Chapter 2
Lagrangian Mechanics

> The reader will find no figures in this work. The methods which I set forth do not require either constructions or geometrical or mechanical reasonings, but only algebraic operations, subject to a regular and uniform rule of procedure.
>
> Joseph-Louis Lagrange in the preface to *Mécanique Analytique*.

2.1 Introduction

In the preceding chapter we outlined the concept of least action and introduced the basis of Lagrangian mechanics. We shall now drop the terminology of the 18th century and develop this concept in a form which can easily be applied in the 21st century.

Our approach will be to first base our development on Newton's Laws and d'Alembert's principle as Lagrange originally did. This will produce the Euler–Lagrange Equations for any system that obeys Newton's Laws. We will follow closely the development of Edmund T. Whittaker[1] in *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* [[125], Chap. II]. We will then show that the Euler–Lagrange Equations result from a variational principle. In this we will introduce, as Lagrange did, the method of Lagrange undetermined multipliers. This will also provide for us a simple method for handling general constraints that otherwise may present insurmountable algebraic difficulties.

The result will be the regular and uniform rule of procedure promised by Lagrange.

[1] Edmund Taylor Whittaker (1873–1956) was professor of mathematics at the University of Edinurgh from 1911 to the end of his career.

C.S. Helrich, *Analytical Mechanics*, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-3-319-44491-8_2

2.2 Kinematics

We consider the motion of a system consisting of a collection of point particles with masses m_i. Because we are treating *classical* mechanics, these particles will not be molecules, atoms, or subatomic particles. They are *classical point particles*, which are small enough, compared to the dimensions of the system, to be considered mathematically as geometrical points. Their description is then completely provided by the position vectors $\boldsymbol{r}_\mathrm{i}$ locating the points i and the velocities $\boldsymbol{v}_\mathrm{i}$ of the moving points. Classical point particles have no rotational energy.

Kinematics is the description of the motion of these classical point particles, which is the mathematical representation of the position and velocity vectors of the point particles. If we were attempting to formulate our problem as a direct translation of Newton's Second Law into vector form we would also require the acceleration. However, as we found in the preceding chapter, the Analytical Mechanics of Euler and Lagrange only requires the positions and velocities of the particles.

We first choose any one of the standard coordinate systems of analytical geometry, which is convenient for the system at hand. For general descriptions we normally simply choose a rectangular Cartesian system. For problems with particular symmetries we naturally choose a system that reflects those symmetries, such as polar or spherical systems. The position and velocity vectors for a point in the system we have chosen is then the kinematic description of each particle.

Each of the coordinate systems of analytical geometry has three orthonormal basis vectors. In the rectangular Cartesian system these basis vectors $\left(\hat{e}_\mathrm{x}, \hat{e}_\mathrm{y}, \hat{e}_\mathrm{z}\right)$ are fixed in space and oriented along the axes (x, y, z). In the cylindrical system the basis vectors are $\left(\hat{e}_\mathrm{r}, \hat{e}_\vartheta, \hat{e}_\mathrm{z}\right)$. Only the basis vector $\hat{e}_\mathrm{z}$ is fixed in space while the orientation of the vectors $\hat{e}_\mathrm{r}$ and $\hat{e}_\vartheta$ change with the motion of the particle. In the spherical system the basis vectors are $\left(\hat{e}_\rho, \hat{e}_\vartheta, \hat{e}_\phi\right)$ none of which is fixed in space as the particle moves.

In the remainder of this section we will formulate the position vector locating the classical point particle and obtain the velocity in each of the coordinate systems. We shall use the standard (Newtonian) dot notation for the time derivative of a function $\dot{f} = \mathrm{d}f/\mathrm{d}t$.

2.2.1 Rectangular Cartesian

We have illustrated the rectangular Cartesian system in Fig. 2.1. The vector $\boldsymbol{r}$ is represented in the basis $\left(\hat{e}_\mathrm{x}, \hat{e}_\mathrm{y}, \hat{e}_\mathrm{z}\right)$. And the triad is right handed, i.e.

$$\hat{e}_\mathrm{x} \times \hat{e}_\mathrm{y} = \hat{e}_\mathrm{z}. \tag{2.1}$$

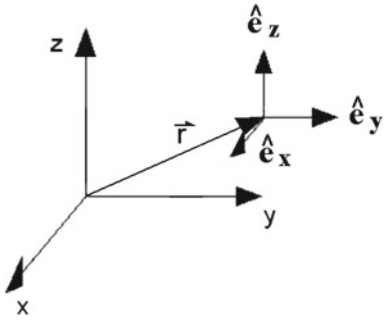


Fig. 2.1 Rectangular Cartesian coordinates

We have drawn the basis vector triad $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ at the tip of the position vector $\boldsymbol{r}$ for illustrative purposes. The orientation of each of the basis vectors $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ is fixed in space.

The vector $\boldsymbol{r}$ is

$$\boldsymbol{r} = x\hat{e}_x + y\hat{e}_y + z\hat{e}_z. \tag{2.2}$$

Because only the components (x, y, z) are time dependent the velocity vector $\boldsymbol{v}$ is

$$\boldsymbol{v} = \dot{x}\hat{e}_x + \dot{y}\hat{e}_y + \dot{z}\hat{e}_z. \tag{2.3}$$

2.2.2 Cylindrical

We have illustrated the cylindrical coordinate system in Fig. 2.2. The vector $\boldsymbol{r}$ is represented in the basis $(\hat{e}_r, \hat{e}_\vartheta, \hat{e}_z)$. And the triad $(\hat{e}_r, \hat{e}_\vartheta, \hat{e}_z)$ is right handed so that

$$\hat{e}_r \times \hat{e}_\vartheta = \hat{e}_z. \tag{2.4}$$

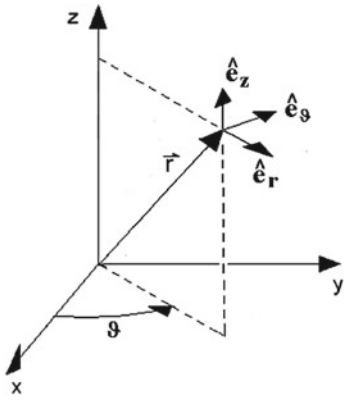


Fig. 2.2 Cylindrical coordinates

We have drawn the vector triad $\left(\hat{e}_{\mathrm{r}}, \hat{e}_{\vartheta}, \hat{e}_{\mathrm{z}}\right)$ at the tip of the position vector $\boldsymbol{r}$. The basis vector $\hat{e}_{\mathrm{z}}$ remains oriented along the vertical axis and is independent of time. The basis vector $\hat{e}_{\mathrm{r}}$ is in the plane of $\boldsymbol{r}$ and $\hat{e}_{\mathrm{z}}$ and oriented perpendicularly to $\hat{e}_{\mathrm{z}}$. The basis vector $\hat{e}_{\vartheta}$ is perpendicular to $\hat{e}_{\mathrm{r}}$ and $\hat{e}_{\mathrm{z}}$. The basis vectors $\hat{e}_{\mathrm{r}}$ and $\hat{e}_{\vartheta}$ then rotate around an axis parallel to $\hat{e}_{\mathrm{z}}$ as the tip of the position vector $\boldsymbol{r}$ moves.

The position vector $\boldsymbol{r}$ is

$$\boldsymbol{r} = r\hat{e}_{\mathrm{r}} + z\hat{e}_{\mathrm{z}}. \tag{2.5}$$

The differential rotation $\mathrm{d}\vartheta$ of the dyad $\left(\hat{e}_{\mathrm{r}}, \hat{e}_{\vartheta}\right)$ around $\hat{e}_{\mathrm{z}}$ results in a differential change $\mathrm{d}\hat{e}_{\mathrm{r}} = \mathrm{d}\vartheta\hat{e}_{\vartheta}$ in the basis vector $\hat{e}_{\mathrm{r}}$. The velocity vector $\boldsymbol{v}$ is then

$$\boldsymbol{v} = \dot{r}\hat{e}_{\mathrm{r}} + r\dot{\vartheta}\hat{e}_{\vartheta} + \dot{z}\hat{e}_{\mathrm{z}}. \tag{2.6}$$

2.2.3 Spherical

We have illustrated the spherical coordinate system in Fig. 2.3. The vector $\boldsymbol{r}$ is represented in the basis $\left(\hat{e}_{\rho}, \hat{e}_{\vartheta}, \hat{e}_{\phi}\right)$. And the triad $\left(\hat{e}_{\rho}, \hat{e}_{\vartheta}, \hat{e}_{\phi}\right)$ is right handed so that

$$\hat{e}_{\rho} \times \hat{e}_{\phi} = \hat{e}_{\vartheta}. \tag{2.7}$$

The length (magnitude) of the position vector $\boldsymbol{r}$ is ρ.

We have drawn the basis vector triad $\left(\hat{e}_{\rho}, \hat{e}_{\vartheta}, \hat{e}_{\phi}\right)$ at the tip of the position vector $\boldsymbol{r}$. The basis vector $\hat{e}_{\rho}$ is oriented along the direction of the position vector $\boldsymbol{r}$ and changes in orientation with the position vector. The basis vector $\hat{e}_{\rho}$ is then independent of the length ρ of the position vector $\boldsymbol{r}$ and depends on the *azimuthal*[2] angle ϑ and the *polar* angle ϕ. An infinitesimal rotation $\mathrm{d}\phi$ of the polar angle produces a differential

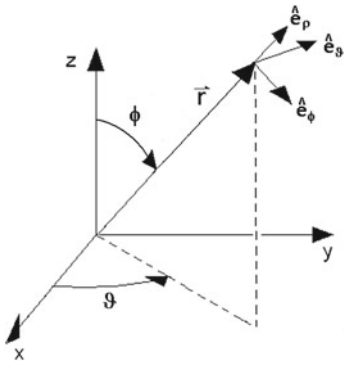


Fig. 2.3 Spherical coordinates

[2] Azimuth comes from the Arabic word as-simt, which means direction, referring to the direction a person faces. In Fig. 2.3 the azimuthal angle ϑ locates the projection of the vestor $\boldsymbol{r}$ on the (x, y) plane with respect to the $x-$axis.

change $\mathrm{d}\hat{e}_\rho = \mathrm{d}\phi\hat{e}_\phi$ in the basis vector $\hat{e}_\rho$, which is analogous to the change $\mathrm{d}\hat{e}_\mathrm{r}$ in the cylindrical basis vector we found above. In the same fashion we deduce the differential change in $\hat{e}_\rho$ resulting from a differential change $\mathrm{d}\vartheta$ in the azimuthal angle as $\mathrm{d}\hat{e}_\rho = \sin\phi\mathrm{d}\vartheta\hat{e}_\vartheta$. The Pfaffian[3] for the basis vector $\hat{e}_\rho$ is then

$$\begin{aligned}\mathrm{d}\hat{e}_\rho &= \left(\frac{\partial\hat{e}_\rho}{\partial\phi}\right)\mathrm{d}\phi + \left(\frac{\partial\hat{e}_\rho}{\partial\vartheta}\right)\mathrm{d}\vartheta \\ &= \mathrm{d}\phi\hat{e}_\phi + \sin\phi\mathrm{d}\vartheta\hat{e}_\vartheta. \end{aligned} \tag{2.8}$$

The position vector $\boldsymbol{r}$ is

$$\boldsymbol{r} = \rho\hat{e}_\rho. \tag{2.9}$$

With (2.8) the velocity vector $\boldsymbol{v}$ is

$$\boldsymbol{v} = \dot{\rho}\hat{e}_\rho + \rho\dot{\vartheta}\sin\phi\hat{e}_\vartheta + \rho\dot{\phi}\hat{e}_\phi. \tag{2.10}$$

2.3 From Newton's Laws

2.3.1 General Formulation

The total force acting on the i^{th} (point) particle we designate as $\boldsymbol{F}_\mathrm{i}$ and consider that this force consists of both forces whose origin is external to the system $\boldsymbol{F}_\mathrm{i}^{\mathrm{ext}}$ and forces arising from interactions among the particles, which we term internal forces $\boldsymbol{F}_\mathrm{i}^{\mathrm{int}}$. That is

$$\boldsymbol{F}_\mathrm{i} = \boldsymbol{F}_\mathrm{i}^{\mathrm{ext}} + \boldsymbol{F}_\mathrm{i}^{\mathrm{int}}.$$

External forces arise from fields, such as gravitational or electromagnetic, and possible contact forces from external bodies constraining the motion.[4] The system of particles we are considering is, therefore, the most general possible in classical terms.

Applying Newton's Second Law to each particle i we obtain

$$m_\mathrm{i}\frac{\mathrm{d}^2}{d\mathrm{t}^2}\boldsymbol{r}_\mathrm{i} = \boldsymbol{F}_\mathrm{i}^{\mathrm{ext}} + \boldsymbol{F}_\mathrm{i}^{\mathrm{int}}. \tag{2.11}$$

Written in (rectangular) Cartesian coordinates (x, y, z) (2.11) is a set of three equations,

---

[3] Johann Friedrich Pfaff (1765–1825) was one of Germany's most eminent mathematicians during the 19th century. He is noted for his work on partial differential equations of the first order, which became part of the theory of differential forms. He was also Carl Friedrich Gauss's formal research supervisor.

[4] We realize, of course, that what are considered as contact forces are the result of electromagnetic forces between the atoms making up the particles and those making up the constraining surfaces.

$$m_i \ddot{x}_i = F_{xi}^{ext} + F_{xi}^{int}, \tag{2.12}$$

$$m_i \ddot{y}_i = F_{yi}^{ext} + F_{yi}^{int}, \tag{2.13}$$

and

$$m_i \ddot{z}_i = F_{zi}^{ext} + F_{zi}^{int}. \tag{2.14}$$

We may sum (2.12)–(2.14) over all particles i to obtain the general form of the application of Newton's Laws to our system. In the summation we use Newton's Third Law , which applies to the internal forces of interaction between all pairs of particles. The summation over all of the internal forces is then zero regardless of whether the forces are from interactions between the particles, making up what we see as a single material body, or include interactions between the particles making up separate bodies. Equations (2.12)–(2.14) then become

$$\sum_i m_i \ddot{x}_i = \sum_i F_{xi}, \tag{2.15}$$

$$\sum_i m_i \ddot{y}_i = \sum_i F_{yi}, \tag{2.16}$$

and

$$\sum_i m_i \ddot{z}_i = \sum_i F_{zi}, \tag{2.17}$$

where we drop the superscript *ext* as superfluous. The set of second order differential equations (2.15)–(2.17) constitute a general description of the mechanical behavior of a group of material bodies, provided we require that the particles obey Newton's Laws.

2.3.2 Generalized Coordinates

In most applications we do not need the full set of Cartesian coordinates because of the constraints on the system we are studying. As an example in Fig. 2.4 we consider

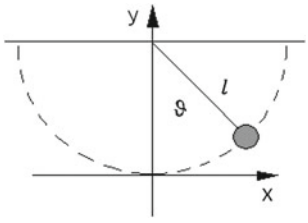


Fig. 2.4 Simple Pendulum with generalized coordinates ℓ and ϑ

a simple pendulum consisting of a mass suspended by a light rod. The motion of the mass is constrained to remain on a circle of radius ℓ. The Cartesian coordinates for this system may be written as

$$x = \ell \sin \vartheta \tag{2.18}$$

and

$$y = \ell \left(1 - \cos \vartheta\right). \tag{2.19}$$

We then have a complete description of the motion in terms of a single variable ϑ. We refer to this single variable, which is completely adequate for the description of the system originally formulated in the two rectangular Cartesian coordinates (x, y), as a *generalized coordinate*. The reduction in coordinates resulted from the single constraint $\ell = \sqrt{x^2 + y^2}$.

Physical constraints on systems generally impose geometric constraints on our representations of those systems. A curve in a plane imposes a relationship between two Cartesian coordinates, which are then no longer independent. We may then reduce the original pair of planar coordinates to a single coordinate. In our example the single independent coordinate was an angle ϑ expressing the relationship between x and y. The situations we encounter in applications will, however, seldom be so simple. For the general case we will require the method of undetermined multipliers introduced by Lagrange.

We may, however, claim that constraints will always impose geometrical relationships between the rectangular coordinates for the particles in our system $(x_\mathrm{i}, y_\mathrm{i}, z_\mathrm{i})$ and the generalized coordinates q_i of the form

$$x_\mathrm{i} = x_\mathrm{i}\left(q, t\right), \tag{2.20}$$

$$y_\mathrm{i} = y_\mathrm{i}\left(q, t\right), \tag{2.21}$$

and

$$z_\mathrm{i} = z_\mathrm{i}\left(q, t\right), \tag{2.22}$$

where we have introduced the shorthand $q = \{q_\mathrm{i}\}$ for the set of generalized coordinates. We must only accept that we may not be capable of writing these relationships in closed form.

From (2.20)–(2.22), Pfaff's[5] differential forms, or Pfaffians,[6] for the rectangular Cartesian coordinates $(x_\mathrm{i}, y_\mathrm{i} z_\mathrm{i})$ are

$$\begin{aligned} \mathrm{d}x_\mathrm{i} &= \sum_\mathrm{k} \frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} \mathrm{d}q_\mathrm{k} + \frac{\partial x_\mathrm{i}}{\partial t} \mathrm{d}t \\ \mathrm{d}y_\mathrm{i} &= \sum_\mathrm{k} \frac{\partial y_\mathrm{i}}{\partial q_\mathrm{k}} \mathrm{d}q_\mathrm{k} + \frac{\partial y_\mathrm{i}}{\partial t} \mathrm{d}t \\ \mathrm{d}z_\mathrm{i} &= \sum_\mathrm{k} \frac{\partial z_\mathrm{i}}{\partial q_\mathrm{k}} \mathrm{d}q_\mathrm{k} + \frac{\partial z_\mathrm{i}}{\partial t} \mathrm{d}t . \end{aligned} \tag{2.23}$$

2.3.3 Virtual Displacement

We introduced the concept of *virtual displacement* of the i^{th} particle $\delta \boldsymbol{r}_\mathrm{i}$ in the preceding chapter. This virtual displacement is consistent with the constraints acting on the system and could be an infinitesimal element of the actual path followed by the system. But we hold the time constant, i.e. $\mathrm{d}t = 0$. The displacement $\delta \boldsymbol{r}_\mathrm{i}$ is then *virtual* in that it does not actually take place. The actual variation would require $\mathrm{d}t \neq 0$. From the Pfaffians (2.23), with $\mathrm{d}t = 0$, we have

$$\boxed{\delta \boldsymbol{r}_\mathrm{i} = \sum_\mathrm{k} \left[(\partial x_\mathrm{i}/\partial q_\mathrm{k})\, \hat{e}_\mathrm{x} + (\partial y_\mathrm{i}/\partial q_\mathrm{k})\, \hat{e}_\mathrm{y} + (\partial z_\mathrm{i}/\partial q_\mathrm{k})\, \hat{e}_\mathrm{z} \right] \delta q_\mathrm{k}} \tag{2.24}$$

for the virtual displacement.

2.3.4 D'Alembert's Principle

Because (2.11) must hold for every particle at each step along the path followed by the system,

$$m_\mathrm{i} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{r}_\mathrm{i} - \boldsymbol{F}_\mathrm{i} = \boldsymbol{0}$$

[5] Johann Friedrich Pfaff (1765–1825) was one of Germany's most eminent mathematicians during the 19th century. He is noted for his work on partial differential equations of the first order, which became part of the theory of differential forms. He was also Carl Friedrich Gauss's formal research supervisor.

[6] Pfaff's differential form for the function $\Psi\,(\xi_1, \ldots, \xi_n)$ is defined as

$$d\Psi = \sum_j^n \left(\frac{\partial \Psi}{\partial \xi_j} \right) d\xi_j \, .$$

at each point and during the next infinitesimal part of the path followed by the particle i. The virtual displacement $\delta \boldsymbol{r}_\mathrm{i}$ is a *possible* next infinitesimal part of the path, consistent with the constraints on the system. If $\delta \boldsymbol{r}_\mathrm{i}$ were the actual path followed we would have

$$\sum_\mathrm{i} \left[m_\mathrm{i} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{r}_\mathrm{i} - \vec{F}_\mathrm{i} \right] \cdot \delta \boldsymbol{r}_\mathrm{i} = 0. \tag{2.25}$$

However, if $\delta \boldsymbol{r}_\mathrm{i}$ is a virtual displacement the product in (2.25) may not be identically zero at all points along $\delta \boldsymbol{r}_\mathrm{i}$. But the difference between the product in (2.25) and zero would only result in higher order terms in δ's. Therefore, to first order in δ's (2.25), which is *d'Alembert's Principle*, is valid for the virtual displacement $\delta \boldsymbol{r}_\mathrm{i}$. [see [65], pp. 88–110]

We note that d'Alembert's Principle is expressed as a scalar equation involving what is termed *virtual work* $\vec{F}_\mathrm{i} \cdot \delta \boldsymbol{r}_\mathrm{i}$, which is the work that would be done on the mass m_i in the virtual displacement $\delta \boldsymbol{r}_\mathrm{i}$. This virtual work is equal to a corresponding virtual change in the kinetic energy of the mass m_i, which is $m_\mathrm{i} \mathrm{d}^2 \boldsymbol{r}_\mathrm{i} / \mathrm{d}t^2 \cdot \delta \boldsymbol{r}_\mathrm{i}$. This formulation includes, in principle, work by dissipative forces[7] as well. Work and energy then replace forces in a formulation of mechanics based on d'Alembert's Principle.

2.3.5 Euler–Lagrange Equations

In this section we provide the details of the derivation of the Euler–Lagrange Equations from d'Alembert's Principle. We will introduce generalized coordinates q into the virtual displacements, which we wrote above in terms of Cartesian coordinates. Beyond the transition from Cartesian coordinates to the generalized coordinates, this section is primarily a mathematical discussion involving some creative use of partial derivative relations. The passage from Newton's Laws to the Euler–Lagrange Equations without a variational principle is, however, a necessary part of the development.

Using (2.24) d'Alembert's Principle, Eq. (2.25), becomes

$$\begin{aligned} 0 = \sum_{\mathrm{i,k}} \Bigg[m_\mathrm{i} \left(\ddot{x}_\mathrm{i} \frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} + \ddot{y}_\mathrm{i} \frac{\partial y_\mathrm{i}}{\partial q_\mathrm{k}} + \ddot{z}_\mathrm{i} \frac{\partial z_\mathrm{i}}{\partial q_\mathrm{k}} \right) \\ - \left(F_{\mathrm{x}i} \frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} + F_{\mathrm{y}i} \frac{\partial y_\mathrm{i}}{\partial q_\mathrm{k}} + F_{\mathrm{z}i} \frac{\partial z_\mathrm{i}}{\partial q_\mathrm{k}} \right) \Bigg] \delta q_\mathrm{k}. \end{aligned} \tag{2.26}$$

[7] Dissipative forces are frictional forces. These ultimately result from contact forces between moving bodies and surfaces or moving bodies and moving fluids, all of which are molecular and electromagnetic in nature. If we claim such detailed knowledge these forces are conservative. We may later insert these in modeled form, if they arise.

Equation (2.26) is of the form

$$0 = \sum_{\mathrm{k}} \alpha_{\mathrm{k}} \delta q_{\mathrm{k}}. \tag{2.27}$$

Since the generalized coordinates q_{k} are independent of one another, the δq_{k} are arbitrary. Therefore (2.27) can only be valid if each α_{k} is independently zero. That is

$$\begin{aligned} &\sum_{\mathrm{i}} m_{\mathrm{i}} \left(\ddot{x}_{\mathrm{i}} \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + \ddot{y}_{\mathrm{i}} \frac{\partial y_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + \ddot{z}_{\mathrm{i}} \frac{\partial z_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \right) \\ &= \sum_{\mathrm{i}} \left(F_{\mathrm{xi}} \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + F_{\mathrm{yi}} \frac{\partial y_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + F_{\mathrm{zi}} \frac{\partial z_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \right) \end{aligned} \tag{2.28}$$

for each component k. Equation (2.28) must then be valid if the system obeys Newton's Laws.

Because the Cartesian coordinates are functions of the generalized coordinates and the time (see (2.20)–(2.22)) the time derivative of the coordinate x_{i} is

$$\dot{x}_{\mathrm{i}} = \frac{\mathrm{d}x_{\mathrm{i}}}{\mathrm{d}t} = \sum_{\mathrm{k}} \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \dot{q}_{\mathrm{k}} + \frac{\partial x_{\mathrm{i}}}{\partial t}. \tag{2.29}$$

If we now take the partial derivative of (2.29) with respect to $\dot{q}_{\mathrm{k}}$ we obtain

$$\frac{\partial \dot{x}_{\mathrm{i}}}{\partial \dot{q}_{\mathrm{k}}} = \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \tag{2.30}$$

Equation (2.30) is often called *cancellation of the dots* because it appears as though we have simply cancelled the dots (time derivatives) in $\partial \dot{x}_{\mathrm{i}} / \partial \dot{q}_{\mathrm{k}}$ to obtain $\partial x_{\mathrm{i}} / \partial q_{\mathrm{k}}$. Mathematically (2.30) is a consequence of the fact that the Cartesian coordinate x_{i} depends only on q and the time t and is independent of the velocities $\dot{q}$.

With (2.30) we can write the terms appearing on the left hand side of (2.28) as

$$\ddot{x}_{\mathrm{i}} \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} = \ddot{x}_{\mathrm{i}} \frac{\partial \dot{x}_{\mathrm{i}}}{\partial \dot{q}_{\mathrm{k}}}. \tag{2.31}$$

Now

$$\ddot{x}_{\mathrm{i}} \frac{\partial \dot{x}_{\mathrm{i}}}{\partial \dot{q}_{\mathrm{k}}} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\dot{x}_{\mathrm{i}} \frac{\partial \dot{x}_{\mathrm{i}}}{\partial \dot{q}_{\mathrm{k}}} \right) - \dot{x}_{\mathrm{i}} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \right), \tag{2.32}$$

using (2.30) in the last term on the right hand side of (2.32). Since $\partial x_{\mathrm{i}} / \partial q_{\mathrm{k}}$ depends on (q, t) as does x_{i}, the time derivative of $\partial x_{\mathrm{i}} / \partial q_{\mathrm{k}}$ has the same form as (2.29). That is

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \right) = \sum_{\mathrm{j}} \frac{\partial^2 x_{\mathrm{i}}}{\partial q_{\mathrm{j}} \partial q_{\mathrm{k}}} \dot{q}_{\mathrm{j}} + \frac{\partial^2 x_{\mathrm{i}}}{\partial t \partial q_{\mathrm{k}}}. \tag{2.33}$$

Also from (2.29) we have

$$\frac{\partial \dot{x}_\mathrm{i}}{\partial q_\mathrm{k}} = \sum_\mathrm{j} \frac{\partial^2 x_\mathrm{i}}{\partial q_\mathrm{j} \partial q_\mathrm{k}} \dot{q}_\mathrm{j} + \frac{\partial^2 x_\mathrm{i}}{\partial t \partial q_\mathrm{k}}, \tag{2.34}$$

since the order of partial differentiation is immaterial.

Therefore from (2.33) and (2.34) we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}}\right) = \frac{\partial \dot{x}_\mathrm{i}}{\partial q_\mathrm{k}}. \tag{2.35}$$

With (2.35) Eq. (2.32) becomes

$$\begin{aligned} \ddot{x}_\mathrm{i} \frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}_\mathrm{i} \frac{\partial \dot{x}_\mathrm{i}}{\partial \dot{q}_\mathrm{k}}\right) - \dot{x}_\mathrm{i} \frac{\partial \dot{x}_\mathrm{i}}{\partial q_\mathrm{k}} \\ &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_\mathrm{k}} - \frac{\partial}{\partial q_\mathrm{k}}\right]\left(\frac{1}{2}\dot{x}_\mathrm{i}^2\right). \end{aligned} \tag{2.36}$$

Then using (2.36) the left hand side of (2.28) becomes

$$\begin{aligned} &\sum_\mathrm{i} m_\mathrm{i}\left(\ddot{x}_\mathrm{i}\frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} + \ddot{y}_\mathrm{i}\frac{\partial y_\mathrm{i}}{\partial q_\mathrm{k}} + \ddot{z}_\mathrm{i}\frac{\partial z_\mathrm{i}}{\partial q_\mathrm{k}}\right) \\ &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_\mathrm{k}} - \frac{\partial}{\partial q_\mathrm{k}}\right]\sum_\mathrm{i}\frac{1}{2} m_\mathrm{i}\left(\dot{x}_\mathrm{i}^2 + \dot{y}_\mathrm{i}^2 + \dot{z}_\mathrm{i}^2\right) \end{aligned} \tag{2.37}$$

We recognize the term $\sum_\mathrm{i} (1/2)\, m_\mathrm{i}\, \left(\dot{x}_\mathrm{i}^2 + \dot{y}_\mathrm{i}^2 + \dot{z}_\mathrm{i}^2\right)$ as the kinetic energy , which, in keeping with the notation of Lagrange and Hamilton,[8] we shall designate as

$$T = \sum_\mathrm{i} \frac{1}{2} m_\mathrm{i}\left(\dot{x}_\mathrm{i}^2 + \dot{y}_\mathrm{i}^2 + \dot{z}_\mathrm{i}^2\right). \tag{2.38}$$

Then (2.28), which is the requirement that the system obeys Newton's Laws, is

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_\mathrm{k}} - \frac{\partial}{\partial q_\mathrm{k}}\right] T = \sum_\mathrm{i}\left(F_{\mathrm{x}\mathrm{i}}\frac{\partial x_\mathrm{i}}{\partial q_\mathrm{k}} + F_{\mathrm{y}\mathrm{i}}\frac{\partial y_\mathrm{i}}{\partial q_\mathrm{k}} + F_{\mathrm{z}\mathrm{i}}\frac{\partial z_\mathrm{i}}{\partial q_\mathrm{k}}\right). \tag{2.39}$$

We now recall that the forces remaining are those arising from external fields. In 18th century notation these forces equal to the positive gradient of the force function U. In modern notation these forces are equal to the *negative* gradient of a scalar

[8] Denoting kinetic energy as T is standard modern notation. The fact that T is also used for thermodynamic temperature, and that the kinetic energy of an ideal gas is proportional to thermodynamic temperature, is incidental.

potential V, which is a function only of spatial coordinates. That is

$$F_{\mathrm{xi}} = -\frac{\partial V}{\partial x_{\mathrm{i}}},\ F_{\mathrm{yi}} = -\frac{\partial V}{\partial x_{\mathrm{i}}},\ F_{\mathrm{zi}} = -\frac{\partial V}{\partial z_{\mathrm{i}}}.$$

Therefore, the right hand side of (2.39) is

$$\sum_{\mathrm{i}} \left(\frac{\partial V}{\partial x_{\mathrm{i}}} \frac{\partial x_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + \frac{\partial V}{\partial y_{\mathrm{i}}} \frac{\partial y_{\mathrm{i}}}{\partial q_{\mathrm{k}}} + \frac{\partial V}{\partial z_{\mathrm{i}}} \frac{\partial z_{\mathrm{i}}}{\partial q_{\mathrm{k}}} \right) = \frac{\partial V}{\partial q_{\mathrm{k}}}, \tag{2.40}$$

using the chain rule. With (2.40) Eq. (2.39) becomes

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{q}_{\mathrm{k}}} - \frac{\partial}{\partial q_{\mathrm{k}}} \right] T = -\frac{\partial V}{\partial q_{\mathrm{k}}}. \tag{2.41}$$

Since the potential V depends only on the coordinates and not on the velocities (2.41) may be written as

$$\left[\frac{\partial}{\partial q_{\mathrm{k}}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{q}_{\mathrm{k}}} \right] (T - V) = 0. \tag{2.42}$$

There is an equation of the form (2.42) for each of the generalized coordinates q_{k}.

The Eqs. (2.42) are the *Euler–Lagrange Equations*. The combination $T - V$ is called the *Lagrangian*

$$L = T - V. \tag{2.43}$$

The Lagrangian is a scalar function of the generalized coordinates q, the time derivatives of the generalized coordinates $\dot{q}$, and possibly the time t. To obtain the Lagrangian we only need the kinetic energies of the interacting bodies and the potential energies of the external fields.

With (2.43) Eq. (2.42) becomes

$$\boxed{\partial L/\partial q_{\mathrm{k}} - \mathrm{d}(\partial L/\partial \dot{q}_{\mathrm{k}})/\mathrm{d}t = 0.} \tag{2.44}$$

The set of Eq. (2.44) is the final form of the Euler–Lagrange Equations.[9]

2.4 Variational Calculus

Our derivation of the Euler–Lagrange Equations in the preceding section was based strictly on the differential calculus and an understanding of the elements of linear algebra (see (2.27)). We had no need of any variational principle, even though we

[9] Some authors prefer to write the Euler–Lagrange equations as $\mathrm{d}(\partial L/\partial \dot{q}_{\mathrm{k}})/\mathrm{d}t - \partial L/\partial q_{\mathrm{k}} = 0$. Our choice is based on the fact that this is the natural order resulting from the variational principle.

know from the history of Analytical Mechanics that the principle of least action was central to the thinking, particularly at the Berlin Academy (see Sects. (1.8.1)–(1.8.3)). The formulation of the science in terms of a variational principle, and an introduction of the rudiments of variational calculus, are, therefore, vitally important to our understanding of Analytical Mechanics. Beyond the philosophical importance, a formulation in terms of a variational principle is also absolutely necessary for the incorporation of general constraints into mechanical problems.

In this section we will introduce the basis of variational calculus and show that the Euler–Lagrange equations result from a variational principle.

2.4.1 Functionals

A *functional* defines an operation on a class of functions $\{y(x)\}$ that returns a real number for each function $y(x)$ [cf. [31]]. For example such a functional may be the definite integral of a quantity $F\left[x, y(x), y'(x)\right]$ dependent on the function $y(x)$, its derivative $y'(x)$ and the independent variable x over the interval $[a, b]$. In this case the functional $J[y]$ is the number

$$J[y] = \int_a^b \mathrm{d}x\, F\left[x, y(x), y'(x)\right], \tag{2.45}$$

which is dependent on the function $y(x)$ chosen from the class $\{y(x)\}$ and the values a and b chosen as the limits of integration.

Example 2.4.1 As a specific example we consider the area of the surface of rotation $A_S[y]$ for a class of functions $\{y(x)\}$. We choose our class of functions to be single-valued, have continuous first derivatives, and pass through the two points (x_0, y_0) and (x_1, y_1). The differential area of the surface formed by the rotation of the curve $y(x)$ about the $x-$axis is shown in Fig. 2.5. The differential area of the surface of revolution $\mathrm{d}A_S\left[x, y(x), y'(x)\right]$ between the points x and $x + \mathrm{d}x$ on the $x-$axis is equal to the product of the circumference of the circular cross section of the surface $2\pi y(x)$ and the distance along the curve $\mathrm{d}s$ resulting from the differential distance $x \to x + \mathrm{d}x$ along the $x-$axis. above. That is

$$\begin{aligned}\mathrm{d}A_S\left[x, y(x), y'(x)\right] &= 2\pi y(x)\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} \\ &= \mathrm{d}x\,[2\pi y(x)]\sqrt{1 + (y'(x))^2}\end{aligned}$$

The total area of revolution is then

$$A_S[y] = \int_{x_2}^{x_1} \mathrm{d}x \left[2\pi y(x)\sqrt{1 + (y'(x))^2}\right]. \tag{2.46}$$

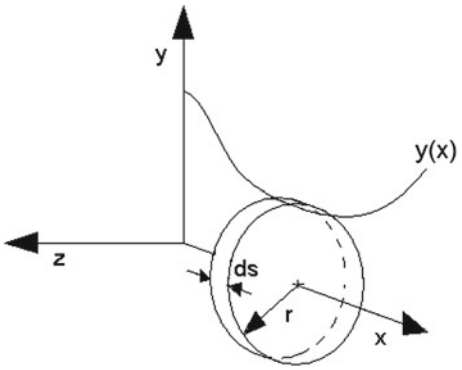


Fig. 2.5 Differential area of the surface of revolution formed by the curve $y = y\,(x)$

In this example the function $F\left[x,\, y\,(x)\,,\, y'\,(x)\right]$ is

$$F\left[x,\, y\,(x)\,,\, y'\,(x)\right] = 2\pi y\,(x)\sqrt{1 + (y'\,(x))^2}$$

We may now ask for the function $y\,(x)$ that results in a maximum or a minimum of the area of rotation in (2.46).

2.4.2 *Extrema of Functionals*

For a function $F\left(x,\, y,\, y'\right)$ there may be a specific function $y\,(x) = \eta\,(x)$ in the class $\{y\,(x)\}$ that results in an extreme value of the functional $J\,[y]$ in (2.45). We then consider all possible functions in the class $\{y\,(x)\}$ that differ only very slightly (infinitesimally) from $\eta\,(x)$. That is we write the function $y\,(x)$, which differs only slightly from $\eta\,(x)$, as

$$y\,(x) = \eta\,(x) + h\,(x)\,, \tag{2.47}$$

where $h\,(x)$ is an infinitesimal function of the variable x contained within the class of functions $\{y\,(x)\}$. We then define $\Delta J\,[h]$ as the functional

$$\Delta J\,[h] = J\,[y] - J\,[\eta]\,. \tag{2.48}$$

This functional is an infinitesimal real number for each infinitesimal function $h\,(x)$ in $\{y\,(x)\}$. The extremum $J\,[\eta]$ is a *minimum* if each $h\,(x)$ increases the value of $J\,[y]$ and

$$\Delta J\,[h] = J\,[y] - J\,[\eta] \geq 0, \tag{2.49}$$

and is a *maximum* if each $h(x)$ decreases the value of $J[y]$ and

$$\Delta J[h] = J[y] - J[\eta] \leq 0. \tag{2.50}$$

Because $h(x)$ is infinitesimal at every point on the interval in x being considered, the function $F\left(x, y, y'\right)$ in (2.45) may be written as a generalized Taylor series of the form

$$F\left(x, y, y'\right) = F\left(x, \eta, \eta'\right) + h \left.\frac{\partial F}{\partial y}\right|_{y=\eta} + h' \left.\frac{\partial F}{\partial y'}\right|_{y=\eta} + \cdots, \tag{2.51}$$

at each point x on the interval of interest (i.e. $x \in [a, b]$). In (2.51) $\partial F/\partial y|_{y=\eta}$ and $\partial F/\partial y'\big|_{y=\eta}$ are the partial derivatives $\partial F/\partial y$ and $\partial F/\partial y'$, which are functions of y and y', evaluated in the limit $h(x) \to 0$ and $h'(x) \to 0$. The general functions $y(x)$ and $y'(x)$ are then replaced by the functions $\eta(x)$ and $\eta'(x)$ in the partial derivatives $\partial F/\partial y$ and $\partial F/\partial y'$. Using (2.51) in $\Delta J[h]$,

$$\begin{aligned}
\Delta J[h] &= J[y] - J[\eta] \\
&= \int_a^b \mathrm{d}x\, F\left(x, \eta + h, \eta' + h'\right) - \int_a^b \mathrm{d}x\, F\left(x, \eta, \eta'\right) \\
&= \int_a^b \mathrm{d}x \left[h(x) \left.\frac{\partial F}{\partial y}\right|_{y=\eta} + h'(x) \left.\frac{\partial F}{\partial y'}\right|_{y=\eta} + \cdots \right].
\end{aligned} \tag{2.52}$$

With the definition

Definition 2.4.1 That contribution to $\Delta J[h]$ which is of order n in $h(x)$, $h'(x)$, or products of $h(x)$ and $h'(x)$ is defined as $\delta^{(n)} J[h]$.

Equation (2.52) may be written as

$$\Delta J[h] = \delta J[h] + \delta^{(2)} J[h] + \cdots. \tag{2.53}$$

The necessary condition for an extremum of the functional is that the first order variation vanishes. That is

$$\delta J[h] = 0 \text{ at an extremum} \tag{2.54}$$

This condition alone is called the *weak extremum* because it does not specify whether the extremum is a maximum or a minimum. Whether the extremum is a maximum or a minimum must be determined by investigating the algebraic sign of the variation $\delta^{(2)} J[h]$.

This is the $\delta-$method developed by Lagrange (see Sect. 1.8.3), provided we also require the the values of the class of functions $\{y(x)\}$ are fixed at the end points. We

must then require that the infinitesimal functions $h\,(x)$ vanish at the end points. This restriction was included by Lagrange.

2.4.3 Euler Problem

The Euler problem is to obtain the conditions on $y\,(x)$ such that the functional (2.45) has a weak extremum, i.e. that $\delta J\,[y] = 0$. [[31], pp. 14, 15] We shall solve this problem using the $\delta-$method of Lagrange, however, since we realize that there are mathematical difficulties at the end points in Euler's method (see Sect. 1.8.2).

We require that the function $F\left(x, y, y'\right)$ has continuous first and second (partial) derivatives with respect to all its arguments. And we require that the function $y\,(x)$ belongs to a class of functions that have continuous first derivatives for $a \leq x \leq b$ (i.e. for $x \in [a, b]$), which satisfies the (fixed) boundary conditions

$$y\,(a) = A, \qquad y\,(b) = B.$$

The functions $y\,(x)$ and $\eta\,(x)$ in (2.47) are then equal at the end points and, therefore, the function $h\,(x)$ must vanish at the end points.

From (2.52) a weak extremum requires that

$$\delta J\,[h] = 0 = \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x \left[h\,(x)\, F_{\mathrm{y}}\big|_{\mathrm{y}=\eta} + h'\,(x)\, F_{\mathrm{y}'}\big|_{\mathrm{y}=\eta} \right], \tag{2.55}$$

where we have introduced the notation $F_{\mathrm{y}} = \partial F/\partial y$ and $F_{\mathrm{y}'} = \partial F/\partial y'$. Once the form of the function $y\,(x)$ is specified, $F\left[x, y\,(x)\,, y'\,(x)\right]$ and its derivatives become functions of x alone. For the extremum specified by (2.55) we have $y\,(x) = \eta\,(x)$. Since $(\mathrm{d}/\mathrm{d}x)\, h\,(x)\, F_{\mathrm{y}'}\,(x) = h'\,(x)\, F_{\mathrm{y}'}\,(x) + h\,(x)\mathrm{d}F_{\mathrm{y}'}\,(x)\,/\mathrm{d}x$ the integral of the second term in (2.55) is

$$\begin{aligned} &\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x h'\,(x)\, F_{\mathrm{y}'}\,(x) \\ &= h\,(x)\, F_{\mathrm{y}'}\,(x)\big|_{\mathrm{x=a}}^{\mathrm{x=b}} - \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x h\,(x)\, \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y}'}\,(x) \end{aligned} \tag{2.56}$$

Because $h\,(x)$ vanishes at the end points $x = a, b$, (2.56) becomes

$$\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x h'\,(x)\, F_{\mathrm{y}'}\,(x) = -\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x h\,(x)\, \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y}'}\,(x) \tag{2.57}$$

and (2.55) is

$$0 = \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x \left[F_{\mathrm{y}} - \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y}'} \right] h\,(x) \tag{2.58}$$

for any arbitrary function $h(x)$ in the class of functions $\{y(x)\}$, which satisfies the boundary conditions $h(a) = h(b) = 0$.

Because $h(x)$ is arbitrary, (2.58) holds if and only if

$$F_{\mathrm{y}} - \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y}'} = 0.$$

We then have the condition for the weak extremum, which we can state as a theorem. [[31], p. 15]

Theorem 2.4.1 *Let $J[y]$ be a functional of the form*

$$J[y] = \int_a^b dx F\left(x, y, y'\right),$$

defined on the set of functions $y(x)$, which have continuous first derivatives in $[a, b]$ and satisfy the boundary conditions $y(a) = A$ and $y(b) = B$. Then a necessary condition for $J[y]$ to have an extremum for a given function $y(x)$ is that $y(x)$ satisfies Euler's Equation

$$F_y - \frac{d}{dx} F_{y'} = 0. \tag{2.59}$$

Proof The proof is given in the preceding development.

We define the $\delta-$variation as Lagrange did.

Definition 2.4.2 Let $J[y]$ be a functional of the form

$$J[y] = \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x F\left(x, y, y'\right),$$

defined on the set of functions $y(x)$, which have continuous first derivatives in $[a, b]$ and satisfy the boundary conditions $y(a) = A$, $y(b) = B$. Then the $\delta-$**variation** of the functional $J[y]$, indicated by $\delta J[y]$, is that for which the variations $h(x)$ in the function $y(x)$ vanish at the end points a and b.

Example 2.4.2 As an example we consider the function

$$F\left(x, y, y'\right) = \frac{1}{2}\left(y'\right)^2 + 4xy, \tag{2.60}$$

and we choose the $x-$ interval to be $[a, b] = [0, 1]$ with the boundary conditions $y(0) = 0$ and $y(1) = 1$. Then (2.59) becomes

$$F_{\mathrm{y}} - \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y}'} = 4x - \frac{\mathrm{d}^2}{\mathrm{d}x^2} y = 0. \tag{2.61}$$

The solution to (2.61) is

$$y\,(x) = \frac{2}{3}x^3 + \frac{1}{3}x. \tag{2.62}$$

And the value of the functional at the weak extremum is

$$\begin{aligned} & J\,[y] \\ &= \int_0^1 \mathrm{d}x \left[\frac{1}{2}\left(y'\right)^2 + 4xy\right] \\ &= \int_0^1 \mathrm{d}x \left[\frac{1}{2}\left(2x^2 + \frac{1}{3}\right)^2 + 4x\left(\frac{2}{3}x^3 + \frac{1}{3}x\right)\right] \\ &= 1.\,655\,6\ldots \end{aligned} \tag{2.63}$$

Since we have only the condition for a weak extremum, we do not know whether (2.62) results in a maximum or a minimum of the functional.

This Theorem 2.4.1 can easily be generalized to a functional defined on a set of m functions. We define

$$J\,[y_1 \cdots y_{\mathrm{m}}] = \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x\, F\left(x, y_1 \cdots y_{\mathrm{m}}, y_1' \cdots y_{\mathrm{m}}'\right)$$

with

$$y_{\mathrm{k}}\,(a) = A_{\mathrm{k}}, \quad y_{\mathrm{k}}\,(b) = B_{\mathrm{k}},$$

fixed for each $k = 1, \ldots, m$. Following the same steps as before, we have for the first variation

$$\begin{aligned} \delta J &= \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x \sum_{\mathrm{k}=1}^{\mathrm{m}} \left[h_{\mathrm{k}}\,(x)\, F_{\mathrm{y_k}}\big|_{y_{\mathrm{k}}=\eta_{\mathrm{k}}} + h_{\mathrm{k}}\,(x)\, F_{\mathrm{y_k'}}\big|_{y_{\mathrm{k}}=\eta_{\mathrm{k}}}\right] \\ &= \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{d}x \sum_{\mathrm{k}=1}^{\mathrm{m}} h_{\mathrm{k}} \left[F_{\mathrm{y_k}} - \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathrm{y'_k}}\right]. \end{aligned}$$

Because each of the functions $h_{\mathrm{k}}\,(x)$ is arbitrary and independent of the others, δJ vanishes if and only if each bracketed term [ ] vanishes independently. That is

$$\frac{\partial}{\partial y_{\mathrm{k}}} F - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial}{\partial y_{\mathrm{k}}'} F = 0 \tag{2.64}$$

for each k.

2.4.4 Hamilton's Principle

If we choose the time t as the independent variable, the generalized coordinates $q_{\mathrm{k}}(t)$ as the functions of interest, and F as the Lagrangian $L = T - V$, we recognize the set of equations (2.64) as the Euler–Lagrange Equations of Analytical Mechanics (2.44). Specifically we have arrived at a set of equations that are completely equivalent to Newton's Second Law from a variational principle. Therefore, if we define the functional

$$\boxed{S[q] = \int_{\mathrm{t_1}}^{\mathrm{t_2}} \mathrm{d}t\, L(q, \dot{q}, t),} \tag{2.65}$$

the condition for a weak extremum, using Lagrange's δ (see the Definition 2.4.2),

$$\delta S = \delta \int_{\mathrm{t_1}}^{\mathrm{t_2}} \mathrm{d}t\, L(q, \dot{q}, t) = 0, \tag{2.66}$$

is that

$$\left[\frac{\partial}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}}\right] L(q, \dot{q}, t) = 0 \tag{2.67}$$

for a single generalized coordinate $q(t)$.

If we consider a set of generalized coordinates $q = \{q_{\mathrm{k}}\}$ the condition that first order variation of the functional

$$S[q] = S[\{q_{\mathrm{k}}\}] = \int_{t_1}^{t_2} \mathrm{d}t\, L(\{q_{\mathrm{k}}\}, \{\dot{q}_{\mathrm{k}}\}, t), \tag{2.68}$$

vanishes, i.e. that the functional $S[\{q_{\mathrm{k}}\}]$ has an extremum, is that the set of equations

$$\left[\frac{\partial}{\partial q_{\mathrm{k}}} - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_{\mathrm{k}}}\right] L(\{q_{\mathrm{k}}\}, \{\dot{q}_{\mathrm{k}}\}, t) = 0 \tag{2.69}$$

is satisfied.

We then have a general theorem of Analytical Mechanics, which is *Hamilton's Principle*.

Theorem 2.4.2 *Let $S[q]$ be the functional*

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

where

$$L(q, \dot{q}, t) = T - V$$

is the Lagrangian function defined on the set of generalized coordinates $q = \{q_k\}$, *which have continuous first time derivatives* $\dot{q} = \{\dot{q}_k\}$ *on the interval* $[t_1, t_2]$, *and fixed values at the end points* t_1 *and* t_2. *Then a necessary condition for* $S[q]$ *to have an extremum for a given set of generalized coordinates* $q = \{q_k\}$ *is that each generalized coordinate* $q_k(t)$ *satisfies the* Euler–Lagrange Equation

$$\frac{\partial}{\partial q_k} L - \frac{d}{dt} \frac{\partial}{\partial \dot{q}_k} L = 0. \tag{2.70}$$

The system then satisfies Newton's Laws.

We note that here the variation $h(t)$ vanishes at the end points and that this is, therefore, the variation $\delta S[q]$ (see the Definition 2.4.2).

Example 2.4.3 As an example we choose the simple pendulum, which we have drawn in Fig. 2.6. In terms of the generalized coordinate ϑ the kinetic energy is

$$T = \frac{1}{2} m\ell^2 \dot{\vartheta}^2$$

and the potential energy is

$$V = mg\ell\,(1 - \cos\vartheta)\,,$$

where the reference is $V = 0$ when $\vartheta = 0$. Then the Lagrangian is

$$L = T - V = \frac{1}{2} m\ell^2 \dot{\vartheta}^2 + mg\ell\,(\cos\vartheta - 1)$$

and

$$S[q] = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{1}{2} m\ell^2 \dot{\vartheta}^2 + mg\ell\,(\cos\vartheta - 1) \right].$$

The Euler–Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(m\ell^2 \dot{\vartheta} \right) + mg\ell \sin\vartheta = 0$$

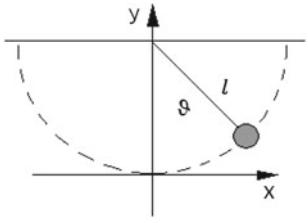


Fig. 2.6 Simple Pendulum with generalized coordinates ℓ and ϑ

or

$$\ddot{\vartheta} = -\left(\frac{g}{\ell}\right)\sin\vartheta.$$

This is a nonlinear equation for $\vartheta(t)$. We may linearize it to obtain the (familiar) equation for the simple pendulum with angular frequency $\omega = \sqrt{g/\ell}$.

2.5 Constraints

The reduction to generalized coordinates depends on the form of the constraints and may not be simple. Indeed in the general case it may be algebraically impossible to obtain a closed form Lagrangian in terms of generalized coordinates.

In this section we shall develop a systematic method for incorporating constraints of any kind into Hamilton's Principle. This will also provide the forces of constraint automatically in the course of the solution of the problem at hand.

We begin our discussion of constraints with a situation for which the constraint can only be formulated in differential terms: the rolling constraint.

2.5.1 Rolling

If a disk of radius R rolls upright and without slipping along a coordinate x the differential angle of rotation of the disk $\mathrm{d}\vartheta$, produces a displacement along the plane, $\mathrm{d}x$, given by

$$\mathrm{d}x = R\mathrm{d}\vartheta$$

or

$$0 = \mathrm{d}x - R\mathrm{d}\vartheta. \tag{2.71}$$

This differential relationship is all the description of rolling gives us. If the rolling is along a straight line we can integrate (2.71) to obtain

$$g_{\text{roll}}(x, \vartheta) = 0 = x - R\vartheta + \text{constant}.$$

Then (2.71) implies the existence of a functional relationship between x and ϑ of the form $g_{\text{roll}}(x, \vartheta) = 0$. And (2.71) is the Pfaffian differential of $g_{\text{roll}}(x, \vartheta)$, i.e.

$$dg_{\text{roll}} = 0 = \left(\frac{\partial g_{\text{roll}}}{\partial x}\right)dx + \left(\frac{\partial g_{\text{roll}}}{\partial \vartheta}\right)d\vartheta. \tag{2.72}$$

Comparing (2.72) with (2.71) we have

$$\left(\frac{\partial g_{\text{roll}}}{\partial x}\right) = 1 \text{ and } \left(\frac{\partial g_{\text{roll}}}{\partial \vartheta}\right) = -R. \tag{2.73}$$

But the integration cannot be performed for the general case of rolling on a surface because the path along which the object is rolling must be obtained from a solution of the equations of motion. And that solution involves the rolling condition as a constraint. We must, therefore, accept that the general rolling condition, or constraint, can only be formulated in differential form.

There is, however, a specific functional relationship between the path followed by the body on a surface and the angle through which the body has rotated, provided there is no slipping. This functional relationship

$$g_{\text{roll}} = g_{\text{roll}}(q),$$

where q is the set of generalized coordinates, only becomes known through solution of the set of dynamical (Euler–Lagrange) equations. Therefore, even though we are only able to write the rolling constraint initially in differential form, we realize that in general this differential form is an exact differential of a function, which we have designated here as g_{roll}.

In complete generality we may then write the rolling constraint as a Pfaffian in N generalized coordinates as

$$\mathrm{d}g_{\text{roll}} = 0 = \sum_{\mathrm{i}}^{\mathrm{N}} \left(\frac{\partial g_{\text{roll}}}{\partial q_{\mathrm{i}}}\right) \mathrm{d}q_{\mathrm{i}}\,. \tag{2.74}$$

And we will always be able to obtain an algebraic formulation of the terms $(\partial g_{\text{roll}}/\partial q_{\mathrm{i}})$, even though we cannot write an expression for g_{roll}.

A Pfaffian is the exact differential of a function. That is the Pfaffian of the function g_{roll} defines the function g_{roll} by specifying the rule for constructing it from infinitesimals. The fact that we cannot write g_{roll} in closed algebraic form is because the interaction among the coordinates in the dynamical system is generally complicated. But this fact in no way denies the existence of the function.[10]

[10]This situation is common in thermodynamics. There we have a Pfaffian for each of the potentials. But we cannot write down a potential for any but the simplest ideal substance because of the complexity of the interdependence of the thermodynamic properties for real substances.

2.5.2 Holonomic and Nonholonomic Constraints

What we have said here regarding the rolling constraint holds for any constraint for which we can only write a differential expression. Such constraints are called *Nonholonomic*.

Constraints for which we can write a general algebraic expression of the form

$$g_k(q) = 0 \tag{2.75}$$

are called *holonomic*. For N generalized coordinates the differential of a holonomic constraint results in the Pfaffian

$$\mathrm{d}g_k = 0 = \sum_i^N \left(\frac{\partial g_k}{\partial q_i}\right) \mathrm{d}q_i \tag{2.76}$$

The Eqs. (2.74) and (2.76) are both Pfaffians and identical in form. If we can show that only the differential of the constraint is of interest to us in our formulation then we can ignore the difference between holonomic and nonholonomic constraints.

In formulating any problem we always choose the coordinates that are the most logical. This results in a set of coordinates we may consider initially to be generalized coordinates. Each additional algebraic equation of constraint, whether holonomic or nonholonomic, reduces the number of necessary coordinates by one. This is true whether or not we can incorporate the constraint algebraically into the Lagrangian.

Some authors have placed various emphases on the distinction between holonomic and nonholonomic constraints. Whittaker's discussion is quite detailed. His example is, however, for the motion of a sphere on a horizontal surface. He points out that for the slipping motion of a sphere on the surface the constraint is holonomic, while for the non-slipping motion on the surface the rolling constraint is nonholonomic [[125], p. 34]. Louis Hand and Janet Finch also have a detailed description of constraints, subdividing the definitions depending on whether or not time is a variable. Their example of a nonholonomic constraint is based on the rolling of a disk on a rough surface [[41], p. 12, 36]. Cornelius Lanczos, who also uses rolling as an example of a nonholonomic constraint points out that the original definition comes from Heinrich Hertz [[65], p. 25].

The distinction between holonomic and nonholonomic constraints is real and interesting. However, in our formulation in terms of *Lagrange Undetermined Multipliers* here the difference has no immediate practical consequences.

2.5.3 *Lagrange Undetermined Multipliers*

We consider that there are n constraints. In principle there are then n algebraic functions $g_k(q) = 0$, as we pointed out at the end of Sect. 2.5.1, although we can actually write these only for the holonomic constraints.

For arbitrary functions of the time $\lambda_k(t)$ $(k = 1, \ldots, n)$ we, therefore, also have n equations

$$\lambda_k(t)\, g_k(q) = 0. \tag{2.77}$$

The integrals of the products (2.77) over any arbitrary time interval must also vanish. Then

$$\int_{t_1}^{t_2} dt\, \lambda_k(t)\, g_k(q) = 0. \tag{2.78}$$

Adding these integrals to the functional $S[q]$ in (2.65) does not change the value of $S[q]$. So Hamilton's Principle requires that the $\delta-$variation (see Definition 2.4.2 in Sect. 2.4.3 of)

$$S[q] = \int_{t_1}^{t_2} dt \left[L + \sum_{k=1}^{n} \lambda_k(t)\, g_k(q) \right] \tag{2.79}$$

must vanish. In performing the variation we must consider also variations in the (arbitrary) functions $\lambda_k(t)$. The result is

$$\begin{aligned} \delta S = 0 = \int_{t_1}^{t_2} dt \Bigg\{ \sum_{j}^{N} \delta q_j \left[\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right) \right. \\ \left. + \sum_{k=1}^{n} \lambda_k(t) \left(\frac{\partial g_k}{\partial q_j} \right) \right] + \sum_{k=1}^{n} \delta \lambda_k g_k(q) \Bigg\}. \end{aligned} \tag{2.80}$$

The integral in (2.80), over any arbitrary time interval, vanishes if and only if the integrand vanishes. That is

$$\begin{aligned} 0 = \sum_{j}^{N} \delta q_j \left[\frac{\partial L}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right) \right. \\ \left. + \sum_{k=1}^{n} \lambda_k(t) \left(\frac{\partial g_k}{\partial q_j} \right) \right] + \sum_{k=1}^{n} \delta \lambda_k g_k(q). \end{aligned} \tag{2.81}$$

Using (2.75) the last term in (2.81) vanishes regardless of variations in the $\lambda_{\mathrm{k}}(t)$. Therefore

$$0=\sum_{\mathrm{j}}^{\mathrm{N}}\delta q_{\mathrm{j}}\left[\frac{\partial L}{\partial q_{\mathrm{j}}}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_{\mathrm{j}}}\right)+\sum_{\mathrm{k}=1}^{\mathrm{n}}\lambda_{\mathrm{k}}(t)\left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}}\right)\right] \tag{2.82}$$

is the condition resulting from Hamilton's Principle .

If all of the coordinates in the set $q=\left\{q_{\mathrm{j}}\right\}_{\mathrm{j}=1}^{\mathrm{N}}$ were independent the variations δq_{j} would be independent and (2.82) would require that each of the square brackets $[\cdots]$ must be zero. But the set of coordinates $q=\left\{q_{\mathrm{j}}\right\}_{\mathrm{j}=1}^{\mathrm{N}}$ must also satisfy the n constraints (2.75). There are then only $N-n$ independent coordinates. So we must take another approach.

The other approach is to first judiciously choose all of the n arbitrary functions $\lambda_{\mathrm{k}}(t)$ so that for the first n expressions in the brackets $[\cdots]$ in (2.82) vanish. That is we choose the arbitrary functions $\lambda_{\mathrm{k}}(t)$ such that

$$\frac{\partial L}{\partial q_{\mathrm{j}}}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_{\mathrm{j}}}\right)+\sum_{\mathrm{k}=1}^{\mathrm{n}}\lambda_{\mathrm{k}}(t)\left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}}\right)\equiv 0 \quad \text{for } j=1\ldots N. \tag{2.83}$$

Solving for the $\lambda_{\mathrm{k}}(t)$ with $(k=1,\ldots,n)$ is, in principle, a straightforward problem in linear algebra. Because the Lagrangian $L=L(q,\dot{q},t)$ is specified, the first term involving L in (2.83) is a known function of $(q,\dot{q},t)$ for each j, which we shall call Φ_{j}. We also know the partial derivatives $\left(\partial g_{\mathrm{k}}/\partial q_{\mathrm{j}}\right)$ for all, including the nonholonomic, constraints as functions of $q=\left\{q_{\mathrm{j}}\right\}_{\mathrm{j}=1}^{\mathrm{N}}$. The Eq. (2.83) are then the n linear algebraic equations for functions $\lambda_{\mathrm{k}}(t)$ for $(k=1,\ldots,n)$,

$$\Phi_{\mathrm{j}}+\sum_{\mathrm{k}=1}^{\mathrm{n}}\lambda_{\mathrm{k}}(t)\left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}}\right)=0,$$

which are soluble in a straightforward fashion.[11] The first n equations (2.83) are then satisfied and the n multipliers $\lambda_{\mathrm{k}}(t)$ are known. There are then $N-n$ equations remaining in the set (2.83). And the set of Eq. (2.82) is reduced to

$$0=\sum_{\mathrm{j}=\mathrm{n}+1}^{\mathrm{N}}\delta q_{\mathrm{j}}\left[\frac{\partial L}{\partial q_{\mathrm{j}}}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_{\mathrm{j}}}\right)+\sum_{\mathrm{k}=1}^{\mathrm{n}}\lambda_{\mathrm{k}}(t)\left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}}\right)\right], \tag{2.84}$$

in which the $N-n$ remaining coordinates q_{j} for $j=n+1,n+2,\ldots,N$ and their variations δq_{j} are independent. Therefore (2.84) requires that

[11] The solution to this set of equations produces $\lambda_{\mathrm{k}}(q,\dot{q},t)$, with t appearing in the event that the Hamiltonian involves t explicitly. The final solution, which produces $q_{\mathrm{j}}=q_{\mathrm{j}}(t)$, results in $\lambda_{\mathrm{k}}(t)$.

$$\frac{\partial L}{\partial q_{\mathrm{j}}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_{\mathrm{j}}}\right) + \sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_{\mathrm{k}}(t)\left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}}\right) = 0 \text{ for } j = n+1 \ldots N. \tag{2.85}$$

We now note that the Eqs. (2.83) and (2.85) are identical. Therefore,

$$\boxed{\partial L/\partial q_{\mathrm{j}} - \mathrm{d}\left(\partial L/\partial \dot{q}_{\mathrm{j}}\right)/\mathrm{d}t + \sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_{\mathrm{k}}(t)\left(\partial g_{\mathrm{k}}/\partial q_{\mathrm{j}}\right) = 0 \quad \forall j.} \tag{2.86}$$

And we need pay no attention to the order of solution. This is what we shall always do in practice. Indeed we may often consider the $\lambda_{\mathrm{k}}(t)$ to be of no interest and never obtain them explicitly.

This is the method of *Lagrange Undetermined Multipliers*. We may use this method to incorporate any and all constraints on our system, whether those constraints are holonomic or nonholonomic, and regardless of the algebraic complexity of the constraints.

Example 2.5.1 We again consider the pendulum and incorporate the constraint that the length is constant using Lagrange Undetermined Multipliers. That is we first write the Lagrangian for the mass m moving freely in space then add the constraint that the radial distance r from the pivot point is fixed and equal to ℓ. We have drawn the situation in Fig. 2.7. The kinetic energy in cylindrical coordinates is

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

The potential energy, referenced to the pivot point, is

$$V = -mgr\ \cos\vartheta.$$

Then

$$L = T - V = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\vartheta}^2\right) + mgr\ \cos\ \vartheta.$$

The constraint is

$$g = r - \ell = 0.$$

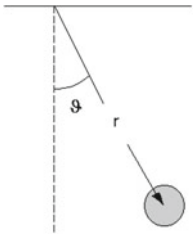


Fig. 2.7 A single generalized coordinate for the simple pendulum is ϑ. Including r and a Lagrange multiplier will allow the calculation of the force in the suspending rod

Then

$$\mathrm{d}g = \mathrm{d}r = 0.$$

and

$$\left(\frac{\partial g}{\partial r}\right) = 1$$

We have the Euler–Lagrange equations (with multiplier λ) for the two coordinates r and ϑ as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{r}\right) - mr\dot{\vartheta}^2 - mg\ \cos\ \vartheta + \lambda = 0$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(mr^2\dot{\vartheta}\right) + mgr\ \sin\ \vartheta = 0.$$

Using the constraint $\ell = r$ we have

$$\lambda = m\ell\dot{\vartheta}^2 + mg\ \cos\ \vartheta.$$

The second equation is, with $\ell = r$,

$$m\ell^2\ddot{\vartheta} + mg\ell\ \sin\ \vartheta = 0,$$

or

$$\ddot{\vartheta} = -\frac{g}{\ell}\ \sin\ \vartheta.$$

The use of Lagrange undetermined multipliers has resulted, then, in the same equation we obtained previously.

2.5.4 *Forces of Constraint*

We define the *canonical momentum* p_j conjugate to the generalized coordinate q_j as

$$p_\mathrm{j} \equiv \frac{\partial L}{\partial \dot{q}_\mathrm{j}}. \tag{2.87}$$

Newton's Second Law equates $\mathrm{d}p_\mathrm{j}/\mathrm{d}t$ to the force F_j applied in the j^{th} direction. With (2.87) we may then write the Eq. (2.86) as

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{\mathrm{j}} = -\frac{\partial V}{\partial q_{\mathrm{j}}} + \sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_{\mathrm{k}}(t) \left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}} \right)$$
$$= -\frac{\partial}{\partial q_{\mathrm{j}}} \left[V - \sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_{\mathrm{k}}(t)\, g_{\mathrm{k}} \right]. \tag{2.88}$$

That is

$$V_{\mathrm{eff}} = V - \sum_{\mathrm{k}}^{\mathrm{n}} \lambda_{\mathrm{k}}(t)\, g_{\mathrm{k}} \tag{2.89}$$

acts as an *effective potential* in which the system moves. The second part of this effective potential produces the reactive forces on the system due to the constraints. The reactive force in the j^{th} direction arising from the constraints is then

$$\boxed{F_{\mathrm{j}}^{\mathrm{react}} = \sum_{\mathrm{k}=1}^{\mathrm{n}} \lambda_{\mathrm{k}}(t) \left(\partial g_{\mathrm{k}} / \partial q_{\mathrm{j}} \right)} \tag{2.90}$$

and the force from the k^{th} constraint alone in the j^{th} direction is

$$F_{\mathrm{kj}}^{\mathrm{react}} = \lambda_{\mathrm{k}}(t) \left(\frac{\partial g_{\mathrm{k}}}{\partial q_{\mathrm{j}}} \right). \tag{2.91}$$

In order to find this reactive force we need to find the undetermined multiplier $\lambda_{\mathrm{k}}(t)$. In practice this usually constitutes an additional step in the solution, which is often unnecessary if we only seek the trajectory of the system. The engineer may, however, be very interested in knowing the forces of constraint on a system. These determine the strength of the structure which confines the system. The Lagrangian formulation provides these constraints in a systematic form.

2.6 Cyclic Coordinates

Let us assume that we have been able to obtain a Lagrangian into which all the constraints have been incorporated algebraically. For such a situation the Euler–Lagrange equations are (2.44). The structure of these equations allows us to make some general statements about the behavior of the system.

If the Lagrangian does not depend explicitly on a certain coordinate it is termed *cyclic* in that coordinate. The conjugate momentum corresponding to that coordinate

is then a constant of the motion, since from (2.87) and (2.44) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_\mathrm{r}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}p_\mathrm{r} = \frac{\partial L}{\partial q_\mathrm{r}} = 0. \tag{2.92}$$

This was first observed by the mathematician Amalie (Emmy) Nöther.[12]

The lack of dependence of the Lagrangian on a particular coordinate indicates a symmetry of the system with respect to that coordinate. For example if there is no dependence in the Lagrangian on the coordinate x the Lagrangian is unchanged by a translation in the $x-$direction and there can, therefore, be no force in the x direction. The canonical momentum p_x is then conserved. A symmetry about a particular axis means the angular momentum about that axis is constant. Later we shall see that a symmetry in the time means energy is conserved. The first two of these are intuitive. But the energy-time relationship is usually a surprise.

2.7 Summary

In this chapter we have laid the foundations of Analytical Mechanics. The end result was the Euler–Lagrange Equations.

We began by showing that Newton's Laws result in the Euler–Lagrange equations using d'Alembert's Principle, which is a virtual work principle based mathematically on a scalar equation. So our formulation was based on scalar energy terms, which promises a simplicity not present in a standard vector formulation.

We then turned to the more elegant formulation in terms of the variational calculus and a variational principle we identified as Hamilton's Principle. This formulation also provided us with the only truly general method to incorporate complicated constraints into mechanical problems: the method of Lagrange Undetermined Multipliers.

The only possible disadvantage in our formulation lies in the fact that the equations we have are second order differential equations in the time. Our approach would be simplified if we could find equivalent first order equations. This will be the subject of our next chapter.

[12] Amale (Emmy) Nöther (1882–1935) was a German mathematician whose specialty was algebra and ring structures. She was Extraordinary Professor at Göttingen from 1922–1933, then she accepted a professorship at Bryn Mawr College, which she held until her death in 1935.

2.8 Exercises

2.1. Cylindrical coordinates are $\{r, \vartheta, z\}$. The position vector from the origin is $\boldsymbol{r} = r\hat{e}_{\mathrm{r}} + z\hat{e}_{\mathrm{z}}$. Show that the velocity vector is

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \dot{r}\hat{e}_{\mathrm{r}} + r\dot{\vartheta}\hat{e}_{\vartheta} + \dot{z}\hat{e}_{\mathrm{z}}.$$

2.2. Spherical coordinates are $\{\rho, \vartheta, \phi\}$. The position vector is $\overrightarrow{r} = \rho\hat{e}_{\rho}$. Show that the velocity vector is

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \dot{\rho}\hat{e}_{\rho} + \rho\dot{\vartheta}\sin\phi\hat{e}_{\vartheta} + \rho\dot{\phi}\hat{e}_{\phi}.$$

Remark 2.8.1 These exercises should serve to provide a familiarity with the coordinates that we will use extensively. The solution may require some hand drawing and the consideration of small variations.

2.3. In this chapter we introduced Lagrange Undetermined Multipliers to add constraints to a variational principle. The reasoning should work just as well if we wish to find the extremum of an algebraic expression subject to a constraint. For example, if we wish to find the minimum distance from the origin to the straight line

$$y = 3x + 2$$

we seek the minimum distance from the origin to a point (x, y) in the plane and then introduce the constraint that the point lies on the line. The calculation will be easier if we minimize the square of the distance from the origin to the point (x, y), which by Pythagoras' Theorem is

$$f\,(x, y) = x^2 + y^2.$$

Carry out the calculation to find the point on the line.

Show that the shortest line between the origin and the straight line $y = 3x + 2$ is perpendicular to the line $y = 3x + 2$.

2.4. Show that D'Alembert's Principle results in conservation of mechanical energy $\mathrm{d}(T + V) = 0$ for impressed forces derivable from a scalar potential as $\boldsymbol{F} = -\mathrm{grad}V$.

2.5. Consider the parabola

$$y_1 = -2x_1^2 - 4$$

and the straight line

$$y_2 = 2x_2 + 1$$

First show that the graphs of these two functions never intersect. Having convinced yourself of this, then go on to find the minimum distance between these two functions.

Show also that the minimum distance is a line perpendicular to the given straight line.

2.6. In statistical mechanics we find that the Gibbs expression for the entropy is

$$S = -k_{\mathrm{B}} \sum_{\mathrm{r}} P_{\mathrm{r}} \ln P_{\mathrm{r}}$$

where k_{B} is the Boltzmann constant and P_{r} is the coefficient of probability for the r^{th} state, which is a measure of the density of states in the system phase space.[13] Thermodynamics teaches us that under conditions of constant energy and volume the *entropy of a system will be a maximum*. That is, we have the constraint that

$$\mathcal{E} = \sum_{\mathrm{r}} P_{\mathrm{r}} \mathcal{E}_{\mathrm{r}},$$

which is that the average energy of the system in the ensemble is a constant. We also realize that there is another constraint in the definition of probability. That is

$$1 = \sum_{r} P_r$$

Show that the probability that a system of atoms will be in a particular state of total energy $\mathcal{E}_{\mathrm{r}}$ is given by

$$P_{\mathrm{r}} = \exp\left[-1 - \alpha - \beta \mathcal{E}_{\mathrm{r}}\right]$$

where α and β are constants. Do this by maximizing the Gibbs entropy subject to the constraints. The α and β are the Lagrange Undetermined Multipliers.

2.7. Study the following functional

$$J\left[y\right] = \int_0^1 \mathrm{d}x \left(y'\right).$$

Determine whether or not it has an extremum. If it does, find that extremum.

2.8. Consider the functional

$$J\left[y\right] = \int_0^1 \mathrm{d}x \left(yy'\right).$$

Determine whether or not this functional has an extremum. If so, find that extremum.

[13] The system phase space has an axis for every canonical coordinate and every canonical momentum of every particle (atom/molecule) in the system. This space is called $\Gamma-$space.

2.9. Study the functional

$$J[y] = \int_0^1 \mathrm{d}x \left(xyy'\right).$$

Determine whether or not this functional has an extremum. If it does, find that extremum.

2.10. Find differential equation for the extremum of the functional

$$J[y] = \int_0^1 \mathrm{d}x \left[y^2 + \left(y'\right)^2 - 2y\sin(x)\right]$$

[answer: $y'' - y = -\sin(x)$]

2.11. Show that the functional of two functions:

$$S[x, y] = \int_{t_1}^{t_2} \mathrm{d}t \Psi[t, x, y, \dot{x}, \dot{y}]$$

has an extremum when Euler–Lagrange equations for each of the functions are satisfied.. That is, show that this functional has an extremum when

$$\frac{\partial}{\partial x}\Psi - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}\Psi = 0$$

and

$$\frac{\partial}{\partial y}\Psi - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{y}}\Psi = 0$$

provided the variations vanish at the end points.

This requires the details of what we indicated in the text.

2.12. Consider that in the functional

$$S[y] = \int_{t_1}^{t_2} \mathrm{d}t F(y, \dot{y})$$

the function F does not depend explicitly on the time t. Show that as a consequence

$$F - \dot{y}\frac{\partial F}{\partial \dot{y}} = \text{constant}.$$

2.13. Using the results of the preceding exercise, i.e.

$$F - \dot{y}\frac{\partial F}{\partial \dot{y}} = \text{constant},$$

show that for F as the Lagrangian $F = T - V$ for a single particle of mass m that the total mechanical energy is constant and that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(F - \dot{y}\frac{\partial F}{\partial \dot{y}}\right) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(T + V\right).$$

2.14. Among all curves joining the points (x_0, y_0) and (x_1, y_1) find that which generates the minimum surface area when rotated around the $x-axis$. Begin with the Pythagorean Theorem that the differential length between two points in the (x, y) −plane is

$$ds = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \mathrm{d}x\sqrt{1 + (y')^2}.$$

At the point x the distance to the curve $y = y(x)$ is equal to the value of y. So the differential area of the surface of rotation defined by the points x and $x + \mathrm{d}x$ on the $x-$axis is

$$\mathrm{d}x\left(2\pi y\sqrt{1 + (y')^2}\right).$$

The area of the surface of revolution between x_0 and $x_1$is

$$A_S = 2\pi\int_{x_0}^{x_1}\mathrm{d}x\left(y\sqrt{1 + (y')^2}\right).$$

[Answer: $y = K\cosh\left[(x + C)/K\right]$, where K and C are (integration) constants]

2.15. A particle is released from rest at a point (x_0, y_0) and slides (without friction) down a curve in the (x, y) plane. Since the differential distance down the plane is

$$\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \mathrm{d}x\sqrt{1 + (y')^2},$$

The speed at which the particle slides is

$$v = \frac{\mathrm{d}s}{\mathrm{d}t} = \sqrt{1 + (y')^2}\frac{\mathrm{d}x}{\mathrm{d}t}.$$

The speed, from energy conservation (no friction) is

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

Then

$$\mathrm{d}t = \mathrm{d}x\frac{\sqrt{1 + (y')^2}}{\sqrt{2gy}}.$$

What must the curve be, down which the particle slides, so that it reaches the vertical line at $x = b\,(> x_0)$ in the shortest time? We then wish to find the extremum of the

functional for the total time

$$T[y] = \int_{x=x_0}^{x=b} dx \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}$$

This is the brachistochrone problem, which was first posed by Johann Bernoulli in 1696.

[Answer: A cycloid]

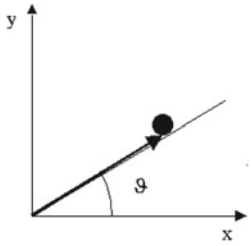
2.16. A particle of mass m moves under no forces in the direction x.

Find the Lagrangian and the canonical momentum. Show that the canonical momentum is conserved. Find the energy and show that its total time derivative is zero so that the energy is a constant.

Do this using the Euler–Lagrange equations.

2.17. Consider a particle of mass m in free fall under the influence of gravity. Find the Lagrangian, the canonical momentum, the Euler–Lagrange equation. Show that the energy is constant and integrate the Euler–Lagrange equation.

2.18. Consider a particle of mass m sliding without friction down an inclined plane. We show this in the figure below.



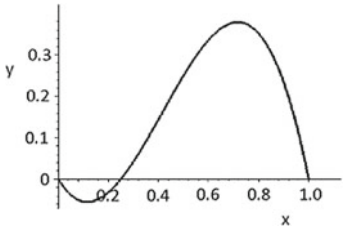
Particle sliding on incline

Find the Lagrangian, the Euler–Lagrange equations, and the energy. Show that the energy is constant and solve the Euler–Lagrange equations. Find the reaction force with the incline using the Lagrange multipliers.

2.19. Consider a mass, m, sliding without friction on the hilly terrain described by the function

$$y = -4x^3 + 5x^2 - x.$$

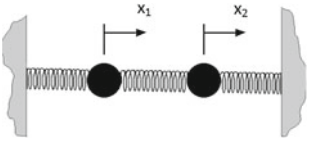
We have shown the hilly terrain in the figure below.



Hilly terrain described by $y = -4x^3 + 5x^2 - x$

Obtain the equations of motion for the particle using Lagrange multipliers.

2.20. In the figure here we have two masses connected by identical springs to one another and to two vertical walls.

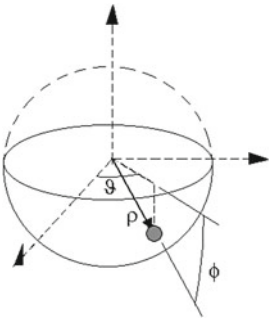


Two masses connected by identical springs to fixed vertical walls.

We neglect gravitational influences.

Study the motion of the system. Find the natural (eigen) frequencies and the corresponding eigenvectors.

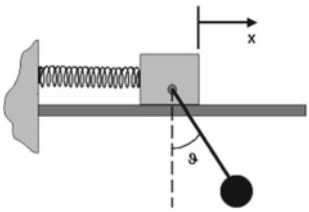
2.21. Consider a mass sliding without friction inside a sphere. We have drawn the picture here.



Mass Inside Sphere. Here a mass slides without friction inside a sphere, which we choose to be of glass so the motion can be observed. The coordinates are spherical (ρ, ϑ, ϕ) with the polar angle ϕ measured from zero in the central horizontal plane.

Study the motion by incorporating all constraints directly into the Lagrangian. Find the equilibrium orbit. Study small perturbations around this orbit

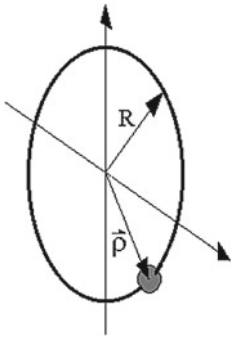
2.22. Consider the block, spring and pendulum system shown here.



Block, spring, and pendulum

Obtain the Euler–Lagrange equations for this system. Then simplify for equal masses ($M = m$). Consider small vibrations (small x and ϑ). Make the Ansatz that the time dependence is $\exp(i\omega t)$ and find the normal modes of motion.

2.23. In the figure below we have drawn a stationary wire loop with a bead of mass m. The bead is free to move with no friction on the wire.



Bead on frictionless, stationary wire loop

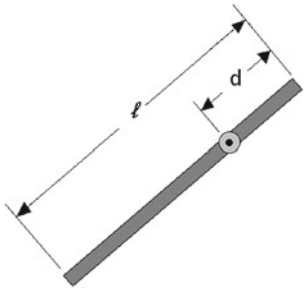
Find the Euler–Lagrange equations. Do not attempt a solution.

2.24. Consider now that the loop in the preceding exercise rotates at a constant angular velocity about the vertical axis. That is $\dot{\vartheta} = \Omega = \text{constant}$ and $\rho = R = \text{constant}$.

Find the Euler–Lagrange equations. What is the equilibrium location of the bead? Show that the equilibrium is stable, that is small deviations form equilibrium result in sinusoidal oscillations around equilibrium, provided

$$1 + \frac{g}{R\Omega^2} - 2\left(\frac{g}{R\Omega^2}\right)^2 > 0.$$

2.25. A physical pendulum is a uniform rod suspended on an axis constrained to move in one plane about a point other than the center of the rod. In the figure below we have shown a physical pendulum.

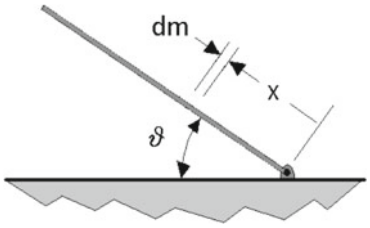


Physical pendulum

Write the Lagrangian for this physical pendulum. Begin by writing the Lagrangian for a differential mass located at a distance x from the axis and then integrate over the rod.

Find the Euler–Lagrange equation for this physical pendulum.

2.26. Consider a uniform rod with linear mass density λ, which is fastened to the floor by a hinge. We have a drawing of the falling rod here.

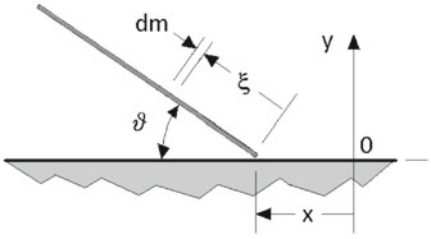


Falling hinged rod

We release the rod from the vertical with a slight nudge so that the angular momentum is initially zero. Obtain the time of fall as an integral. Do not attempt the integration.

[Hint: $\ddot{\vartheta} = (1/2)\, d\dot{\vartheta}^2/\mathrm{d}\vartheta$]

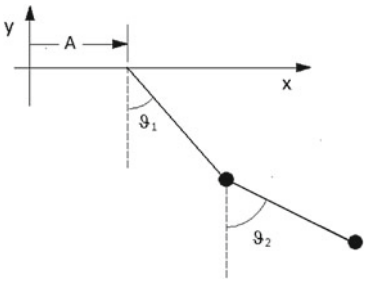
2.27. Now consider the falling rod as in the preceding exercise, except that instead of being hinged the end of the rod is free and the floor is frictionless. We again release the rod at $\vartheta = \pi/2$ with a slight nudge. We have drawn the rod in the figure below.



Free rod falling

Find the Euler–Lagrange equations for the falling rod.

2.28. We have drawn a double pendulum in the figure here.



Double pendulum with equal lengths and bobs of equal mass.

Both pendulum lengths are ℓ and the masses of the pendulum bobs are both m. We consider the masses rods connecting the bobs to be zero.

Obtain the Euler–Lagrange equations, linearize these for small angles and find the normal modes of oscillation.

[Answers: for the Euler–Lagrange equations

$$\begin{aligned} &-2m\ell^2\ddot{\vartheta}_1 - m\ell^2\ddot{\vartheta}_2\,(\cos\vartheta_1\cos\vartheta_2 + \sin\vartheta_1\sin\vartheta_2) \\ &-m\ell^2\dot{\vartheta}_2^2\,(-\cos\vartheta_1\sin\vartheta_2 + \sin\vartheta_1\cos\vartheta_2) \\ &-2mg\ell\sin\vartheta_1 \\ =\ &0 \end{aligned}$$

and

$$\begin{aligned} &-m\ell_1^2\ddot{\vartheta}\,(\cos\vartheta_1\cos\vartheta_2 + \sin\vartheta_1\sin\vartheta_2) - m\ell_2^2\ddot{\vartheta} \\ &-m\ell^2\dot{\vartheta}_1^2\,(-\sin\vartheta_1\cos\vartheta_2 + \cos\vartheta_1\sin\vartheta_2) \\ &-mg\ell\sin\vartheta_2 \\ =\ &0. \end{aligned}$$

For the linearized equations

$$2\ddot{\vartheta}_1 + 2\omega_0^2\vartheta_1 + \ddot{\vartheta}_2 = 0$$

$$\ddot{\vartheta}_1 + \ddot{\vartheta}_2 + \omega_0^2\vartheta_2 = 0,$$

where $\omega_0 = \sqrt{g/\ell}$. For the normal modes

$$\omega = \pm\omega_0\sqrt{2+\sqrt{2}}$$

$$\omega = \pm\omega_0\sqrt{2-\sqrt{2}}$$

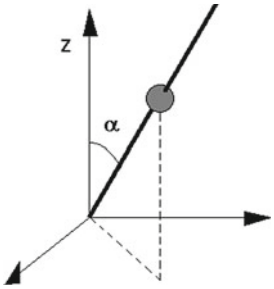
the eigenvectors are

$$\text{For } \omega = \omega_0\sqrt{\left(2+\sqrt{2}\right)}: \quad \begin{bmatrix}\Theta_1\\ \Theta_2\end{bmatrix} = \frac{2}{\sqrt{6}}\begin{bmatrix}-\frac{1}{2}\sqrt{2}\\ 1\end{bmatrix}$$

and

$$\text{For } \omega = \omega_0\sqrt{\left(2-\sqrt{2}\right)}: \quad \begin{bmatrix}\Theta_1\\ \Theta_2\end{bmatrix} = \frac{2}{\sqrt{6}}\begin{bmatrix}\frac{1}{2}\sqrt{2}\\ 1\end{bmatrix}.$$

2.29. In the figure below we have a bead moving without friction on a wire. The wire makes an angle α with the vertical and is free to rotate about the vertical axis, also without friction.



Bead on a frictionless wire

We neglect the mass of the wire.

Show that motion can be described as that of a particle moving in an (effective) potential well

$$V_{\text{eff}} = -\frac{1}{2}\frac{\ell^2}{mr^2} + mg\frac{r}{\tan\alpha}$$

where ℓ is the angular momentum of the bead.

Is there a position of stable equilibrium? This requires consideration of both the radial velocity $\dot{r}$ and the radial acceleration $\ddot{r}$.

Show that the Lagrange Undetermined multiplier is

$$\lambda = -\frac{\ell^2}{mr^3\left(1+\tan^2\alpha\right)} - \frac{mg\tan\alpha}{1+\tan^2\alpha}$$

and that the forces of the wire on the bead are then

$$f_{\text{r}} = -\frac{\ell^2}{mr^3\left(1+\tan^2\alpha\right)} - \frac{mg\tan\alpha}{1+\tan^2\alpha}$$

and

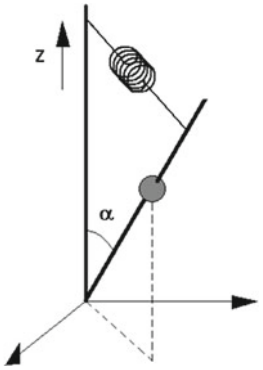
$$f_{\text{z}} = \frac{\ell^2\tan\alpha}{mr^3\left(1+\tan^2\alpha\right)} + \frac{mg\tan^2\alpha}{1+\tan^2\alpha}.$$

Comment on the time dependence of the λ.

2.30. In the preceding exercise the wire was free to rotate about the vertical axis, while the angle remained constant. We now choose to drive the wire at a constant angular velocity Ω about the vertical axis. The angle with the axis will still remain constant at α.

Incorporate the angular constraint and the constant angular velocity using Lagrange undetermined multipliers.

2.31. Consider the situation above with the wire mounted at the origin in a fashion that allows frictionless motion around the vertical and about the pivot point so that the angle to the vertical α varies. Let there be a vertical post erected from the origin. A spring retains the wire so that the pivot angle relative to this vertical post is limited. The spring is mounted at a distance h above the ground on the vertical post by a collar that permits rotation around the post. Consider small vibrations so that the spring remains parallel to the floor. The situation is shown below

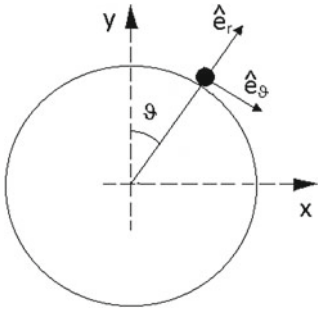


Bead on a frictionless, massless wire with a spring tying the wire to a central pole. The height of the spring is h

Assume that the spring is massless.

Study the motion. Is there an equilibrium orbit for the bead?

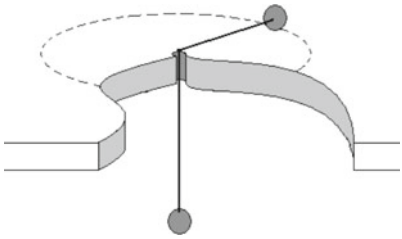
2.32. In the figure below we have drawn a cylinder of radius R lying on a laboratory table. Assume that the surface of the cylinder is frictionless. Its axis is parallel to the table top and the ground. The cylinder remains fixed. We then place a small mass m on the uppermost part of the cylinder. If we release the mass and nudge it slightly it will slide without friction on the cylinder.



Small mass on a frictionless cylinder

At some point the small mass will fall off the cylinder. Find this point.

2.33. Consider two balls connected by a string of length b. One is suspended through a hole in a table and the other moves on the (frictionless) top of the table. We have drawn the situation in the figure here.



Two balls connected by a string. One is suspended below a frictionless table and the other freely moves on the frictionless table

Investigate the motion. Use Lagrange multipliers. Find an equilibrium point, if there is one. Study the general form of the motion. If you find a point of dynamic equilibrium, consider small oscillations about that point. Determine if the orbit is open or closed.

2.34. In a rocket engine the thermal energy of the burning fuel and oxidant is converted in the nozzle into kinetic energy. This high energy gas is expelled. The momentum carried away by this expelled gas results in an increase in momentum of the rocket. Consider a rocket for which

$$\begin{aligned} m_{\mathrm{r}} &= \text{mass of the rocket excluding fuel} \\ m_{\mathrm{f}} &= \text{mass of the fuel at any instant} \\ m_{\mathrm{e}} &= \text{mass of exhaust gases in the nozzle at any time} \\ \dot{m} &= \text{rate at which fuel is burned.} \end{aligned}$$

Let

$$
\begin{aligned}
v &= \text{ velocity of the rocket} \\
u &= \text{ velocity of exhaust gases in space} \\
\dot{v} &= \text{ acceleration of the rocket.}
\end{aligned}
$$

If we consider that the rocket is in a region of space in which all forces may be neglected, obtain the Euler–Lagrange equation for the rocket. This will be the standard propulsion equation

$$\frac{\mathrm{d}}{\mathrm{d}t} p = M\dot{v} - \dot{m}U = 0,$$

where

$$M = m_{\mathrm{r}} + m_{\mathrm{f}} + m_{\mathrm{e}}.$$

and U is the velocity of the exhaust gas relative to the rocket. Note that the kinetic energy of the rocket, including unburned fuel and the exhaust gases in the nozzle is

$$T = \frac{1}{2} m_{\mathrm{r}} v^2 + \frac{1}{2} m_{\mathrm{f}} v^2 + \frac{1}{2} m_{\mathrm{e}} (v - u)^2 .$$

The exhaust gases are considered part of the rocket until they exit the nozzle.

http://www.springer.com/978-3-319-44490-1

Analytical Mechanics

Helrich, C.S.

2017, XV, 349 p. 58 illus., Softcover
ISBN: 978-3-319-44490-1