpha^2 = \mathbf{v}_\alpha \cdot \mathbf{v}_\alpha = (\mathbf{v} + \omega \times \mathbf{r}_\alpha) \cdot (\mathbf{v} + \omega \times \mathbf{r}_\alpha) \tag{3}$$

$$= \mathbf{v} \cdot \mathbf{v} + 2\mathbf{v} \cdot (\omega \times \mathbf{r}_\alpha) + (\omega \times \mathbf{r}_\alpha) \cdot (\omega \times \mathbf{r}_\alpha)$$

$$= \mathbf{v}^2 + 2\mathbf{v}(\omega \times \mathbf{r}_\alpha) + (\omega \times \mathbf{r}_\alpha)^2. \tag{4}$$

Then the total energy is

$$\begin{aligned} T &= \textstyle\sum_\alpha T_\alpha = \sum_\alpha \frac{1}{2} m_\alpha \mathbf{v}^2 + \sum_\alpha m_\alpha \left[\mathbf{v}\cdot(\omega \times \mathbf{r}_\alpha)\right] + \\ &\qquad + \tfrac{1}{2} \textstyle\sum_\alpha m_\alpha (\omega \times \mathbf{r}_\alpha)^2 \; ; \\ T &= \tfrac{1}{2} M \mathbf{v}^2 + \mathbf{v}\cdot\left[\omega \times \textstyle\sum_\alpha m_\alpha \mathbf{r}_\alpha\right] + \tfrac{1}{2} \textstyle\sum_\alpha m_\alpha \left(\omega \times \mathbf{r}_\alpha\right)^2 . \end{aligned}$$

If the origin is fixed to the solid body, we can take it in the center of mass. Thus,

$$\mathbf{R} = \frac{\sum_\alpha m_\alpha \mathbf{r}_\alpha}{M} = 0,$$

and therefore we get

$$T = \frac{1}{2} M \mathbf{v}^2 + \frac{1}{2} \sum_\alpha m_\alpha \left(\omega \times \mathbf{r}_\alpha\right)^2 \tag{5}$$

$$T = T_{trans} + T_{rot} \tag{6}$$

where

$$T_{trans} = \frac{1}{2} \sum_\alpha m_\alpha \mathbf{v}^2 = \frac{1}{2} M \mathbf{v}^2 \tag{7}$$

$$T_{rot} = \frac{1}{2} \sum_\alpha m_\alpha \left(\omega \times \mathbf{r}_\alpha\right)^2 . \tag{8}$$

In Eq. (8) we use the vectorial identity

$$(\mathbf{A} \times \mathbf{B})^2 = \mathbf{A}^2 \mathbf{B}^2 - (\mathbf{A} \cdot \mathbf{B})^2 \tag{9}$$

to get the following form of the equation

$$T_{rot} = \frac{1}{2} \sum_\alpha m_\alpha \left[\omega^2 \mathbf{r}^2 - (\omega \cdot \mathbf{r}_\alpha)^2\right] \; ,$$

which in terms of the components of ω and $\mathbf{r}$

$$\omega = (\omega_1, \omega_2, \omega_3) \quad \text{and} \quad \mathbf{r}_\alpha = (x_{\alpha 1}, x_{\alpha 2}, x_{\alpha 3})$$

can be written as follows

$$T_{rot} = \frac{1}{2} \sum_\alpha m_\alpha \left\{ \left(\sum_i \omega_i^2 \right) \left(\sum_k x_{\alpha k}^2 \right) - \left(\sum_i \omega_i x_{\alpha i} \right) \left(\sum_j \omega_j x_{\alpha j} \right) \right\} .$$

Now, we introduce

$$\omega_i = \sum_j \delta_{ij} \omega_j$$

$$T_{rot} = \frac{1}{2}\sum_{\alpha}\sum_{ij} m_\alpha \left\{ \omega_i\omega_j\delta_{ij}\left(\sum_k x_{\alpha k}^2\right) - \omega_i\omega_j x_{\alpha i}x_{\alpha j}\right\} \tag{10}$$

$$T_{rot} = \frac{1}{2}\sum_{ij}\omega_i\omega_j \sum_{\alpha} m_\alpha \left[\delta_{ij}\sum_k x_{\alpha k}^2 - x_{\alpha i}x_{\alpha j}\right]. \tag{11}$$

We can write T_{rot} as follows

$$T_{rot} = \frac{1}{2}\sum_{ij} I_{ij}\omega_i\omega_j \tag{12}$$

where

$$I_{ij} = \sum_{\alpha} m_\alpha \left[\delta_{ij}\sum_k x_{\alpha k}^2 - x_{\alpha i}x_{\alpha j}\right]. \tag{13}$$

The nine quantities I_{ij} are the components of a new mathematical entity, denoted by $\{I_{ij}\}$ and called *tensor of inertia*. It can be written in a convenient way as a (3×3) matrix

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

$$= \begin{pmatrix} \sum_\alpha m_\alpha(x_{\alpha 2}^2 + x_{\alpha 3}^2) & -\sum_\alpha m_\alpha x_{\alpha 1}x_{\alpha 2} & -\sum_\alpha m_\alpha x_{\alpha 1}x_{\alpha 3} \\ -\sum_\alpha m_\alpha x_{\alpha 2}x_{\alpha 1} & \sum_\alpha m_\alpha(x_{\alpha 1}^2 + x_{\alpha 3}^2) & -\sum_\alpha m_\alpha x_{\alpha 2}x_{\alpha 3} \\ -\sum_\alpha m_\alpha x_{\alpha 3}x_{\alpha 1} & -\sum_\alpha m_\alpha x_{\alpha 3}x_{\alpha 2} & \sum_\alpha m_\alpha(x_{\alpha 1}^2 + x_{\alpha 2}^2) \end{pmatrix}. \tag{14}$$

We note that $I_{ij} = I_{ji}$, and therefore $\{I_{ij}\}$ is a *symmetric tensor*, implying that only six of the components are independent. The diagonal elements of $\{I_{ij}\}$ are called *moments of inertia* with respect to the axes of coordinates, whereas the negatives of the nondiagonal elements are called the *products of inertia*. For a continuous distribution of mass, of density $\rho(\mathbf{r})$, $\{I_{ij}\}$ is written in the following way

$$I_{ij} = \int_V \rho(\mathbf{r})\left[\delta_{ij}\sum_k x_k^2 - x_i x_j\right] dV. \tag{15}$$

<u>*EXAMPLE*</u>:
Find the elements I_{ij} of the tensor of inertia $\{I_{ij}\}$ for a cube of uniform density of side b, mass M, with one corner placed at the origin.

$$I_{11} = \int_V \rho\left[x_1^2 + x_2^2 + x_3^2 - x_1x_1\right] dx_1dx_2dx_3 = \rho\int_0^b\int_0^b\int_0^b (x_2^2 + x_3^2)dx_1dx_2dx_3\ .$$

The result of the three-dimensional integral is $I_{11} = \frac{2}{3}(\rho b^3)^2 = \frac{2}{3}Mb^2$.

$$I_{12} = \int\limits_V \rho(-x_1 x_2)dV = -\rho \int\limits_0^b \int\limits_0^b \int\limits_0^b (x_1 x_2)dx_1 dx_2 dx_3 = -\frac{1}{4}\rho b^5 = -\frac{1}{4}Mb^2 \ .$$

We see that all the other integrals are equal, so that

$$I_{11} = I_{22} = I_{33} = \frac{2}{3}Mb^2$$

$$\underset{i\neq j}{I_{ij}} = -\frac{1}{4}Mb^2 \ ,$$

leading to the following form of the matrix

$$\{I_{ij}\} = \begin{pmatrix} \frac{2}{3}Mb^2 & -\frac{1}{4}Mb^2 & -\frac{1}{4}Mb^2 \\ -\frac{1}{4}Mb^2 & \frac{2}{3}Mb^2 & -\frac{1}{4}Mb^2 \\ -\frac{1}{4}Mb^2 & -\frac{1}{4}Mb^2 & \frac{2}{3}Mb^2 \end{pmatrix} .$$

3.4 Angular Momentum

The angular momentum of a RB made of N particles of masses m_α is given by

$$\mathbf{L} = \sum_\alpha \mathbf{r}_\alpha \times \mathbf{p}_\alpha \ , \tag{16}$$

where

$$\mathbf{p}_\alpha = m_\alpha \mathbf{v}_\alpha = m_\alpha(\omega \times \mathbf{r}_\alpha) \ . \tag{17}$$

Substituting (17) in (16), we get

$$\mathbf{L} = \sum_\alpha m_\alpha \mathbf{r}_\alpha \times (\omega \times \mathbf{r}_\alpha) \ .$$

Employing the vectorial identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{A}) = (\mathbf{A} \cdot \mathbf{A})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{A} = \mathbf{A}^2\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{A} \ ,$$

leads to

$$\mathbf{L} = \sum_\alpha m_\alpha(\mathbf{r}_\alpha^2 \omega - \mathbf{r}_\alpha(\omega \cdot \mathbf{r}_\alpha).$$

Considering the ith component of the vector $\mathbf{L}$

$$L_i = \sum_\alpha m_\alpha \left(\omega_i \sum_k x_{\alpha k}^2\right) - x_{\alpha i}\left(\sum_j x_{\alpha j}\omega_j\right) \ ,$$

and introducing the equation

$$\omega_i = \sum_j \omega_j \delta_{ij} \ ,$$

we get

$$\begin{aligned} L_i &= \textstyle\sum_\alpha m_\alpha \left(\sum_j \delta_{ij} \omega_j \sum_k x_{\alpha k}^2 \right) - \left(\sum_j x_{\alpha j} x_{\alpha j} \omega_j \right) && (18) \\ &= \textstyle\sum_\alpha m_\alpha \sum_j \omega_j \delta_{ij} \left(\sum_k x_{\alpha k}^2 - x_{\alpha i} x_{\alpha j} \right) && (19) \\ &= \textstyle\sum_j \omega_j \sum_\alpha m_\alpha \left(\delta_{ij} \sum_k x_{\alpha k}^2 - x_{\alpha i} x_{\alpha j} \right) \ . && (20) \end{aligned}$$

Comparing with the equation (13) leads to

$$L_i = \sum_j I_{ij} \omega_j \ . \tag{21}$$

This equation can also be written in the form

$$\mathbf{L} = \{ I_{ij} \} \, \omega \ , \tag{22}$$

or

$$\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} . \tag{23}$$

The rotational kinetic energy, T_{rot}, can be related to the angular momentum as follows: first, multiply the equation (21) by $\frac{1}{2}\omega_i$

$$\omega_i \frac{1}{2} L_i = \frac{1}{2} \omega_i \sum_j I_{ij} \omega_j \ , \tag{24}$$

and next summing over all the i indices, gives

$$\sum_i \frac{1}{2} L_i \omega_i = \frac{1}{2} \sum_{ij} I_{ij} \omega_i \omega_j \ .$$

Compararing this equation with (12), we see that the second term is T_{rot}. Therefore

$$T_{rot} = \sum_I \frac{1}{2} L_i \omega_i = \frac{1}{2} \mathbf{L} \cdot \omega \ . \tag{25}$$

Now, we substitute (22) in the equation (25), getting the relationship between T_{rot} and the tensor of inertia

$$T_{rot} = \frac{1}{2} \omega \cdot \{ I_{ij} \} \cdot \omega . \tag{26}$$

3.5 Principal axes of inertia

Taking the tensor of inertia $\{I_{ij}\}$ diagonal, that is $I_{ij} = I_i \delta_{ij}$, the rotational kinetic energy and the angular momentum are expressed as follows

$$T_{rot} = \frac{1}{2} \sum_{ij} I_{ij} \omega_i \omega_j$$

$$= \frac{1}{2} \sum_{ij} \delta_{ij} I_i \omega_i \omega_j$$

$$T_{rot} = \frac{1}{2} \sum_{i} I_i \omega_i^2 \tag{27}$$

and

$$L_i = \sum_j I_{ij} \omega_j$$

$$= \sum_j \delta_{ij} I_i \omega_j = I_i \omega_i$$

$$\mathbf{L} = \mathbf{I}\omega. \tag{28}$$

To seek a diagonal form of $\{I_{ij}\}$ is equivalent to finding a new system of three axes for which the kinetic energy and the angular momentum take the form given by (27) and (28). In this case the axes are called *principal axes of inertia*. That means that given an inertial reference system within the body, we can pass from it to the principal axes by a particular orthogonal transformation, which is called *transformation to the principal axes*.
Making equal the components of (22) and (28), we have

$$L_1 \quad = I\omega_1 = I_{11}\omega_1 + I_{12}\omega_2 + I_{13}\omega_3 \tag{29}$$
$$L_2 \quad = I\omega_2 = I_{21}\omega_1 + I_{22}\omega_2 + I_{23}\omega_3 \tag{30}$$
$$L_3 \quad = I\omega_3 = I_{31}\omega_1 + I_{32}\omega_2 + I_{33}\omega_3 \ . \tag{31}$$

This is a system of equations that can be rewritten as

$$\begin{aligned} (I_{11} - I)\omega_1 + I_{12}\omega_2 + I_{13}\omega_3 \quad &= 0 \\ I_{21}\omega_1 + (I_{22} - I)\omega_2 + I_{23}\omega_3 \quad &= 0 \\ I_{31}\omega_1 + I_{32}\omega_2 + (I_{33} - I)\omega_3 \quad &= 0 \ . \end{aligned} \tag{32}$$

To get nontrivial solutions, the determinant of the system should be zero

$$\begin{vmatrix} (I_{11}-I)\omega_1 & I_{12}\omega_2 & I_{13}\omega_3 \\ I_{21}\omega_1 & (I_{22}-I)\omega_2 & I_{23}\omega_3 \\ I_{31}\omega_1 & I_{32}\omega_2 & (I_{33}-I)\omega_3 \end{vmatrix} = 0 \ . \tag{33}$$

This determinant leads to a polynomial of third order in I, known as *the characteristic polynomial*. The equation (33) is called the *secular equation* or *characteristic equation*. In practice, the principal moments of inertia, being the eigenvalues of $\mathbf{I}$, are obtained as solutions of the secular equation.

EXAMPLE:
Determine the principal axes of inertia for the cube of the previous example.

Substituting the values obtained in the previous example in the equation (33) we get:

$$\left| \begin{pmatrix} (\frac{2}{3}\beta - I) & -\frac{1}{4}\beta & -\frac{1}{4}\beta \\ -\frac{1}{4}\beta & (\frac{2}{3}\beta - I) & -\frac{1}{4}\beta \\ -\frac{1}{4}\beta & -\frac{1}{4}\beta & (\frac{2}{3}\beta - I) \end{pmatrix} \right| = 0 \ ,$$

where $\beta = Mb^2$. Thus, the characteristic equation will be

$$\left(\frac{11}{12}\beta - I\right)\left(\frac{11}{12}\beta - I\right)\left(\frac{1}{6}\beta - I\right) = 0 \ .$$

The solutions, i.e., the principal moments of inertia are:

$$I_1 = \frac{1}{6}\beta, \quad I_2 = I_3 = \frac{11}{12}\beta \ ,$$

whose corresponding eigenvalues are given by

$$I = \frac{1}{6}\beta \leftrightarrow \frac{1}{\sqrt[2]{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad I_2, I_3 = \frac{11}{12}\beta \leftrightarrow \frac{1}{\sqrt[2]{2}} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\} \ .$$

The matrix that diagonalizes $\{I_{ij}\}$ is:

$$\lambda = \sqrt[2]{\frac{1}{3}} \begin{pmatrix} 1 & -\sqrt[2]{\frac{3}{2}} & -\sqrt[2]{\frac{3}{2}} \\ 1 & \sqrt[2]{\frac{3}{2}} & 0 \\ 1 & 0 & \sqrt[2]{\frac{3}{2}} \end{pmatrix} \ .$$

The diagonalized $\{I_{ij}\}$ will be

$$\{I_{ij}\}_{diag} = (\lambda)^* \{I_{ij}\} \lambda = \begin{pmatrix} \frac{1}{6}\beta & 0 & 0 \\ 0 & \frac{11}{12}\beta & 0 \\ 0 & 0 & \frac{11}{12}\beta \end{pmatrix}.$$

3.6 The theorem of parallel axes

We suppose that the system x_1, x_2, x_3 has the origin in the center of mass of the RB. A second system X_1, X_2, X_3, has the origin in another position w.r.t. the first system. The only imposed condition on them is to be parallel. We define the vectors $\mathbf{r} = (x_1, x_2, x_3)$, $\mathbf{R} = (X_1, X_2, X_3)$ y $\mathbf{a} = (a_1, a_2, a_3)$ in such a way that $\mathbf{R} = \mathbf{r} + \mathbf{a}$, or in component form

$$X_i = x_i + a_i. \tag{34}$$

Let J_{ij} be the components of the tensor of inertia w.r.t. the system $X_1X_2X_3$,

$$J_{ij} = \sum_\alpha m_\alpha \left[\delta_{ij} \sum_k X_{\alpha k}^2 - X_{\alpha i} X_{\alpha j}\right] . \tag{35}$$

We substitute (34) in (35),

$$J_{ij} = \sum_\alpha m_\alpha \left[\delta_{ij} \sum_k (x_{\alpha k} + a_k)^2 - (x_{\alpha i} + a_i)(x_{\alpha j} + a_j)\right]$$

$$\begin{aligned} &= \left[\textstyle\sum_\alpha m_\alpha \left(\delta_{ij} \sum_k (x_{\alpha k})^2 - x_{\alpha i} x_{\alpha j}\right)\right] + \sum_\alpha m_\alpha \left(\delta_{ij} \sum_k a_k^2 - a_i a_j\right) \\ &+ \left[\textstyle\sum_k 2a_k \delta_{ij} \left(\sum_\alpha m_\alpha x_{\alpha k}\right) - a_j \left(\sum_\alpha m_\alpha x_{\alpha j}\right) - a_i \left(\sum_\alpha m_\alpha x_{\alpha i}\right)\right] . \end{aligned} \tag{36}$$

Since the center of mass coordinate is defined as

$$\bar{x} = \frac{\sum_\alpha m_\alpha x_\alpha}{M}$$

we take into account that we have already set the origin in the center of mass, i.e.,

$$(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (0, 0, 0) .$$

Now, if we also compare the first term in (36) with the equation (13), we have

$$J_{ij} = I_{ij} + M(a^2 \delta_{ij} - a_i a_j) \tag{37}$$

and therefore the elements of the tensor of inertia I_{ij} for the center of mass system will be given by:

$$I_{ij} = J_{ij} - M(\delta_{ij}a^2 - a_i a_j) \ . \tag{38}$$

This is known as the *theorem of the parallel axes.*

EXAMPLE:
Find I_{ij} for the previous cube w.r.t. a reference system parallel to the system in the first example and with the origin in the center of mass.
We already know from the previous example that:

$$\{J_{ij}\} = \begin{pmatrix} \frac{2}{3}\beta & -\frac{1}{4}\beta & -\frac{1}{4}\beta \\ -\frac{1}{4}\beta & \frac{2}{3}\beta & -\frac{1}{4}\beta \\ -\frac{1}{4}\beta & -\frac{1}{4}\beta & \frac{2}{3}\beta \end{pmatrix} \ .$$

Now, since the vector $\mathbf{a} = (\frac{b}{2}, \frac{b}{2}, \frac{b}{2})$ and $\mathbf{a}^2 = \frac{3}{4}b^2$, we can use the equation (38) and the fact that $\beta = Mb^2$ to get,

$$I_{11} = J_{11} - M(a^2 - a_1^2) = \tfrac{1}{6}Mb^2 \tag{39}$$
$$I_{22} = J_{22} - M(a^2 - a_2^2) = \tfrac{1}{6}Mb^2 \tag{40}$$
$$I_{33} = J_{33} - M(a^2 - a_3^2) = \tfrac{1}{6}Mb^2 \tag{41}$$
$$I_{12} = J_{12} - M(-a_1 a_2) = 0 \tag{42}$$
$$I_{12} = I_{13} = I_{23} = 0 \ . \tag{43}$$

Therefore

$$\{I\} = \begin{pmatrix} \frac{1}{6}Mb^2 & 0 & 0 \\ 0 & \frac{1}{6}Mb^2 & 0 \\ 0 & 0 & \frac{1}{6}Mb^2 \end{pmatrix} .$$

EXAMPLE:
We consider the case for which the vector $\mathbf{a} = (0, \frac{b}{2}, \frac{b}{2})$ and $a^2 = \frac{b^2}{2}$. Then, the new tensor of inertia will be:

$$I_{11} = J_{11} - M(a^2 - a_1^2) = \left(\tfrac{2}{3}Mb^2\right) - M\left(\tfrac{b^2}{2} - 0\right) = \tfrac{1}{6}Mb^2 \tag{44}$$
$$I_{22} = J_{22} - M(a^2 - a_2^2) = \left(\tfrac{2}{3}Mb^2\right) - M\left(\tfrac{b^2}{2} - \tfrac{b^2}{4}\right) = \tfrac{5}{12}Mb^2 \tag{45}$$
$$I_{33} = J_{33} - M(a^2 - a_3^2) = \left(\tfrac{2}{3}Mb^2\right) - M\left(\tfrac{b^2}{2} - \tfrac{b^2}{4}\right) = \tfrac{5}{12}Mb^2 \tag{46}$$
$$I_{12} = J_{12} - M(-a_1 a_2) = \left(-\tfrac{1}{4}Mb^2\right) - M(0) = -\tfrac{1}{4}Mb^2 \tag{47}$$

$$I_{13} \quad = J_{13} - M(-a_1 a_3) = \left(-\tfrac{1}{4}Mb^2\right) - M(0) = -\tfrac{1}{4}Mb^2 \tag{48}$$

$$I_{23} \quad = J_{23} - M(-a_2 a_3) = \left(-\tfrac{1}{4}Mb^2\right) - M(\tfrac{1}{4}Mb^2) = 0\ . \tag{49}$$

It follows that $\{I_{ij}\}$ is equal to:

$$\{I_{ij}\} = \begin{pmatrix} \frac{1}{6}Mb^2 & -\frac{1}{4}Mb^2 & -\frac{1}{4}Mb^2 \\ -\frac{1}{4}Mb^2 & \frac{5}{12}Mb^2 & 0 \\ -\frac{1}{4}Mb^2 & 0 & \frac{5}{12}Mb^2 \end{pmatrix}\ .$$

3.7 The dynamics of the rigid body

The rate of change in time of the angular momentum $\mathbf{L}$ is given by:

$$\left(\frac{d\mathbf{L}}{dt}\right)_{inertial} = \mathbf{N}^{(e)}. \tag{50}$$

For the description w.r.t. the system fixed to the body we have to use the operator identity

$$\left(\frac{d}{dt}\right)_{inertial} = \left(\frac{d}{dt}\right)_{body} + \omega \times\ . \tag{51}$$

Applying this operator to the equation (50)

$$\left(\frac{d\mathbf{L}}{dt}\right)_{inertial} = \left(\frac{d\mathbf{L}}{dt}\right)_{body} + \omega \times \mathbf{L}. \tag{52}$$

Then, instead of (50) we shall have

$$\left(\frac{d\mathbf{L}}{dt}\right)_{body} + \omega \times \mathbf{L} = \mathbf{N}. \tag{53}$$

Now we project the equation (53) onto the principal axes of inertia, that we call (x_1, x_2, x_3); then T_{rot} and $\mathbf{L}$ take by far simpler forms, e.g.,

$$L_i = I_i \omega_i\ . \tag{54}$$

The ith component of (53) is

$$\frac{dL_i}{dt} + \epsilon_{ijk}\omega_j L_k = N_i\ . \tag{55}$$

Now projecting onto the principal axes of inertia and using the equation (54), one can put (55) in the form:

$$I_i \frac{d\omega_i}{dt} + \epsilon_{ijk}\omega_j\omega_k I_k = N_i \tag{56}$$

since the principal elements of inertia are independent of time. Thus, we obtain the following system of equations known as Euler's equations

$$\begin{aligned} I_1\dot{\omega}_1 + \omega_2\omega_3(I_2 - I_3) &= N_1 \\ I_2\dot{\omega}_2 + \omega_3\omega_1(I_3 - I_1) &= N_2 \\ I_3\dot{\omega}_3 + \omega_1\omega_2(I_1 - I_2) &= N_3 \ . \end{aligned} \tag{57}$$

EXAMPLE:
For the rolling and sliding of a billiard ball, prove that after a horizontal kick the ball slips a distance

$$x_{_1} = \frac{12u_0^2}{49\mu g} \ ,$$

where u_0 is the initial velocity. Then, it starts rolling without gliding at the time

$$t_1 = \frac{2u_0}{7\mu g}.$$

SOLUTION: When the impulsive force stops, the initial conditions are

$$\begin{aligned} x_0 &= 0, \qquad \dot{x}_0 = u_0 \\ \phi &= 0, \qquad \dot{\phi} = 0 \ . \end{aligned}$$

The friction force is given by

$$\mathbf{F}_f = -\mu g \hat{\mathbf{e}}_1 \ ,$$

and the equation of motion reads

$$\ddot{x} = -\mu g M. \tag{58}$$

The equation for $\mathbf{L}$ is

$$\frac{dL_3}{dt} = I_3\ddot{\phi} = N_3 \tag{59}$$

where I_3 is

$$I_3 = \int \rho(\mathbf{r}) \left[x_1^2 - x_2^2\right] dx_1 dx_2 dx_3 = \frac{2}{5} M a^2$$

and

$$N_3 = F_f a = \mu M g a \ .$$

Substituting I_3 and N_3 in (59), one gets

$$a\ddot{\phi} = \frac{5}{2}\mu g. \tag{60}$$

Integrating once both (58) and (60), we get

$$\dot{x} = -\mu g t + C_1 \tag{61}$$

$$a\dot{\phi} = \frac{5}{2}\mu g t + C_2 \ . \tag{62}$$

To these equations we apply the initial conditions to put them in the form

$$\dot{x}(t) = -\mu g t + u_0 \tag{63}$$

$$a\dot{\phi}(t) = \frac{5}{2}\mu g t. \tag{64}$$

The condition of pure rolling (no friction) is

$$\dot{x}(t) = a\dot{\phi}(t). \tag{65}$$

From (64) and (65) evaluated at t_1 we get

$$\frac{5}{2}\mu g t_1 = -\mu g t_1 + u_0$$

$$\Rightarrow \ t_1 = \frac{2u_0}{7\mu g}. \tag{66}$$

Now we integrate (63) once again and applying the initial conditions we get

$$x(t) = -\mu g \frac{t^2}{2} + u_0 t \ . \tag{67}$$

By evaluating (67) and (63) at time t_1 we are led to

$$x = \frac{12u^2}{49\mu g}$$

$$\dot{x} = \frac{5}{7} u_0 \ .$$

3.8 Symmetrical top free of torques

A symmetric top is any solid of revolution. If the moments of inertia are

$$I_1 = I_2 = I_3 \qquad \text{spherical top}$$

$$I_1 = I_2 \neq I_3 \qquad \text{symmetric top}$$

$$I_1 \neq I_2 \neq I_3 \qquad \text{asymmetric top.}$$

Let us consider the symmetric top $I_1 = I_2 \neq I_3$. In this case the axis X_3 is the axis of symmetry. The Euler equations projected onto the principal axes of inertia read

$$I_1\dot{\omega}_1 + \omega_2\omega_3(I_2 - I_3) = N_1 \tag{68}$$

$$I_2\dot{\omega}_2 + \omega_3\omega_1(I_3 - I_1) = N_2 \tag{69}$$

$$I_3\dot{\omega}_3 + \omega_1\omega_2(I_1 - I_2) = N_3. \tag{70}$$

Since the system we consider here is free of torques

$$N_1 = N_2 = N_3 = 0 \ , \tag{71}$$

we use $I_1 = I_2$ in (71) to get

$$I_1\dot{\omega}_1 + \omega_2\omega_3(I_2 - I_3) = 0 \tag{72}$$

$$I_2\dot{\omega}_2 + \omega_3\omega_1(I_3 - I_1) = 0 \tag{73}$$

$$I_3\dot{\omega}_3 = 0. \tag{74}$$

The equation (74) implies that

$$\omega_3 = \text{const.}$$

The equations (72) and (73) are rewritten as follows:

$$\dot{\omega}_1 = -\Omega\omega_2 \qquad \text{where } \Omega = \omega_3\left(\frac{I_3 - I_1}{I_1}\right) \tag{75}$$

$$\dot{\omega}_2 = -\Omega\omega_1 \ . \tag{76}$$

Multiplying (76) by i and summing it to (75), we have

$$\begin{aligned}(\dot{\omega}_1 + i\dot{\omega}_2) &= -\Omega(\omega_2 - i\omega_1)\\ (\dot{\omega}_1 + i\dot{\omega}_2) &= i\Omega(\omega_1 + i\omega_2).\end{aligned}$$

If we write $\eta(t) = \dot{\omega}_1(t) + i\dot{\omega}_2(t)$, then

$$\dot{\eta}(t) - i\Omega\eta(t) = 0 \ .$$

The solution is

$$\eta(t) = A\exp(i\Omega t) \ .$$

This implies

$$(\omega_1 + i\omega_2) = A\cos(\Omega t) + i\sin(\Omega t) \ .$$

Thus,

$$\omega_1 = A\cos(\Omega t) \tag{77}$$

$$\omega_2 = A\sin(\Omega t). \tag{78}$$

The modulus of the vector ω does not change in time

$$\omega = ||\omega|| = \sqrt[2]{\omega_1 + \omega_2 + \omega_3} = \sqrt[2]{A^2 + \omega_3^2} = const \ .$$

This vector performs a precessional motion of precession frequency Ω given by

$$\Omega = \omega_3\left(\frac{I_3 - I_1}{I_1}\right) \ .$$

Moreover, we notice that Ω is constant.
If we denote by λ the angle between ω and X_3 the equations (77) and (78) take the form

$$\omega_1 = \omega\sin\lambda\cos(\Omega t)$$

$$\omega_2 = \omega\sin\lambda\sin(\Omega t)$$

$$\omega_1 = \omega\cos\lambda \ ,$$

where $A = \omega\sin\lambda$.
For a flattened body of revolution $I_1 = I_2 = I_{12}$ and $I_3 > I_1$. For example, for the case of the earth

$$\Omega_{\bigoplus} = \omega_3\left(\frac{I_3 - I_{12}}{I_{12}}\right) \simeq \frac{\omega_3}{305}.$$

The observations point to a mean value of fourteen months $\simeq$ 450 days. (This is due to the fact that the earth is not strictly a RB; there is also an internal liquid structure).

3.9 Euler angles

As we already know, a rotation can be described by a rotation matrix λ by means of the equation

$$\mathbf{x} = \lambda \mathbf{x}^{'}. \tag{79}$$

$\mathbf{x}$ represents the set of axes of the system rotated w.r.t. the system whose axes are represented by $\mathbf{x}^{'}$. The rotation λ can be accomplished through a set of "partial" rotations $\lambda = \lambda_1 \lambda_2 ... \lambda_n$. There are many possibilities to choose these λ´s. One of them is the set of angles ϕ, θ and φ called *Euler angles*. The partial rotations are in this case the following:

- A rotation around the $X_3^{'}$ axis of angle φ (in the positive trigonometric sense). The corresponding matrix is:

$$\lambda_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

- A rotation of angle θ around the $X_1^{'}$ axis (positive sense).The associated matrix is:

$$\lambda_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} .$$

- A rotation of angle ϕ around the $X_3^{'}$ axis (positive sense); the assoiated matrix is:

$$\lambda_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The full transformation of the system of axes $\{X_1^{'}, X_2^{'}, X_3^{'}\}$ to the system of axes $\{X_1, X_2, X_3\}$ is given by (79), where

$$\lambda = \lambda_{_\phi} \lambda_\theta \lambda_{_\varphi} \ .$$

Doing the product of matrices, we get

$$\lambda_{11} = \cos\varphi \cos\phi - \cos\theta \sin\phi \sin\varphi$$

$$\lambda_{21} = -\sin\varphi\cos\phi - \cos\theta\sin\phi\cos\varphi$$
$$\lambda_{31} = \sin\theta\sin\phi$$
$$\lambda_{12} = \cos\varphi\sin\phi + \cos\theta\cos\phi\sin\varphi$$
$$\lambda_{22} = -\sin\varphi\sin\phi + \cos\theta\cos\phi\sin\varphi$$
$$\lambda_{32} = -\sin\varphi\cos\phi$$
$$\lambda_{13} = \sin\varphi\cos\phi$$
$$\lambda_{23} = \cos\varphi\sin\theta$$
$$\lambda_{33} = \cos\theta$$

where

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

Now, we take into account that:

- $\dot{\phi}$ is along the X'_3 (fixed) axis.
- $\dot{\theta}$ is along the so-called *line of nodes*.
- $\dot{\varphi}$ is along the X_3 axis (of the body).

This allows to write the three components of each of the three vectors in the system $\{X_1, X_2, X_3\}$ as follows:

$$\begin{array}{ccc} \dot{\phi}_1 = \dot{\phi}\sin\theta\sin\varphi, & \dot{\theta}_1 = \dot{\theta}\cos\varphi & \dot{\varphi}_1 = 0 \\ \dot{\phi}_2 = \dot{\phi}\sin\theta\cos\varphi, & \dot{\theta}_2 = -\dot{\theta}\sin\varphi & \dot{\varphi}_2 = 0 \\ \dot{\phi}_3 = \dot{\phi}\cos\theta, & \dot{\theta}_3 = 0 & \dot{\varphi}_3 = \dot{\varphi}\ . \end{array}$$

Then,

$$\begin{aligned} \omega &= \dot{\phi} + \dot{\theta} + \dot{\varphi} \\ &= \left[\left(\dot{\phi}_1 + \dot{\theta}_1 + \dot{\varphi}_1\right), \left(\dot{\phi}_2 + \dot{\theta}_2 + \dot{\varphi}_2\right), \left(\dot{\phi}_3 + \dot{\theta}_3 + \dot{\varphi}_3\right)\right]. \end{aligned}$$

Thus, we are led to the following components of ω:

$$\begin{aligned} \omega_1 &= \dot{\phi}\sin\theta\sin\varphi + \dot{\theta}\cos\varphi \\ \omega_2 &= \dot{\phi}\sin\theta\cos\varphi - \dot{\theta}\sin\varphi \\ \omega_3 &= \dot{\phi}\cos\theta + \dot{\varphi}. \end{aligned}$$

3.10 Symmetrical top with a fixed point

As a more complicated example of the methods used to describe the dynamics of the rigid body, we shall consider the motion of a symmetric body in a unifom gravitational field when a point of the axis of symmetry is fixed in the space.

The axis of symmetry is of course one of the principal axes, and we shall take it as the z axis of the body-fixed reference system. Since there is a fixed point, the configuration of the top will be determined by the three Euler angles: θ measuring the deviation of z from the vertical, ϕ, giving the azimuth of the top w.r.t. the vertical, and φ, which is the rotation angle of the top w.r.t. its proper z. The distance from the center of gravity to the fixed point will be denoted by l. To get a solution to the motion of the top we shall use the method of Lagrange instead of the Euler equations.

The kinetic energy is:

$$T = \frac{1}{2} I_1(\omega_1^2 + \omega_2^2) + \frac{1}{2} I_3 \omega_3^2 \ ,$$

or, in terms of the Euler angles:

$$T = \frac{1}{2} I_1(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2) + \frac{1}{2} I_3(\dot{\phi}\cos\theta + \dot{\varphi})^2 \ .$$

According to an elementary theorem, in a constant gravitational field the potential energy of a body is the same with that of a material point of equal mass concentrated in its center of mass. A formal proof is as follows. The potential energy of the body is the sum of the potential energies of all its particles:

$$V = -m_i \mathbf{r}_i \cdot \mathbf{g} \ , \tag{80}$$

where $\mathbf{g}$ is the constant acceleration of gravity. According to the definition of the center of mass, this is equivalent to

$$V = -M\mathbf{R}_i \cdot \mathbf{g}, \tag{81}$$

thus proving the theorem. The potential energy is a function of the Euler angles:

$$V = Mgl\cos\theta, \tag{82}$$

and the Lagrangian will be

$$L = \frac{1}{2} I_1(\dot{\phi}^2 \sin^2\theta + \dot{\theta}^2) + \frac{1}{2} I_3(\dot{\phi}\cos\theta + \dot{\varphi})^2 - Mgl\cos\theta. \tag{83}$$

We note that ϕ and φ are cyclic coordinates, and therefore p_ϕ and p_φ are constants of motion.

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = I_3(\dot{\varphi} + \dot{\phi}\cos\theta) = const \tag{84}$$

and

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = I_1\dot{\phi}\sin^2\theta + I_3(\dot{\phi}\cos^2\theta + \dot{\varphi}\cos\theta) = const. \tag{85}$$

From the equation (84) we get $\dot{\varphi}$

$$\dot{\varphi} = \frac{p_\varphi - I_3\dot{\phi}\cos\theta}{I_3}\ , \tag{86}$$

that we substitute in (85)

$$\begin{aligned} p_\phi &= \frac{\partial L}{\partial \dot{\phi}} = I_1\dot{\phi}\sin^2\theta + I_3(\dot{\phi}\cos^2\theta + \frac{p_\varphi - I_3\dot{\phi}\cos\theta}{I_3}\cos\theta) = const. \\ p_\phi &= I_1\dot{\phi}\sin^2\theta + p_\varphi\cos\theta\ , \end{aligned}$$

where from we get

$$\dot{\phi} = \frac{p_\phi - p_\varphi\cos\theta}{I_1\sin^2\theta}\ . \tag{87}$$

Substituting it in (86) one gets

$$\dot{\varphi} = \frac{p_\varphi}{I_3} - \frac{p_\phi - p_\varphi\cos\theta}{I_1\sin^2\theta}\cos\theta. \tag{88}$$

Now, since the system is conservative, another integral of motion is the energy

$$E = T + V = \frac{1}{2}I_1(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) + \frac{1}{2}I_3(\dot{\phi}\cos\theta + \dot{\varphi})^2 + Mgl\cos\theta.$$

The quantity $I_3\omega_3 = p_\varphi$ is an integral of motion. Multiplying this constant by $p_\varphi\omega_3$ we get

$$\begin{aligned} I_3 p_\varphi\omega_3^2 &= p_\varphi^2\omega_3 \\ I_3^2\omega_3^3 &= p_\varphi^2\omega_3 \\ \frac{1}{2}I_3\omega_3^2 &= \frac{1}{2}\frac{p_\varphi^2}{I_3}\ . \end{aligned}$$

The quantity $\frac{1}{2}I_3\omega_3^2$ is a constant. Therefore, we can define the quantity

$$\begin{aligned} E' &= E - \frac{1}{2}I_3\omega_3^2 = const. \\ &= \frac{1}{2}I_1\dot{\theta}^2 + \frac{1}{2}I_1\dot{\phi}^2\sin^2\theta + Mgl\cos\theta \ , \end{aligned}$$

wherefrom we can identify

$$V(\theta) = \frac{1}{2}\dot{\phi}^2\sin^2\theta + Mgl\cos\theta$$

$$V(\theta) = \frac{1}{2}I_1\left(\frac{p_\phi - p_\varphi\cos\theta}{I_1\sin^2\theta}\right)^2\sin^2\theta + Mgl\cos\theta. \tag{89}$$

Thus, E' is:

$$E' = \frac{1}{2}I_1\dot{\theta}^2 + V(\theta) \ .$$

From this equation we get $\dot{\theta} \equiv \frac{d\theta}{dt} = \left[\frac{2}{I_1}\left(E' - V(\theta)\right)\right]^{1/2}$, which leads to

$$t(\theta) = \int \frac{d\theta}{\sqrt[2]{\left(\frac{2}{I_1}\right)\left(E' - V(\theta)\right)}} \ . \tag{90}$$

Performing the integral in (90) one gets $t = f(\theta)$, and therefore, in principle, one can get $\theta(t)$. Then, $\theta(t)$ is replaced by $\dot{\phi}$ and $\dot{\varphi}$ (in eqs. (87) and (88)) and integrating them we can obtain the complete solution of the problem.

References

- H. Goldstein, *Classical Mechanics*, (Addison-Wesley, 1992).
- L. D. Landau & E. M. Lifshitz, *Mechanics,* (Pergammon, 1976).
- J. B. Marion & S.T. Thornton, *Classical Dynamics of Particles and Systems*, (Harcourt Brace, 1995).
- W. Wrigley & W.M. Hollister, *The Gyroscope: Theory and application*, Science 149, 713 (Aug. 13, 1965).

4. SMALL OSCILLATIONS

Forward: A familiar type of motion in mechanical and many other systems are the small oscillations (vibrations). They can be met as atomic and molecular vibrations, electric circuits, acoustics, and so on. In general, any motion in the neighborhood of stable equilibria is vibrational.

CONTENTS:

4.1 THE SIMPLE HARMONIC OSCILLATOR

A system is at stable equilibrium when its potential energy $U(q)$ is at minimum; when the system is slightly displaced from the equilibrium position, a force $-dU/dq$ occurs which acts to restore the equilibrium. Let q_0 be the value of the generalized coordinate corresponding to the equilibrium position. Expanding $U(q)-U(q_0)$ in a Taylor series of $q-q_0$ for small deviations from the equilibrium

$$U(q) - U(q_0) \cong \frac{1}{2}k(q-q_0)^2 \ ,$$

donde:

$$\begin{aligned} \frac{\partial U}{\partial q} &= 0 \\ U(q) &= 0 \ , \end{aligned}$$

This means that there are no external forces acting on the system and the zero has been chosen at the equilibrium position; moreover, higher-order terms have been neglected. The coefficient k represents the value of the second derivative of $U(q)$ for q=q_0. For simplicity reasons we denote

$$x = q - q_0$$

for which the potential energy can be written as:

$$U(x) = \frac{1}{2}kx^2 \ . \tag{1}$$

For simplicity reasons we denote

$$x = q - q_0$$

for which the potential energy can be written as:

$$U(x) = \frac{1}{2}kx^2 \ . \tag{2}$$

The kinetic energy of a system is

$$T = \frac{1}{2}m\,\dot{x}^2 \ , \tag{3}$$

and using (2) and (3) we get the Lagrangian of a system performing linear oscillations (so-called linear oscillator):

$$L = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2 \ . \tag{4}$$

The equation of motion corresponding to this L is:

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

or

$$\ddot{x} + w^2 x = 0 \ , \tag{5}$$

where $w^2 = \sqrt{k/m}$. This differential equation has two independent solutions: $\cos wt$ and $\sin wt$, from which one can form the general solution:

$$x = c_1 \cos wt + c_2 \sin wt \ , \tag{6}$$

or, we can also write the solution in the form:

$$x = a \cos(wt + \alpha) \ . \tag{7}$$

Since $\cos(wt+\alpha) = \cos wt \cos\alpha - \sin wt \sin\alpha$, by comparing with (6), one can see that the arbitrary constants a and α are related to c_1 and c_2 as follows:

$$a = \sqrt{(c_1^2 + c_2^2)}, \quad \text{y} \quad \tan\alpha = -c_1/c_2 \ .$$

Thus, any system in the neighborhood of the stable equilibrium position performs harmonic oscillatory motion. The a coefficient in (7) is the amplitude of the oscillations, whereas the argument of the cosine function is the phase of the harmonic oscillation; α is the initial value of the phase, which depends on the chosen origin of time. The quantity w is the angular frequency of the oscillations, which does not depend on the initial conditions of the system, being a proper characteristic of the harmonic oscillations.
Quite often the solution is expressed as the real part of a complex quantity

$$x = \mathrm{Re}\left[A \exp(iwt)\right]$$

where A is the complex amplitude, whose modulus gives the ordinary amplitude:

$$A = a \exp(i\alpha) \ .$$

The energy of a system in small oscillatory motion is:

$$E = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} k x^2 \ ,$$

or by substituting (7)

$$E = \frac{1}{2} m w^2 a^2 \ .$$

Now, we consider the case of n degrees of freedom. In this case, taking the sum of exterior forces as zero, the generalized force will be given by

$$Q_i = -\frac{\partial U}{\partial q_i} = 0 \ . \tag{8}$$

Repeating the procedure for the case of a single degree of freedom, we expand the potential energy in Taylor series taking the minimum of the potential energy at $q_i = q_{i0}$. Introducing small oscillation coordinates

$$x_i = q_i - q_{i0} \ ,$$

we can write the series as follows

$$U(q_1, q_2, ..., q_n) = U(q_{10}, q_{20}, ..., q_{n0}) + \sum \left(\frac{\partial U}{\partial q_i} \right)_0 x_i + \frac{1}{2!} \sum \left(\frac{\partial^2 U}{\partial q_i \partial q_j} \right)_0 x_i x_j + ... \tag{9}$$

Under the same considerations as given for (2), we obtain:

$$U(q_1, q_2, ..., q_n) = U = \frac{1}{2} \sum_{i,j} k_{ij} x_i x_j \ . \tag{10}$$

From (9) one notes that $k_{ij} = k_{ji}$, i.e., they are symmetric w.r.t. their subindices.

Let us look now to the kinetic energy, which, in general, is of the form

$$\frac{1}{2} a_{ij}(q) \dot{x}_i \dot{x}_j \ ,$$

where the a_{ij} are functions of the coordinates only. Denoting them by $a_{ij} = m_{ij}$ the kinetic energy will be

$$T = \frac{1}{2} \sum_{i,j} m_{ij} \dot{x}_i \dot{x}_j \ . \tag{11}$$

We can pass now to the Lagrangian for the system of n degrees of freedom

$$L = T - U = \frac{1}{2}\sum_{i,j}(m_{ij}\ \dot{x}_i\dot{x}_j\ - k_{ij}x_ix_j)\ . \tag{12}$$

This Lagrangian leads to the following set of simultaneous differential equations of motion

$$\frac{d}{dt}\frac{\partial L}{\partial\ \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \tag{13}$$

or

$$\sum(m_{ij}\ \ddot{x}_j\ + k_{ij}x_j) = 0\ . \tag{14}$$

This is a linear system of homogeneous equations, which can be considered as the n components of the matricial equation

$$(M)(\ddot{X}) + (K)(X) = 0\ , \tag{15}$$

where the matrices are defined by:

$$(M) = \begin{pmatrix} m_{11} & m_{12} & ... & m_{1n} \\ m_{21} & m_{22} & ... & m_{2n} \\ \vdots & & & \vdots \\ m_{n1} & m_{n2} & ... & m_{nn} \end{pmatrix} \tag{16}$$

$$(K) = \begin{pmatrix} k_{11} & k_{12} & ... & k_{1n} \\ k_{21} & k_{22} & ... & k_{2n} \\ \vdots & & & \vdots \\ k_{n1} & k_{n2} & ... & k_{nn} \end{pmatrix} \tag{17}$$

$$(\ddot{X}) = \frac{d^2}{dt^2}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{18}$$

$$(X) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\ . \tag{19}$$

Similarly to the one dof system, we look for n unknown functions $x_j(t)$ of the form

$$x_j = A_j \exp(iwt) \ , \tag{20}$$

where A_j are constants to be determined. Substituting (20) in (14) and dividing by $\exp(iwt)$, one gets a linear system of algebraic and homogeneous equations, which should be fulfilled by A_j.

$$\sum_j (-w^2 m_{ik} + k_{ik}) A_k = 0 \ . \tag{21}$$

This system has nponzero solutions if the determinant of its coefficients is zero.

$$\left| k_{ij} - w^2 m_{ij} \right|^2 = 0 \ . \tag{22}$$

This is the characteristic equation of order n w.r.t. w^2. In general, it has n different real and positive solutions w_α ($\alpha = 1, 2, ..., n$). The w_α are called proper frequencies of the system. Multiplying by A_i^* and summing over i one gets

$$\sum_j (-w^2 m_{ij} + k_{ij}) A_i^* A_j = 0 \ ,$$

where from

$$w^2 = \sum k_{ij} A_i^* A_i / \sum m_{ij} A_i^* A_i \ .$$

Since the coefficients k_{ij} and m_{ij} are real and symmetric, the quadratic forms of the numerator and denominator are real, and being essentially positive one concludes that w^2 are equally positive.

EXAMPLE

As an example we model the equations of motion of a double pendulum. The potential energy of this system with two degrees of freedom is

$$U = m_1 g l_1 (1 - \cos\theta_1) + m_2 g l_1 (1 - \cos\theta_1) + m_2 g l_2 (1 - \cos\theta_2) \ .$$

Applying (9), one gets

$$U = \frac{1}{2}(m_1 + m_2) g l_1 \theta_1^2 + \frac{1}{2} m_2 g l_2 \theta_2^2 \ .$$

Comparing with (10), we identify

$$\begin{aligned} k_{11} &= (m_1 + m_2) l_1^2 \\ k_{12} &= k_{21} = 0 \\ k_{22} &= m_2 g l_2 \ . \end{aligned}$$

For the kinetic energy one gets

$$T = \frac{1}{2}(m_1 + m_2) l_1^2 \, \dot{\theta}_1^2 + \frac{1}{2} m_2 l_2^2 \, \dot{\theta}_2^2 + m_2 l_1 l_2 \, \dot{\theta}_1 \dot{\theta}_2 \ \ .$$

Identifying terms from the comparison with (11) we find

$$\begin{aligned} m_{11} &= (m_1 + m_2) l_1^2 \\ m_{12} &= m_{21} = m_2 l_1 l_2 \\ m_{22} &= m_2 l_2^2 \ . \end{aligned}$$

Substituting the energies in (12) one obtains the Lagrangian for the double pendulum oscillator and as the final result the equations of motion for this case:

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{pmatrix} + \begin{pmatrix} k_{11} & 0 \\ 0 & k_{22} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = 0 \ .$$

4.2 FORCED HARMONIC OSCILLATOR

If an external weak force acts on an oscillator system the oscillations of the system are known as forced oscillations.
Besides its proper potential energy the system gets a supplementary potential energy $U_e(x,t)$ due to the external field. Expanding the latter in a Taylor series of small amplitudes x we get:

$$U_e(x,t) \cong U_e(0,t) + x \left[\frac{\partial U_e}{\partial x} \right]_{x=0} \ .$$

The second term is the external force acting on the system at its equilibrium position, that we denote by $F(t)$. Then, the Lagrangian reads

$$L = \frac{1}{2} m \, \dot{x}^2 - \frac{1}{2} k x^2 + x F(t) \ . \tag{23}$$

The corresponding equation of motion is

$$m\,\ddot{x} + kx = F(t)\ ,$$

or

$$\ddot{x} + w^2 x = F(t)/m\ , \tag{24}$$

where w is the frequency of the proper oscillations. The general solution of this equation is the sum of the solution of the homogeneous equation and a particular solution of the nonhomogeneous equation

$$x = x_h + x_p\ .$$

We shall study the case in which the external force is periodic in time of frequency γ forma

$$F(t) = f\cos(\gamma t + \beta)\ .$$

The particular solution of (24) is sought in the form $x_1 = b\cos(\gamma t + \beta)$ and by substituting it one finds that the relationship $b = f/m(w^2 - \gamma^2)$ should be fulfilled. Adding up both solutions, one gets the general solution

$$x = a\cos(wt + \alpha) + \left[f/m(w^2 - \gamma^2)\right]\cos(\gamma t + \beta)\ . \tag{25}$$

This result is a sum of two oscillations: one due to the proper frequency and another at the frequency of the external force.
The equation (24) can in general be integrated for an arbitrary external force. Writing it in the form

$$\frac{d}{dt}(\dot{x} + iwx) - iw(\dot{x} + iwx) = \frac{1}{m}F(t)\ ,$$

and making $\xi = \dot{x} + iwx$, we have

$$\frac{d}{dt}\xi - iw\xi = F(t)/m\ .$$

The solution to this equation is $\xi = A(t)\exp(iwt)$; for $A(t)$ one gets

$$\dot{A} = F(t)\exp(-iwt)/m\ .$$

Integrating it leads to the solution

$$\xi = \exp(iwt) \int_0^t \frac{1}{m} F(t) \exp(-iwt) dt + \xi_o \ . \tag{26}$$

This is the general solution we look for; the function x(t) is given by the imaginary part of the general solution divided by w.

EXAMPLE

We give here an example of employing the previous equation.
Determine the final amplitude of oscillations of a system acted by an extenal force $F_0 = const.$ during a limited time T. For this time interval we have

$$\xi = \frac{F_0}{m} \exp(iwt) \int_0^T \exp(-iwt) dt \ ,$$

$$\xi = \frac{F_0}{iwm} [1 - \exp(-iwt)] \exp(iwt) \ .$$

Using $|\xi|^2 = a^2 w^2$ we obtain

$$a = \frac{2F_0}{mw^2} \sin(\frac{1}{2} wT) \ .$$

4.3 DAMPED HARMONIC OSCILLATOR

Until now we have studied oscillatory motions in free space (the vacuum), or when the effects of the medium through which the oscillator moves are negligeable. However, when a system moves through a medium its motion is retarded by the reaction of the latter. There is a dissipation of the energy in heat or other forms of energy. We are interested in a simple description of the dissipation phenomena.
The reaction of the medium can be imagined in terms of friction forces. When they are small we can expand them in powers of the velocity. The zero-order term is zero because there is no friction force acting on a body at rest. Thus, the lowest order nonzero term is proportional to the velocity, and moreover we shall neglect all higher-order terms

$$f_r = -\alpha \ \dot{x} \ \ ,$$

where x is the generalized coordinate and α is a positive coefficient; the minus sign shows the oposite direction to that of the moving system. Adding this force to the equation of motion we get

$$m \ \ddot{x} = -kx - \alpha \ \dot{x} \ \ ,$$

or

$$\ddot{x} = -kx/m - \alpha\, \dot{x}\, /m\ . \tag{27}$$

Writing $k/m = w_o^2$ and $\alpha/m = 2\lambda$; where w_o is the frequency of free oscillations of the system and λ is the damping coefficient. Therefore

$$\ddot{x} + 2\lambda\, \dot{x} + w_o^2 x = 0\ .$$

The solution of this equation is sought of the type $x = \exp(rt)$, which we substitute back in the equation to get the characteristic equation for r. Thus

$$r^2 + 2\lambda + w_o^2 = 0\ ,$$

where from

$$r_{1,2} = -\lambda \pm \sqrt{(\lambda^2 - w_o^2)}\ .$$

We are thus led to the following general solution of the equation of motion

$$x = c_1 \exp(r_1 t) + c_2 \exp(r_2 t)\ .$$

Among the roots r we shall look at the following particular cases:

(i) $\lambda < w_o$. One gets complex conjugate solutions. The solution is

$$x = \mathrm{Re}\left\{ A exp\left[-\lambda t + i\sqrt{(w_o^2 - \lambda^2)} \right] \right\}\ ,$$

where A is an arbitrary complex constant. The solution can be written of the form

$$x = a \exp(-\lambda t) \cos(wt + \alpha), \quad \text{where} \quad w = \sqrt{(w_o^2 - \lambda^2)}\ , \tag{28}$$

where a and α are real constants. Thus, one can say that a damped oscillation is a harmonic oscillation with an exponentially decreasing amplitude. The rate of decreasing of the amplitude is determined by the exponent λ. Moreover, the frequency w is smaller than that of free oscillations.

(ii) $\lambda > w_o$. Then, both r are real and negative. The general form of the solution is:

$$x = c_1 \exp\left\{ -\left[\lambda - \sqrt{(\lambda^2 - w_o^2)} \right] t \right\} + c_2 \exp\left\{ -\left[\lambda + \sqrt{(\lambda^2 - w_o^2)} \right] t \right\}\ .$$

If the friction is large, the motion is just a monotone decaying amplitude asymptotically ($t \to \infty$) tending to the equilibrium position (without oscillations). This type of motion is called aperiodic.

(iii) $\lambda = w_o$. Then $r = -\lambda$, whose general solution is

$$x = (c_1 + c_2 t) \exp(-\lambda t) \ .$$

If we generalize to systems of n degrees of freedom, the generalized friction forces corresponding to the coordinates x_i are linear functions of the velocities

$$f_{r,i} = \sum_j \alpha_{ij} \ \dot{x}_i \quad . \tag{29}$$

Using $\alpha_{ik} = \alpha_{ki}$, one can also write

$$f_{r,i} = -\frac{\partial F}{\partial \ \dot{x}_i} \ ,$$

where $F = \frac{1}{2} \sum_{i,j} \alpha_{ij} \ \dot{x}_i \dot{x}_j$ is called the dissipative function. The differential equation is obtained by adding up all these forces to (14)

$$\sum (m_{ij} \ \ddot{x}_j + k_{ij} x_j) = -\sum_j \alpha_{ij} \ \dot{x}_i \quad . \tag{30}$$

Employing

$$x_k = A_k \exp(rt)$$

in (30) and deviding by $\exp(rt)$, one can obtain the following system of linear algebraic equations for the constants A_j

$$\sum_j (m_{ij} r^2 + \alpha_{ij} r + k_{ij}) A_j = 0 \ .$$

Making equal to zero the determinant of this system, one gets the corresponding characteristic equation

$$\left| m_{ij} r^2 + \alpha_{ij} r + k_{ij} \right| = 0 \ . \tag{31}$$

This is an equation for r of degree $2n$.

4.4 NORMAL MODES

Before defining the normal modes, we rewrite (15) as follows

$$M\left|\ddot{X}\right\rangle + K\left|X\right\rangle = 0 \ ,$$

where $|X\rangle$ is the n-dimensional vector whose matrix representation is (19); M and K are two operators having the matrix representation given by (16) and (17), respectively. We have thus an operatorial equation. Since M is a nonsingular and symmetric operator, the inverse operator M^{-1} and the operators $\mathrm{M}^{1/2}$ and $\mathrm{M}^{-1/2}$ are well defined. In this case, we can express the operatorial equation in the form

$$\frac{d^2}{dt^2}M^{1/2}\left|X\right\rangle = -M^{-1/2}KM^{-1/2}M^{1/2}\left|X\right\rangle \ ,$$

or more compactly

$$\frac{d^2}{dt^2}\left|\bar{X}\right\rangle = -\lambda\left|\bar{X}\right\rangle \ , \tag{32}$$

where

$$\left|\bar{X}\right\rangle = M^{1/2}\left|X\right\rangle$$

and

$$\lambda = M^{-1/2}KM^{-1/2} \ .$$

Since $\mathrm{M}^{-1/2}$ and K are symmetric operators, then λ is also symmetric. If we use orthogonal eigenvectors as a vectorial base (for example, the three-dimensional Euclidean space), the matrix representation of the operator can be diagonal, e.g.,

$$\lambda_{ij} = \lambda_i\delta_{ij} \ .$$

Let us consider the following eigenvalue problem

$$\lambda\left|\rho_i\right\rangle = \lambda_i\left|\rho_i\right\rangle \ , \tag{33}$$

where $|\rho_i\rangle$ is an orthogonal set of eigenvectors. Or

$$M^{-1/2}KM^{-1/2}\left|\rho_i\right\rangle = \lambda_i\left|\rho_i\right\rangle \ .$$

The eigenvalues are obtained by multiplying both sides by $\langle \rho_i |$, leading to

$$\lambda_i = \frac{\langle \rho_i | \, M^{-1/2} K M^{-1/2} \, | \rho_i \rangle}{\langle \rho_i \, | \rho_i \rangle} \ .$$

Since the potential and kinetic energies are considered positive quantities, one should take

$$\langle \rho_i | \, M^{-1/2} K M^{-1/2} \, | \rho_i \rangle \rangle 0$$

and therefore

$$\lambda_i > 0 \ .$$

This leads to the set

$$\lambda_i = w_i^2 \ .$$

If we express the vector $\left| \bar{X} \right\rangle$ in terms of these eigenvectors of λ,

$$\left| \bar{X} \right\rangle = \sum_i y_i \left| \bar{X} \right\rangle \ ,$$

where

$$y_i = \langle \rho_i \left| \bar{X} \right\rangle \ . \tag{34}$$

Inserting this result in the equation of motion (32), we obtain

$$\frac{d^2}{dt^2} \sum_i y_i \, | \rho_i \rangle = -\lambda \left| \bar{X} \right\rangle = - \sum_i \lambda_i y_i \, | \rho_i \rangle \ .$$

The scalar product of this equation with the constant eigenvector $\langle \rho_j |$ leads to the equation of motion for the generalized coordinate y_j

$$\frac{d^2}{dt^2} y_j = -w_j^2 y_j \ .$$

The solution of this equation reads

$$y_j = A_j \cos(w_j t + \phi_j) \ . \tag{35}$$

Use of these new generalized harmonic coordinates lead to a set of independent equations of motion. The relationship between y_j and $\bar{x}_i$ is given by (34)

$$y_j = \rho_{j1}\,\bar{x}_1 + \rho_{j2}\,\bar{x}_2 + ... + \rho_{jn}\,\bar{x}_n \quad .$$

The components ρ_{jl} $(l = 1, 2, .., n)$ are determined by solving the eigenvalue problem given by (33). The new coordinates are called normal coordinates and the w_j are known as the normal frequencies. The equivalent matrix form (35) is

$$\begin{pmatrix} \bar{x}_1^{(j)} \\ \bar{x}_2^{(j)} \\ \vdots \\ \bar{x}_n^{(j)} \end{pmatrix} = A_j \cos(w_j t + \phi_j) \begin{pmatrix} \rho_{j1} \\ \rho_{j2} \\ \vdots \\ \rho_{jn} \end{pmatrix} \quad . \tag{36}$$

These are the normal vibrational modes of the system. One reason for introducing the coordinates y_j is found from the expression for the kinetic energy, which is seen to be invariant under the rotation to the new axes.

$$T = \frac{1}{2} \sum_{j=1}^{n} M_j\, \dot{y}_j^2 \quad .$$

EXAMPLE

Apply the matricial procedure as already shown, given the following equations of motion

$$\frac{d^2}{dt^2} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = - \begin{pmatrix} 5 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 5 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} \quad .$$

Comparing with (32), we identify the operator λ. To find the eigenvectores, we use (33) getting

$$\begin{pmatrix} 5 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 5 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = \lambda_i \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} \quad .$$

The characteristic equation for λ_i is

$$\det(\lambda - \lambda_i I) = 0 \ ,$$

and by substituting the values

$$\begin{vmatrix} 5-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 5-\lambda \end{vmatrix} = 0 \ .$$

Solving the equation one gets $\lambda_i = 2, 4, 6$. For $\lambda = 4$

$$\begin{pmatrix} 5 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 5 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = 4 \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix}$$

we have the following set of equations

$$\begin{aligned} (5-4)\rho_1 + \rho_3 &= 0 \\ 2\rho_2 - 4\rho_2 &= 0 \\ \rho_1 + (5-4)\rho_3 &= 0 \ . \end{aligned}$$

Taking into account the normalization condition, one is led to the following values

$$\begin{aligned} \rho_1 &= -\rho_3 = \frac{1}{\sqrt{2}} \\ \rho_2 &= 0 \ . \end{aligned}$$

Therefore

$$|\rho_{\lambda=4}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \ .$$

By the same means one gets

$$\begin{aligned} |\rho_{\lambda=6}\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ |\rho_{\lambda=2}\rangle &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \ . \end{aligned}$$

Thus, the new vectorial space is determined by

$$|\rho_i\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} ,$$

where from

$$\langle\rho_i| = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} .$$

Thus, the normal coordinates are given by (34)

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} .$$

4.5 PARAMETRIC RESONANCE

The important phenomenon of parametric resonance shows up for systems initially at rest in un unstable equilibrium position, say $x = 0$; thus, the slightest deviation from this position produces a displacement growing rapidly (exponentially) in time. This is different from the ordinary resonances, where the displacement grows only linearly in time.
The parameters of a linear system are the coefficients m and k of the Lagrangian (4); if they are functions of time, the equation of motion is:

$$\frac{d}{dt}(m\,\dot{x}) + kx = 0 \ . \tag{37}$$

If we take a constant mass, the previous equation can be written in the form

$$\frac{d^2x}{dt^2} + w^2(t)x = 0 \ . \tag{38}$$

The function $w(t)$ is given by the problem at hand. Assuming it a periodic function of frequency γ (of period $T = 2\pi/\gamma$), i.e.,

$$w(t+T) = w(t) \ ,$$

any equation of the type (38) is invariant w.r.t. the transformation $t \to t+T$. Thus, if $x(t)$ is one of its solutions, $x(t+T)$ is also a solution. Let $x_1(t)$ and

$x_2(t)$ be two independent solutions of 4.6.2). They should change to itselves in a linear combination when $t \to t+T$. Thus, one gets

$$\begin{aligned} x_1(t+T) &= \mu_1 x(t) \\ x_2(t+T) &= \mu_2 x(t) \ , \end{aligned} \tag{39}$$

or, in general

$$\begin{aligned} x_1(t) &= \mu_1^{t/T} F(t) \\ x_2(t) &= \mu_2^{t/T} G(t) \ , \end{aligned}$$

where $F(t)$ and $G(t)$ are periodical functions in time of period T. The relationship between these constants can be obtained by manipulating the following equations

$$\begin{aligned} \ddot{x}_1 + w^2(t) x_1 &= 0 \\ \ddot{x}_2 + w^2(t) x_2 &= 0 \ . \end{aligned}$$

Multiplying by x_2 and x_1, respectively, and substracting term by term, we get

$$\ddot{x_1}\ x_2 - \ddot{x}_2\ x_1 = \frac{d}{dt}(\dot{x_1}\ x_2 - \dot{x}_2\ x_1) = 0 \ ,$$

or

$$\dot{x_1}\ x_2 - \dot{x}_2\ x_1 = const. \ .$$

Substituting t by $t+T$ in the previous equation, the right hand side is multiplied by $\mu_1\mu_2$ (see eqs. (39)); thus, it is obvious that the following condition holds

$$\mu_1 \mu_2 = 1 \ , \tag{40}$$

where one should take into account (38) and the fact that the coefficients are real. If $x(t)$ is one integral of this equation, then $x^*(t)$ is also a solution. Therefore μ_1, μ_2 should coincide with μ_1^*, μ_2^*. This leads to either $\mu_1 = \mu_2^*$ or μ_1 and μ_2 both real. In the first case, based on (40) one gets $\mu_1 = 1/\ \mu_1^*$,

that is equivalent to $|\mu_1|^2 = |\mu_2|^2 = 1$. In the second case, the two solutions are of the form

$$\begin{aligned} x_1(t) &= \mu^{t/T} F(t) \\ x_2(t) &= \mu^{-t/T} G(t) \ . \end{aligned}$$

One of these functions grows exponentially in time, which is the characteristic feature of the parametric resonance.

REFERENCES AND FURTHER READING

* H. Goldstein, *Classical mechanics*, Second ed. (Addison-Wesley, 1981).

* L. D. Landau & E. M. Lifshitz, *Mechanics*, (Pergammon, 1976).

* W. Hauser, *Introduction to the principles of mechanics*, (Wesley, 1965).

* E.I. Butikov, *Parametric Resonance*, Computing in Science & Engineering, May/June 1999, pp. 76-83 (http://computer.org).

5. CANONICAL TRANSFORMATIONS

Forward: The main idea of canonical transformations is to find all those coordinate systems in the phase space for which the form of the Hamilton eqs is invariant for whatever Hamiltonian. In applications one chooses the coordinate system that allows a simple solution of the problem at hand.

CONTENTS:

5.1 Definitions, Hamiltonians and Kamiltonians

For the time-independent and time-dependent cases, respectively, one defines a canonical transformation as follows

Definition 1: A time-independent transformation $Q = Q(q,p)$, and $P = P(q,p)$ is called canonical if and only if there is a function $F(q,p)$ such that

$$dF(q,p) = \sum_i p_i dq_i - \sum_i P_i(q,p) dQ_i(q,p) \ .$$

Definition 2: A time-dependent transformation $Q = Q(q,p,t)$, and $P = P(q,p,t)$ is called canonical if and only if there is a function $F(q,p,t)$ such that for an arbitrary fixed time $t = t_0$

$$dF(p,q,t_0) = \sum_i p_i dq_i - \sum_i P_i(q,p,t_0) dQ_i(p,q,t_0) \ ,$$

where

$$dF(p,q,t_0) = \sum_i \frac{\partial F(p,q,t_0)}{\partial q_i} dq_i + \sum_i \frac{\partial F(p,q,t_0)}{\partial p_i} dp_i$$

and

$$dQ(p,q,t_0) = \sum_i \frac{\partial Q(p,q,t_0)}{\partial q_i} dq_i + \sum_i \frac{\partial Q(p,q,t_0)}{\partial p_i} dp_i$$

Example: Prove that the following transformation is canonical

$$\begin{aligned} P &= \frac{1}{2}(p^2+q^2) \\ Q &= Tan^{-1}\left(\frac{q}{p}\right) \ . \end{aligned}$$

Solution: According to the first definition we have to check that $pdq - PdQ$ is an exact differential. Substituting P and Q in the definition we get

$$pdq - PdQ = pdq - \frac{1}{2}(p^2+q^2)\frac{pdq - qdq}{p^2+q^2} = d\left(\frac{pq}{2}\right) \ .$$

We can see that indeed the given transformation is canonical. We know that a dynamical system is usually characterized by a Hamiltonian $H = H(q,p,t)$, where $q = q(q_1, q_2, ..., q_n)$, and $p = p(p_1, p_2, ..., p_n)$. Therefore, the dynamics of the system fulfills a set of $2n$ first-order differential eqs

(Hamilton's eqs.)

$$\dot{q_i} = \frac{\partial H}{\partial p_i} \tag{1}$$

$$-\dot{p_i} = \frac{\partial H}{\partial q_i} \ . \tag{2}$$

Let us denote the coordinate transformations in the phase space by

$$Q_j = Q_j(q, p, t) \tag{3}$$

$$P_j = P_j(q, p, t) \ . \tag{4}$$

According to the aforementioned principle for the set of canonical transformations denoted by (3) and (4), analogously to (1) and (2), there is a function $K = K(Q, P, t)$ such that we can write

$$\dot{Q_i} = \frac{\partial K}{\partial P_i} \tag{5}$$

$$-\dot{P_i} = \frac{\partial K}{\partial Q_i} \ . \tag{6}$$

The relationship between the Hamiltonian H and the Kamiltonian K[1] can be obtained arguing as follows[2].

According to Hamilton's principle, the real trajectory of a classical system can be obtained from the variation of the action integral

$$\delta \int (\sum_i p_i dq_i - H dt) = 0 \ . \tag{7}$$

If the transformation is canonical, the Kamiltonian K should fulfill a relationship similar to (7). In other words, for the new set of variables Q and P we still have

$$\delta \int (\sum_i P_i dQ_i - K dt) = 0 \ . \tag{8}$$

[1]Here we follow the terminology of Goldstein by referring to $K = K(Q, P, t)$, which is different from the Hamiltonian $H = H(p, q, t)$ by an additive time derivative, as the *Kamiltonian.*

[2]An alternative derivation has been given by G. S. S. Ludford and D. W. Yannitell, Am. J. Phys. 36, 231 (1968).

Moreover, according to the Legendre transformation, $\sum_i p_i dq_i - Hdt = L(q, \dot{q}, t)dt$, (7) - like (8)- is equivalent to

$$\delta \int_{t_1}^{t_2} L(q, \dot{q}, t)dt = 0 \ . \tag{9}$$

In addition, (9) does not change if L is replaced by $\mathcal{L} = L + \frac{dF(q,t)}{dt}$ because in this case

$$\delta \int_{t_1}^{t_2} \mathcal{L}dt = \delta \int_{t_1}^{t_2} (L + \frac{dF(q,t)}{dt})dt \ , \tag{10}$$

or similarly

$$\delta \int_{t_1}^{t_2} \mathcal{L}dt = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)dt + \delta F(q_{(2)}, t_2) - \delta F(q_{(1)}, t_1) \ . \tag{11}$$

Thus, (10) and (11) differ only by constant terms whose variation is zero in Hamilton's principle.
It follows that the Hamiltonian H and the Kamiltonian K are related by the equation [3]

$$p_i \dot{q}_i - H = P_i \dot{Q}_i - K + \frac{dF}{dt} \ . \tag{12}$$

The function F is the so-called *generating function*. It can be expressed as a function of any arbitrary set of independent variables. However, some very convenient results are obtained if F is expressed as a function of the n old variables and the n new ones, plus the time. The results are especially convenient if the n old variables are exactly the n q_i - or the n p_i-, and if the new variables are all of them the n Q_i - or the n P_i.

Using these coordinates, the possible combinations of n old variables and n new variables -including t- in the generating function are[4]

$$\begin{aligned} F_1 &= F_1(Q, q, t) \\ F_2 &= F_2(P, q, t) \\ F_3 &= F_3(Q, p, t) \\ F_4 &= F_4(P, p, t) \ . \end{aligned} \tag{13}$$

[3]Some authors add to the right hand side of this equation a constant multiplicative factor A that does not change (9). Here we use $A = 1$, that is, we decided to work with the so-called *reduced canonical transformations*, since this simpler case is sufficient to illustrate the structure of the canonical transformations.

[4]We shall use the convention of Goldstein to denote each of the different combinations of the new and old variables in the generating function.

On the other hand, if we multiply (12) by dt we get:

$$p_i dq_i - H dt = P dQ_i - K dt + dF \ . \tag{14}$$

Making the change $F \to F_1$ above, and recalling that dQ_i, dq_i, and dt are independent variables, we get:

$$\begin{aligned} P_i &= -\frac{\partial F_1}{\partial Q_i} \\ p_i &= \frac{\partial F_1}{\partial q_i} \\ K &= H + \frac{\partial F_1}{\partial t} \ . \end{aligned}$$

Using now some algebraic manipulation it is possible to obtain analogous expressions to the previous one for the rest of the generating functions. The results are the following:

$$F_2 : \quad Q_i = \frac{\partial F_2}{\partial P_i} \quad p_i = \frac{\partial F_2}{\partial q_i} \quad K = H + \frac{\partial F_2}{\partial t}$$

$$F_3 : \quad P_i = -\frac{\partial F_3}{\partial Q_i} \quad q_i = -\frac{\partial F_3}{\partial p_i} \quad K = H + \frac{\partial F_3}{\partial t}$$

$$F_4 : \quad Q_i = \frac{\partial F_4}{\partial P_i} \quad q_i = -\frac{\partial F_4}{\partial p_i} \quad K = H + \frac{\partial F_4}{\partial t} \ .$$

In practice, one usually applies a useful theorem (see below) that allows, together with the definitions we gave in the introduction for canonical transformations, to solve any mechanical problem of interest [5].

Theorem 5.1 We consider a system acted by a given external force. We also suppose that the dynamical state of the system is determined by a set of variables $q, p = q_1, q_2, ..., q_n, p_1, p_2, ..., p_n$ and that the Hamiltonian of the system is $H = H(q, p, t)$. The time evolution of the variables q and p is given by Hamilton's eqs.

$$\begin{aligned} \dot{q_i} &= \frac{\partial H(q,p,t)}{\partial p_i} \\ \dot{p_i} &= -\frac{\partial H(q,p,t)}{\partial q_i} \ . \end{aligned}$$

If we now perform a transformation to the new variables

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

[5] For an example, see the final section of this chapter.

and if the transformation is canonical, i.e., there exists a function $F(q,p,t)$ such that for a fixed arbitrary time $t=t_0$ we have

$$dF(q,p,t_0)=\sum_i y_i dx_i-\sum_i Y_i dX_i\ ,$$

where $x_i,y_i=q_i,p_i$ or $p_i,-q_i$ y $X_i,Y_i=Q_i,P_i$, or $P_i,-Q_i$, then the equations of motion in terms of the variables Q and P are

$$\begin{aligned}\dot{Q}_i &= \frac{\partial K(Q,P,t)}{\partial P_i}\\ \dot{P}_i &= -\frac{\partial K(Q,P,t)}{\partial Q_i}\ ,\end{aligned}$$

where

$$K\equiv H+\frac{\partial F(q,p,t)}{\partial t}+\sum_i Y_i\frac{\partial X_i(q,p,t)}{\partial t}\ .$$

Moreover, if the determinant of the matrix $[\frac{\partial X_i}{\partial y_j}]$ is different of zero, then the latter equation takes the form

$$K\equiv H+\frac{\partial F(x,X,t)}{\partial t}\ .$$

5.2 Necessary and sufficient conditions for a transformation to be canonical

We have already mentioned that by a canonical transformation we mean a transformation, which, independently of the form of the Hamiltonian, keeps unchanged the form of Hamilton's equations. However, one should be very careful with this issue because some transformations fulfill this requirement only for *a particular Hamiltonian* [6]. Some authors call such transformations *canonical transformations w.r.t. H* [7].

To illustrate this point we use the following example, given in the paper of J. Hurley: Let us consider a *particular physical system* whose Hamiltonian is

$$H=\frac{p^2}{2m} \tag{15}$$

[6] See, for example, J. Hurley Am. J. Phys. **40**, 533 (1972).

[7] See, for example, R. A. Matzner and L. C. Shepley, *Classical Mechanics* (Prentice Hall, 1991).

and the following transformations

$$\begin{aligned} P &= p^2 \\ Q &= q \ . \end{aligned} \tag{16}$$

It is easy to show that the Kamiltonian K given by

$$K = \frac{2P^{3/2}}{3m}$$

leads to

$$\dot{P} = 2p\dot{p} = 0 = -\frac{\partial K}{\partial Q}$$

and

$$\dot{Q} = \dot{q} = \frac{p}{m} = \frac{P^{1/2}}{m} = \frac{\partial K}{\partial P} \ . \tag{17}$$

On the other hand, if we choose the Hamiltonian

$$H = \frac{p^2}{2m} + q^2 \ ,$$

then it is possible to find a Kamiltonian K for which the usage of the transformation equations (16) maintains unchanged the form of Hamilton's equations. Thus, the equations (16) keeps unchanged the form of Hamilton's equations only for a *particular Hamiltonian*.
It can be shown that the necessary and sufficient conditions for the canonicity of transformations of the form (3) and (4), that is, to keep unchanged the form of Hamilton's equations *whatever the Hamiltonian*, are the following

$$[Q_i, P_j] = \alpha \tag{18}$$

$$[P_i, P_j] = 0 \tag{19}$$

$$[Q_i, Q_j] = 0 \ , \tag{20}$$

where α is an arbitrary constant related to scale changes. Finally, we would like to make a few important comments before closing this section. First, we should keep in mind that Q and P *are not variables defining the configuration of the system*, i.e., they are not in general a set of generalized coordinates [8]. To distinguish Q and P from the generalized coordinates q and p, one

[8] Except for the trivial case in which the canonical transformation is $Q = q$ and $P = p$.

calls them *canonical variables*. In addition, the equations of motion -similar in form to the Hamiltonian ones- for the generalized coordinates q and p- that one gets for Q and P are called *canonical Hamilton equations*. Second, although we did not check here, if the transformation $Q = Q(q, p, t)$ and $P = P(q, p, t)$ is canonical, then its inverse $q = q(Q, P, t)$ and $p = p(Q, P, t)$ is also canonical [9].

5.3 Example of application of TC

As we already mentioned in the introduction, the main idea in performing a canonical transformation is to find a phase space coordinate system for which the form of the Hamilton eqs is maintaind whatever the Hamiltonian and to choose the one that makes easy the solution of the problem. We illustrate this important fact with the following example.

EXAMPLE:

The Hamiltonian of a physical system is given by $H = \omega^2 p(q+t)^2$, where ω is a constant. Determine q as a function of time.

Solution:

1. *Solving the Hamilton equations for the variables q and p.* Applying (1) and (2) to the given Hamiltonian we get

$$\omega^2(q+t)^2 = \dot{q}, \qquad 2\omega^2 p(q+t) = -\dot{p} \ .$$

This system is not easy to solve. However, we can get the solution by means of an appropriate canonical transformation as we show in the following.

2. *Using $Q = q + t$, $P = p$.* According to the theorem given in section (5.1), since

$$\begin{aligned} \frac{\partial Q}{\partial p} &= 0 \\ \frac{\partial P}{\partial(-q)} &= 0 \ , \end{aligned}$$

then the Kamiltonian K of the system is given by

$$K = H + \frac{\partial F(q,p,t)}{\partial t} + P\frac{\partial Q}{\partial t} - Q\frac{\partial P}{\partial t} \ . \tag{21}$$

[9] For a proof see, for example, E. A. Desloge, *Classical Mechanics, Volume 2* (John Wiley & Sons, 1982).

The form of the function $F(q,p,t)$ can be obtained from the canonical transformation given in section 5.1 (the case corresponds to a time-dependent canonical transformation). Therefore, we substitute $Q = q + t$, $P = p$ in

$$dF(q,p,t) = pdq - PdQ \ ,$$

to get without difficulty

$$F(q,p,t) = c, \quad \text{c= constant} \ .$$

On the other hand,

$$\begin{aligned} \frac{\partial P}{\partial t} &= 0 \\ \frac{\partial Q}{\partial t} &= 1 \ . \end{aligned}$$

Finally, substituting these results in (21) (and also $Q = q + t$, $P = p$ in H) we get

$$K = P(\omega^2 Q^2 + 1) \ .$$

Moreover, from (5) we find

$$\dot{Q} = \omega^2 Q^2 + 1 \ .$$

This differential equation is now easy to solve, leading to

$$q = \frac{1}{\omega}\tan(\omega t + \phi) - t \ ,$$

where ϕ is an arbitrary phase.

6. POISSON BRACKETS

Forward: The Poisson brackets are very useful analytical tools for the study of any dynamical system. Here, we define them, give some of their properties, and finally present several applications.

CONTENTS:

1. Definition and properties of Poisson brackets

If u and v are any two quantities that depend on the dynamical state of a system, i.e., on p and q) and possibly on time, the Poisson bracket of u and v w.r.t. a set of canonical variables q and p [10] is defined as follows

$$[u,v] \equiv \sum_i \left(\frac{\partial u(q,p,t)}{\partial q_i} \frac{\partial v(q,p,t)}{\partial p_i} - \frac{\partial u(q,p,t)}{\partial p_i} \frac{\partial v(q,p,t)}{\partial q_i} \right) . \qquad (1)$$

The Poisson brackets have the following properties (for u, v, and w arbitrary functions of q, p, and t; a is an arbitrary constant, and r is any of q_i, p_i or t) [11]:

1. $[u,v] \equiv -[v,u]$

2. $[u,u] \equiv 0$

3. $[u,v+w] \equiv [u,v] + [u,w]$

4. $[u,vw] \equiv v[u,w] + [u,v]w$

5. $a[u,v] \equiv [au,v] \equiv [u,av]$

6. $\frac{\partial [u,v]}{\partial r} \equiv [\frac{\partial u}{\partial r}, v] + [u, \frac{\partial v}{\partial r}]$

7. The Jacobi identity, $[u,[v,w]] + [v,[w,u]] + [w,[u,v]] \equiv 0$.

Another very important property of PBs is the content of the following theorem

Theorem 6.1 If the transformation $Q = Q(q,p,t)$, $P = P(q,p,t)$ is a canonical transformation, the PB of u and v w.r.t. the variables q, p is equal to the PB of u and v w.r.t. the set of variables Q, P, i.e.,

$$\sum_i \left(\frac{\partial u(q,p,t)}{\partial q_i} \frac{\partial v(q,p,t)}{\partial p_i} - \frac{\partial u(q,p,t)}{\partial p_i} \frac{\partial v(q,p,t)}{\partial q_i} \right) =$$

$$\sum_i \left(\frac{\partial u(q,p,t)}{\partial Q_i} \frac{\partial v(q,p,t)}{\partial P_i} - \frac{\partial u(q,p,t)}{\partial P_i} \frac{\partial v(q,p,t)}{\partial Q_i} \right) .$$

[10] As in the previous chapter, by q and p we mean $q = q_1, q_2, ..., q_n$ y $p = p_1, p_2, ..., p_n$.

[11] The proof of these properties can be obtained by using the definition of the PBs in order to express each term of these identities by means of partial derivatives of u, v, and w, and noticing by inspection that the resulting equations do hold.

2. Poisson formulation of the equations of motion

In the following, we outline as theorem-like statements the most important results on the PB formulation of the eqs of motion of the dynamical systems [12]:

Theorem 6.2 Consider a system whose dynamical state is defined by the canonical variables q, p and whose dynamical behaviour is defined by the Hamiltonian $H = H(q, p, t)$. Let F be an arbitrary quantity depending on the dynamical state of the system, i.e., on q, p, and possibly on t. The rate of change in time of F is given by

$$\dot{F} = [F, H] + \frac{\partial F(q, p, t)}{\partial t} \ ,$$

where $[F, H]$ is the PB of F and H.

Theorem 6.3 (Poisson formulation of the eqs. of motion). Consider a system described in terms of the canonical variables q, p, and whose Hamiltonian is $H = H(q, p, t)$. The motion of the system is governed in this case by the equations

$$\begin{aligned} \dot{q}_i &= [q_i, H] \\ \dot{p}_i &= [p_I, H] \ . \end{aligned}$$

3. Constants of motion in Poisson's formulation

We shall use again a theorem-like sketch of the basic results on the constants of motion in Poisson's formulation. These results are the following.

Theorem 6.4 If one dynamical quantity F is not an explicit function of time and if the PB of F and H is zero, $[F, H] = 0$, then F is a constant/integral of motion as one can see from the theorem 6.2.

Corollary 6.4. If the Hamiltonian is not an explicit function of time, then it is a constant of motion.

[12] The proofs have been omitted as being well known. See, for example, E. A. Desloge, *Classical Mechanics*, Volume 2 (John Wiley & Sons, 1982).

7. HAMILTON-JACOBI EQUATIONS

Forward: It is known from the previous chapters that in principle it is possible to reduce the complexity of many dynamical problems by choosing an appropriate canonical transformation. In particular, we can try to look for those canonical transformations for which the Kamiltonian K is zero, a situation leading to the Hamilton-Jacobi equations.

CONTENTS:

7.1 Introduction

In order to reach the goals of this chapter we need to make use of the following result allowing us to find the set of canonical variables for which the Kamiltonian takes a particular form.

Theorem 7.1 Consider a system whose dynamical state is defined by p, q and whose behaviour under the action of a given force is governed by the Hamiltonian $H = H(q, p, t)$. Let $K = K(Q, P, t)$ be a *known function* of the canonical variables Q, P, and time. Then, any function $F(q, Q, t)$ that satisfies the partial differential equation

$$K\left[Q, -\frac{\partial F(q,Q,t)}{\partial Q}, t\right] = H\left[q, \frac{\partial F(q,Q,t)}{\partial q}, t\right] + \frac{\partial F(q,Q,t)}{\partial t}$$

and also the condition

$$\left|\frac{\partial^2 F(q,Q,t)}{\partial q_j \partial Q_j}\right| \neq 0$$

is a generating function for a canonical transformation of q, p to Q, P, and the corresponding Kamiltonian is $K = K(Q, P, t)$.

In the following sections we shall use this theorem to find those canonical transformations whose Kamiltonian is zero[13], that leads us to the Hamilton-Jacobi equations.

7.2 Time-dependent HJ equations.

As a consequence of Theorem 7.1 and of requiring a zero Kamiltonian we get the following theorem

Theorem 7.2 Consider a system of f degrees of freedom defined by the set of variables q, p and of Hamiltonian $H = H(q, p, t)$. If we build the partial differential equation

$$H\left[q, \frac{\partial S(q,t)}{\partial q}, t\right] + \frac{\partial S(q,t)}{\partial t} = 0 \tag{1}$$

and if we are able to find a solution of the form

$$S = S(q, \alpha, t)\ ,$$

[13] More exactly, $K\left[Q, -\frac{\partial F(q,Q,t)}{\partial Q}, t\right] = 0$.

where $\alpha = \alpha_1, \alpha_2, ..., \alpha_f$ is a set of constants and if in addition the solution satisfies the condition

$$\left| \frac{\partial^2 S(q, \alpha, t)}{\partial q_i \partial \alpha_i} \right| \neq 0 \ ,$$

then $q(t)$ can be obtained from the equations

$$\frac{\partial S(q, \alpha, t)}{\partial \alpha_i} = \beta_i \ , \tag{2}$$

where $\beta = \beta_1, \beta_2, ..., \beta_f$ is a set of constants. The set of equations (2) provide us with f algebraic equations in the f unknown variables $q_1, q_2, ..., q_f$. The values of the constants α and β are determined by the boundary conditions. Moreover, if $q(t)$ is given it is possible to find $p(t)$ starting from

$$p_i \ = \ \frac{\partial S(q, \alpha, t)}{\partial q_i} \ . \tag{3}$$

The partial differential equation (1) is called *the time-dependent Hamilton-Jacobi equation*. The function $S(q, \alpha, t)$ is known as *Hamilton's principal function*.

To achieve a better meaning of the theorem, as well as of the constants α and β, we proceed with its proof.

Proof of the Theorem 7.2. According to the Theorem 7.1, any function $F(q, Q, t)$ satisfying the partial differential equation

$$H \left[q, \frac{\partial F(q, Q, t)}{\partial q}, t \right] + \frac{\partial F(q, Q, t)}{\partial t} = 0$$

and also the condition

$$\left| \frac{\partial^2 F(q, Q, t)}{\partial q_j \partial Q_j} \right| \neq 0$$

should be a generating function of a set of canonical variables Q, P for which the Kamiltonian K is zero, i.e., $K(Q, P, t) = 0$. The function

$$F(q, Q, t) = [S(q, \alpha, t)]_{\alpha = Q} \equiv S(q, Q, t)$$

belongs to this class. Then $S(q, Q, t)$ is the generating function for a canonical transformation leading to the new set of canonical variables Q, P, for

which the Kamiltonian K is identically zero. The transformation equations associated to $S(q, Q, t)$ are

$$p_i = \frac{\partial S(q, Q, t)}{\partial q_i} \tag{4}$$

$$P_i = -\frac{\partial S(q, Q, t)}{\partial Q_i} \tag{5}$$

and because $K(Q, P, t) \equiv 0$, the equations of motion are

$$\begin{aligned} \dot{Q}_i &= \frac{\partial K(Q,P,t)}{\partial P_i} = 0 \\ \dot{P}_i &= -\frac{\partial K(Q,P,t)}{\partial Q_i} = 0 \end{aligned} \quad .$$

From these equations we infer that

$$Q_i = \alpha_i \tag{6}$$

$$P_i = -\beta \ , \tag{7}$$

where α_i and β_i are constants. The choice of the negative sign for β in (7) is only a convention. If now we substitute the equations (6) and (7) in (5) we get

$$-\beta_i = -\left[\frac{\partial S(q, Q, t)}{\partial Q_i}\right]_{Q=\alpha} = -\frac{\partial S(q, \alpha, t)}{\partial \alpha_i} \ ,$$

which reduces to (2). If, in addition, we substitute (6) in (4) we get (2). This complets the proof.

7.3 Time-independent HJ equations

If the Hamiltonian does not depend explicitly on time, we can partially solve the time-dependent Hamilton-Jacobi equation. This result can be spelled out as the following theorem

Theorem 7.3 Consider a system of f degrees of freedom defined in terms of q, p, and whose behaviour under a given force is governed by the time-independent Hamiltonian $H(q, p)$.
If we build the partial differential equation

$$H\left[q, \frac{\partial W(q)}{\partial q}\right] = E \ , \tag{8}$$

where E is a constant whose value for a particular set of conditions is equal to the value of the integral of motion $H(q,p)$ for the given boundary conditions, and if we can find a solution to this equation of the form

$$W = W(q,\alpha) \ ,$$

where $\alpha \equiv \alpha_1, \alpha_2, ..., \alpha_f$ is a set of constants that explicitly or implicitly include the constant E, i.e., $E = E(\alpha)$, and if the solution satisfies the condition

$$\left| \frac{\partial^2 W(q,\alpha)}{\partial q_i \partial \alpha_j} \right| \neq 0 \ ,$$

then the equations of motion are given by

$$\frac{\partial S(q,\alpha,t)}{\partial \alpha_i} = \beta_i \tag{9}$$

where

$$S(q,\alpha,t) \equiv W(q,\alpha) - E(\alpha)t$$

and $\beta = \beta_1, \beta_2, ..., \beta_f$ is a set of constants. The set of equations (9) provide f algebraic equations in the f unknown variables $q_1, q_2, ..., q_f$. The values of the constants α and β are determined by the boundary conditions. The partial differential equation (8) is *the time-independent Hamilton-Jacobi equation*, and the function $W(q,\alpha)$ is known as *the characteristic Hamilton function.*

7.4 Generalization of the HJ equations

The Hamilton-Jacobi equation can be generalized according to the following theorem allowing sometimes the simplification of some Hamilton-Jacobi problems.

Theorem 7.4 Consider a system of f degrees of freedom whose dynamics is defined by x, y, where $x_i, y_i = q_i, p_i$ or $p_i, -q_i$, and whose behaviour under the action of a given force is governed by the Hamiltonian $H(x,y,t)$. If we write the partial differential equation

$$H\left[x, \frac{\partial S(x,t)}{\partial x}, t\right] + \frac{\partial S(x,t)}{\partial t} = 0$$

and if we can find a solution of this equation of the form

$$S = S(x,\alpha,t)$$

where $\alpha \equiv \alpha_1, \alpha_2, ..., \alpha_f$ is a set of constants and in addition the solution satisfies the condition

$$\left| \frac{\partial^2 S(x, \alpha, t)}{\partial x_j \partial \alpha_j} \right| \neq 0 \ ,$$

then the laws of motion of the system can be obtained from the equations

$$\frac{\partial S(x, \alpha, t)}{\partial x_i} = y_i \tag{10}$$

$$\frac{\partial S(x, \alpha, t)}{\partial \alpha_i} = \beta_i \ , \tag{11}$$

where $\beta \equiv \beta_1, \beta_2, ..., \beta_f$ is a set of constants.

7.5 Example of application of the HJ equations

We shall solve the problem of the one-dimensional harmonic oscillator of mass m, using the Hamilton-Jacobi method.
We know that the Hamiltonian of the system is

$$H = \frac{p^2}{2m} + \frac{kx^2}{2} \tag{12}$$

According to Theorem 7.2 the Hamilton-Jacobi equation for the system is

$$\frac{1}{2m} \left(\frac{\partial F}{\partial q} \right) + \frac{kq^2}{2} + \frac{\partial F}{\partial t} = 0 \tag{13}$$

We assume a solution of (13) of the form $F = F_1(q) + F_2(t)$. Therefore, (13) converts to

$$\frac{1}{2m} \left(\frac{dF_1}{dq} \right)^2 + \frac{kq^2}{2} = -\frac{dF_2}{dt} \tag{14}$$

Making each side of the previous equation equal to α, we find

$$\frac{1}{2m} \left(\frac{dF_1}{dq} \right)^2 + \frac{kq^2}{2} = \alpha \tag{15}$$

$$\frac{dF_2}{dt} = -\alpha \tag{16}$$

For zero constants of integration, the solutions are

$$F_1 = \int \sqrt{2m(\alpha - \frac{kq^2}{2})} dq \tag{17}$$

$$F_2 = -\alpha t \tag{18}$$

Thus, the generating function F is

$$F = \int \sqrt{2m(\alpha - \frac{kq^2}{2})}dq - \alpha t \ . \tag{19}$$

According to (2), $q(t)$ is obtained starting from

$$\beta = \frac{\partial}{\partial \alpha}\left\{\int \sqrt{2m(\alpha - \frac{kq^2}{2})}dq - \alpha t\right\} \tag{20}$$

$$= \frac{\sqrt{2m}}{2}\int \frac{dq}{\sqrt{\alpha - \frac{kq^2}{2}}} - t \tag{21}$$

and effecting the integral we get

$$\sqrt{\frac{m}{k}}\sin^{-1}(q\sqrt{k/2\alpha}) = t + \beta \ , \tag{22}$$

from which q is finally obtained in the form

$$q = \sqrt{\frac{2\alpha}{k}}\sin\sqrt{k/m}(t+\beta) \ . \tag{23}$$

In addition, we can give a physical interpretation to the constant α according to the following argument.
The factor $\sqrt{\frac{2\alpha}{k}}$ should correspond to the amplitude A of the oscillator. On the other hand, the total energy E of a one-dimensional harmonic oscillator of amplitude A is given by

$$E = \frac{1}{2}kA^2 = \frac{1}{2}k\left(\sqrt{\frac{2\alpha}{k}}\right)^2 = \alpha \ .$$

In other words, α is physically the total energy E of the one-dimensional harmonic oscillator.

FURTHER READING

C.C. Yan, *Simplified derivation of the HJ eq.*, Am. J. Phys. 52, 555 (1984)

N. Anderson & A.M. Arthurs, *Note on a HJ approach to the rocket pb.*, Eur. J. Phys. 18, 404 (1997)

M.A. Peterson, *Analogy between thermodynamics and mechanics*, Am. J. Phys. 47, 488 (1979)

Y. Hosotani & R. Nakayama, *The HJ eqs for strings and p-branes*, hep-th/9903193 (1999)

8. ACTION-ANGLE VARIABLES

Forward: The Hamilton-Jacobi equation provides a link to going from a set of canonical variables q, p to a second set Q, P, where both are constants of motion.
In this chapter, we briefly present another important procedure by which one goes from an initial pair of canonical variables to a final one, where not both variables are simultaneously constants of motion.

CONTENS:

8.1 Separable systems

Separable systems are those ones for which the Hamiltonian is not an explicit function of time, i.e.,

$$H = H(q, p) \ ,$$

allowing, in addition, to find a solution of the time-independent Hamilton-Jacobi of the form

$$W(q, \alpha) = \sum_i W_i(q_i, \alpha) \ .$$

8.2 Cyclic systems

We know that the dynamical state of a system is characterized by a set of generalized coordinates $q \equiv q_1, q_2, ..., q_f$ and momenta $p \equiv p_1, p_2, ..., p_f$. A system in motion will describe an orbit in the phase space q, p At the same time, there is an orbit in each of the subspaces q_i, p_i. In every plane q_i, p_i, the orbit can be represented by an equation of the form $p_i = p_i(q_i)$ or a pair of equations $p_i = p_i(t)$, $q_i = q_i(t)$. If for each value of i, the orbit $p_i = p_i(q_i)$ is a closed curve in the plane $q_i - p_i$, then we say that the system is cyclic. In the figure 8 we show the two possibilities for a system to be cyclic. In 8.1a, the system is cyclic because q_i oscillates between the limits defined by $q_i = a$ and $q_i = b$, whereas in figure 8.1b, the system is cyclic because q_i moves from $q_i = a$ to $q_i = b$, and repeats the same motion afterwards.

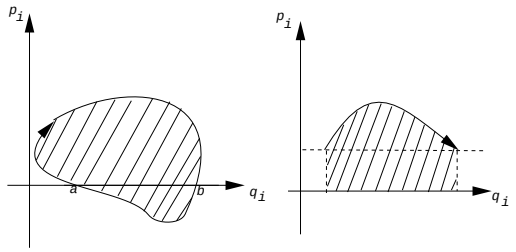


Fig.8.1

At this point, it is worthwhile to make two helpful remarks.

Remark 1: The cyclic term has been introduced only for simplifying the notation in the next sections. One should not interpret this term as if

the system is cyclic in each subspace q_i, pi. The system should come back to its initial state only in the global space q, p.

Remark 2: If the cyclic system has only one degree of freedom, the time required by the system to accomplish the cycle in $q-p$ is constant; therefore the motion in the space $q-p$ will be periodic in time. If the system has more degrees of freedom, then, in general, the time required for a particular cycle in one of the subspaces q_i, p_i will not be a constant, but will depend on the motion of the other coordinates. As a result, the motion in the subspace q_i, p_i will not be periodic in time. One should be careful with this point, since *not all the motions in the subspaces q_i, p_i are periodic.*

8.3 Action-angle variables

We consider now a cyclic system of f degrees of freedom, whose dynamical state is characterized by the canonical set q, p. Let $H(q,p)$ be the Hamiltonian of the system and let

$$W(q,\alpha) \equiv \sum_i W_i(q_i,\alpha) \ ,$$

(where $\alpha = \alpha_1, \alpha_2, ..., \alpha_f$ are constants) be a solution of the time-independent Hamilton-Jacobi equation

$$H(q, \frac{\partial W}{\partial q}) = E \ .$$

Let $J \equiv J_1, J_2, ..., J_f$ be the set of constants defined by the equations

$$J_i(\alpha) = \oint \frac{\partial W_i(q_i,\alpha)}{\partial q_i} dq_i \ , \tag{1}$$

where the integral is along a complete cycle in the variable q_i. If we use the function

$$\begin{aligned} W(q,\alpha) &\equiv W[q,\alpha(J)] \\ &\equiv \sum_i W_i[q_i,\alpha(J)] \\ &\equiv \sum_i W_i(q_i,J) \end{aligned}$$

as the generating function of a canonical transformation of q, p to a new set of coordinates w and momenta J, i.e., if we define the variables w and J by

the transformation equations

$$p_i = \frac{\partial W(q,\alpha)}{\partial q_i} = \frac{\partial W_i(q_i,J)}{\partial q_i} \tag{2}$$

$$w_i = \frac{\partial W(q,J)}{\partial J_i} \ , \tag{3}$$

then the new coordinates $w_1, w_2, ..., w_f$ are called *angle variables*, and the new momenta $J_1, J_2, ..., J_f$ are called *action variables.*
From (2) we get

$$p_i(q_i,\alpha) = \frac{\partial W_i[q_i, J(\alpha)]}{\partial q_i} = \frac{\partial W_i(q_i,\alpha)}{\partial q_i} \ . \tag{4}$$

Substituting (4) in (1) one gets

$$J_i(\alpha) = \oint p_i(q_i,\alpha) dq_i \ . \tag{5}$$

The equation $p_i = p_i(q_i,\alpha)$ gives the projected orbit $p = p(q)$ on the subspace p_i, q_i. The integral in the right hand side of the equation (5) is thus the area bordered by the closed orbit, or beneath the orbit, as shown in figure 8.1. Thus, the function $J_i(\alpha)$ has a geometric interpretation as the area covered in the subspace q_i, p_i during a complete cycle in the subspace. This area depends on the constants α or equivalently on the initial conditions and can be arbitrary [14].

8.4 Motion in terms of action-angle variables

We give the following theorem-like statement for the motion of a system in terms of action-angle variables.
Theorem 8.4
Consider a separable cyclic system of f degrees of freedom whose motion is described by the variables $q, p \equiv q_1, q_2, ..., q_f, p_1, p_2, ..., p_f$, together with the Hamiltonian $H(q,p)$. If we transform the motion to the action-angle variables J, w, then the Hamiltonian H is a function of J alone, i.e.,

$$H = H(J)$$

[14] Historically, the first intents to pass from the classical mechanics to quantum mechanics was related to the assumption that the value of J_i could be only a multiple of $h/2\pi$, where h is Planck's constant.

and the equations of motion will be

$$\begin{aligned} J_i &= \gamma_i \\ w_i &= \nu_i t + \phi_i \ , \end{aligned}$$

where γ_i y ϕ_i are constants determined by the initial conditions, whereas the ν_i are constants known as *the frequencies of the system* being defined as follows

$$\nu_i = \left[\frac{\partial H(J)}{\partial J_i}\right]_{J=\gamma_i} \ .$$

8.5 Importance of the action-angle variables

The importance of the action-angle variables resides in providing a powerful technique for directly getting the frequencies of periodic motions without solving for the equations of motion of the system.
This important conclusion can be derived through the following argument. Consider the change of w when q describes a complete cycle

$$\Delta w = \oint \frac{\partial w}{\partial q} dq$$

On the other hand, we know that

$$w = \frac{\partial W}{\partial J} \ ,$$

and therefore

$$\begin{aligned} \Delta w &= \oint \frac{\partial^2 W}{\partial q \partial J} dq \\ &= \frac{d}{dJ} \oint \frac{\partial W}{\partial q} dq \\ &= \frac{d}{dJ} \oint p dq \\ &= 1 \ . \end{aligned}$$

This result shows that w changes by a unity when q varies within a complete period.
From the relationship

$$w = \nu t + \phi \ ,$$

we infer that in a period τ

$$\begin{aligned} \Delta w &= 1 \\ &= \nu\tau \ . \end{aligned}$$

This means that we can identify the constant ν with the inverse of the period

$$\nu = \frac{1}{\tau} \ .$$

8.6 Example: The simple harmonic oscillator

Using the action-angle formalism prove that the frequency ν of the simple one-dimensional harmonic oscillator is given by $\nu = \sqrt{k/m}/2\pi$.
Since H is a constant of motion, the orbit in the space $q - p$ is given

$$\frac{p^2}{2m} + \frac{kq^2}{2} = E \ ,$$

where E is the energy. This is the equation of an ellipse of semiaxes $\sqrt{2mE}$ and $\sqrt{2E/k}$. The area enclosed by the ellipse is equal to the action J. Therefore,

$$J = \pi\sqrt{2mE}\sqrt{\frac{2E}{k}} = 2\pi\sqrt{\frac{m}{k}}E \ .$$

It follows that

$$H(J) = E = \frac{\sqrt{k/m}}{2\pi}J \ .$$

Thus, the frequency will be

$$\nu = \frac{\partial H(J)}{\partial J} = \frac{\sqrt{k/m}}{2\pi} \ .$$

9. CANONICAL PERTURBATION THEORY

Forward: The great majority of problems that we have to solve in Physics are not exactly solvable. Because of this and taking into account that we live in the epoch of computers the last decades have seen a lot of progress in developing techniques leading to approximate solutions. The perturbation method is used for not exactly solvable Hamiltonian problems when the Hamiltonian differs slightly from an exactly solvable one. The difference between the two Hamiltonians is known as the perturbation Hamiltonian. All perturbation methods are based on the smallness of the latter with respect to both Hamiltonians.

CONTENTS:

9.1 Time-dependent perturbation theory

The most appropriate formulation of classical mechanics for the development of perturbation methods is the Hamilton-Jacobi approach. Thus, we take $H_0(p, q, t)$ as the Hamiltonian corresponding to the solvable (unperturbed) problem and consider the Hamilton principal function $S(q, \alpha_0, t)$, as the generating function of a canonical transformation of (p, q) to the new canonical pair (α_0, β_0) for which the new Hamiltonian (or Kamiltonian) K_0 of the unperturbed system is zero

$$\frac{\partial S}{\partial t} + H_0(\frac{\partial S}{\partial q}, q, t) = K_0 = 0 \ . \tag{1}$$

This is the Hamilton-Jacobi equation, where we used $p = \partial S/\partial q$. The whole set of new canonical coordinates (α_0, β_0) are constant in the unperturbed case because $K_0 = 0$ and:

$$\begin{aligned} \dot{\alpha}_0 &= -\frac{\partial K_0}{\partial \beta_0}, \\ \dot{\beta}_0 &= \frac{\partial K_0}{\partial \alpha_0}. \end{aligned} \tag{2}$$

We now consider the Hamiltonian of the perturbed system written as follows:

$$H(q, p, t) = H_0(q, p, t) + \Delta H(q, p, t); \qquad (\Delta H \ll H_0). \tag{3}$$

Although (α_0, β_0) are still canonical coordinates (since the transformation generated by S is independent of the particular form of the Hamiltonian), they will not be constant and the Kamiltonian K of the perturbed system will not be zero. In order not to forget that in the perturbed system the transformed coordinates are not constant, we denote them by α and β instead of α_0 and β_0, which are the corresponding constants in the unperturbed system. Thus, for the perturbed system we have:

$$K(\alpha, \beta, t) = H + \frac{\partial S}{\partial t} = (H_0 + \frac{\partial S}{\partial t}) + \Delta H = \Delta H(\alpha, \beta, t). \tag{4}$$

The equations of motion for the transformed variables in the perturbed system will be:

$$\begin{aligned} \dot{\alpha}_i &= -\frac{\partial \Delta H(\alpha,\beta,t)}{\partial \beta_i} \\ \dot{\beta}_i &= \frac{\partial \Delta H(\alpha,\beta,t)}{\partial \alpha_i}, \end{aligned} \tag{5}$$

where $i = 1, 2, ..., n$ and n is the number of degrees of freedom of the system. These are rigorous equations. If the system of $2n$ equations could be solved for α_i and β_i as functions of time, the transformation equation $(p, q) \rightarrow (\alpha, \beta)$ would give p_i and q_i as functions of time and the problem would be solved. However, the exact solution of the equations (5) is not easier to get w.r.t. the original equations. From (5) we see that when α and β are not constant their variation in time is slow if we assume that ΔH changes infinitesimally w.r.t. α y β. A first approximation for the temporal variations of (α, β) can be obtained by substituting α y β in the second terms of (5) by their constant unperturbed values

$$\begin{aligned} \dot{\alpha}_{i1} &= -\left.\frac{\partial \Delta H(\alpha,\beta,t)}{\partial \beta_i}\right|_0 \\ \dot{\beta}_{i1} &= \left.\frac{\partial \Delta H(\alpha,\beta,t)}{\partial \alpha_i}\right|_0, \end{aligned} \tag{6}$$

where α_{i1} and β_{i1} are the first-order solutions, i.e., in the first power of the perturbation to α_i and β_i, and the vertical bars with zero subindices show that after the derivation one should substitute α and β by their constant unperturbed values. Once this is done, the equations (6) can be integrated leading to α_i and β_i as functions of time (in the first order). Next, using the equations of transformation one can get p and q as functions of time in the same first-order approximation. The second-order approximation can be obtained now by substituting in the second terms of (6) the first-order approximation of α and β w.r.t. time. In general, the perturbation solution of order N is obtained by integrating the following equations

$$\begin{aligned} \dot{\alpha}_{iN} &= -\left.\frac{\partial \Delta H(\alpha,\beta,t)}{\partial \beta_i}\right|_{N-1} \\ \dot{\beta}_{iN} &= \left.\frac{\partial \Delta H(\alpha,\beta,t)}{\partial \alpha_i}\right|_{N-1}. \end{aligned} \tag{7}$$

Example 1

Let us consider the simple case of a free particle that next is subject to a harmonic perturbation. Although this example is trivial can be used to illustrate the aforementioned procedure. The unperturbed Hamiltonian is

$$H_0 = \frac{p^2}{2m}. \tag{8}$$

Since $H_0 \neq H_0(x)$, x is a cyclic variable and $p = \alpha_0$ is a constant of motion in the unperturbed system. Recalling that $p = \partial S/\partial x$, we substitute in (1):

$$\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + \frac{\partial S}{\partial t} = 0. \tag{9}$$

Since the system is conservative, it is convenient to consider the principal function of the form

$$S = \mathbf{S}(x) + F(t). \tag{10}$$

This type of separation of variables is quite useful when the Hamiltonian does not depend explicitly on time. Then, one writes $F(t) = -Et$, where E is the total energy of the system [15]. Putting (10) in (9) we obtain

$$\frac{1}{2m}\left(\frac{d\mathbf{S}}{dx}\right)^2 = E, \qquad \mathbf{S} = \sqrt{2mE}x = \alpha_0 x. \tag{11}$$

Substituting (11) in (10), together with the fact that in this case the Hamiltonian is equal to the energy, we can write the principal function of Hamilton as follows

$$S = \alpha_0 x - \frac{\alpha_0^2 t}{2m}. \tag{12}$$

If the transformed momentum is α_0, the transformed coordinate (which is also constant in the unperturbed system) is

$$\beta_0 = \frac{\partial S}{\partial \alpha_0} = x - \frac{\alpha_0 t}{m}\ .$$

Therefore, the transformation generated by S is given by the following eqs

[15] See, M.R. Spiegel, *Theoretical Mechanics*, pp. 315, 316.

$$
\begin{aligned}
p &= \alpha_0, \\
x &= \frac{\alpha_0 t}{m} + \beta_0.
\end{aligned} \tag{13}
$$

They represent the solution for the motion of free particle. What we have given up to now is only the procedure to obtain the equations of motion using the Hamilton-Jacobi formulation. Now, we introduce the perturbation

$$
\Delta H = \frac{kx^2}{2} = \frac{m\omega^2 x^2}{2}, \tag{14}
$$

or, in terms of the transformed coordinates, using (13)

$$
\Delta H = \frac{m\omega^2}{2}\left(\frac{\alpha t}{m} + \beta\right)^2. \tag{15}
$$

Notice that we have renounced at the subindices 0 for the transformed coordinates since we already study the perturbed system.

Substituting (15) in (5) we get

$$
\begin{aligned}
\dot{\alpha} &= -m\omega^2\left(\frac{\alpha t}{m} + \beta\right), \\
\dot{\beta} &= \omega^2 t\left(\frac{\alpha t}{m} + \beta\right).
\end{aligned} \tag{16}
$$

As one may expect, these equations have an exact solution of the harmonic type. To be sure of that we perform the time derivative of the first equation allowing us to conclude that α has a simple harmonic variation. The same holds for x as a consequence of the transformations (13), which are invariant in form in the perturbed system (up to the subindices of the transformed coordinates). However, we are interested to illustrate the perturbation method, so that we consider that k (the elastic constant) is a small parameter. We seek approximate solutions in different perturbative orders, without missing the fact that the transformed variables (α, β) in the perturbed system are not constants of motion. In other words, even though (α, β) contain information on the unperturbed system, the effect of the perturbation is to make these parameters varying in time.

The first-order perturbation is obtained in general as given by (6). Thus, we

have to substitute α and β by their unperturbed values in the second terms of (16). To simplify, we take $x(t=0)=0$ and therefore $\beta_0 = 0$, leading to

$$\begin{aligned} \dot{\alpha}_1 &= -\omega^2 \alpha_0 t, \\ \dot{\beta}_1 &= \alpha_0 \frac{\omega^2 t^2}{m}, \end{aligned} \tag{17}$$

which integrated reads

$$\begin{aligned} \alpha_1 &= \alpha_0 - \frac{\omega^2 \alpha_0 t^2}{2}, \\ \beta_1 &= \frac{\alpha_0 \omega^2 t^3}{3m}. \end{aligned} \tag{18}$$

The first-order solutions for x and p are obtained by putting α_1 and β_1 in the transformation eqs. (13), where from

$$\begin{aligned} x &= \frac{\alpha_0}{m\omega}\left(\omega t - \frac{\omega^3 t^3}{6}\right), \\ p &= \alpha_0 \left(1 - \frac{\omega^2 t^2}{2}\right). \end{aligned} \tag{19}$$

To obtain the approximate solution in the second perturbative order we have to find $\dot{\alpha_2}$ and $\dot{\beta}_2$, as was indicated in (7), by substituting in the second terms of (16), α and β by α_1 and β_1 as given in (18). Integrating $\dot{\alpha_2}$ and $\dot{\beta}_2$ and using again the transformation eqs. (13), we get the second-order solutions for x and p:

$$\begin{aligned} x &= \frac{\alpha_0}{m\omega}\left(\omega t - \frac{\omega^3 t^3}{3!} + \frac{\omega^5 t^5}{5!}\right), \\ p &= \alpha_0 \left(1 - \frac{\omega^2 t^2}{2!} + \frac{\omega^4 t^4}{4!}\right). \end{aligned} \tag{20}$$

In the limit in which the perturbation order N tends to infinity, we obtain the expected solutions compatible with the initial conditions

$$x \to \frac{\alpha_0}{m\omega}\sin\omega t, \qquad p \to \alpha_0 \cos\omega t. \tag{21}$$

The transformed variables (α, β) contain information about the unperturbed

orbit parameters. For example, if we consider as unperturbed system that corresponding to the Kepler problem, a convenient coordinate pair (α, β) could be the (J, δ) variables, which are the action and the phase angle of the angle w, respectively (remember that $w = \nu t + \delta$, where ν is the frequency). These variables are related to the set of orbital parameters such as semimajor axis, eccentricity, inclination, and so on.
The effect of the perturbation is to produce a time variation of all these parameters. If the perturbation is small, the variation of the parameters during a period of the unperturbed motion will also be small. In such a case, for small time intervals, the system moves along the so-called *osculating* orbit, having the same functional form as the orbit of the unperturbed system; the difference is that the parameters of the osculating curve vary in time.

The osculating parameters can vary in two ways

- Periodic variation: if the parameter comes back to its initial value after a time interval that in first approximation is usually the unperturbed period. These periodic effects of the perturbation do not alter the mean values of the parameters. Therefore, the trajectory is quite similar to the unperturbed one. These effects can be eliminated by taking the average of the perturbations in a period of the unperturbed motion.

- Secular variation: At the end of every succesive orbital period there is a net increment of the value of the parameter. Therefore, at the end of many periods, the orbital parameters can be very different of their unperturbed values. The instantaneous value of the variation of a parameter, for example the frequency, is seldom of interest, because its variation is very small in almost all cases in which the perturbation formalism works. (This variation is so small that it is practically impossible to detect it in a single orbital period. This is why the secular variation is measured after at least several periods.)

Example 2

From the theory of the Kepler two-body problem it is known that if we add a potential $1/r^2$, the orbit of the motion of negative energy is a rotating ellipse whose periapsis is precessing. In this example, we find the precession velocity for a more general perturbation

$$V = -\frac{k}{r} - \frac{h}{r^n} \,, \tag{22}$$

where $n \geq 2$ is an entire number, and h is such that the second term of the potential is a small perturbation of the first one. The perturbative Hamiltonian is

$$\Delta H = -\frac{h}{r^n} \, . \tag{23}$$

In the unperturbed problem, the angular position of the periapsis in the orbit plane is given by the constant $\omega = 2\pi w_2$. In the perturbed case, we have

$$\dot{\omega} = 2\pi \frac{\partial \Delta H}{\partial J_2} = \frac{\partial \Delta H}{\partial l} \, , \tag{24}$$

where we have used $J_2 = 2\pi l$. Moreover, J_2 and w_2 are two of the five integrals of motion that can be obtained when one uses the action-angle variables to solve the Kepler problem.

We need to know the mean of $\dot{\omega}$ in a period τ of the unperturbed orbit

$$\langle \dot{\omega} \rangle \equiv \frac{1}{\tau} \int_0^\tau \frac{\partial \Delta H}{\partial l} dt = \frac{\partial}{\partial l} \left(\frac{1}{\tau} \int_0^\tau \Delta H \, dt \right) = \frac{\partial \langle \Delta H \rangle}{\partial l} \, . \tag{25}$$

The temporal mean of the unperturbed Hamiltonian is

$$\langle \Delta H \rangle = -h \langle \frac{1}{r^n} \rangle = -\frac{h}{\tau} \int_0^\tau \frac{dt}{r^n} \, . \tag{26}$$

On the other hand, since $l = mr^2(d\theta/dt)$, we can get dt and plunge it in (27). This leads to

$$\begin{aligned} \langle \Delta H \rangle &= -\frac{mh}{l\tau} \int_0^{2\pi} \frac{d\theta}{r^{n-2}} \\ &= -\frac{mh}{l\tau} \left(\frac{mk}{l^2} \right)^{n-2} \int_0^{2\pi} [1 + e \cos(\theta - \eta)]^{n-2} d\theta \, . \end{aligned} \tag{27}$$

where η is a constant phase, e is the eccentricity, and where we expressed r as a function of θ using the general equation of the orbit with the origin in one of the focal points of the corresponding conic

$$\frac{1}{r} = \frac{mk}{l^2}\Big[1 + e\cos(\theta - \eta)\Big] \tag{28}$$

.

For $n = 2$:

$$\begin{aligned} \langle \Delta H \rangle &= -\frac{2\pi m h}{l\tau} \,, \\ \langle \dot{\omega} \rangle &= \frac{2\pi m h}{l^2 \tau} \,. \end{aligned} \tag{29}$$

For $n = 3$:

$$\begin{aligned} \langle \Delta H \rangle &= -\frac{2\pi m^2 h k}{l^3 \tau} \,, \\ \langle \dot{\omega} \rangle &= \frac{6\pi m^2 h k}{l^4 \tau} \,. \end{aligned} \tag{30}$$

The latter case, $n = 3$, is of particular importance because the theory of General Relativity predicts a correction of the Newtonian motion precisely of r^{-3} order. This prediction is related to the famous problem of the precession of the orbit of Mercury. Substituting the appropriate values of the period, mass, great semiaxis, that is included in h, and so on, (30) predicts a mean precession velocity of

$$\langle \dot{\omega} \rangle = 42.98 \ \text{ arcsec./century} \,.$$

The mean value is by far larger than the aforementioned one (by a factor larger than one hundred). But before making any comparison one should eliminate from the mean value the contributions due to the following factors: *a)* the effect known as the precession of the equinoxes (the motion of the reference point of longitudes w.r.t. the Milky Galaxy), *b)* the perturbations of the Mercury orbit due to the interaction with the other planets. Once these are eliminated, of which the first is the most significant, one may hopefully obtain the contribution of the relativistic effect. In 1973, this contribution has been estimated as 41.4 ± 0.9 arcsec./century. This is consistent with the prediction given by (30).

9.2 Time-independent perturbation theory

While in the time-dependent perturbation theory one seeks the time dependence of the parameters of the unperturbed system initially considered as

constant, the aim of the time-independent approach is to find the constant quantities of the perturbed system. This theory can be applied only to conservative periodic systems (both in the perturbed and unperturbed state). For example, it can be applied to planetary motion when one introduces any type of conservative perturbation to the Kepler problem, in which case it is known as the von Zeipel or Poincaré method.

Here, we consider here the case of a periodic system of one degree of freedom and time-independent Hamiltonian of the form

$$H = H(p, q, \lambda), \tag{31}$$

where λ is a small constant specifying the strength of the perturbation. We assume that

$$H_0(p, q) = H(p, q, 0) \tag{32}$$

corresponds to a system that has an exact (closed-form) unperturbed solution in the action-angle variables (J_0, w_0), i.e.,

$$\begin{aligned} H_0(p, q) &= K_0(J_0) \\ \nu_0 &= \dot{w}_0 = \frac{\partial K_0}{\partial J_0} \; ; \qquad (w_0 = \nu_0 t + \delta_0). \end{aligned} \tag{33}$$

Since the canonical transformation from (p, q) to (J_0, w_0) is independent of the particular form of the Hamiltonian, the perturbed Hamiltonian $H(p, q, \lambda)$ can be written as $H(J_0, w_0, \lambda)$. Due to the fact that the perturbed Hamiltonian depends on w_0, J_0, it is not constant any more. On the other hand, in principle, one can get new action-angle variables (J, w) that may be more appropriate for the perturbed system, such as

$$\begin{aligned} H(p, q, \lambda) &= E(J, \lambda) \\ \nu &= \dot{w} = \frac{\partial E}{\partial J} \\ \dot{J} &= -\frac{\partial E}{\partial w} = 0 \; ; \quad (J = constant). \end{aligned} \tag{34}$$

Since the transformation connecting (p, q) to (J_0, w_0) is known, we have to find the canonical transformation S connecting (J_0, w_0) to (J, w). If we assume that λ is small the transformation we look for should not differ much from the identity transformation. Thus, we write the following expansion

$$S = S(w_0, J, \lambda) = S_0(w_0, J) + \lambda S_1(w_0, J) + \lambda^2 S_2(w_0, J) + ... \ . \qquad (35)$$

For $\lambda = 0$ we ask S to provide an identity transformation that leads to

$$S_0 = w_0 J \ . \qquad (36)$$

The canonical transformations generated by S read

$$\begin{aligned} w &= \frac{\partial S}{\partial J} = w_0 + \lambda \frac{\partial S_1}{\partial J}(w_0, J) + \lambda^2 \frac{\partial S_2}{\partial J}(w_0, J) + ... \ . \\ J_0 &= \frac{\partial S}{\partial w_0} = J + \lambda \frac{\partial S_1}{\partial w_0}(w_0, J) + \lambda^2 \frac{\partial S_2}{\partial w_0}(w_0, J) + ... \ . \end{aligned} \qquad (37)$$

Due to the fact that w_0 is an angle variable of the unperturbed system we know that $\Delta w_0 = 1$ over a cycle. On the other hand, we know that the canonical transformations have the property to conserve the phase space volume. Therefore, we can write:

$$J = \oint p dq = \oint J_0 dw_0 \ . \qquad (38)$$

Integrating the second equation of (37) along an orbit of the perturbed system, we get

$$\oint J_0 dw_0 = \oint J dw_0 + \sum_{n=1} \lambda^n \oint \frac{\partial S_n}{\partial w_0} dw_0, \qquad (39)$$

and substituting (39) in (38) leads to

$$J = J\Delta w_0 + \sum_{n=1} \lambda^n \oint \frac{\partial S_n}{\partial w_0} dw_0. \qquad (40)$$

Since $\Delta w_0 = 1$, one gets

$$\sum_{n=1} \lambda^n \oint \frac{\partial S_n}{\partial w_0} dw_0 = 0, \qquad (41)$$

or

$$\oint \frac{\partial S_n}{\partial w_0} dw_0 = 0. \tag{42}$$

Moreover, the Hamiltonian can be expanded in λ as a function of w_0 and J_0:

$$H(w_0, J_0, \lambda) = K_0(J_0) + \lambda K_1(w_0, J_0) + \lambda^2 K_2(w_0, J_0) + \dots , \tag{43}$$

where the K_i are known because H is a known function of w_0 and J_0 for a given λ. On the other hand, one can write

$$\begin{aligned} H(p, q, \lambda) &= H(w_0, J_0, \lambda) \\ &= E(J, \lambda) , \end{aligned} \tag{44}$$

which is the expression for the energy in the new action-angle coordinates (where J will be constant and w will be a linear function of time).

E can also be expanded in powers of λ:

$$E(J, \lambda) = E_0(J) + \lambda E_1(J) + \lambda^2 E_2(J) + \dots . \tag{45}$$

Taking into account (44) we can obtain equalities for the coefficients of the same powers of λ in (43) and (45). However, these expressions for the energy depend on two different sets of variables. To solve this issue we express H_0 in terms of J by writing a Taylor expansion of $H(w_0, J_0, \lambda)$ w.r.t. J_0 in the infinitesimal neighborhood of J:

$$H(w_0, J_0, \lambda) = H(w_0, J, \lambda) + (J_0 - J)\frac{\partial H}{\partial J} + \frac{(J_0 - J)^2}{2}\frac{\partial^2 H}{\partial J^2} + \dots , \tag{46}$$

The derivatives of this Taylor expansion, which in fact are the derivatives w.r.t. J_0 calculated for $J_0 = J$, can also be written as derivatives w.r.t. J, once we substitute J_0 by J in $H_0(J_0)$. In the previous equation, all the terms containing J_0 should be rewritten in terms of J by using the transformation defined by (37) connecting the coordinates (J_0, w_0) and (J, w). Thus, from the second equation in (37) we obtain $(J_0 - J)$, which introduced in (46) gives:

$$H(w_0,J_0,\lambda)=H(w_0,J,\lambda)+\frac{\partial H}{\partial J}\left(\lambda\frac{\partial S_1}{\partial w_0}+\lambda^2\frac{\partial S_2}{\partial w_0}+...\right)+\frac{1}{2}\frac{\partial^2 H}{\partial J^2}\lambda^2\left(\frac{\partial S_1}{\partial w_0}\right)^2+O(\lambda^3)\ . \quad (47)$$

Next, we can make use of (43) to write $H(w_0,J,\lambda)=H(w_0,J_0,\lambda)\mid_{J_0=J}$, that we put in (47) in order to get

$$\begin{aligned} H(w_0,J_0,\lambda) &= K_0(J)+\lambda K_1(w_0,J)+\lambda^2 K_2(w_0,J)+... \\ &+ \lambda\frac{\partial S_1}{\partial w_0}\left(\frac{\partial K_0(J)}{\partial J}+\lambda\frac{\partial K_1(w_0,J)}{\partial J}+...\right) \\ &+ \lambda^2\left[\frac{\partial K_0(J)}{\partial J}\frac{\partial S_2}{\partial w_0}+\frac{1}{2}\frac{\partial^2 K_0}{\partial J^2}\left(\frac{\partial S_1}{\partial w_0}\right)^2+...\right] \\ &\equiv E(J,\lambda) \quad (48) \\ &= E_0(J)+\lambda E_1(J)+\lambda^2 E_2(J)+...\ . \end{aligned}$$

Now, we can solve the problem in terms of the coefficients $E_i(J)$ allowing us to calculate the frequency of the perturbed motion in various perturbation orders. Since the expansion of the terms of E_i does not imply a dependence on w_0, then the occurrence of w_0 in (48) is artificial. The $K_i(w_0,J)$ of (48) are known functions, whereas the $S_i(w_0,J)$ and $E_i(J)$ are the unknown quantities.

Making equal the corresponding powers of λ, we get

$$\begin{aligned} E_0(J) &= K_0(J) \\ E_1(J) &= K_1(w_0,J)+\frac{\partial S_1}{\partial w_0}\frac{\partial K_0(J)}{\partial J} \\ E_2(J) &= K_2(w_0,J)+\frac{\partial S_1}{\partial w_0}\frac{\partial K_1(w_0,J)}{\partial J} \\ &+ \frac{1}{2}\left(\frac{\partial S_1}{\partial w_0}\right)^2\frac{\partial^2 K_0(J)}{\partial J^2}+\frac{\partial S_2}{\partial w_0}\frac{\partial K_0(J)}{\partial J}\ . \quad (49) \end{aligned}$$

We can see that to obtain E_1 we need to know not only K_1 but also S_1. Moreover, we should keep in mind that the E_i are constant, being functions of J only, which is a constant of motion. We also have to notice that $\partial K_0/\partial J$ does not depend on w_0 (since $K_0=K_0(J)=K_0(J_0)\mid_{J_0=J}$). Averaging over w_0 on both sides of the second equation in (49), one gets

$$
\begin{aligned}
E_1 &= \langle E_1 \rangle \\
&= \langle K_1 \rangle + \frac{\partial K_0}{\partial J} \langle \frac{\partial S_1}{\partial w_0} \rangle \; . \qquad (50)
\end{aligned}
$$

But we have already seen that $\langle \partial S_i / \partial w_0 \rangle = \oint (\partial S_i / \partial w_0) dw_0 = 0$. Therefore,

$$
E_1 = \langle E_1 \rangle = \langle K_1 \rangle \; . \qquad (51)
$$

Introducing (51) in the left hand side of the second equation of (49) we get $(\partial S_1 / \partial w_0)$ as follows

$$
\frac{\partial S_1}{\partial w_0} = \frac{\langle K_1 \rangle - K_1}{\nu_0(J)} \; , \qquad (52)
$$

where we used $\nu_0 = \partial K_0 / \partial J$.

The solution for S_1 can now be found by a direct integration. In general, once we assume that we already have E_{n-1}), the procedure to obtain E_n is the following

- Perform the average on both sides of the nth equation of (49).
- Introduce the obtained mean value of E_n, in the complete equation for E_n given by (49) (the one before averaging).
- The only remaining unknown S_n can be obtained by integrating the following relationship

$$
\frac{\partial S_n}{\partial w_0} = \text{ known function of } w_0 \text{ and } J \; .
$$

- Substitute S_n in the complete equation for E_n.

Once all this has been done, the procedure can be repeated for $n + 1$.

As one can see, getting the energy at a particular order n is possible if and only if one has obtained the explicit form of S_{n-1}. On the other hand, S_n can be obtained only when E_n has been already found.

The time-independent perturbation theory is very similar to the Rayleigh-Schroedinger perturbation scheme in wave mechanics, where one can get

E_n only if the wavefunction is known at the $n-1$ order. Moreover, the wavefunction of order n can be found only if E_n has been calculated.

Bibliography

- H. Goldstein, *Classical Mechanics*, Second ed., Spanish version, Editorial Reverté S.A., 1992.

- R.A. Matzner and L.C. Shepley, *Classical Mechanics*, Prentice-Hall Inc., U.S.A., 1991.

- R. Murray Spiegel, *Theoretical Mechanics*, Spanish version, McGraw-Hill S.A. de C.V. México, 1976.

- L.D. Landau and E.M. Lifshitz, *Mechanics*, 3rd. ed. Course of Theoretical Physics, Volume 1. Pergammon Press, Ltd. 1976.

10. ADIABATIC INVARIANTS

Forward: An adiabatic invariant is a function of the parameters and constants of motion of a system, which remains almost constant in the limit in which the parameters change infinitesimally in time, even though in the end they may change by large amounts.

CONTENTS:

10.1 BRIEF HISTORY

The notion of adiabatic invariance goes back to the early years of quantum mechanics. Around 1910 people studying the emission and absorption of radiation noticed that the atoms could live in quasi-stable states in which their energy was alsmost constant. At the 1911 Solvay Congress, the problem of adiabatic invariance became widely known due to Einstein that draw the attention of the physics community to the adiabatic invariant E/ν of the one-dimensional pendulum of slowly varying length. He suggested that similar invariants could occur for atomic systems setting their stability limits. Later, Ehrenfest was able to find such adiabatic invariants and their employment led to the first quantum approach of Bohr and Sommerfeld.

The method of adiabatic invariants resurged after several decades in the area of magnetospheric physics. The Scandinavian scientists were especially interested in boreal aurora phenomena, i.e., the motion of electrons and ions in the terrestrial magnetosphere. One of them, H. Alfven, showed in his book *Cosmic Electrodynamics* that under appropriate conditions a certain combination of dynamical parameters of the charged particles remains constant in the first order . Apparently, Alfven did not realized that he found an adiabatic invariant. This was pointed out by L. Landau and E. Lifshitz who discussed in detail this issue.

10.2 GENERALITIES

We first show in a simple way which is the adiabatic invariant in the case of the one-dimensional harmonic oscillator. The employed method will be to prove that in the limit of the infinitesimally slow variation of the parameters, the adiabatic invariant of the one-dimensional system goes to a quantity that is exactly conserved in a corresponding two-dimensional system.

Consider a particle at the end of a rope of negligible mass that rotates uniformly on a table. Let a be the radius of the circle, m the mass of the particle, and ω the angular frequency. If the origin is in the center of the circle the projection of the motion on the x axis corresponds to the motion of a simple harmonic oscillator. The motion continues to be harmonic even when the rope is slowly made shorter by pulling it through a small hole drilled at the origin. The problem is to find the invariant quantity in this case. It is by far more difficult to get the answer for the circular motion in comparison with the corresponding projected oscillator. Since there are only central forces even when the rope is slowly shortened the conserved quantity is the angular momentum l. Therefore, in the slow limit one can write

$$l = m\omega a^2 \ . \tag{1}$$

From this equation one can see that although l is an *exact* invariant $m\omega a^2$ is only a slow limit, i.e., adiabatic invariant. More exactly, it is only in the slow limit that we can claim that the two quantities are equal and we are allowed to think of the right hand side as an invariant. In the slow limit, the projected one-dimensional oscillator has a slowly varying amplitude, a frequency $\nu = \omega/2\pi$, and a total energy $E = \frac{1}{2}m\omega^2 a^2$. Therefore, the adiabatic invariance of the right hand side of (1) implies the adiabatic invariance of the quotient E/ω for the projected oscillator.

A short analytic proof of this fact is the following. One writes the two-dimensional Hamiltonian

$$H = (2m)^{-1}(p_x^2 + p_y^2) + \frac{1}{2}m\omega^2(x^2 + y^2) \ . \tag{2}$$

Hamilton's equations lead to

$$\ddot{x} + \omega^2 x = o, \qquad \ddot{y} + \omega^2 y = 0 \ . \tag{3}$$

Changing to polar coordinates we note that $l = m\dot{\theta}r^2$. The previous Hamilton equations can be written as a single complex equation $\ddot{z} + \omega^2 z = 0$, where $z = x + iy$. We consider a solution of the form $z = a \exp i(\omega t + b)$, where a, b, and ω are real functions that are slowly varying in time. The real and imaginary parts are $x = a\cos(\omega t + b)$ and $y = a\sin(\omega t + b)$, respectively. Since ω, a, and b vary slowly, l can be written in a very good approximation as $l = m\omega a^2$. The energy of the projected oscillator is $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2$. In the slow varying limit of the parameters, we can substitute x in E by the cosine function to get $E = \frac{1}{2}m\omega^2 a^2$. Thus, if $m\omega a^2$ is an adiabatic invariant then E/ω is an adiabatic invariant for the one-dimensional oscillator.

We now clarify the meaning of the term almost constant in the definition of the adiabatic invariant in the forward to this chapter.
Consider a system of one degree of freedom, which initially is consevative and periodic and depends on an initially constant parameter a. The slow variation of this parameter due, for example, to a low amplitude perturbation does not alter the periodic nature of the motion. By a slow variation we mean the one for which a varies slowly during a period τ of the motion:

$$\tau(da/dt) \ll a \ . \tag{4}$$

However, even when the variations of a are small during a given period, after a sufficently long time the motion may display large changes.

When the parameter a is constant, the system will be described by the action-angle variables (w_0, J_0) and the Hamiltonian $H = H(J_0, a)$. Assume now that the generating function of the transformation $(q, p) \to (w_0, J_0)$ is of the form $W^*(q, w_0, a)$.

When a is allowed to vary in time, (w_0, J_0) will still be valid canonical variables, but W^* turns into a function of time through a. Then, neither J_0 will be a constant nor w_0 a linear function of time. The appropriate Hamiltonian will be

$$\begin{aligned} K(w_0, J_0, a) &= H(J_0, a) + \frac{\partial W^*}{\partial t} \\ &= H(J_0, a) + \dot{a}\frac{\partial W^*}{\partial a} \ . \end{aligned} \tag{5}$$

The second term in (5) can be seen as a perturbation. Then, the temporal dependence of J_0 is given by

$$\dot{J}_0 = -\frac{\partial K}{\partial w_0} = -\dot{a}\frac{\partial}{\partial w_0}\left(\frac{\partial W^*}{\partial a}\right) \ . \tag{6}$$

Proceeding similarly to the time-dependent perturbation theory, we seek the variation to first-order of the mean value of $\dot{J}_0$ during the period of the unperturbed motion. Since a varies slowly, we can think of it as constant during this interval. Thus, we can write

$$\langle \dot{J}_0 \rangle = -\frac{1}{\tau}\int_\tau \dot{a}\frac{\partial}{\partial w_0}\left(\frac{\partial W^*}{\partial a}\right) dt = -\frac{\dot{a}}{\tau}\int_\tau \frac{\partial}{\partial w_0}\left(\frac{\partial W^*}{\partial a}\right) dt + O(\dot{a}^2, \ddot{a}). \tag{7}$$

One can prove that W^* is a periodic function of w_0, and therefore, can be written, together with its derivative w.r.t. a, as a Fourier series

$$\frac{\partial W^*}{\partial a} = \sum_k A_k e^{2\pi i k w_0}. \tag{8}$$

Substituting (8) in (7) we get

$$\langle \dot{J}_0 \rangle = -\frac{\dot{a}}{\tau}\int_\tau \sum_k 2\pi i k A_k e^{2\pi i k w_0} dt + O(\dot{a}^2, \ddot{a}). \tag{9}$$

Since the integrant does not contain any constant term the integral is zero. This leads to

$$\langle \dot{J}_0 \rangle = 0 + O(\dot{a}^2, \ddot{a}). \tag{10}$$

Thus, $\langle \dot{J}_0 \rangle$ will not display secular variations in the first order, i.e., in $\dot{a}$, which is one of the basic properties of the adiabatic invariance. In this way, the term *almost constant* in the definition of the adiabatic invariant should be interpreted as *constant in the first order*.

Further reading

- L. Parker, *Adiabatic invariance in simple harmonic motion*, Am. J. Phys. 39 (1971) pp. 24-27.

- A.E. Mayo, *Evidence for the adiabatic invariance of the black hole horizon area*, Phys. Rev. D58 (1998) 104007 [gr-qc/9805047].

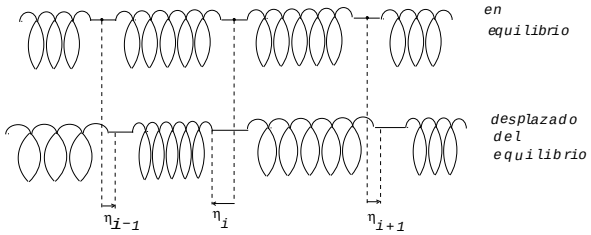
11. MECHANICS OF CONTINUOUS SYSTEMS

Forward: All the formulations of mechanics up to now have dealt with systems having a finite number of degrees of freedom or infinitely countable. However, many physical systems are continuous; for example, an elastic solid in vibrational motion. Every point of such a solid participates in the oscillation(s) and the total motion can be described only by specifying the coordinates of all points. It is not so difficult to modify the previous formulations in order to get a formalism that works for continuous media. The most direct method consists in approximating the continuum by a collection of discrete subunits ('particles') and then study how the equations of motion change when one goes to the continuous limit.

CONTENTS:

11.1 Lagrangian formulation: from discrete to the continuous case

As one of the most simple case for which one can go easily from a discrete system to a continuous counterpart we consider an infinitely long elastic rod doing longitudinal vibrations, i.e., oscillations of its points along the rod axis. A system made of discrete particles that may be considered as a discrete approximation of the rod is an infinite chain of material points separated by the same distance a and connected through identical massless resorts of elastic constant k (see the figure).



We suppose that the material points can move only along the chain. Thus, the discrete system is an extension of the polyatomic lineal molecule presented in chapter 6 of Goldstein's textbook. Thus, the equations of motion of the chain can be obtained by applying the common techniques used in the study of small oscillations. Denoting by η_i the displacement of the ith particle with respect to its equilibrium position, the kinetic energy is

$$T = \frac{1}{2}\sum_i m\,\dot{\eta}_i^2, \tag{1}$$

where m is the mass of each particle. The corresponding potential energy is the sum of the potential energies of each of the resorts:

$$V = \frac{1}{2}\sum_i k\left(\eta_{i+1} - \eta_i\right)^2. \tag{2}$$

From the equations (1) and (2) we get the Lagrangian of the system

$$L = T - V = \frac{1}{2}\sum_i \left(m\,\dot{\eta}_i^2 - k\left(\eta_{i+1} - \eta_i\right)^2\right), \tag{3}$$

which can also be written in the form

$$L = \frac{1}{2}\sum_i a\left[\frac{m}{a}\,\dot{\eta}_i^2 - ka\left(\frac{\eta_{i+1} - \eta_i}{a}\right)^2\right] = \sum_i aL_i, \tag{4}$$

where a is the equilibrium distance between the points. The Euler-Lagrange equations of motion for the η_i coordinates read

$$\frac{m}{a}\,\ddot{\eta}_i - ka\left(\frac{\eta_{i+1} - \eta_i}{a^2}\right) + ka\left(\frac{\eta_i - \eta_{i-1}}{a^2}\right) = 0. \tag{5}$$

The particular form of L in equation (4) and the corresponding equations of motion have been chosen as being the appropriate ones for passing to the continuous limit by means of $a \to 0$. it is clear that m/a turns into mass per lenght unit μ of the continuous system, but the limiting value of ka is not so obvious. We recall that in the case of an elastic rod for which Hooke's law holds, the enlargement of the rod per unit of length is proportional to the force (tension) according to

$$F = Y\xi,$$

where ξ is the enlargement per length unit and Y is the Young modulus. On the other hand, the relative enlargement of the length a of a discrete system is given by $\xi = (\eta_{i+1} - \eta_i)/a$. The corresponding force required to act on the resort will be

$$F = k(\eta_{i+1} - \eta_i) = ka\left(\frac{\eta_{i+1} - \eta_i}{a}\right),$$

and therefore κa should correspond to the Young modulus de Young of the continuous rod. When one passes from the discrete to the continuous case the integer index i for the particular material point of the system turns into the continuous position coordinate x; the discrete variable η_i is replaced by $\eta(x)$. Moreover, the following quantity

$$\frac{\eta_{i+1} - \eta_i}{a} = \frac{\eta(x+a) - \eta(x)}{a}$$

entering L_i goes obviously to the limit

$$\frac{d\eta}{dx},$$

when $a \to 0$. Finally, the sum over the number of discrete particles turns into an integral over x, the length of the rod, and the Lagrangian (4) takes the form

$$L = \frac{1}{2}\int\left(\mu\,\dot{\eta}^2 - Y\left(\frac{d\eta}{dx}\right)^2\right)dx. \tag{6}$$

In the limit $a \to 0$, the last two terms in the equation of motion (5) are

$$\mathrm{Lim}_{a\to 0}\ \frac{-Y}{a}\left\{\left(\frac{d\eta}{dx}\right)_x - \left(\frac{d\eta}{dx}\right)_{x-a}\right\},$$

and taking again the limit $a \to 0$ the last expression defines the second derivative of η. Thus, the equation of motion for the elastic rod will be

$$\mu\frac{d^2\eta}{dt^2} - Y\frac{d^2\eta}{dx^2} = 0, \tag{7}$$

which is the known one-dimensional wave equation of propagational velocity

$$\upsilon = \sqrt{\frac{Y}{\mu}}. \tag{8}$$

The equation (8) is the well-known formula for the sound velocity (the propagational velocity of the elastic longitudinal waves).
This simple example is sufficient to illustrate the main features of the transition from a discrete to a continuous system. The most important fact that we have to understand is the role of the coordinate x. It is not a generalized coordinate; it is merely a continuous index substituting the discrete one i. To any value of x corresponds a generalized coordinate $\eta(x)$. Since η also depends on the continuous variable t, maybe we have to write it in a more precise way than $\eta(x.t)$ by showing that x, equally to t, can be viewed as a parameter of the Lagrangian. If the system would have been 3-dimensional and not one-dimensional, the generalized coordinates would have to be distinguished by three continuous indices x, y, z and write in the form $\eta(x, y, z, t)$. We note that the quantities x, y, z, t are totally independent and show up only in η as explicit variables. The derivatives of η with respect to any of them could thus be written as total derivatives without any ambiguity. The equation (6) also shows the Lagrangian is an integral over the continuous index x; in the three-dimensional case, the Lagrangian would write

$$L = \int\int\int \mathcal{L} dx dy dz, \tag{9}$$

where $\mathcal{L}$ is called the Lagrangian density. In the case of the longitudinal vibrations of the continuous rod, the Lagrangian density is

$$\mathcal{L} = \frac{1}{2}\left\{\mu\left(\frac{d\eta}{dt}\right)^2 - Y\left(\frac{d\eta}{dx}\right)^2\right\}, \tag{10}$$

and corresponds to the continuous limit of L_i appearing in equation (4). Thus, it is more the Lagrangian density than the Lagrangian that we used to describe the motion in this case.

11.2 Lagrangian formulation for continuous systems

We note that in eq. (9) the $\mathcal{L}$ for the elastic rod dependes on $\dot{\eta} = \partial\eta/\partial t$, the spatial derivative of η, $\partial\eta/\partial x$; x and t play a role similar to its parameters. If besides the interactions between first neighbours there would have been some local forces, $\mathcal{L}$ would have been a function of η. In general, for any continuous system $\mathcal{L}$ can be an explicit function of x and t. Therefore, the Lagrangian density for any one-dimensional continuous system should be of the form

$$\mathcal{L} = \mathcal{L}\left(\eta, \frac{d\eta}{dx}, \frac{d\eta}{dt}, x, t\right). \tag{11}$$

The total Lagrangian, following eq.(9), will be

$$L = \int \mathcal{L} dx,$$

and Hamilton's principle in the limit of the continuous system takes the form

$$\delta I = \delta \int_1^2 \int \mathcal{L} dx dt = 0. \tag{12}$$

From Hamilton's principle for the continuous system one would expect to get the continuous limit of the equations of motion. For this (as in section 2-2 of Goldstein) we can use a varied path for a convenient integration, by choosing η from a family of functions of η depending on a parameter as follows

$$\eta(x,t;\alpha) = \eta(x,t;0) + \alpha\zeta(x,t). \tag{13}$$

where $\eta(x,t;0)$ is the correct function satisfying Hamilton's principle and ζ is an arbitrary function of 'good' behaviour that is zero in the extreme points of t and x. If we consider I as a function of α, then in order to be an extremal solution for the derivative of I with respect to α it should become zero in $\alpha = 0$. By directly deriving I we get

$$\frac{dI}{da} = \int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left\{ \frac{\partial \mathcal{L}}{\partial \eta} \frac{\partial \eta}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \frac{\partial}{\partial \alpha} \left(\frac{d\eta}{dt} \right) + \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) \frac{\partial}{\partial \alpha} \left(\frac{d\eta}{dx} \right) dt \right\}. \tag{14}$$

Since the variation of η, $\alpha\zeta$, should be zero at the end points, by integrating by parts on x and t we obtain the relationships

$$\int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \frac{\partial}{\partial \alpha} \left(\frac{d\eta}{dt} \right) dt = - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) \frac{d\eta}{d\alpha} dt,$$

and

$$\int_{x_1}^{x_2} \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \frac{\partial}{\partial \alpha} \left(\frac{d\eta}{dx} \right) dx = - \int_{x_1}^{x_2} \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) \frac{d\eta}{d\alpha} dx.$$

From this, Hamilton's principle could be written as follows

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} dx dt \left\{ \frac{\partial \mathcal{L}}{\partial \eta} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) \right\} \left(\frac{\partial \eta}{\partial \alpha} \right)_0 = 0 \ . \tag{15}$$

Since the varied path is arbitrary the expression in the curly brackets is zero:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} \right) + \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} \right) - \frac{\partial \mathcal{L}}{\partial \eta} = 0. \tag{16}$$

This equation is precisely the right equation of motion as given by Hamilton's principle.

In the particular case of longitudinal vibrations along an elastic rod, the form of the Lagrangian density given by equation (10) shows that

$$\frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dt}} = \mu \frac{d\eta}{dt} \ , \qquad \frac{\partial \mathcal{L}}{\partial \frac{d\eta}{dx}} = -Y \frac{d\eta}{dx} \ , \qquad \frac{\partial \mathcal{L}}{\partial \eta} = 0.$$

Thus, as we would have liked, the Euler-Lagrange equation (16) can be reduced to the equation of motion (7).
The Lagrange formulation that we presented up to now is valid for continuous systems. It can be easily generalized to two, three, and more dimensions. It is convenient to think of a four-dimensional space of coordinates $x_o = t, x_1 = x, x_2 = y, x_3 = z$.
In addition, we introduce the following notation

$$\eta_{\rho,\nu} \equiv \frac{d\eta_\rho}{dx_\nu}; \qquad \eta_{,j} \equiv \frac{d\eta}{dx_j}; \qquad \eta_{i,\mu\nu} \equiv \frac{d^2\eta_i}{dx_\mu dx_\nu}. \tag{17}$$

Employing this notation and the four x coordinates the Lagrangian density (11) takes the form:

$$\mathcal{L} = \mathcal{L}\left(\eta_\rho, \eta_{\rho,\nu}, \S_\nu\right). \tag{18}$$

Thus, the total Lagrangian is an integral extended to all three-dimensional space:

$$L = \int \mathcal{L}\left(dx_i\right). \tag{19}$$

In the case of Hamilton's principle the integral is extended to a region of the four-dimensional space

$$\delta I = \delta \int \mathcal{L}\left(dx_\mu\right) = 0, \tag{20}$$

where the variation of the η_ρ nullify on the surface S entailing the integration region. The symbolic calculation needed to obtain the corresponding Euler-Lagrange equations of motion is similar to the previous symbolic exercise. Let a set of variational functions be

$$\eta_\rho\left(x_\nu; \alpha\right) = \eta_\rho\left(x_\nu\right) + \alpha\zeta\left(x_\nu\right) \ .$$

They depend on a single parameter and reduce to $\eta_\rho\left(x_\nu\right)$ when the parameter α goes to zero. The variation of I is equivalent to put to zero the derivative of I with respect to α, i.e.:

$$\frac{dI}{d\alpha} = \int \left(\frac{\partial \mathcal{L}}{\partial \eta_\rho}\frac{\partial \eta_\rho}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}}\frac{\partial \eta_{\rho,\nu}}{\partial \alpha}\right)\left(dx_\mu\right) = 0. \tag{21}$$

Integrating by parts equation (21), we get

$$\frac{dI}{d\alpha} = \int \left[\frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx_\nu}\left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}}\right)\right]\frac{\partial \eta_\rho}{\partial \alpha}\left(dx_\mu\right) + \int \left(dx_\mu\right)\frac{d}{dx_\nu}\left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}}\frac{\partial \eta_{\rho,\nu}}{\partial \alpha}\right) = 0,$$

and taking the limit $\alpha \to 0$ the previous expression turns into:

$$\left(\frac{dI}{d\alpha}\right)_0 = \int \left(dx_\mu\right)\left[\frac{\partial \mathcal{L}}{\partial \eta_\rho} - \frac{d}{dx_\nu}\left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}}\right)\right]\left(\frac{\partial \eta_\rho}{\partial \alpha}\right)_0 = 0. \tag{22}$$

Since the variations of each η_ρ is arbitrary and independent equation (22) is zero when each term in the brackets is zero separately:

$$\frac{d}{dx_\nu}\left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}}\right) - \frac{\partial \mathcal{L}}{\partial \eta_\rho} = 0. \tag{23}$$

The equations (23) are a system of partial differential equations for the field quantities, with as many equations as ρ's are.

Example: Given the Lagrangian density of an acoustical field

$$\mathcal{L} = \frac{1}{2}\left(\mu_0 \dot{\vec{\eta}}^2 + 2P_0 \nabla \cdot \vec{\eta} - \gamma P_0 \left(\nabla \cdot \vec{\eta}\right)^2\right).$$

μ_0 is the equilibrium mass density and P_0 is the equilibrium pressure of the gas. The first term of $\mathcal{L}$ is the kinetic energy density, while the rest of the terms represent the change in the potential energy of the gas per volume unit due to the work done on the gas o por el curso de las contracciones y expansiones que son la marca de las vibraciones acústicas, γ es el cociente entre los calores molares a presión y a volumen constante obtener las ecuaciones de movimiento.

Solution:

In the four-dimensional notation, the form of the Lagrangian density is

$$\mathcal{L} = \frac{1}{2}\left(\mu_0 \eta_{i,0}\eta_{i,0} + 2P_0\eta_{i,i} - \gamma P_0 \eta_{i,i}\eta_{j,j}\right) . \tag{24}$$

From the equation (23) the following equations of motion are obtained

$$\mu_0 \eta_{j,00} - \gamma P_0 \eta_{i,ij} = 0, \qquad j = 1, 2, 3. \tag{25}$$

Coming back to the vectorial notation the equations (25) can be written as follows

$$\mu_0 \frac{d^2}{\vec{\eta}} dt^2 - \gamma P_0 \nabla\nabla \cdot \vec{\eta} = 0. \tag{26}$$

Using the fact that the vibrations are of small amplitude the relative variation of the gas density is given by the relationship

$$\sigma = -\nabla \cdot \vec{\eta} \, .$$

Applying the divergence and using the previous equation we obtain

$$\nabla^2 \sigma - \frac{\mu_0}{\gamma P_0}\frac{d^2\sigma}{dt^2} = 0$$

which is a three-dimensional wave equation, where

$$\upsilon = \sqrt{\frac{\gamma P_0}{\mu_0}}$$

is the sound velocity in gases.

11.3 Hamiltonian formulation and Poisson brackets

11.3.1 Hamiltonian formulation

Hamilton's formulation for continuous systems is similar to that for discrete systems. To show the procedure we go back to the chain of material points we considered at the beginning of the chapter, where for each η_i one introduces a canonical momentum

$$p_i = \frac{\partial L}{\partial\, \dot{\eta}_i} = a\frac{\partial L_i}{\partial\, \dot{\eta}_i}. \tag{27}$$

The Hamiltonian of the system will be

$$H \equiv p_i\, \dot{\eta}_i - L = a\frac{\partial L_i}{\partial\, \dot{\eta}_i}\, \dot{\eta}_i - L, \tag{28}$$

or

$$H = a\left(\frac{\partial L_i}{\partial\, \dot{\eta}_i}\, \dot{\eta}_i - L_i\right)\ . \tag{29}$$

Recalling that in the limit $a \to 0$, $L \to \mathcal{L}$ and the sum in the equation (29) turns into an integral the Hamiltonian takes the form:

$$H = \int dx \left(\frac{\partial \mathcal{L}}{\partial\, \dot{\eta}}\dot{\eta} - \mathcal{L}\right)\ . \tag{30}$$

The individual canonical momenta p_i, given by equation (27), go to zero in the continuous limit, nevertheless we can define a momentum density π that remains finite:

$$\mathrm{Lim}_{a\to 0}\frac{p_i}{a} \equiv \pi = \frac{\partial \mathcal{L}}{\partial\, \dot{\eta}}\ . \tag{31}$$

The equation (30) has the form of a spatial integral of the Hamiltonian density $\mathcal{H}$ defined as

$$\mathcal{H} = \pi\, \dot{\eta} - \mathcal{L}\ . \tag{32}$$

Even when one can introduce a Hamiltonian formulation in this direct way for classical fields, we should keep in mind the the procedure has to give a special treatment to the time variable. This is different of the Lagrangian formulation where all the independent variables were considered on the same foot. This is why the Hamilton method will be treated in a distinct manner. The obvious way to do a three-dimensional generalizeation of the field η_ρ is the following.
We define a canonical momentum

$$\pi_{_\rho}\left(x_{_\mu}\right) = \frac{\partial \mathcal{L}}{\partial\, \dot{\eta}_{_\rho}}. \tag{33}$$

where $\eta_{_\rho}\left(x_i, t\right), \pi_{_\rho}\left(x_i, t\right)$ together, define the phase space of infinite dimensions describing the classical field and its time evolution.

Similarly to a discrete system we can seek a conservation theorem for π that look like the corresponding canonical momentum of the discrete systems. If a given field η_ρ is a cyclic variable ($\mathcal{L}$ does not present an explicit dependence on η_ρ), the Lagrange field equation has the form of a conservation of a current:

$$\frac{d}{dx_{_\mu}}\frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,\mu}}} = 0$$

that is

$$\frac{d\pi_{_\rho}}{dt} - \frac{d}{dx_{_i}}\frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,i}}} = 0\ . \tag{34}$$

Thus, if η_ρ is cyclic there is a conservative integral quantity

$$\Pi_\rho = \int dV \pi_{_\rho}\left(x_i, t\right).$$

The generalization for the density (eq. (32)) in the case of the Hamiltonian density is

$$\mathcal{H}\left(\eta_\rho, \eta_{_{\rho,i}}, \pi_\rho, x_\mu\right) = \pi_{_\rho}\ \dot{\eta}_{_\rho} - \mathcal{L}, \tag{35}$$

where it is assumed that the functional dependence of $\dot{\eta}_\rho$ can be eliminated by inverting the eqs. (33). From this definition, one gets

$$\frac{\partial \mathcal{H}}{\partial \pi_\rho} = \dot{\eta}_{_\rho} + \pi_{_\lambda}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \pi_{_\rho}} - \frac{\partial \mathcal{L}}{\partial\ \dot{\eta}_{_\lambda}}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \pi_{_\rho}} = \dot{\eta}_{_\rho} \tag{36}$$

as a consequence of eq. (33). In a similar way, we obtain

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = \pi_{_\lambda}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \eta_{_\rho}} - \frac{\partial \mathcal{L}}{\partial\ \dot{\eta}_{_\lambda}}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \eta_{_\rho}} - \frac{\partial \mathcal{L}}{\partial \eta_{_\rho}} = -\frac{\partial \mathcal{L}}{\partial \eta_{_\rho}}. \tag{37}$$

Now, using Lagrange equations, eq. (37) turns into

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} = -\frac{d}{dx_{_\mu}}\left(\frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,\mu}}}\right) = -\ \dot{\pi}_{_\rho} - \frac{d}{dx_{_i}}\left(\frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,i}}}\right). \tag{38}$$

Due to the occurrence of $\mathcal{L}$, we still do not have a useful form. However, by a similar procedure as used for getting the terms $\frac{\partial \mathcal{H}}{\partial \pi_\rho}$ and $\frac{\partial \mathcal{H}}{\partial \eta_\rho}$ for $\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}}$ we have

$$\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} = \pi_{_\lambda}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \eta_{_{\rho,i}}} - \frac{\partial \mathcal{L}}{\partial\ \dot{\eta}_{_\lambda}}\frac{\partial\ \dot{\eta}_{_\lambda}}{\partial \eta_{_{\rho,i}}} - \frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,i}}} = -\frac{\partial \mathcal{L}}{\partial \eta_{_{\rho,i}}}\ . \tag{39}$$

Thus, by substituting (39) in (38) we get

$$\frac{\partial \mathcal{H}}{\partial \eta_\rho} - \frac{d}{dx_i}\left(\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}}\right) = -\ \dot{\pi}_{_\rho}\ \ . \tag{40}$$

The equations (36) y (40) can be rewritten using a notation closer to the Hamilton ones for a discrete system. This is possible by employing the concept of functional derivative

$$\frac{\delta}{\delta\psi} = \frac{\partial}{\partial\psi} - \frac{d}{dx_i}\frac{\partial}{\partial\psi_{,i}}. \tag{41}$$

Since $\mathcal{H}$ is not a function of $\pi_{\rho,i}$ the equations (36) and (40) can be written

$$\dot{\eta}_\rho = \frac{\delta\mathcal{H}}{\delta\pi_\rho}, \qquad \dot{\pi}_\rho = -\frac{\delta\mathcal{H}}{\delta\eta_\rho}. \tag{42}$$

Now, by employing the same notation, the Lagrange eqs. (23) can be written as follows

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\,\dot{\eta}_\rho}\right) - \frac{\delta\mathcal{L}}{\delta\eta_\rho} = 0. \tag{43}$$

It is fair to say that the almost unique advantage of the functional derivative is the similarity of the formulas with the discrete case. Moreover, one can see the parallel treatment of space and time variables.

11.3.2 Poisson brackets

We can get other properties of $\mathcal{H}$ by developing the total time derivative of eq. (35), remembering that $\dot{\eta}_\rho$ es función de $\eta_\rho, \eta_{\rho,j}, \pi_\rho$ y π_μ. Thus, we have that

$$\frac{d\mathcal{H}}{dt} = \dot{\pi}_\rho\dot{\eta}_\rho + \pi_\rho\frac{d\,\dot{\eta}_\rho}{dt} - \frac{\partial\mathcal{L}}{\partial\eta_\rho}\,\dot{\eta}_\rho - \frac{\partial\mathcal{L}}{\partial\,\dot{\eta}_\rho}\frac{d\,\dot{\eta}_\rho}{dt} - \frac{\partial\mathcal{L}}{\partial\eta_{\rho,i}}\frac{d\eta_{\rho,i}}{dt} - \frac{\partial\mathcal{L}}{\partial t}.$$

In this expression, the second and the forth terms nullifie each other because of the definition (33). The derivative simplifies to

$$\frac{d\mathcal{H}}{dt} = \dot{\pi}_\rho\dot{\eta}_\rho - \frac{\partial\mathcal{L}}{\partial\eta_\rho}\,\dot{\eta} - \frac{\partial\mathcal{L}}{\partial\eta_{\rho,i}}\frac{d\eta_{\rho,i}}{dt} - \frac{\partial\mathcal{L}}{\partial t}. \tag{44}$$

On the other hand, considering $\mathcal{H}$ as a function of $\eta_\rho, \eta_{\rho,j}, \pi_\rho$ and π_μ, the total time derivative is

$$\frac{d\mathcal{H}}{dt} = \dot{\pi}_\rho\,\frac{\partial\mathcal{H}}{\partial\pi_\rho} + \frac{\partial\mathcal{H}}{\partial\eta_\rho}\,\dot{\eta}_\rho + \frac{\partial\mathcal{H}}{\partial\eta_{\rho,i}}\frac{d\eta_{\rho,i}}{dt} + \frac{\partial\mathcal{H}}{\partial t}\ , \tag{45}$$

where we wrote the expression in such a manner to get an easy comparison with the second terms of eq. (44), and where using the eqs. (36), (37) and (39) we obtain

$$\frac{\partial\mathcal{H}}{\partial t} = -\frac{\partial\mathcal{L}}{\partial t}, \tag{46}$$

which is an analog to the corresponding one for discrete systems.

On the other hand, the equality of the total and partial time derivatives does not hold. Using Hamilton's equations of motion (eq. (36) yand(40)) and interchanging the order of derivation, the eq.(45) can be written as follows

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial \pi_\rho} \frac{d}{dx_i} \left(\frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) + \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \frac{d\,\dot{\eta}_\rho}{dx_i} + \frac{\partial \mathcal{H}}{\partial t} \ .$$

Now, employing eq. (46) and combining the terms we finally have

$$\frac{d\mathcal{H}}{dt} = \frac{d}{dx_i} \left(\dot{\eta}_\rho \ \frac{\partial \mathcal{H}}{\partial \eta_{\rho,i}} \right) + \frac{\partial \mathcal{H}}{\partial t}, \tag{47}$$

which is the closest we can approximate to the corresponding equations for discrete systems.

When $\mathcal{L}$ does not depend explicitly on time t, it will not be in $\mathcal{H}$ as well. This implies the existence of a conservative current and consequently the conservation of an integral quantity, which in this case is

$$H = \int \mathcal{H} dV \ . \tag{48}$$

Thus, if $\mathcal{H}$ is not an explicit function of time, the conserved quantity is not $\mathcal{H}$, but the integral H.

The Hamiltonian is just an example of functions that are volume integrals of densities. A general formalism can be provided for the time derivatives of such integral quantities. Consider a given density $\mathcal{U}$ and let it be a function of the coordinates of the phase space (η_ρ, π_ρ), of its spatial gradients and possibly on x_μ:

$$\mathcal{U} = \mathcal{U} \left(\eta_\rho, \pi_{_\rho}, \eta_{\rho,i}, \pi_{_{\rho,i}}, x_\mu \right) \ . \tag{49}$$

The corresponding integral quantity is

$$U\left(t\right) = \int \mathcal{U} dV \tag{50}$$

where the volume integral is extended to all the space limited by the contour surface on which η_ρ y π_ρ take zero values. Doing the time derivative of U we have in general,

$$\frac{dU}{dt} = \int \left\{ \frac{\partial \mathcal{U}}{\partial \eta_\rho} \ \dot{\eta}_{_\rho} + \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \ \dot{\eta}_{\rho,i} + \frac{\partial \mathcal{U}}{\partial \pi_\rho} \ \dot{\pi}_\rho + \frac{\partial \mathcal{U}}{\partial \pi_{\rho,i}} \ \dot{\pi}_{\rho,i} + \frac{\partial \mathcal{U}}{\partial t} \right\} dV \ . \tag{51}$$

Let us consider the term

$$\int dV \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \ \dot{\eta}_{\rho,i} = \int dV \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \frac{d\,\dot{\eta}_\rho}{dx_i} \ .$$

Integrating by parts and taking into account the nullity of η_ρ andd its derivatives on the contour surface, we have

$$\int dV \frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \ \dot{\eta}_{\rho,i} = - \int dV \ \dot{\eta}_{_\rho} \ \frac{d}{dx_i} \left(\frac{\partial \mathcal{U}}{\partial \eta_{\rho,i}} \right).$$

For the term in $\dot{\pi}_{\rho,i}$ one uses a similar technique. Substituting the obtained expressions and grouping appropriately the coefficients of $\dot{\eta}$ and $\dot{\pi}_\rho$, respectively, and using the functional derivative notation equation (51) is reduced to

$$\frac{dU}{dt} = \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \dot{\eta}_\rho + \frac{\delta \mathcal{U}}{\delta \pi_\rho} \dot{\pi}_\rho + \frac{\partial \mathcal{U}}{\partial t} \right\} . \tag{52}$$

Finally, introducing the canonical equations of motion (42), we have

$$\frac{dU}{dt} = \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \frac{\delta \mathcal{H}}{\delta \pi_\rho} - \frac{\delta \mathcal{H}}{\delta \eta_\rho} \frac{\delta \mathcal{U}}{\delta \pi_\rho} \right\} + \int dV \frac{\partial \mathcal{U}}{\partial t} . \tag{53}$$

The first integral in the rhs corresponds clearly to the Poisson brackets. If $\mathcal{U}$ and $\mathcal{W}$ are two functions of density, these considerations allows us to take as the definition of the Poisson bracket for integral quantities as

$$[U, W] = \int dV \left\{ \frac{\delta \mathcal{U}}{\delta \eta_\rho} \frac{\delta \mathcal{W}}{\delta \pi_\rho} - \frac{\delta \mathcal{W}}{\delta \eta_\rho} \frac{\delta \mathcal{U}}{\delta \pi_\rho} \right\} . \tag{54}$$

We define the partial time derivative of U through the following expression

$$\frac{\partial U}{\partial t} = \int dV \frac{\partial \mathcal{U}}{\partial t} . \tag{55}$$

Thus, the eq. (53) podrá could be written

$$\frac{dU}{dt} = [U, H] + \frac{\partial U}{\partial t} , \tag{56}$$

which exactly corresponds, in this notation, to the equation for the discrete systems. Since by definition the Poisson bracket of H with itself is zero, the eq. (46) turns into

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} , \tag{57}$$

which is the integral form of eq. (47). Thus, the Poisson bracket formalism ocurr as a consequence of the Hamiltonian formulation. However, one cannot perform a description in terms of Poisson brackets for field theories by a step by step correspondence with the discrete case.

Nevertheless there is one way to work out the classical fields which includes almost all the ingredients of Hamilton's formulation and Poisson brackets of the discrete case. The basic idea is to replace the continuous spatial variable or the continuous index by a countable discrete index.

The requirement that η be zero at the end points is a contour condition that might be achieved physically only by fixing the rod between two rigid walls. Then, the amplitud of oscillation can be represented by means of a Fourier series:

$$\eta(x) = \sum_{n=0}^{\infty} q_n \sin \frac{2\pi n (x - x_1)}{2L} . \tag{58}$$

Instead of the continuous index x we have the discrete η. We could use this representation of x only when $\eta(x)$ is a regular function, which happens for many field quantities.
We assume that there is only one real field η that can be developed in a three-dimensional Fourier series

$$\eta(\overrightarrow{r}, t) = \frac{1}{V^{1/2}} \sum_{k=0} q_k(t) \exp\left(i \overrightarrow{k} \cdot \overrightarrow{r}\right) \tag{59}$$

Here, $\vec{k}$ is a wave vector that can take only discrete modulus and directions, so that in only one lineal dimension there will be an integer (or sometimes, half-integer) wavelengths. We say that $\vec{k}$ has a discrete spectrum. The scalar subindex k represents a certain order of the set of integer subindices that are used to enumerate the discrete values of $\vec{k}$; V is the volume of the system, which appear as a normalization factor.
The orthogonality of the exponentials in the entire volume can be stated through the relationship

$$\frac{1}{V} \int e^{i\left(\vec{k}-\vec{k}'\right).\vec{r}} dV = \delta_{k,k'} \ . \tag{60}$$

As a matter of fact, the allowed values of k are those for which the condition (60) is satisfied, and the coefficients $q_k(t)$ are given by

$$q_k(t) = \frac{1}{V^{1/2}} \int e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \eta(\overrightarrow{r}, t)\, dV \ . \tag{61}$$

Similarly, for the density of the canonical momentum we have

$$\pi(\overrightarrow{r}, t) = \frac{1}{V^{1/2}} \sum_{k} p_k(t)\, e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \tag{62}$$

with $p_k(t)$ defined as

$$p_k(t) = \frac{1}{V^{1/2}} \int e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \pi(\overrightarrow{r}, t)\, dV \ . \tag{63}$$

Both q_k and p_k are integral quantities. Thus, we can look for their Poisson brackets. Since the exponentials do not contain the fields we have, according to (54)

$$\begin{aligned} [q_k, p_{k'}] &= \frac{1}{V} \int dV e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \left\{ \frac{\delta\eta}{\delta\eta} \frac{\delta\pi}{\delta\pi} - \frac{\delta\pi}{\delta\eta} \frac{\delta\eta}{\delta\pi} \right\} \\ &= \frac{1}{V} \int dV e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \end{aligned}$$

that is, by equation (60),

$$[q_k, p_{k'}] = \delta_{k,k'} \ . \tag{64}$$

From the definition of the Poisson brackets it is obvious that

$$\left[q_k, q_{k'}\right] = \left[p_k, p_{k'}\right] = 0 \ . \tag{65}$$

The time dependence of q_k is sought starting from

$$\dot{q}_k \ (t) = [q_k, H] = \frac{1}{V^{1/2}} \int dV e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \left\{ \frac{\delta \eta}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \mathcal{H}}{\delta \eta} \frac{\delta \eta}{\delta \pi} \right\}$$

that is

$$\dot{q}_k \ (t) = \frac{1}{V^{1/2}} \int e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \frac{\delta \mathcal{H}}{\delta \pi} \ . \tag{66}$$

On the other hand, we have

$$\frac{\partial H}{\partial p_k} = \int dV \frac{\partial \mathcal{H}}{\partial \pi} \frac{\partial \pi}{\partial p_k} \tag{67}$$

and therefore we get

$$\frac{\partial \pi}{\partial p_k} = \frac{1}{V^{1/2}} e^{-i \overrightarrow{k} \cdot \overrightarrow{r}} \ . \tag{68}$$

Comparando las ecuaciones (67) y (66) tenemos

$$\dot{q}_k \ (t) = \frac{\partial H}{\partial p_k} \ . \tag{69}$$

In a similar way we can obtain the equation of motion for p_k

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} \ . \tag{70}$$

Thus, p_k y q_k, obey the Hamilton equations of motion.

11.4 Noether's theorem

We have already seen that the properties of the Lagrangian (or of the Hamiltonian) imply the existence of conservative quantities. Thus, if the Lagrangian does not contain explicitly a particular displacement coordinate, the corresponding caonical momentum is conserved. The absence of an explicit dependence on a coordinate means that the Lagrangian is not changed by a transformation that alter the value of that coordinate; we say that it is invariant or symmetric for that transformation.
The symmetry under a coordinate transformation refers to the effects of an infinitesimal transformation as follows

$$x_\mu \rightarrow x_\mu^{'} = x_\mu + \delta x_\mu, \tag{71}$$

where the variation δx_μ can be a function of all the other x_ν. Noether's theorem deals with the effect of the transformation of the field quantities itselves. Such a transformation can be written

$$\eta \left(x_\mu\right) \rightarrow \eta_\rho^{'} \left(x_\mu^{'}\right) = \eta_\rho \left(x_\mu\right) + \delta \eta_\rho \left(x_\mu\right) . \tag{72}$$

Here $\delta\eta_\rho\left(x_\mu\right)$ is a measure of the effect of the variations of x_μ and of η_ρ. It can be a function of all the other fields η_λ. The variation of one of the field variables in a particular point of space x_μ is a different quantity $\overline{\delta}\eta_\rho$:

$$\eta_\rho^{'}\left(x_\mu^{'}\right)=\eta_\rho\left(x_\mu\right)+\overline{\delta}\eta_\rho\left(x_\mu\right). \tag{73}$$

The characterization of the transformations by means of infinitesimal variations, starting from untransformed quantities means that we consider only continuous transformations. Therefore, the symmetry under inversion of the three-dimensional space is not a symmetry of the continuous type to which Noether's theorem can be applied. As a consequence of the transformations both in the coordinates and the fields, the Lagrangian will, in general, appear as a different function of the field and spacetime coordinates:

$$\mathcal{L}\left(\eta_\rho\left(x_\mu\right),\eta_{\rho,\nu}\left(x_\mu\right),x_\mu\right)\rightarrow\mathcal{L}^{'}\left(\eta_\rho^{'}\left(x_\mu^{'}\right),\eta_{\rho,\nu}^{'}\left(x_\mu^{'}\right),x_\mu^{'}\right). \tag{74}$$

The version of Noether's theorem that we shall present is not of the most general form possible, but makes easier the proof, without loosing too much of its generality and the usefulness of the conclusions. We shall suppose the following three conditions:

1. The spacetime is Euclidean, meaning that the relativity is reduced to the Minkowski space, which is complex but flat.

2. The Lagrangian density is of the same functional form for the transformed quantities as for the original ones, that is

$$\mathcal{L}^{'}\left(\eta_\rho^{'}\left(x_\mu^{'}\right),\eta_{\rho,\nu}^{'}\left(x_\mu^{'}\right),x_\mu^{'}\right)=\mathcal{L}\left(\eta_\rho^{'}\left(x_\mu^{'}\right),\eta_{\rho,\nu}^{'}\left(x_\mu^{'}\right),x_\mu^{'}\right). \tag{75}$$

3. The value of the action integral is invariant under the transformation

$$I^{'}\equiv\int_{\Omega^{'}}\left(dx_\mu\right)\mathcal{L}^{'}\left(\eta_\rho^{'}\left(x_\mu^{'}\right),\eta_{\rho,\nu}^{'}\left(x_\mu^{'}\right),x_\mu^{'}\right)=\int_\Omega\mathcal{L}\left(\eta_\rho\left(x_\mu\right),\eta_{\rho,\nu}\left(x_\mu\right),x_\mu\right). \tag{76}$$

Combining the equations (75) and (76) we get the condition

$$\int_{\Omega^{'}}\left(dx_\mu\right)\mathcal{L}\left(\eta_\rho^{'}\left(x_\mu\right),\eta_{\rho,\nu}^{'}\left(x_\mu\right),x_\mu\right)-\int_\Omega\mathcal{L}\left(\eta_\rho\left(x_\mu\right),\eta_{\rho,\nu}\left(x_\mu\right),x_\mu\right)=0\ . \tag{77}$$

From the invariance condition, the equation (77) becomes

$$\begin{aligned}&\int_{\Omega^{'}}dx_\mu\mathcal{L}\left(\eta^{'},x_\mu\right)-\int_\Omega dx_\mu\mathcal{L}\left(\eta,x_\mu\right) \\ =\ &\int_\Omega dx_\mu\left[\mathcal{L}\left(\eta^{'},x_\mu\right)-\mathcal{L}\left(\eta,x_\mu\right)\right]+\int_s\mathcal{L}\left(\eta\right)\delta x_\mu dS_\mu=0\ .\end{aligned} \tag{78}$$

Here, $\mathcal{L}\left(\eta,x_\mu\right)$ is a shorthand notation for the total functional dependence, S is the three-dimensional surface of the region Ω and δx_μ is the difference

vector between the points of S and the corresponding points of the transformed surface $S^{'}$. The last integral can be transformedthrough the theorem of four-dimensional divergence. This leads to the following invariance condition

$$0 = \int_{\Omega} dx_{\mu} \left\{ \left[\mathcal{L}\left(\eta^{'}, x_{\mu}\right) - \mathcal{L}\left(\eta, x_{\mu}\right) \right] + \frac{d}{dx_{\mu}} \left(\mathcal{L}\left(\eta, x_{\mu}\right) \delta x_{\nu} \right) \right\}. \tag{79}$$

Now, using the equation (73), the term in the brackets can be written in the first-order of approximation as follows

$$\mathcal{L}\left(\eta^{'}_{\rho}\left(x_{\mu}\right), \eta^{'}_{\rho,\nu}\left(x_{\mu}\right), x_{\mu}\right) - \mathcal{L}\left(\eta_{\rho}\left(x_{\mu}\right), \eta_{\rho,\nu}\left(x_{\mu}\right), x_{\mu}\right) = \frac{\partial \mathcal{L}}{\partial \eta_{\rho}} \overline{\delta} \eta_{\rho} + \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \overline{\delta} \eta_{\rho,\nu}.$$

Using the Lagrange field equations

$$\mathcal{L}\left(\eta', x_{\mu}\right) - \mathcal{L}\left(\eta, x_{\mu}\right) = \frac{d}{dx_{\nu}} \left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \overline{\delta} \eta_{\rho} \right).$$

Then, the invariance condition (79) ocurr as

$$\int \left(dx_{\mu}\right) \frac{d}{dx_{\nu}} \left\{ \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \overline{\delta} \eta_{\rho} - \mathcal{L} \delta x_{\nu} \right\} = 0, \tag{80}$$

which already has the form of an equation for the conservation of a current. It is useful to develop more the condition giving the form of the infinitesimal transformation as a function of the R infinitesimal parameters $\varepsilon_{r,}r = 1, 2, ..., R$, such that the variations of x_{μ} y η_{ρ} be lineal in ε_{r}:

$$\delta x_{\nu} = \varepsilon_{r} X_{r\nu}, \qquad \delta \eta_{\rho} = \epsilon_{r} \Psi_{r\rho}. \tag{81}$$

By substituting these conditions in eq. (80) we get

$$\int \epsilon_{r} \frac{d}{dx_{\nu}} \left\{ \left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\sigma} - \mathcal{L} \delta_{\nu\sigma} \right) X_{r\sigma} - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \Psi_{r\rho} \right\} \left(dx_{\mu}\right) = 0 \ .$$

Since the ε_{r} parameters are arbitrary, there are $r$conservative currents as solutions of the differential conservation theorems:

$$\frac{d}{dx_{\nu}} \left\{ \left(\frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \eta_{\rho,\sigma} - \mathcal{L} \delta_{\nu\sigma} \right) X_{r\sigma} - \frac{\partial \mathcal{L}}{\partial \eta_{\rho,\nu}} \Psi_{r\rho} \right\} = 0 \ . \tag{82}$$

The equations (82) are the main conclusion of Noether's theorem, telling that if the system has symmetry properties fulfilling the conditions (1) and (2) for transformations of the type given by the equations (81), then there exist r conserved quantities.

Further reading

R.D. Kamien, *Poisson bracket formulation of nematic polymer dynamics*, cond-mat/9906339 (1999)