Chapter 4
Elements of Classical Mechanics*

Based on d'Alembert's principle, we formulate Hamilton's integral principle by defining a function $L = L(q_i, \dot{q}_i, t)$ such that for any mechanical system the variation of the integral $\int_{t_0}^{t_1} L \mathrm{d}t$ vanishes along any real path (Fig. 4.1) so that

$$\delta \int_{t_0}^{t_1} L(t)\mathrm{d}t = 0. \tag{4.1}$$

Here, the variables $(q_i, \dot{q}_i, t)$ are the coordinates and velocities, respectively, and t is the independent variable time. We may expand this function and get

$$\delta \int_{t_0}^{t_1} L \mathrm{d}t = \int \sum_i \frac{\partial L}{\partial q_i} \delta q_i \mathrm{d}t + \int \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \mathrm{d}t \,. \tag{4.2}$$
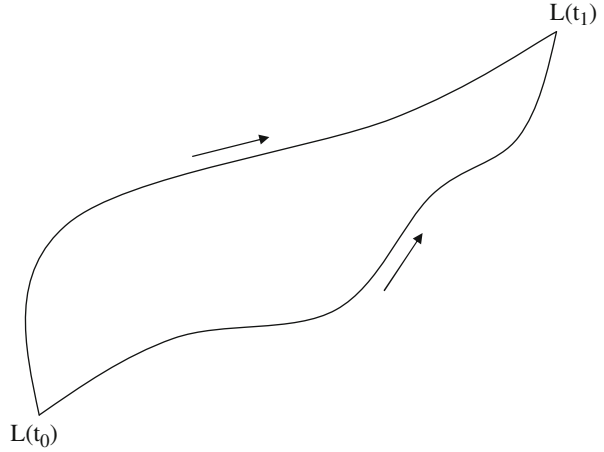
The second term can be modified using the assumption of the variational theorem which requires that $\delta q_i = 0$ at the beginning and end of the path. The second term can be integrated by parts and is

$$\int \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \mathrm{d}t = \int \frac{\partial L}{\partial \dot{q}_i} \frac{\mathrm{d}}{\mathrm{d}t} \delta q_i \mathrm{d}t = \underbrace{\frac{\partial L}{\partial \dot{q}_i} \frac{\mathrm{d}}{\mathrm{d}t} \delta q_i \bigg|_{t_0}^{t_1}}_{=0} - \int \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} \delta q_i \mathrm{d}t. \tag{4.3}$$

Both terms can now be combined for

$$\delta \int_{t_0}^{t_1} L \mathrm{d}t = \int \sum_i \left(\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i \mathrm{d}t = 0. \tag{4.4}$$

Fig. 4.1 Variational principle

This integral is zero for any arbitrary path if and only if the integrand vanishes for each component i independently. The resulting equations are called the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0. \tag{4.5}$$

Bypassing a more accurate discussion [1], we guess at the nature of the Euler-Lagrange equations by considering a falling mass m. The kinetic energy is $T = \frac{1}{2}mv^2$ and the potential energy $V = gx$, where g is the gravitational force. If we set $L = T - V = \frac{1}{2}mv^2 - gx$ and apply (4.5), we get $m\dot{v} = g$ which is the well known classical equation of motion for a falling mass in a gravitational field. The time independent Lagrangian can be defined by

$$L = T - V \tag{4.6}$$

and the Lagrange function therefore has the dimension of an energy. Furthermore, in analogy with basic mechanics like a falling mass, we can define the momenta of a system by

$$P_i = \frac{\partial L}{\partial \dot{q}_i} \tag{4.7}$$

and call them the generalized canonical momenta. We use a capital P for the canonical momentum to distinguish it from the ordinary momentum p. Both are different only when electromagnetic fields are involved.

4.1 How to Formulate a Lagrangian?

To formulate an expression for the Lagrangian is a creative process of physics. Whatever expression one might propose, it should be independent of a particular reference system and therefore Lorentz invariant. Earlier, we have learned that the product of two 4-vectors is Lorentz invariant and the product of two, not necessarily different, 4-vectors is therefore a good choice to form a Lagrangian. We investigate, for example, the product of the momentum-energy $\left(cp_x^*, cp_y^*, cp_z^*, iE^*\right) = \left(0, 0, 0, imc^2\right)$ and the differential space-time 4-vectors $(dx^*, dy^*, dz^*, icd\tau)$ in the particle rest frame and get

$$\frac{1}{c}\left(dx^*, dy^*, dz^*, icd\tau\right)\left(cp_x^*, cp_y^*, cp_z^*, iE^*\right) = -mc^2 d\tau = -mc^2 \sqrt{1-\beta^2} dt. \tag{4.8}$$

This expression has the dimension of an energy and is Lorentz invariant. We consider therefore this as the Lagrangian for a particle at rest being observed from a relatively moving laboratory system

$$L = -mc^2 \sqrt{1-\beta^2}. \tag{4.9}$$

The conjugate momentum is from (4.7) for the x-component

$$P_x = -m\frac{-v_x}{\sqrt{1-\beta^2}} = \gamma m v_x \tag{4.10}$$

and the equation of motion $\frac{d}{dt}\frac{\partial L}{\partial v_x} - \frac{\partial L}{\partial x}$ becomes

$$\frac{dP_x}{dt} = 0 \tag{4.11}$$

indicating that the particle is in uniform motion with velocity β.

The Lagrangian (4.9) is consistent with classical experience if we set $\beta \ll 1$ and $L = -mc^2\sqrt{1-\beta^2} \approx -mc^2 + \frac{1}{2}mv^2$. Since we use only derivatives of the Lagrangian, we may ignore the constant $-mc^2$ and end up with the kinetic energy of the free particle.

4.1.1 The Lagrangian for a Charged Particle in an EM-Field

The interaction between charged particle and electromagnetic field depends only on the particle charge and velocity and on the field. We try therefore the product of field and velocity 4-vector. Formulating this product in the laboratory system, where the

fields have been generated, we get.

$$e \left(A_x, A_y, A_z, i\phi \right) \gamma \left(v_x, v_y, v_z, i \right) = e\gamma \left(A\boldsymbol{v} - \phi \right). \tag{4.12}$$

Noting that $\gamma d\tau = dt$, the extension to the Lagrange function in the presence of electromagnetic fields is

$$L = -mc^2 \sqrt{1 - \beta^2} + eA\boldsymbol{v} - e\phi. \tag{4.13}$$

The canonical momentum is from (4.7)

$$\boldsymbol{P} = \frac{m\boldsymbol{v}}{\sqrt{1 - \beta^2}} + e\boldsymbol{A} = \gamma m\boldsymbol{v} + e\boldsymbol{A} = \boldsymbol{p} + e\boldsymbol{A}, \tag{4.14}$$

where $\boldsymbol{p}$ is the ordinary momentum. Equation (4.13) is consistent with $L = T - V$, where the potential $V = e\phi - eA\boldsymbol{v}$.

4.2 Lorentz Force

The conjugate momenta in Cartesian coordinates $\boldsymbol{r} = (x, y, z)$ can be derived from (4.5) with (4.13)

$$\dot{\boldsymbol{P}} = \frac{\partial L}{\partial \boldsymbol{r}} = e\nabla \left(A\boldsymbol{v} \right) - e\nabla\phi = e \left(\boldsymbol{v}\nabla \right) \boldsymbol{A} + e \left[\boldsymbol{v} \times \left(\nabla \times \boldsymbol{A} \right) \right] - e\nabla\phi, \tag{4.15}$$

where we used the algebraic relation (A.18). Insertion into

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\boldsymbol{r}}} = \frac{d\boldsymbol{P}}{dt} = \frac{d}{dt} \left(\boldsymbol{p} + e\boldsymbol{A} \right) = e \left(\boldsymbol{v}\nabla \right) \boldsymbol{A} + e \left[\boldsymbol{v} \times \left(\nabla \times \boldsymbol{A} \right) \right] - e\nabla\phi$$

results with $\dot{\boldsymbol{r}} = \boldsymbol{v}$ and $\frac{dA}{dt} = \frac{\partial A}{\partial t} + (\boldsymbol{v}\nabla)A$ in an expression for the ordinary momentum $\boldsymbol{p}$

$$\frac{d\boldsymbol{p}}{dt} = -e\frac{\partial \boldsymbol{A}}{\partial t} + e \left[\boldsymbol{v} \times \left(\nabla \times \boldsymbol{A} \right) \right] - e\nabla\phi. \tag{4.16}$$

Converting potentials to fields, we recover the Lorentz force $\boldsymbol{F}_{\mathrm{L}} = \frac{d\boldsymbol{p}}{dt}$ or

$$\boldsymbol{F}_{\mathrm{L}} = e\boldsymbol{E} + e \left(\boldsymbol{v} \times \boldsymbol{B} \right). \tag{4.17}$$

4.3   Frenet-Serret Coordinates

A particle trajectory follows a path described by

$$r(z) = r_0(z) + \delta r(z). \tag{4.18}$$

Here $r_0(z)$ is the ideal path for beam dynamics and an orthogonal coordinate system moves along this path with its origin at $r_0(z)$ as shown in Fig. 4.2. For this Frenet-Serret coordinate system we define three vectors

$$
\begin{array}{ll}
u_x(z) & \text{unit vector } \perp \text{ to trajectory} \\
u_z(z) = \frac{dr_0(z)}{dz} & \text{unit vector } \parallel \text{ to trajectory} \\
u_y(z) = u_z(z) \times u_x(z) & \text{unit binormal vector}
\end{array}
\tag{4.19}
$$

to form an orthogonal coordinate system moving along the trajectory with a reference particle at $r_0(z)$. In beam dynamics, we identify the plane defined by vectors u_x and $u_z(z)$ as the horizontal plane and the plane orthogonal to it as the vertical plane, parallel to u_y. Change in vectors are determined by curvatures.

$$\frac{du_x(z)}{dz} = \kappa_x u_z(z), \qquad \text{and} \qquad \frac{du_y(z)}{dz} = \kappa_y u_z(z), \tag{4.20}$$

where $\left(\kappa_x, \kappa_y\right)$ are the curvatures in the horizontal and vertical plane respectively. The particle trajectory can now be described by

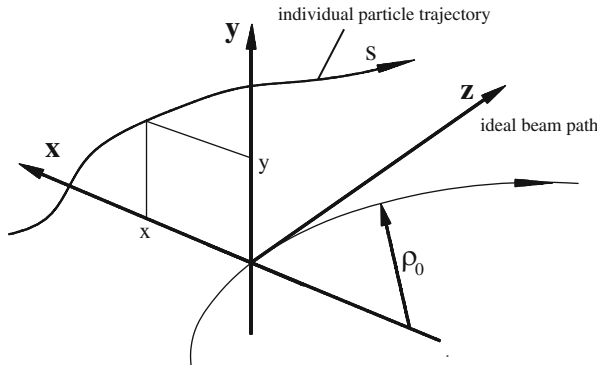$$r(x, y, z) = r_0(z) + x(z)u_x(z) + y(z)u_y(z), \tag{4.21}$$

Fig. 4.2 Frenet-Serret coordinate system

where $\boldsymbol{r}_0(z)$ is the location of the coordinate system's origin (reference particle) and (x, y) are the deviations of a particular particle from $\boldsymbol{r}_0(z)$. The derivative with respect to z is then

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{r}(x, y, z) = \frac{\mathrm{d}\boldsymbol{r}_0}{\mathrm{d}z} + x(z)\frac{\mathrm{d}\boldsymbol{u}_x(z)}{\mathrm{d}z} + y(z)\frac{\mathrm{d}\boldsymbol{u}_y(z)}{\mathrm{d}z} + x'(z)\boldsymbol{u}_x(z) + y'(z)\boldsymbol{u}_y(z) \quad (4.22)$$

or with (4.19) and (4.20)

$$\mathrm{d}\boldsymbol{r} = \boldsymbol{u}_x\mathrm{d}x + \boldsymbol{u}_y\mathrm{d}y + \boldsymbol{u}_z h\mathrm{d}z, \quad (4.23)$$

where

$$h = 1 + \kappa_{0x}x + \kappa_{0y}y. \quad (4.24)$$

Using these Frenet-Serret coordinates, we are able to describe particle trajectories much more efficient than we could do in Cartesian coordinates. Essentially, we have transformed away the ideal path or the geometry of the design beam transport line which is already well known to us from the placement of beam guidance elements. The new coordinates measure directly the deviation of any particles from the reference particle.

We may use these relations to introduce a transformation, from the Cartesian coordinate system to curvilinear Frenet-Serret coordinates, in the Lagrangian $L = -mc^2\sqrt{1 - \beta^2} + e\dot{\boldsymbol{r}}\boldsymbol{A} - e\phi$. In the new coordinates, $\sqrt{1 - \beta^2} = \sqrt{1 - \frac{1}{c^2}(\dot{x}^2 + \dot{y}^2 + h^2\dot{z}^2)}$, $\dot{\boldsymbol{r}}\boldsymbol{A} = \dot{x}A_x + \dot{y}A_y + h\dot{z}A_z$ and the Lagrangian becomes in curvilinear coordinates of beam dynamics

$$L = -mc^2\sqrt{1 - \tfrac{1}{c^2}(\dot{x}^2 + \dot{y}^2 + h^2\dot{z}^2)} + e\left(\dot{x}A_x + \dot{y}A_y + h\dot{z}A_z\right) - e\phi. \quad (4.25)$$

4.4   Hamiltonian Formulation

Like any other mechanical system, particle beam dynamics in the presence of external electromagnetic fields can be described and studied very generally through the Hamiltonian formalism. The motion of particles in beam transport systems, expressed in normalized coordinates, is that of a harmonic oscillator and deviations caused by nonlinear restoring forces appear as perturbations of the harmonic oscillation. Such systems have been studied extensively in the past and powerful mathematical tools have been developed to describe the dynamics of harmonic oscillators under the influence of perturbations. Of special importance is the Hamiltonian formalism which we will apply to the dynamics of charged particles. Although this theory is well documented in many text books, for example in [1, 2], we will shortly recall the Hamiltonian theory with special attention to the application in charged particle dynamics.

The canonical variables in the Hamiltonian theory are the coordinates and momenta rather than coordinates and velocities used in the Lagrangian. We use a coordinate transformation $(q_i, \dot{q}_i, t) \implies (q_i, P_i, t)$ through the definition of the momenta $P_i = \partial L / \partial \dot{q}_i$ and define the Hamiltonian function by

$$H(q_i, p_i) = \sum \dot{q}_i P_i - L(q_i, \dot{q}_i). \tag{4.26}$$

In analogy to the Lagrangian, we find that $\dot{q}_i P_i = 2T$ and the Hamiltonian which does not depend on the time explicitly is therefore the sum of kinetic and potential energy

$$H = T + V. \tag{4.27}$$

This will become useful later since we often know forces acting on particles which can be derived from a potential. Similar to the Euler-Lagrange equations, we define Hamiltonian equations by

$$\frac{\partial H}{\partial q_i} = -\dot{P}_i, \qquad \text{and} \qquad \frac{\partial H}{\partial P_i} = +\dot{q}_i. \tag{4.28}$$

With $L = -mc^2 \sqrt{1 - \beta^2} + e\boldsymbol{A}\boldsymbol{v} - e\phi$ and replacing velocities with momenta the Hamiltonian becomes

$$H(q_i, P_i) = \sum \dot{q}_i P_i + mc^2 \sqrt{1 - \beta^2} - e\boldsymbol{A}\dot{\boldsymbol{q}} + e\phi, \tag{4.29}$$

where $\boldsymbol{q} = (q_1, q_2, .., q_i, ..)$ and $\boldsymbol{A} = (A_1, A_2, .., A_i, ..)$, etc. and the canonical momentum is defined in (4.14). The canonical momentum $\boldsymbol{P}$ is from (4.14) the combination of the ordinary particle momentum $\boldsymbol{p} = \gamma m \dot{\boldsymbol{q}}$ and field momentum $e\boldsymbol{A}$. Insertion into the Hamiltonian and reordering gives $(H - e\phi)^2 = m^2c^4 + c^2 (\boldsymbol{P} - e\boldsymbol{A})^2$, or

$$c^2 (\boldsymbol{P} - e\boldsymbol{A})^2 - (H - e\phi)^2 = -m^2c^4, \tag{4.30}$$

The Hamiltonian (4.30) is equal to the square of the length of the energy momentum 4-vector $[c\boldsymbol{P}, iE]$, where $E = H - e\phi$, and is therefore Lorentz invariant. A more familiar form is

$$H = e\phi + \sqrt{c^2 (\boldsymbol{P} - e\boldsymbol{A})^2 + m^2c^4}. \tag{4.31}$$

In nonrelativistic mechanics, the Hamiltonian becomes with $\beta \ll 1$ and ignoring the constant mc^2

$$H_{\text{class}} \approx \tfrac{1}{2} mv^2 + e\phi, \tag{4.32}$$

which is the sum of kinetic and potential energy.

4.4.1 Cyclic Variables

The solution of the equations of motion become greatly simplified in cases, where
the Hamiltonian does not depend on one or more of the coordinates or momenta.
In this case one or more of the Hamiltonian equations (4.28) are zero and the
corresponding conjugate variables are constants of motion. Of particular interest
for particle dynamics or harmonic oscillators are the cases where the Hamiltonian
does not depend on say the coordinate q_i but only on the momenta P_i. In this case
we have

$$H = H(q_1, \dots q_{i-1}, q_{i+1} \dots, P_1, P_2 \dots, P_i, \dots) \tag{4.33}$$

and the first Hamiltonian equation becomes

$$\frac{\partial H}{\partial q_i} = -\dot{P}_i = 0 \qquad \text{or} \qquad P_i = \text{const}. \tag{4.34}$$

Coordinates q_i which do not appear in the Hamiltonian are called cyclic coordinates
and their conjugate momenta are constants of motion. From the second Hamiltonian
equation we get with $P_i = \text{const}$.

$$\frac{\partial H}{\partial p_i} = \dot{q}_i = a_i = \text{const},$$

which can be integrated immediately for

$$q_i(t) = a_i t + c_i, \tag{4.35}$$

where c_i is the integration constant. It is obvious that the complexity of a mechanical
system can be greatly reduced if by a proper choice of canonical variables some or
all dependence of the Hamiltonian on space coordinates can be eliminated. We will
derive the formalism that allows the transformation of canonical coordinates into
new ones, where some of them might be cyclic.

   Example: Assume that the Hamiltonian does not depend explicitly on the time,
then $\frac{\partial H}{\partial t} = 0$ and the momentum conjugate to the time is a constant of motion.
From the second Hamilton equation, we have $\frac{\partial H}{\partial p_i} = \frac{d}{dt}t = 1$ and the momentum
conjugate to the time is therefore the total energy $p_i = H = \text{const}$. The total energy
of a system with a time independent Hamiltonian is constant and equal to the value
of the Hamiltonian.

4.4.2 Canonical Transformations

For mechanical systems which allow in principle a formulation in terms of cyclic
variables, we need to derive rules to transform one set of variables to another

set, while preserving their property of being conjugate variables appropriate to formulate the Hamiltonian for the system. In other words, the coordinate transformation must preserve the variational principle (4.1). Such transformations are called canonical transformations $\bar{q}_k = f_k(q_i, P_i, t)$ and $\bar{P}_k = g_k(q_i, P_i, t)$, where (q_i, P_i, t) are the old and $(\bar{q}_k, \bar{P}_k, t)$ the new coordinates. The variational principle reads now

$$\delta \int \left(\sum_k \dot{q}_k P_k - H \right) dt = 0 \qquad \text{and} \qquad \delta \int \left(\sum_k \dot{\bar{q}}_k \bar{P}_k - \overline{H} \right) dt = 0. \qquad (4.36)$$

The new Hamiltonian $\overline{H}$ need not be the same as the old Hamiltonian H nor need both integrands be the same. Both integrands can differ, however, only by a total time derivative of an otherwise arbitrary function G

$$\sum_k \dot{q}_k P_k - H = \sum_k \dot{\bar{q}}_k \bar{P}_k - \overline{H} + \frac{dG}{dt}. \qquad (4.37)$$

After integration $\int \frac{dG}{dt} dt$ becomes a constant and the variation of the integral obviously vanishes under the variational principle (Fig. 4.1). The arbitrary function G is called the generating function and may depend on some or all of the old and new variables

$$G = G(q_k, \bar{q}_k, P_k, \bar{P}_k, t) \qquad \text{with} \quad 0 \leq k \leq N. \qquad (4.38)$$

The generating functions are functions of only $2N$ variables, coordinates and momenta. Of the $4N$ variables only $2N$ are independent because of another $2N$ transformation equations (4.36). We may now choose any two of four variables to be independent keeping only in mind that one must be an old and one a new variable. Depending on our choice for the independent variables, the generating function may have one of four forms

$$\begin{aligned} G_1 &= G_1(q, \bar{q}, t), & G_3 &= G_3(P, \bar{q}, t), \\ G_2 &= G_2(q, \bar{P}, t), & G_4 &= G_4(P, \bar{P}, t), \end{aligned} \qquad (4.39)$$

where we have set $q = (q_1, q_2, \ldots q_N)$ etc. We take, for example, the generating function G_1, insert the total time derivative

$$\frac{dG_1}{dt} = \sum_k \frac{\partial G_1}{\partial q_k} \frac{\partial q_k}{\partial t} + \sum_k \frac{\partial G_1}{\partial p_k} \frac{\partial P_k}{\partial t} + \frac{\partial G_1}{\partial t} \qquad (4.40)$$

in (4.37) and get after some sorting

$$\sum_k \dot{q}_k \left(P_k - \frac{\partial G_1}{\partial q_k} \right) - \sum_k \dot{\bar{q}}_k \left(\bar{P}_k + \frac{\partial G_1}{\partial \bar{q}_k} \right) - \left(H - \overline{H} + \frac{\partial G_1}{\partial t} \right) = 0. \qquad (4.41)$$

Both, old and new variables are independent and the expressions in the brackets must therefore vanish separately leading to the defining equations

$$P_k = \frac{\partial G_1}{\partial q_k}, \qquad \bar{P}_k = -\frac{\partial G_1}{\partial \bar{q}_k}, \qquad H = \bar{H} - \frac{\partial G_1}{\partial t}. \tag{4.42}$$

Variables for which (4.42) hold are called canonical variables and the transformations (4.36) are called canonical.

Generating functions for other pairings of new and old canonical variables can be obtained from G_1 by Legendre transformations of the form

$$G_2(q, \bar{P}, t) = G_1(q, \bar{q}, t) + q\bar{P}. \tag{4.43}$$

Equations (4.42) can be expressed in a general form for all four different types of generating functions. We write the general generating equation as $G = G(x_k, \bar{x}_k, t)$, where the variables x_k and $\bar{x}_k$ can be either coordinates or momenta. Furthermore, x_k and $\bar{x}_k$ are the old and new coordinates or momenta respectively and the $(y_k, \bar{y}_k)$ are the conjugate coordinates or momenta to $(x_k, \bar{x}_k)$. Then

$$y_k = \pm \frac{\partial}{\partial x_k} G(x_k, \bar{x}_k, t),$$

$$\bar{y}_k = \mp \frac{\partial}{\partial \bar{x}_k} G(x_k, \bar{x}_k, t), \tag{4.44}$$

$$H = \bar{H} - \frac{\partial}{\partial t} G(x_k, \bar{x}_k, t).$$

The upper signs are to be used if the derivatives are taken with respect to coordinates and the lower signs if the derivatives are taken with respect to momenta. It is not obvious which type of generating function should be used for a particular problem. However, the objective of canonical transformations is to express the problem at hand in as many cyclic variables as possible. Any form of generating function that achieves this goal is therefore appropriate. To illustrate the use of generating functions for canonical transformations, we will discuss a few very examples. For an identity transformation we use a generating function of the form

$$G = q_1 \bar{P}_1 + q_2 \bar{P}_2 + \ldots \tag{4.45}$$

and get with (4.44) and $i = 1, 2, \ldots N$ the identities

$$P_i = -\frac{\partial G}{\partial q_i} = \bar{P}_i, \qquad \text{and} \qquad \bar{q}_i = +\frac{\partial G}{\partial \bar{P}_i} = q_i. \tag{4.46a}$$

A transformation from rectangular (x, y, z) to cylindrical (r, φ, z) coordinates is defined by the generating function

$$G(P, \bar{q}) = -P_x r \cos \varphi - P_y r \sin \varphi - P_z z \tag{4.47}$$

and the transformation relations are

$$
\begin{array}{ll}
x = -\frac{\partial G}{\partial p_x} = r \cos \varphi, & P_r = -\frac{\partial G}{\partial r} = +P_x \cos \varphi + P_y \sin \varphi, \\
y = -\frac{\partial G}{\partial p_y} = r \sin \varphi, & P_\varphi = -\frac{\partial G}{r \partial \varphi} = -P_x \sin \varphi + P_y \cos \varphi, \\
z = -\frac{\partial G}{\partial p_z} = z, & P_z = -\frac{\partial G}{\partial z} = P_z.
\end{array}
\tag{4.48}
$$

Similarly, relations for the transformation from rectangular to polar coordinates can be derived from the generating function

$$G = -P_x r \cos \varphi \sin \vartheta - P_y r \sin \varphi \sin \vartheta - P_z r \cos \vartheta. \tag{4.49}$$

It is not always obvious if a coordinate transformation is canonical. To identify a canonical transformation, we use Poisson brackets [1] defined by

$$\left[f_k(q_i, P_j), g_k(q_i, P_j) \right] = \sum_i \left(\frac{\partial f_k}{\partial q_i} \frac{\partial g_k}{\partial P_j} - \frac{\partial f_k}{\partial P_j} \frac{\partial g_k}{\partial q_i} \right). \tag{4.50}$$

It can be shown [1] that the new variables $\bar{q}_k$, $\bar{P}_k$ or (4.36) are canonical if and only if the Poisson brackets

$$[\bar{P}_i, \bar{P}_j] = 0 \qquad [\bar{q}_i, \bar{q}_j] = 0 \qquad [\bar{q}_i, \bar{P}_j] = \lambda \delta_{ij}, \tag{4.51}$$

where δ_{ij} is the Kronecker symbol and the factor λ is a scale factor for the transformation. To preserve the scale in phase space, the scale factor must be equal to unity, $\lambda = 1$. While the formalism for canonical transformation is straight-forward, we do not get a hint as to the optimum set of variables for a particular mechanical system. In the next sections we will see, however, that specific transformations have been identified and developed which prove especially useful for a whole class of mechanical systems.

4.4.3 Curvilinear Coordinates

The choice of a particular coordinate system, of course, must not alter the physical result and from this point of view any coordinate system could be used. However, it soon becomes clear that the pursuit of physics solutions can be mathematically much easier in one coordinate system that in another. For systems which are symmetric about a point we would use polar coordinates, for systems which are symmetric

about a straight line we use cylindrical coordinates. In beam dynamics there is no such symmetry, but we have a series of magnets and other components aligned along some, not necessarily straight, line. The collection of these elements is what we call a beam line. The particular arrangement of elements is in most cases not determined by physics but other more practical considerations. The matter of fact is that we know about the "ideal" path and that all particle should travel along a path being defined by the physical centers of the beam line elements. In a Cartesian coordinate system fixed to the stars the result of "ideal" beam dynamics would be a complicated mathematical expression trying to describe the "ideal" path in which we have no interest, since we already know where it is. What we are interested in is the deviation a particular particle might have from the ideal path. The most appropriate coordinate system would therefore be one which moves along the ideal path. In Sect. 4.3 we have introduced such a curvilinear reference system also known as the Frenet-Serret reference system. The transformation from Cartesian to Frenet-Serret coordinates can be derived from the generating function formed from the old momenta and the new coordinates

$$G(z, x, y, P_{c,z}, P_{c,x}, P_{c,y}) = -(c\mathbf{P}_c - ec\mathbf{A}_c)\left[\mathbf{r}_0(z) + x\mathbf{u}_x(z) + y\mathbf{u}_y(z)\right]. \quad (4.52)$$

The momenta and fields in the old Cartesian coordinate system are designated with the index $_c$ and the new canonical momenta P in the Frenet-Serret system are then in both systems while noting that the transverse momenta are the same

$$(cP_z - ecA_zh) = -\frac{\partial G}{\partial z} = (cP_z - ecA_z)_c\ h,$$

$$(cP_x - ecA_x) = -\frac{\partial G}{\partial x} = (cP_x - ecA_x)_c\,, \quad (4.53)$$

$$\left(cP_y - ecA_y\right) = -\frac{\partial G}{\partial y} = \left(cP_y - ecA_y\right)_c\,,$$

with h as defined in (4.24). The Hamiltonian $H_c = e\phi + c\sqrt{m^2c^2 + (\mathbf{P} - e\mathbf{A})_c^2}$ in Cartesian coordinates transforms to the one in curvilinear coordinates of beam dynamics

$$H = e\phi + c\sqrt{m^2c^2 + \frac{(P_z - eA_zh)^2}{h^2} + (P_x - eA_x)^2 + \left(P_y - eA_y\right)^2}. \quad (4.54)$$

For a particle travelling through a uniform field B_y, we have $\mathbf{A} = (0, 0, A_z) = \left(0, 0, -B_yx\right)$, $P_{x,y} = p_{x,y}$, and the Hamiltonian is with $A_z = A_{c,z}h$

$$H_h = e\phi + c\sqrt{m^2c^2 + p_x^2 + p_y^2 + \frac{1}{h^2}\left(P_z + eB_yhx\right)^2}. \quad (4.55)$$

The distinction, we make here on fields in curvilinear and Cartesian coordinates stems from the practice to build magnets in a certain way. Dipole magnets are designed carefully to have a uniform field in the beam area along the curved path, which is not consistent with the transformation of a uniform dipole field in Cartesian coordinates.

4.4.4 Extended Hamiltonian

The Hamiltonian as derived so far depends on the canonical variables (q_i, P_i) and the independent variable t or z defined for individual particles. This separate treatment of the independent variable can be eliminated by formulating an extended Hamiltonian in which all coordinates are treated the same.

Starting with $H(q_1, q_2 \ldots q_f, P_1, P_2, P_3 \ldots P_f, t)$, we introduce the independent variables (q_0, P_0) by setting

$$q_0 = t \quad \text{and} \quad P_0 = -H \tag{4.56}$$

and obtain a new Hamiltonian

$$\mathcal{H}(q_0, q_1, q_2 \ldots q_f, P_0, P_1, P_2, P_3 \ldots P_f) = H + P_0 = 0 \tag{4.57}$$

and Hamilton's equations are then

$$\left. \begin{array}{l} \frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial P_i} \\ \frac{dP_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i} \end{array} \right\} \quad \text{for} \quad i = 0, 1, 2 \ldots \tag{4.58}$$

In particular for $i = 0$ the equations are

$$\frac{dq_0}{dt} = 1 \rightarrow q_0 = t + C_1 \tag{4.59}$$

and

$$\frac{dP_0}{dt} = -\frac{\partial \mathcal{H}}{\partial q_0} = -\frac{d\mathcal{H}}{dt} \implies P_0 = -\mathcal{H} + C_2 . \tag{4.60}$$

The momentum conjugate to the time is equal to the Hamiltonian and since $\mathcal{H} \neq \mathcal{H}(t)$ for static fields, it follows that

$$\frac{dP_0}{dt} = 0 \implies \mathcal{H} = \text{const.} \tag{4.61}$$

Now, the independent variable is no more distinguishable from all other coordinates, the Hamiltonian is expressed as a function of coordinates and momenta only.

4.4.5 Change of Independent Variable

Since no particular coordinate is designated as the independent variable, we may use any of the coordinates as that. For example, we prefer often to use the longitudinal coordinate z as the independent variable rather than the time t. More generally, consider to change the independent variable from q_i to q_j. Defining, for example, q_3 as the new independent variable, we solve $\mathcal{H}$ for P_3

$$P_3 = -K(q_0, q_1, q_2, q_3 \ldots q_{\mathrm{f}}, \ P_0, P_1, P_2, P_4, \ldots P_{\mathrm{f}}) \tag{4.62}$$

and define a new extended Hamiltonian

$$\mathcal{K} = P_3 + K = 0. \tag{4.63}$$

Then the equations

$$\frac{\partial \mathcal{K}}{\partial P_3} = \frac{dq_3}{dq_3} = 1, \tag{4.64a}$$

$$-\frac{\partial \mathcal{K}}{\partial q_3} = \frac{dP_3}{dq_3} = -\frac{\partial K}{\partial q_3}, \tag{4.64b}$$

$$\frac{\partial \mathcal{K}}{\partial P_{i \neq 3}} = \frac{dq_{i \neq 3}}{dq_3} = \frac{\partial K}{\partial P_{i \neq 3}}, \tag{4.64c}$$

$$-\frac{\partial \mathcal{K}}{\partial q_{i \neq 3}} = \frac{dP_{i \neq 3}}{dP_3} = -\frac{\partial K}{\partial q_{i \neq 3}} \tag{4.64d}$$

with the Hamiltonian

$$K = -p_3. \tag{4.65}$$

As an example, to use the longitudinal coordinate z rather than the time t as the independent variable, we start with (4.54)

$$H(x, y, z, t) = e\phi + \sqrt{\frac{1}{h^2}(cP_z - ecA_z h)^2 + c^2 p_\perp^2 + m^2 c^4}, \tag{4.66}$$

where $p_\perp^2 = p_x^2 + p_y^2$. The longitudinal momentum is

$$cP_z = ceA_z h + h\sqrt{(H - e\phi)^2 - (cp_\perp)^2 - m^2 c^4} = ceA_z h + h\sqrt{c^2 p^2 - c^2 p_\perp^2},$$

$$(4.67)$$

where $E^2 = (H - e\phi)^2 = (cp)^2 + (mc^2)^2$ has been used. We further normalize to the momentum p and use trajectory slopes, $x' = \mathrm{d}x/\mathrm{d}z = p_x/p_z$ etc. rather than momenta. With this, the new Hamiltonian is $K(x, x', y, y', z) = -P_z/p$ or using $P_z/p = eA_z/p + h\sqrt{1 - p_\perp^2/p^2}$ and $p_\perp^2/p^2 \approx x'^2 + y'^2$

$$K(x, x', y, y', z) = -\frac{eA_z h}{p} - h\sqrt{1 - x'^2 - y'^2}. \qquad (4.68)$$

In beam dynamics, we restrict ourselves to paraxial beams, where $x' \ll 1$ and $y' \ll 1$, and the momentum $p \approx p_z$. Note, p may not be the canonical momentum if there is an electromagnetic field present, but $P = p + eA$ is canonical. In this last step, we seem to have lost terms involving transverse vector potential components. This meets with the requirements of almost all beam transport lines, where we use predominantly transverse fields which can be derived from the A_z-component only. This is not true when we consider, for example, solenoid fields which occur rather seldom and will be treated separately as perturbations. Finally, we separate the ideal particle momentum p_0 from the momentum deviation $\delta = \Delta p/p_0$ and while ignoring higher order terms in δ replace $1/p = 1/[p_0(1 + \delta)] \approx \frac{1}{p_0}(1 - \delta)$ in the Hamiltonian for

$$K(x, x', y, y', z) \approx -\frac{eA_z h}{p_0}(1 - \delta) - h\sqrt{1 - x'^2 - y'^2}. \qquad (4.69)$$

As discussed before, magnetic fields for particle beam dynamics can be derived from a single component A_z of the vector potential and the task to determine equations of motion is now reduced to that of determining the vector potential for the magnets in use. The equations of motion are from (4.69)

$$\begin{aligned}
\frac{\partial K}{\partial x} &= -x'' = -\frac{ec}{cp_0}\frac{\partial A_z h}{\partial x}(1 - \delta) - \kappa_{0x}\sqrt{1 - x'^2 - y'^2}, \\
\frac{\partial K}{\partial y} &= -y'' = -\frac{ec}{cp_0}\frac{\partial A_z h}{\partial y}(1 - \delta) - \kappa_{0y}\sqrt{1 - x'^2 - y'^2}.
\end{aligned} \qquad (4.70)$$

With $hB_y = -\frac{\partial A_z h}{\partial x}$ and $hB_x = \frac{\partial A_z h}{\partial y}$ the equations of motion become finally in paraxial approximation

$$\begin{aligned}
x'' + \frac{ec}{cp_0}B_y h(1 - \delta) - \kappa_{0x} &= 0, \\
y'' - \frac{ec}{cp_0}B_x h(1 - \delta) - \kappa_{0y} &= 0.
\end{aligned} \qquad (4.71)$$

These are the equations of motion in curvilinear coordinates under the influence of the magnetic field (B_x, B_y).

Problems

4.1 (S). Show that the Hamiltonian transforms like $\mathcal{H}_\varphi = \frac{dt}{d\varphi} \mathcal{H}_t$, if the independent variable is changed from t to φ.

4.2 (S). Derive from the Lagrangian (4.25) the equation of motion.

4.3. Show that the transformations [a.), c.) for upper signs, d.) for $\epsilon = 0$] are canonical and [b.), c.) for lower signs, d.) for $\epsilon \neq 0$] are not:

a.) $\begin{aligned} q_1 &= x_1 \; p_1 = \dot{x}_1 \\ q_2 &= x_2 \; p_2 = \dot{x}_2 \end{aligned}$ b.) $q = r\cos\psi, \quad p = r\sin\psi$

c.) $\begin{aligned} q_1 &= x_1, \qquad p_1 = \dot{x}_1 \pm \dot{x}_2, \\ q_2 &= x_1 \pm x_2, \; p_2 = \dot{x}_2 \end{aligned}$ d.) $q = q_0 e^\epsilon, \quad p = p_0 e^\epsilon$

Show the formalism you use.

References

1. H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 1950)
2. L.D. Landau, E.M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1976)