HAMILTONIAN THEORY AS A TOOL FOR ACCELERATOR PHYSICISTS

H.L. Hagedoorn, J.I.M. Botman, W.J.G.M. Kleeven
Eindhoven University of Technology,
Cyclotron Laboratory, Eindhoven, The Netherlands

Abstract

Hamiltonian theory is treated for the purpose of the description of orbit motion in accelerators. The fundamental theory is presented leading to Lagrange's equations of motion, the canonical equations of motion, flow lines in phase space, canonical transformations, conservation laws, etc. The theoretical principles are applied to a number of practical cases.

INTRODUCTION

Classical Hamiltonian theory has proven to be an adequate tool for analytically solving complicated motions of charged particles in electromagnetic fields as exist in accelerators.
The main subject of these lectures is not to go into details of the foundations of classical mechanics but more to show that the results can be used fruitfully. For acquiring deeper insight a number of excellent books is available [1-6]. Furthermore the contribution of Bell [7], proceedings of the CERN Accelerator School 1985, should be read.

As a rather vague remark one can state that the Hamiltonian theory can be seen by the user as a convenient way to "manage" the mathematics leading to a deep insight in the phenomena and yielding practical model descriptions for the particle orbits. In this respect one should remember that the matrix theory for ion optics, which in itself is a consequence of the Hamiltonian theory, behaves roughly in the same way: simple representations and meaningful quantities as matrix elements, as well as a well organized way of calculation.

The Hamiltonian theory should not be seen as a replacement of numerical methods: for orbits in realistic fields it gives understanding but generally not exact solutions in a simple way. It helps in finding errors in the numerical programs, in finding the right way of solving the

equations, in indicating what the right presentation of output quantities should be, etc.

In the first section Hamilton's principle, Lagrange's equations of motion and phase space properties will be treated. The second section deals with applications of the variational integral (circulation integral, principle of least action, etc.). The harmonic oscillator will be used as a representation for an ion optical lens. In part 3 the basic principles will be applied to explain the canonical transformations. Finally in part 4 some examples of transformations are shown together with the use of action and angle variables and a qualitative indication of the adiabatic invariance of the action integral.

1. BASIC PRINCIPLES AND CONSEQUENCES

1.1. Lagrange equations of motion

In real space a function $L(q_k,\dot{q}_k,t)$, called the Lagrangian, is assumed, where q_k means a coordinate, $\dot{q}_k$ the time derivative (i.e. the velocity) and t the time. The index k indicates that one deals with all relevant degrees of freedom.

From variational calculus it is found that if the variation of the line integral for fixed end points and times is zero, i.e.:

$$\delta \int_1^2 L(q_k,\dot{q}_k,t)dt = 0 \ , \tag{1}$$

then

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \ . \tag{2}$$

The variations may be taken between two fixed points (space, time), see Fig. 1.

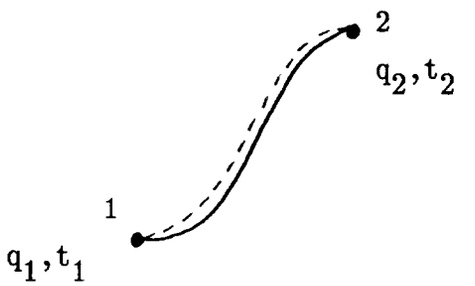

Figure 1. A trajectory in real space between two fixed points and times and a small variation.

Until now we have not said anything about physics. However, the equations of motion in classical mechanics follow the equations (2), the Lagrange equations of motion. Consequently the quantity L has a physical meaning; L is the Lagrange function or the Lagrangian. Equation (1) is called "Hamilton's principle", while the integral is called "Hamilton's principal function". In the case of motion in a potential field, L has a simple form

$$L = T - V \ , \tag{3}$$

where T is the kinetic energy and V the potential energy.

The equations have a representation which is coordinate independent and thus may lead to general results and also to an easy choice of a coordinate system that fits the problem in the best way.

If the end points are not fixed but varied with δq_1, δq_2 one gets

$$\delta \int_{t_1}^{t_2} L dt = \left[\frac{\partial L}{\partial \dot{q}_k} \delta q_k\right]_1^2 .$$

1.2. Hamiltonian function

The Lagrange equations of motion can be rewritten with the definition:

$$p_k = \frac{\partial L}{\partial \dot{q}_k} \ , \tag{4}$$

to

$$\frac{d}{dt} p_k - \frac{\partial L}{\partial q_k} = 0 \ , \tag{2a}$$

where p_k is called the canonical momentum. Note that $\frac{\partial L}{\partial q_k}$ acts as a gradient of a potential function, i.e. as a kind of force.

A function $H(p_k,q_k,t)$, called the Hamiltonian function or the Hamiltonian, is defined as:

$$H(p_k,q_k,t) = p_k\dot{q}_k - L(q_k,\dot{q}_k,t). \tag{5}$$

(Note that in a product like $p_k\dot{q}_k$, we mean the summation over the index k. Sometimes even the index will be omitted; at places where one dimension is

used, it will be explicitly mentioned.)

On the left hand side of Eq. (5) we have a function of p_k, q_k, t, on the right hand side a function of $p_k, q_k, \dot{q}_k, t$ which gives some confusion. However, in the Hamiltonian formalism the (p_k, q_k) should be taken as the two independent variables in the same way as we have the $(q_k, \dot{q}_k)$ in the Lagrangian. The transformation from $(q_k, \dot{q}_k)$ to (p_k, q_k) with Eq. (4) as the relation between p_k and $(q_k, \dot{q}_k)$ is known as a Legendre transformation.

Taking variations at both sides of Eq. (5) we get

$$\delta H = \frac{\partial H}{\partial q_k} \delta q_k + \frac{\partial H}{\partial p_k} \delta p_k + \frac{\partial H}{\partial t} \delta t$$

$$= p_k \delta \dot{q}_k + \dot{q}_k \delta p_k - \frac{\partial L}{\partial q_k} \delta q_k - \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k - \frac{\partial L}{\partial t} \delta t.$$

We had already, by definition (Eq. (4)), $p_k = \frac{\partial L}{\partial \dot{q}_k}$. Therefore the coefficient of $\delta \dot{q}_k$ cancels. Further the coefficients of δq_k, δp_k, δt must be equal on both sides. This yields

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \frac{\partial L}{\partial q_k} = - \frac{\partial H}{\partial q_k}, \quad \frac{\partial H}{\partial t} = - \frac{\partial L}{\partial t}. \tag{6}$$

The Lagrangian equations (2a) show that $\dot{p}_k = \frac{\partial L}{\partial q_k}$. Thus

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = - \frac{\partial H}{\partial q_k}. \tag{7}$$

These are the canonical equations of motion. The Hamiltonian H represents the energy of the mechanical system. In the simple example of a particle in a potential V we have

$$H = \frac{p^2}{2m} + V.$$

Relativistically we have $H = W + V$,

where W the total energy $= (E_r^2 + p^2c^2)^{1/2}$, with E_r the rest energy of the

particle and c the velocity of light:

$$H = \{E_r^2 + p^2c^2\}^{1/2} + V.$$

Having a system which obeys the canonical equations, one finds a Hamiltonian function Eq. (7), via Eq. (5) a Lagrangian function and the Lagrangian equations of motion and thus also Hamilton's principle, the variational integral Eq. (1), is valid (this is illustrated in Fig. 2).

Canonical equations	(7)
↓ ↑	
Hamilton function	(7)
↓ ↑	
Lagrangian function	(5)
↓ ↑	
Lagrangian equation of motion	(2a)
↓ ↑	
Hamilton's principle	(1)

Figure 2. Formulations of the laws of mechanics.

Now Hamilton's principal function can be written as

$$\int_1^2 (p_k\dot{q}_k - H)dt = \int_1^2 (p_k dq_k - Hdt).$$

A variation of the endpoints as well as the end time thus yields

$$[p_k\delta q_k - H\delta t]_1^2.$$

1.3. Incorporation of a magnetic field

The Hamiltonian of a charged particle in an electromagnetic field can be written as

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + eV, \tag{8}$$

where $\mathbf{p}$ is the canonical momentum, $\mathbf{A}$ the vector potential of the magnetic field and V the electric potential ($\mathbf{p}$ and $\mathbf{A}$ are vector quantities). As the kinetic momentum $\mathbf{P}$ equals $(\mathbf{p} - e\mathbf{A})$, the Hamiltonian function still represents the total energy. The derivation of Eq. (8) will be done by first

setting up the equations of motion and then writing these in such a way that they are in the canonical form. The Hamiltonian then follows directly.

At first a simple case is treated where the vector potential of the magnetic field is represented by one component only: A_x. The components of the magnetic induction now are

$$B_z = - \frac{\partial A_x}{\partial y} , \qquad B_y = \frac{\partial A_x}{\partial z} .$$

Further, as $\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t} = - \frac{\partial}{\partial t} (\nabla \times \mathbf{A})$, where $\mathbf{E}$ is the electric field strength generated by a time variation of the magnetic induction, we have

$$\mathbf{E} = - \frac{\partial \mathbf{A}}{\partial t} .$$

The equations of motion become:

$$\dot{x} = \frac{P_x}{m} , \qquad \dot{y} = \frac{P_y}{m} , \qquad \dot{z} = \frac{P_z}{m} ,$$

$$\dot{P}_x = - \frac{e\partial V}{\partial x} + \frac{e}{m} (P_y B_z - P_z B_y) - \frac{e\partial A_x}{\partial t} ,$$

$$\dot{P}_y = - \frac{e\partial V}{\partial y} + \frac{e}{m} (- P_x B_z),$$

$$\dot{P}_z = - \frac{e\partial V}{\partial z} + \frac{e}{m} (P_x B_y).$$

Substituting in the last three equations $\dot{y}$ for $\frac{P_y}{m}$ etcetera and $B_z = - \frac{\partial A_x}{\partial y}$ etc. and rewriting the magnetic field part

$$\frac{e}{m} (P_y B_z - P_z B_y) - e \frac{\partial A_x}{\partial t} =$$

$$- e (\dot{y} \frac{\partial A_x}{\partial y} + \dot{z} \frac{\partial A_x}{\partial z} + \frac{\partial A_x}{\partial t}) =$$

$$- e (\dot{x} \frac{\partial A_x}{\partial x} + \dot{y} \frac{\partial A_x}{\partial y} + \dot{z} \frac{\partial A_x}{\partial z} + \frac{\partial A_x}{\partial t}) + e\dot{x} \frac{\partial A_x}{\partial x} ,$$

we get $\dot{P}_x = - e \frac{\partial V}{\partial x} - e\dot{A}_x + e \dot{x} \frac{\partial A_x}{\partial x}$, $\qquad P_x = m\dot{x}$,

where $\dot{A}_x$ is the total time derivative of A_x. We define a new (canonical) momentum $p_x = P_x + eA_x$. Then the two equations of motion in the x-dimension are

$$\dot{p}_x = - e \frac{\partial V}{\partial x} + e\dot{x} \frac{\partial A_x}{\partial x} ,$$

$$\dot{x} = \frac{p_x}{m} - \frac{eA_x}{m} ,$$

or substituting $\dot{x}$ in the first equation

$$\dot{p}_x = - e \frac{\partial V}{\partial x} + \frac{e}{m} (p_x - eA_x) \frac{\partial A_x}{\partial x} \stackrel{?}{=} - \frac{\partial H}{\partial x} ,$$

$$\dot{x} = \frac{p_x}{m} - \frac{eA_x}{m} = \frac{1}{m} (p_x - eA_x) \stackrel{?}{=} \frac{\partial H}{\partial p_x} .$$

These equations of motion follow from a Hamiltonian

$$H = \frac{1}{2m} (p_x - eA_x)^2 + \frac{1}{2m} p_y^2 + \frac{1}{2m} p_z^2 + eV.$$

The other two sets of equations (for y and z) also follow from this Hamiltonian.

Next the derivation in vector notation: $\mathbf{A},\mathbf{P},\mathbf{x}$ are vector quantities. Then

$$\frac{d\mathbf{P}}{dt} = - \nabla V - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{P} * (\nabla * \mathbf{A}), \qquad \frac{d\mathbf{x}}{dt} = \mathbf{P},$$

(where for simplicity e,m = 1 is taken). The operator * represents the cross product. With $\mathbf{P} * (\nabla * \mathbf{A}) = \nabla(\mathbf{P} \cdot \mathbf{A}) - \mathbf{P} \cdot \nabla\mathbf{A}$ it follows that

$$\frac{dP}{dt} = - \nabla V - \underbrace{\frac{\partial A}{\partial t} - P \cdot \nabla A}_{-\dot{A}} + \nabla(P \cdot A),$$

$$P = \frac{dx}{dt} .$$

Take a new momentum $p = P + A$.

$$\frac{dp}{dt} = \frac{d(P+A)}{dt} = - \nabla V + \nabla(P \cdot A) \stackrel{?}{=} - \nabla H, \qquad (9)$$

$$\frac{dx}{dt} = p - A \qquad \stackrel{?}{=} \nabla_p H, \qquad (10)$$

where by definition $\nabla_p H = (\frac{\partial}{\partial p_x} , \frac{\partial}{\partial p_y} , \frac{\partial}{\partial p_z})H$.

As a remark we note that in the old system $\nabla(P \cdot A)$ works on the vector component. In the new system we have to take

$$\nabla(P \cdot A) = \nabla((p - A) \cdot A) + \frac{1}{2} \nabla(A \cdot A). \qquad !$$

From Eq. (9) we find $H = V - (p - A) \cdot A - \frac{1}{2} A \cdot A + f(p)$. From Eq. (10) we find $H = \frac{1}{2} (p - A)^2 + g(x)$, where $f(p)$ and $g(x)$ are arbitrary functions of p and x respectively.
It follows that

$$H = V(x) + \frac{1}{2} (p - A)^2.$$

1.4. Coordinate system

Using the equations (1) and (5), Hamilton's principle can be written as

$$\delta \int(p_k \dot{q}_k - H)dt = 0,$$

$$\text{or} \quad \delta \int p_k dq_k \; - H dt \; = 0,$$

$$\text{or} \quad \delta \int \mathbf{p} \cdot \mathbf{ds} \; - H dt \; = 0.$$

From this variational integral all equations of motion are derived. The formulation is such that it is coordinate system independent. For cartesian coordinates we have

$$\mathbf{p} \cdot \mathbf{ds} = \; p_x dx + p_y dy + p_z dz;$$

for cylindrical coordinates

$$\mathbf{p} \cdot \mathbf{ds} = p_r dr + (\frac{p_\theta}{r})\; r d\theta + p_z dz;$$

for curvilinear coordinates (curvature in a plane)

$$\mathbf{p} \cdot \mathbf{ds} = p_x dx + (\frac{p_s}{1+kx})\; (1 + kx)\; ds + p_z dz.$$

Setting up the Hamiltonian one has to realize that the components of the vector $\mathbf{p}$ are $(p_r, \frac{p_\theta}{r}, p_z)$ and $(p_x, \frac{p_s}{1+kx}, p_z)$ in the case of the cylindrical and curvilinear coordinates. The coefficient k is the curvature $(= \frac{1}{\rho})$ and is generally a function of s.

The Hamiltonians for cylindrical and the above given curvilinear coordinates are (taking again m = e = 1)

$$H = \tfrac{1}{2}\, p_r^2 + \tfrac{1}{2}\, (\frac{p_\theta}{r})^2 + p_z^2 + V(r,\theta,z),$$

$$H = \tfrac{1}{2}\, p_x^2 + \tfrac{1}{2}\, \frac{p_s^2}{(1+kx)^2} + \tfrac{1}{2}\, p_z^2 + V(x,s,z),$$

and e.g. with a vector potential for the last representation

$$H = \tfrac{1}{2}\, (p_x - A_x)^2 + \tfrac{1}{2}\, (\frac{p_s}{1+kx} - A_s)^2 + \tfrac{1}{2}\, (p_z - A_z)^2 + V(x,s,z).$$

1.5. Some simple properties of the Hamiltonian

1.5.1 A purely time-dependent part in the Hamiltonian can be omitted.

$$H = H_1 + f(t),$$
$$\dot{q} = \frac{\partial H}{\partial p} = \frac{\partial H_1}{\partial p},$$
$$\dot{p} = -\frac{\partial H}{\partial q} = -\frac{\partial H_1}{\partial q}.$$

Thus f(t) has no meaning for these equations of motion and thus may be skipped. As a natural consequence a constant can also be skipped.

1.5.2 If H does not depend explicitly on time it is a constant (constant of motion or integral of motion):

$$H = H(p,q), \qquad \frac{dH}{dt} = \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial t},$$
$$\frac{\partial H}{\partial t} = 0 \text{ by assumption}, \quad \dot{q} = +\frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q},$$

thus $\frac{dH}{dt} = 0$, H = constant.

1.5.3 If a coordinate is not explicitly present in the Hamiltonian, then its conjugated momentum is a constant of motion:

Suppose q_i is not present

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = 0 \quad \rightarrow \quad p_i = \alpha = \text{constant}.$$

q_i is called a cyclic coordinate. In one degree of freedom we have

$$\dot{q}_i = +\frac{\partial H}{\partial p_i} = \frac{\partial H}{\partial \alpha} = \text{constant},$$

taking α as a parameter.

Therefore $q_i = Ct.$

1.5.4 If a Hamiltonian can be split in separate parts, each part depending only on one coordinate and its conjugate momentum, then each part can be treated as a separate Hamiltonian:

$$H = \sum_i H_i(p_i,q_i,t),$$

$$\dot{p}_i = - \frac{\partial H_i}{\partial q_i} , \quad \dot{q}_i = \frac{\partial H}{\partial p_i} .$$

1.5.5 For one degree of freedom a constant Hamiltonian means that the problem is solved:

$$H = H(q_i, p_i) = \text{constant}.$$

For each q_i, the corresponding p_i is found algebraically. The aim of the problem treatment in orbit dynamics is often to achieve a presentation for H that is explicitly time independent.

1.6 Scaling

The canonical equations of motion remain canonical if we scale the coordinates and momenta by a constant factor. For example $\bar{q} = \alpha q$ and $\bar{p} = p$. We have

$$\dot{q} = \frac{\partial H}{\partial p} , \qquad \dot{p} = - \frac{\partial H}{\partial q} ,$$

$$\dot{\bar{q}} = \alpha \dot{q} = \frac{\partial \alpha H}{\partial p} , \quad \dot{\bar{p}} = \dot{p} = - \frac{\partial \alpha H}{\partial \alpha q} .$$

Thus the new Hamiltonian $\bar{H}$ becomes

$$\bar{H} = \alpha H.$$

In the same way we find for $\bar{q} = q$, $\bar{p} = \beta p$ that $\bar{H} = \beta H$.

Scaling the independent variable also leads to a change of the Hamiltonian:

$$\bar{t} = \omega t \text{ then } \bar{H} = H/\omega.$$

So for $\bar{q} = \alpha q$, $\bar{p} = \beta p$, $\bar{t} = \omega t$ one finds

$$\bar{H} = \frac{\alpha \beta}{\omega} H.$$

As a special case one may put

$$d\bar{t} = \omega(t)\, dt,$$

then

$$\dot{q} = \frac{dq}{dt} = \omega(t)\,\frac{dq}{d\bar{t}} = \frac{\partial H}{\partial p}\ ,$$

and we have in the new time

$$\dot{q} = \frac{\partial H/\omega}{\partial p}\ , \qquad \dot{p} = -\,\frac{\partial H/\omega}{\partial q}\ ,$$

so

$$\bar{H} = \frac{H}{\omega(t)}\ .$$

This may sometimes be helpful if the revolution frequency depends on time (e.g. in a microtron).

1.7. Choice of the independent variable

The variational integral is symmetric in coordinates and time, in momenta and - H (minus H):

$$\delta \int p_k dq_k - H dt = 0.$$

The equations of motion follow from this variational principle. If we take - H as p_4 and t as q_4 we have

$$\delta \int p_\kappa dq_\kappa = 0, \qquad \kappa = 1,\ 2,\ 3,\ 4.$$

In fact any coordinate can be used as new independent variable ("time") and the negative of its conjugated momentum as the new Hamiltonian.

This property is especially useful for Hamiltonians that are not explicitly dependent on time: i.e. $\frac{\partial H}{\partial t} = 0$. Then

$$H = H(p,q) = \text{const.}$$

Solving algebraically for p_i we get

$$K = -\,p_i = -\,p_i(p,q,q_i,H).$$

From the Hamiltonian H we get a set of six differential equations. In the Hamiltonian K we only have a set of four equations, thus simplifying our problem. Often the new independent variable can be physically interpreted

such as, e.g., the position along a central trajectory (closed orbit, optical axis).

The procedure of choosing a new independent variable often already amounts to solving the problem altogether (an example is given in Section 2.4).

1.8. The meaning of a first degree part in the Hamiltonian

A first degree part in a Hamiltonian that can be described by a power series in the canonical variables has a close connection with a central trajectory around which small variations can be described. Let the Hamiltonian be (one degree of freedom)

$$H = a_{ij} x^i p^j \ , \qquad i,j = 0,1, \ldots\ldots$$

For simplicity the coefficients are taken as constants. In that case a special solution can quickly be given. However, it often occurs that the coefficients are time dependent, for example in circular accelerators with a certain periodicity. The derivation of a central orbit in that case does require some extra work (an example is given in Section 4.1).

As a first remark we note that a_{00} can be omitted. The first degree part is represented by

$$a_{10} x + a_{01} p.$$

A special solution is found from

$$\dot{x} = \frac{\partial H}{\partial p} = j a_{ij} x^i p^{j-1} = 0,$$

$$\dot{p} = - \frac{\partial H}{\partial x} = - i \, a_{ij} x^{i-1} p^j = 0.$$

From these equations x and p can be solved algebraically, which gives the special time independent solution x_0, p_0.

A new Hamiltonian can now be constructed without a first degree part. We take new variables

$$\bar{x} = x - x_0 \ , \ \bar{p} = p - p_0 \ .$$

Further $\bar{x}$ and $\bar{p}$ are substituted in H. The new first degree part in H is

$$\ddot{x} = \dot{p} = - a - \omega^2 x \ ,$$

or

$$\ddot{x} + \omega^2 x = - a,$$

being the differential equation for a free oscillation around the equilibrium point, $x_0 = - \frac{a}{\omega^2}$.

1.9. Phase space and flow lines

The coordinates q and the conjugated momenta p are represented in a six dimensional space, the phase space. The motion is represented by a flow line in this space. In any case two flow lines can not cross. This follows from the canonical equations

$$\dot{q} = \frac{\partial H}{\partial p} \ , \qquad \dot{p} = - \frac{\partial H}{\partial q} \ .$$

In any point the derivatives, i.e. the local direction of the flow lines, are unambiguously defined by the coordinates and momenta of that point.

The flow in phase space can be seen as the flow of an <u>incompressible fluid</u>. If the velocity in the six dimensional phase fluid is denoted by v_6, then

$$\nabla \cdot v_6 = \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \ .$$

As

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 H}{\partial q_i \partial p_i} \text{ and } \frac{\partial \dot{p}_i}{\partial p_i} = - \frac{\partial^2 H}{\partial q_i \partial p_i}$$

it follows that $\nabla \cdot v_6 = 0$, which shows that the flow behaves as an incompressible fluid.

<u>This is Liouville's theorem: The phase space volume remains constant</u>.

For a two-dimensional phase space (q_x, p_x) the study of flow lines reveals a lot about the orbit dynamical properties. Two flow lines lying near to each other will show a higher phase velocity if they come closer to each other, due to the incompressibility (see Fig. 3). By definition the phase space velocity is given by the canonical equations.

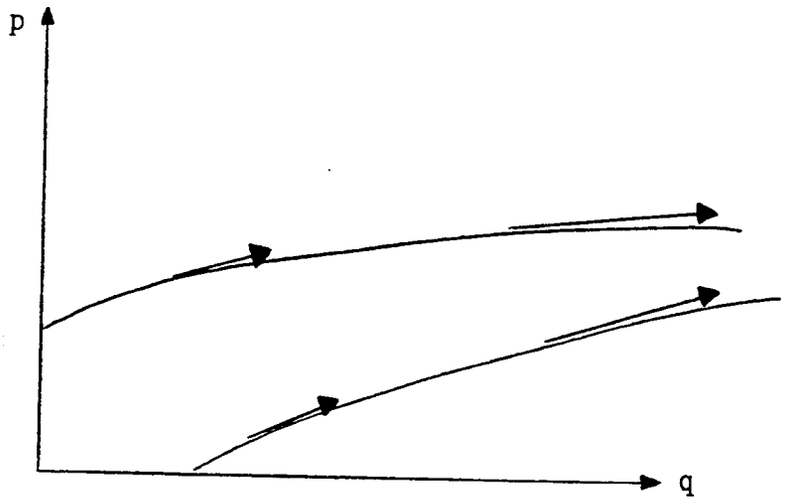

Figure 3. Two flow lines in phase space.

The study of flow lines shows whether a motion is stable or unstable; the stability region and the stable and unstable fixed points can easily be found. A few simple examples illustrate this.

1) $H = \frac{1}{2} p_x^2 + \frac{1}{2} \omega^2 x^2$, ellipse, stable motion.

Scaling $\bar{x} = \omega x$

$\bar{H} = \frac{1}{2} \omega \bar{p}_x^2 + \frac{1}{2} \omega \bar{x}^2$, circle, stable motion.

Scaling time $\bar{t} = \omega t$

$\bar{H} = \frac{1}{2} \bar{p}_x^2 + \frac{1}{2} \bar{x}^2$, circle, stable motion, unit frequency.

2) $\bar{H} = \frac{1}{2} \bar{p}_x^2 - \frac{1}{2} \bar{x}^2$, hyperbola, unstable motion,

as $\bar{H}$ = constant, only the difference of $\bar{p}^2$ and $\bar{x}^2$ is constant. They both can grow to infinity.

3) $H = \frac{1}{2} (\nu-1)(x^2 + y^2) + D(x^3 - 3xy^2)$,

where x and y are the coordinate and canonical momentum respectively (for cyclotrons they represent the coordinates of the orbit centre). This

Hamiltonian describes a third-order resonance excited by a third-harmonic perturbation (used e.g. for extraction purposes in synchrotrons). The flow lines are given in Fig. 4.

(1) Stable fixed point
(2) (3) (4) unstable fixed points

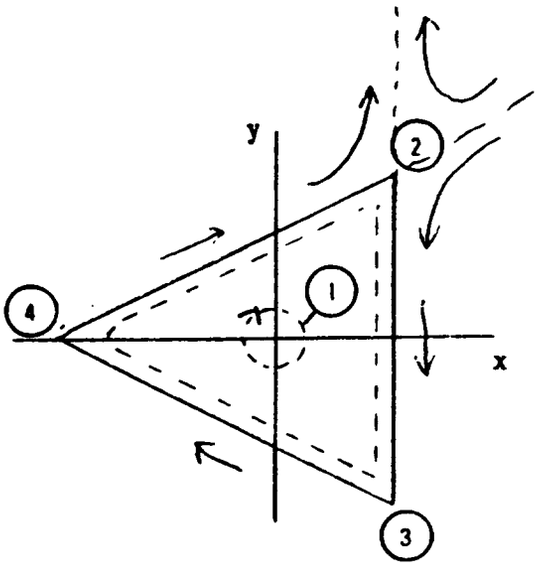

Figure 4. Flow lines, separatrix and fixed points in the case of a third-order resonance.

The flow lines are triangularly shaped. We see a stable region, where the flow lines are closed lines and an unstable region where they go to infinity (in reality higher degree terms in a Hamiltonian prevent this; however, these are not incorporated in the above given H).
First : for small values of x,y the second-degree term is most important, showing a harmonic oscillator motion which is stable (region 1).
Second : there are four fixed points. These are found by putting

$$\frac{\partial H}{\partial x} = 0, \qquad \frac{\partial H}{\partial y} = 0.$$

The question arises which point is stable:
For a stable fixed point we must be able to find a closed curve around it with a value of H which is for all points on it slightly larger, than the value of H in the fixed point, or for all points on it slightly smaller. So in a three dimensional landscape a stable point is either the top of a mountain or the lowest point of a valley. For an unstable fixed point one observes a saddle shape. If $\frac{\partial^2 H}{\partial x^2}$ and $\frac{\partial^2 H}{\partial y^2}$ have different signs, then surely there is an unstable fixed point. If these second derivatives both have

the same sign it is not absolutely sure that the point is a stable fixed point.

In any case in the example given above (see Fig. 4) point (1) is a stable fixed point. Points (2), (3) and (4) are unstable fixed points.

The coordinates of the fixed points are

$$\begin{aligned}
&(1)\ x = 0, && y = 0,\\
&(2)\ x = \frac{\nu-1}{6D}, && y = \frac{\nu-1}{2\sqrt{3}D},\\
&(3)\ x = \frac{\nu-1}{6D}, && y = -\frac{\nu-1}{2\sqrt{3}D},\\
&(4)\ x = -\frac{\nu-1}{3D}, && y = 0.
\end{aligned}$$

Third : the boundary between the stable and unstable regions is called the separatrix: In the figure it is the triangle. The equation of the separatrix is found by substituting in the Hamiltonian the values of x and y belonging to an unstable fixed point. Then the value of H is found and one gets the equation for the separatrix

$$H(x_i,y_i) = \text{constant} = H(x,y).$$

In the example it follows that

$$(x + ay + b)(x - ay + b)\ (x - \frac{b}{2}) = 0,$$

where $a = \sqrt{3}$, $b = \frac{\nu-1}{3D}$. This equation represents the three straight lines of the separatrix.

2. THE VARIATIONAL INTEGRAL AND SOME APPLICATIONS

2.1. The circulation theorem

A number of particles start at t_1 all lying on a closed curve. We follow those particles until t_2 where they again will lie on a closed curve (see Fig. 5). From Hamilton's principle we know that for each particle the integral

$$S = \int_1^2 (pdq - Hdt),$$

is stationary.

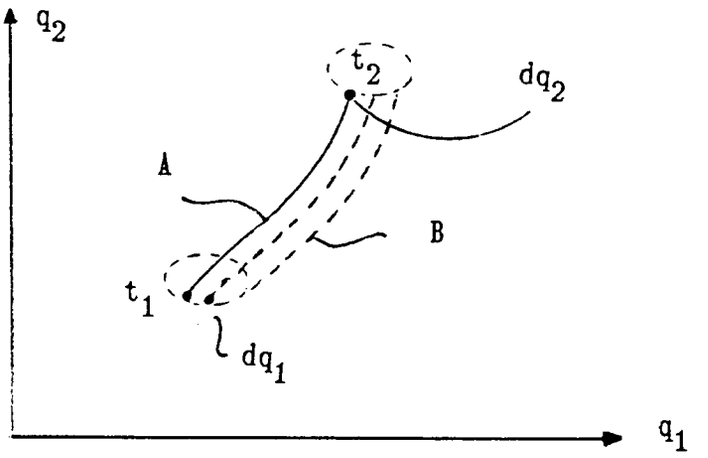

Figure 5. Particle trajectories in real space, starting and ending on a closed curve.

We now make a variation of the starting point, keeping the time t_1 fixed. Because t does not enter in this variation we have

$$S_B = S_A + p_2 dq_2 - p_1 dq_1,$$

for two nearby trajectories.

Moving the start point along a closed curve at t_1 until a trajectory N which coincides with the first trajectory A we have

$$S_N = S_A + \left(\oint pdq\right)_2 - \left(\oint pdq\right)_1,$$

and by definition $S_N = S_A$.

Therefore $\oint pdq$ = constant as a function of time. This is the circulation theorem. In fact the variation dt may be applied also, as long as we come back to the original point and time: $\oint (pdq - Hdt)$ = constant.

The quantity $S = \int_1^2 pdq$ is called the action integral, or quite often 'the

now given by

$$a_{10}\bar{x} + i\, a_{ij} x_0^{i-1} p_0^j\, \bar{x} + a_{01}\bar{p} + j\, a_{ij} x_0^i\, p_0^{j-1}\bar{p}\ ,$$

in which the original first degree part has been taken out of the summation. By the special solution of the canonical equations this first degree part equals zero. The new Hamiltonian $H(\bar{x},\bar{p})$ does not have a first degree part. Inspection of the canonical equations of motion following from $H(\bar{x},\bar{p})$ shows that a solution $\dot{\bar{x}} = 0 = \dot{\bar{p}} = \bar{x} = \bar{p}$ exists. So all solutions are now described around the central one (x_0, p_0).

As said before, time dependent coefficients give some extra work: the simple statement $\dot{x} = 0 = \dot{p}$ for finding x_0 and p_0 does not work. Also the new Hamiltonian can not be found by simply substituting $x = \bar{x} + x_0(t)$ and $p = \bar{p} + p_0(t)$. Canonical transformations (see section 4.1) should be applied.

(In the treatment above $\dot{\bar{x}} = \dot{x} = \dfrac{\partial H(x,p)}{\partial p} = \dfrac{\partial H(\bar{x} + x_0,\ \bar{p} + p_0)}{\partial \bar{p}}$; therefore if x_0, p_0 are constants, the new Hamiltonian is found by simply substituting $x = \bar{x} + x_0$, $p = \bar{p} + p_0$ in the old one.)

As a simple example we now consider the following Hamiltonian

$$H = \alpha x + \frac{1}{2} p^2 + \frac{1}{2} \omega^2 x^2.$$

The equations of motion are $\dot{x} = p$, $\dot{p} = -\alpha - \omega^2 x$. The special solution $\dot{x} = 0 = \dot{p}$ yields $x_0 = -\dfrac{\alpha}{\omega^2}$, $p_0 = 0$, and the new Hamiltonian becomes

$$H = \frac{1}{2} \bar{p}^2 + \frac{1}{2} \omega^2 \bar{x}^2,$$

where a constant term $\dfrac{\alpha^2}{2\omega^2}$ has been omitted. Combining the two original first order differential equations we have

optical path length' (see Section 2.3). (Sometimes also $\int_1^2 (pdq - Hdt)$ is called the action integral.)

2.2. An application of the circulation theorem

We assume a magnetic field B(r). At t_1 we have a number of ions lying on a circle with radius $r_1 = p_1/(eB(t_1))$ with p_1 the kinetic momentum. We are interested in the radius r_2 at t_2 if we very slowly change the magnetic field. At first one would think

$$r_2 = \frac{p_1}{eB(t_2)} ,$$

thus in case of a homogeneous field, one would expect for a decrease in B by a factor 2 an increase of r by a factor 2. This is wrong because of the betatron acceleration which always has its influence, no matter how slowly we change the magnetic field.

From the circulation theorem we know that $\oint p_{can} ds$ is a constant, where p_{can} is the canonical momentum:

$$p_{can} = p_{kin} - eA.$$

Starting all particles on a circle, there is no reason why the particles should not lie on a circle at the end of the time interval during which we changed the magnetic field. Thus

$$\oint_{(t_1)} p_{can} ds = \oint_{(t_2)} p_{can} ds,$$

$$\oint_{(t_1)} (p_{kin} - eA) ds = \oint_{(t_2)} (p_{kin} - eA) ds,$$

$$2\pi r_1 p_1 - e\phi_1 = 2\pi r_2 p_2 - e\phi_2 = \text{constant},$$

where ϕ is the enclosed magnetic flux. For a homogeneous field $\phi = \pi r^2 B$, so that with p = eBr we have

$$\pi r^2 B = \text{constant}.$$

Thus by a factor two decrease in B the enclosed area of the particles increases with a factor 2 and the radius $\sqrt{2}$.

Thus $$r_2 = r_1 \sqrt{\frac{B_1}{B_2}} = \frac{p_1/e}{\sqrt{B_1 B_2}}.$$

The betatron condition follows if we require the radius of the particle orbit to be constant for all values of the magnetic field. For very small values ($\rightarrow$ zero) the flux will go to zero and also the kinetic momentum. Thus the circulation integral is zero

$$2\pi r p - e\phi = 0.$$

With p = eBr we have $$B = \frac{1}{2\pi r^2}\phi,$$

which shows the betatron condition, stating that the magnetic induction at the orbit equals half the average induction over the enclosed area.

2.3. The principle of least action and focusing

If the Hamiltonian is time independent we may state Hamilton's principle as

$$S = \int_1^2 p\,dq = \text{stationary}.$$

S is called the action integral and has a minimum value as long as there is no image of the two end points in between them (principle of least action). In this way it equals the statement in optics that a light ray will follow a path that has the least optical path length (Fermat's principle).

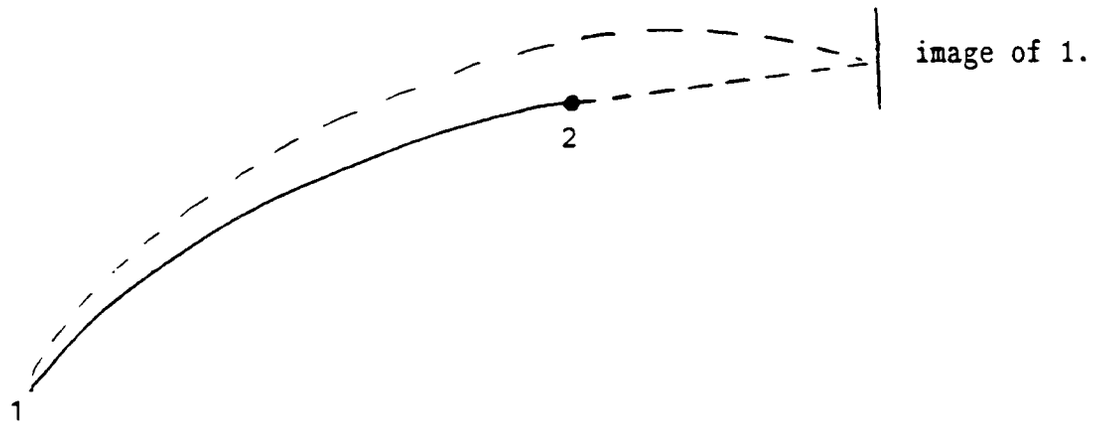

Figure 6. Two trajectories in real space.

In the case of imaging we know that many trajectories start from a source point and come together at the image point (see Fig. 6). Thus all trajectories have the same optical path length. In Fig. 7 two trajectories are drawn crossing a magnetic field. From the action integral we know that, if a and b start at (1) and come together at (2), the optical path lengths are equal. Of course we have to consider nearby trajectories.

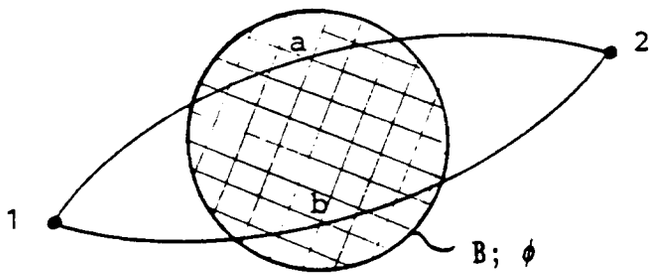

Figure 7. Two trajectories in real space, enclosing magnetic flux.

Thus

$$S_a = S_b,$$

$$\int_a (p_{kin} - eA)ds = \int_b (p_{kin} - eA)ds.$$

Because we assumed no explicit time dependency the kinetic momentum is a constant. It follows that

$$\int_a p_{kin} ds - \int_b p_{kin} ds = e \int_a A ds - e \int_b A ds,$$

$$\oint p_{kin} ds = e \oint A ds,$$

$$p_{kin} \Delta l = e\phi,$$

where Δl is the difference in geometrical pathlength between a and b and ϕ the enclosed flux. A general focusing condition therefore is

$$\Delta l = \frac{\phi}{p/e} .$$

Let us apply the general focusing condition on a thin lens (see Fig. 8).

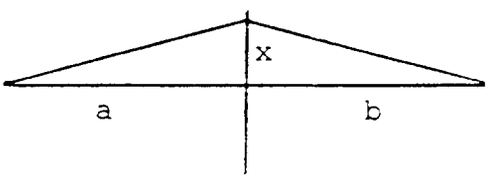

Figure 8. Axial and off-axis trajectory through a thin lens.

$$\Delta l = (x^2 + a^2)^{1/2} + (x^2 + b^2)^{1/2} - a - b,$$

$$= a(1 + \frac{x^2}{a^2})^{1/2} + b\ (1 + \frac{x^2}{b^2})^{1/2} - a - b,$$

for $x \ll a,b$

$$\Delta l = \frac{1}{2} x^2\ (\frac{1}{a} + \frac{1}{b}) = \frac{1}{2} x^2 \cdot \frac{1}{F} ,$$

with F the focal length.

Therefore

$$\frac{1}{F} = \frac{\phi/(\frac{1}{2} x^2)}{p/e} .$$

For a quadrupole with length L

$$\phi = \frac{1}{2} L \frac{dB}{dx} x^2 , \qquad \frac{1}{F} = \frac{L \frac{dB}{dx}}{p/e} .$$

For a sector magnet with an angle α

$$\phi = \frac{1}{2}\,\alpha x^2 B, \quad \frac{1}{F} = \frac{\alpha B}{p/e} = \frac{\alpha}{R}\ ,$$

with R the radius of curvature.

2.4. An orbit dynamical example

Ions moving in a cylindrically symmetrical magnetic and electric field will be considered. The magnetic field is pointing in the z-direction, the electric field has a radial component. Thus in the median plane we have $B_z \neq 0$, $B_r = B_\theta = 0$ and $E_r \neq 0$, $E_z = E_\theta = 0$. In this problem several of the treated concepts will be used:

- Choice of the independent variable (Section 1.7)
- Elimination of a first degree part (Section 1.8)
- Scaling (Section 1.6)

The magnetic field can be represented by one component of a vector potential: $A_\theta(r)$. The Hamiltonian in cylindrical coordinates becomes

$$H = \frac{1}{2m}\left\{ \left(\frac{p_\theta}{r} - eA_\theta\right)^2 + p_r^2 + p_z^2\right\} + eV(r)\ .$$

The Hamiltonian is time independent, therefore it is useful to choose a new independent variable. The Hamiltonian H is equal to the total energy of the particle. We put $H = \frac{p_0^2}{2m}$. For the new independent variable we take the azimuth θ. Thus the new Hamiltonian is

$$K = -\,p_\theta = -\,r\,\{p_0^2 - p_r^2 - 2\,meV(r)\}^{1/2} - eA_\theta r,$$

omitting for simplicity the vertical motion. Observe that p_r^2, $2\,meV(r) \ll p_0^2$ (the electric field is taken as a small correction for isochronism or for focusing).

The second step is scaling of the coordinates and momenta:

$$x = \frac{r - r_0}{r_0}\ , \qquad r = r_0\,(1+x), \qquad \pi = \frac{p_r}{p_0}\ ,$$

in which r_0 is a reference radius, not yet defined.

Then $$\bar{K} = \frac{K}{r_0 p_0} .$$

Furthermore we define a normalized, dimensionless electric potential $\phi(r)$ by:

$$\frac{2\ meV(r)}{p_0^2} = \phi(r).$$

The electric potential is expanded around the reference radius r_0:

$$\phi = \phi_0 + x\phi_0' + \frac{1}{2} x^2 \phi_0''$$

where $\phi_0 = \phi(r_0)$, $\phi_0' = r_0(\frac{d\phi}{dr})_{r_0}$, etc. The magnetic induction is given by $B = B_0(1 + nx)$, with n the field index. A vector potential is defined by

$$eA_\theta r = - er_0^2 B_0 \ (x + \frac{1}{2} (n + 1) \ x^2).$$

(This representation is chosen such that for $x = 0$, A_θ equals zero. One has to note that this is only allowed for time independent fields!)

The third step is an expansion of the square root, keeping terms up to the second degree and taking $\phi_0 = 0$, the potential at $r = r_0$. The Hamiltonian becomes

$$K = \frac{1}{2} \pi^2 + \frac{1}{2} x^2 \ (1 + n + \delta + \phi_0' + \frac{1}{4} \ (\phi_0')^2 + \frac{1}{2} \phi_0'' - x + (1 + \delta)x + \frac{1}{2} x \ \phi_0',$$

in which $\frac{er_0B_0}{p_0} = 1 + \delta$; due to the electric field the radius of the orbit will not be given by $er_0B_0 = p_0$.

The fourth step is the choice of δ such that there is no first degree part.

$$\delta = - \frac{1}{2} \phi_0'.$$

Thus $$K = \frac{1}{2} \pi^2 + \frac{1}{2} x^2 \ (1 + n + \frac{1}{2} \phi_0' + \frac{1}{2} \phi_0''),$$

in which $\frac{1}{4} \ (\phi_0')^2$ being small, has been omitted. The problem is now solved; it shows a harmonic oscillator with frequency

$$\nu_x^2 = 1 + n + \frac{1}{2}\phi_0' + \frac{1}{2}\phi_0''.$$

n → normal magnetic focusing

$\frac{1}{2}\phi_0'$ → electric focusing, equivalent to the magnetic focusing as it gives the radial derivative of the electric field strength

$\frac{1}{2}\phi_0''$ → a special term, which may be seen as an "energy focusing". It is due to the fact that particles loose or gain energy if they move radially in the electric field. This causes a different radius of curvature outside the equilibrium orbit. The same type of focusing occurs in the Wien filter.

2.5. A special feature of the harmonic oscillator

The Hamiltonian for a particle trajectory is quite often represented by

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 x,$$

where the independent variable is the pathlength s. Further $\omega = \omega(s)$. The equations of motion are

$$\dot{x} = p, \quad \dot{p} = -\omega^2 x.$$

It follows that

$$x + \omega^2 x = 0.$$

For a small path element we have $\Delta\dot{x} = -\omega^2 x \Delta s$.
If $\omega^2(s)$ only deviates from zero over a small path length where x is not varying significantly, we have

$$\Delta\dot{x} = -x \cdot \int \omega^2 ds \equiv -\frac{x}{F}.$$

The lens strength is given by

$$\frac{1}{F} = \int \omega^2 ds \quad \text{(thin lens approximation)}.$$

Thus for a system which is described by the simple Hamiltonian given above, one may always consider that system as a lens (thus following the approximate derivation above as a combination of many thin lenses, thus as an optical system). The second-order differential equation leads to the ion optical linear matrix theory, which is not treated further here.

3. CANONICAL TRANSFORMATIONS

3.1. Canonical transformations

The canonical transformations are in many cases a powerful method for getting analytical solutions of rather complicated particle trajectories. The aim generally is to transform the coordinates and momenta, often via several steps, such that the final Hamiltonian is time independent, easy to understand (e.g. via the study of flow lines in phase space) and/or shows originally coupled motion as a system where uncoupling exists for well chosen variables. In all cases it is wise to construct the trans- formations in such a way that the resulting new variables still have some physical meaning, as close as possible to realistic practical quantities.

The Hamiltonian equations follow from Hamilton's principle

$$\delta \int_1^2 L(q_k,\dot{q}_k,t)\ dt = 0.$$

In a transformed system we want to have the same principle

$$\delta \int_1^2 \bar{L}\ (\bar{q}_k,\dot{\bar{q}}_k,t)dt = 0.$$

The difference between the two Lagrangians must be a function that does not depend on the special path between the points (1) and (2). So we may put

$$L(q,\dot{q},t) - \bar{L}\ (\bar{q},\dot{\bar{q}},t) = \frac{d\phi(q,\bar{q},t)}{dt}\ .$$

The integral $\int_1^2 \frac{d\phi}{dt}\ dt$ only depends on the initial and final coordinates only and thus

$$\bar{L} = L - \frac{d\phi}{dt}\ ,$$

satisfies Hamilton's principle.

Let us make small variations δq, $\delta\dot{q}$, $\delta\bar{q}$, $\delta\dot{\bar{q}}$. These variations are not independent of each other (only two of them can be chosen freely).

$$\frac{\partial L}{\partial q}\,\delta q+\frac{\partial L}{\partial \dot{q}}\,\delta \dot{q}+\frac{\partial L}{\partial t}\,\delta t-\frac{\partial \bar{L}}{\partial \bar{q}}\,\delta \bar{q}-\frac{\partial \bar{L}}{\partial \dot{\bar{q}}}\,\delta \dot{\bar{q}}-\frac{\partial \bar{L}}{\partial t}\,\delta t=\frac{d}{dt}\left(\frac{\partial \phi}{\partial q}\,\delta q+\frac{\partial \phi}{\partial \bar{q}}\,\delta \bar{q}+\frac{\partial \phi}{\partial t}\,\delta t\right).$$

Knowing that $\quad p=\dfrac{\partial L}{\partial \dot{q}}\,, \qquad \bar{p}=\dfrac{\partial \bar{L}}{\partial \dot{\bar{q}}}\,,$

we find that, after equating the coefficients of the same variations, the coefficients for $\delta\dot{q}$, $\delta\dot{\bar{q}}$ yield

$$p=\frac{\partial \phi}{\partial q} \quad \text{and} \quad \bar{p}=-\frac{\partial \phi}{\partial \bar{q}}\,. \qquad \text{(Type 1)} \tag{11}$$

The other coefficients do not give new information, as the reader can find out for himself.

Both systems follow the Hamiltonian mechanics. Thus

$$\bar{H}(\bar{q},\bar{p},t)=\bar{p}\dot{\bar{q}}-\bar{L}=\bar{p}\dot{\bar{q}}-L+\frac{d\phi}{dt}=\bar{p}\dot{\bar{q}}-L+\frac{\partial \phi}{\partial q}\,\dot{q}+\frac{\partial \phi}{\partial \bar{q}}\,\dot{\bar{q}}+\frac{\partial \phi}{\partial t}\,.$$

$$\frac{\partial \phi}{\partial q}\downarrow = p \qquad \frac{\partial \phi}{\partial \bar{q}}\downarrow = -\bar{p}$$

Thus it follows that $\quad \bar{H}(\bar{q},\bar{p},t)=p\dot{q}-L+\dfrac{\partial \phi}{\partial t}\,,$

or $$\bar{H}=H+\frac{\partial \phi}{\partial t}\,(q,\bar{q},t). \tag{12}$$

The relations (11) and (12) together form the rules of the canonical transformation. Knowing ϕ the relation between the old and new variables follow from Eq. (11). The function ϕ is called "the generating function".

An easier method is writing the principal function as $\int (pdq - Hdt)$.

Then $$\int(pdq-Hdt)-\int(\bar{p}d\bar{q}-\bar{H}dt)=\int d\phi \quad , \quad \phi=\phi(q,\bar{q})$$

$$=\int\left(\frac{\partial \phi}{\partial q}\,dq+\frac{\partial \phi}{\partial \bar{q}}\,d\bar{q}+\frac{\partial \phi}{\partial t}\,dt\right).$$

Then $$p=\frac{\partial \phi}{\partial q}\,, \quad \bar{p}=-\frac{\partial \phi}{\partial \bar{q}}\,, \quad \bar{H}=H+\frac{\partial \phi}{\partial t}\,.$$

The following forms of ϕ also obey a canonical transformation:

$$\phi = \phi(q, \bar{p}, t), \text{ (Type 2)} \tag{13}$$

$$p = \frac{\partial \phi}{\partial q}, \qquad \bar{q} = \frac{\partial \phi}{\partial \bar{p}}, \qquad \bar{H} = H + \frac{\partial \phi}{\partial t}$$

$$\phi = \phi(p, \bar{q}, t), \text{ (Type 3)} \tag{14}$$

$$q = -\frac{\partial \phi}{\partial p}, \qquad \bar{p} = -\frac{\partial \phi}{\partial \bar{q}}, \qquad \bar{H} = H + \frac{\partial \phi}{\partial t}$$

$$\phi = \phi(p, \bar{p}, t), \text{ (Type 4)} \tag{15}$$

$$q = -\frac{\partial \phi}{\partial p}, \qquad \bar{q} = \frac{\partial \phi}{\partial \bar{p}}, \qquad \bar{H} = H + \frac{\partial \phi}{\partial t}.$$

The transformations keep the equations of motion Hamiltonian. Therefore Liouville is preserved in each system. Later we will see that the phase space volume is also conserved from one system to another system, transformed into each other by the above given canonical transformations.

3.2. Description of a particle trajectory by a canonical transformation

Suppose we have a trajectory as given in Fig. 9. At all times the system is canonical. Then if we know the equations describing the trajectory, we may always say that the variables at t_2 are a known function of those at t_1, the time interval $\tau = t_2 - t_1$, and t_1. So the transformation that relates the variables between a time $t_1 = t$ and a later time $t_2 = t + \tau$ should be a canonical transformation.

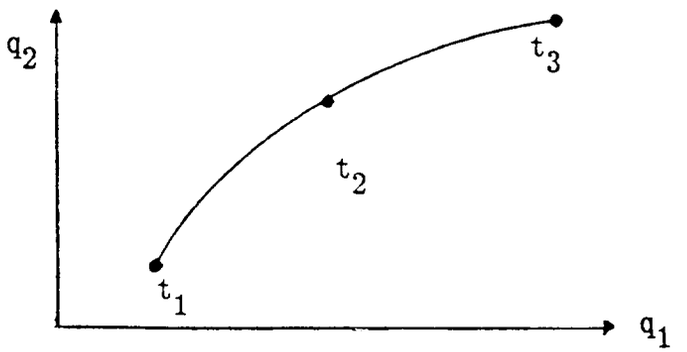

Figure 9. A particle trajectory in real space.

Therefore we must have

$$L(t + \tau) - L(t) = - \frac{d\phi}{dt} ,$$

$$\int_{t_0}^{t} [L(t + \tau) - L(t)]dt = - (\phi(t) - \phi(t_0)).$$

Now, $$\int_{t_0}^{t} L(t + \tau)dt = \int_{t_0}^{t} L(t)dt + \int_{t}^{t + \tau} L(t)dt - \int_{t_0}^{t_0 + \tau} L(t)dt.$$

Thus we have $$\int_{t}^{t + \tau} L(t)dt - \int_{t_0}^{t_0 + \tau} L(t)dt = - \phi(t) + \phi(t_0),$$

and we may choose $$\phi(t) = - \int_{t}^{t + \tau} L(t)dt.$$

The generating function $\phi(t)$ which yields the canonical transformation between the variables at t and those at $t + \tau$ equals Hamilton's principal function between the two times, t and $t + \tau$.

Starting with initial coordinates p_i, q_i we see that the final coordinates p_f,q_f are related with them via

$$p_i = \frac{\partial \phi}{\partial q_i} \qquad , \qquad p_f = - \frac{\partial \phi}{\partial q_f} .$$

So Hamilton's principle function - if known - describes the trajectory.
A slightly different reasoning starts from the variation of Hamilton's principal function $\int_{1}^{2} (p_k dq_k - Hdt)$, with τ the time interval.

We get $$p_2 dq_2 - H_2 dt_2 - (p_1 dq_1 - H_1 dt_1) = d\phi,$$

where $d\phi$ is the difference in value of the varied function w.r.t. the unvaried function. The quantities p_2, q_2, H_2 are the new momenta, coordinates and Hamiltonian coordinates. This has the same shape as was found for the construction of the generating function. Therefore ϕ, in this case $\int_{1}^{2} Ldt$ is the required generating function.

3.3. Lagrangian invariant

In Section 2.1 we have seen that the circulation integral is a constant during the motion

$$\oint p\,dq = \text{constant}.$$

The same constancy arises when transforming from one Hamiltonian system to a second one via the canonical transformations, which originated from

$$L(q,\dot{q},t) - \bar{L}(\bar{q},\dot{\bar{q}},t) = \frac{d\phi}{dt}\ .$$

Writing $$L = p\dot{q} - H$$

we find $$p\,dq - H\,dt - \bar{p}\,d\bar{q} - \bar{H}\,dt = d\phi.$$

As ϕ is a function which only depends on begin and end points we can state that

$$\oint (p\,dq - H\,dt) = \oint (\bar{p}\,d\bar{q} - \bar{H}\,dt) = \text{constant},$$

and thus invariant with respect to canonical transformations.
Keeping t fixed we have at that time:

$$\oint p\,dq \text{ is invariant w.r.t.canonical transformations.}$$

The circulation integral can be seen as an infinite summation of circulations around small areas (see Fig. 10). In this way it can be represented as a surface integral. Total circulation = sum of elementary circulations.

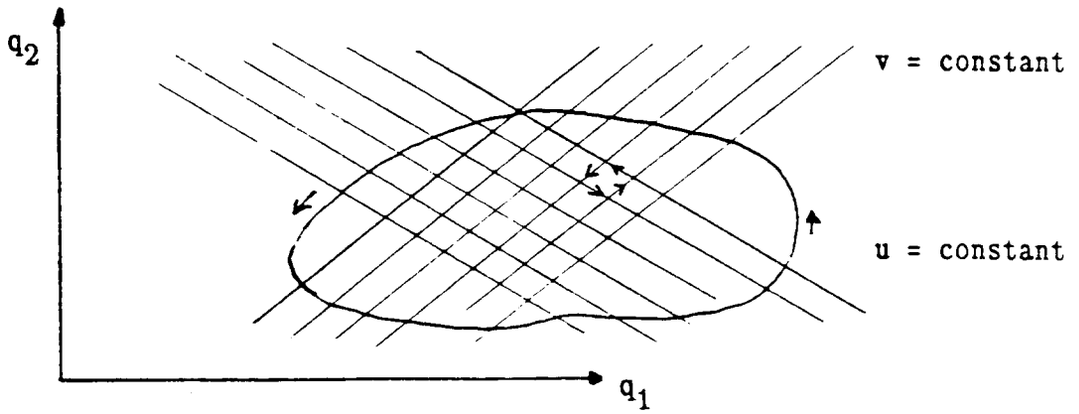

Figure 10. The circulation integral decomposed into small elementary circulations.

In Fig. 10 lines of constancy for two variables u and v are drawn. A variation δ_v, keeping v constant gives a variation in p and q:

$$\delta_v q_i \quad , \quad \delta_v p_i .$$

In the same way a variation in v yields

$$\delta_u q_i \quad , \quad \delta_u p_i .$$

It follows after some reflection that

$$\delta_v p_i \delta_u q_i - \delta_u p_i \delta_v q_i = \text{constant}, \tag{16}$$

and equals the elementary circulation. Integrating over the total area gives

$$\iint \left(\frac{\partial p_i}{\partial u}\frac{\partial q_i}{\partial v} - \frac{\partial p_i}{\partial v}\frac{\partial q_i}{\partial u}\right)dudv \quad = \text{constant},$$

and thus also

$$\left(\frac{\partial p_i}{\partial u}\frac{\partial q_i}{\partial v} - \frac{\partial p_i}{\partial v}\frac{\partial q_i}{\partial u}\right) = \text{constant}. \tag{17}$$

(Note that we mean summation over i, i = 1, 2, 3.) Equation (16) is called the "Lagrange bilinear invariant". The Lagrange invariant (Eq. (17)) is often written as

$$[v,u] = \sum_i \left(\frac{\partial p_i}{\partial u}\frac{\partial q_i}{\partial v} - \frac{\partial p_i}{\partial v}\frac{\partial q_i}{\partial u}\right),$$

[v,u] is called the "Lagrange bracket".
(Note: The Lagrange brackets are directly related to the Poisson brackets. The invariance of the one implies the invariance of the other.

Given a variable F the total time derivative of F is given by

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial p_i}\dot{p}_i \ .$$

If p and q are canonically conjugated variables the total time derivative may be rewritten as

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q_i} \ ,$$

or as

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + (F,H),$$

where $$(F,H) = \frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q_i} .$$

(F,H) is called the "Poisson bracket". As the total time derivate of F must be invariant with respect to the special choice of the system of canonically conjugated variables the Poisson bracket (F,H) must be invariant w.r.t. canonical transformations. As a consequence any (F,G) is invariant under canonical transformations as we can always interpret an arbitrary function F or G as a representation of a Hamiltonian function.

Having a number of dynamical variables u_k, we can construct a matrix (u,u) with the elements given by the Poisson brackets (u_i,u_j). In the same way a matrix [u,u] can be constructed using the Lagrange brackets $[u_i,u_j]$. The relation between these two matrices is given by

$$(u,u) = - [u,u]^{-1}.$$

Therefore the invariance of the Poisson brackets implies directly the invariance of the Lagrange brackets and reversely.)

3.4. Symplectic conditon

Taking two small variations x_u and x_v represented by column vectors

$$x_u = \begin{pmatrix} \delta_u q_1 \\ \delta_u p_1 \\ \delta_u q_2 \\ \delta_u p_2 \\ \delta_u q_3 \\ \delta_u p_3 \end{pmatrix} \qquad x_v = \begin{pmatrix} \delta_v q_1 \\ \delta_v p_1 \\ \delta_v q_2 \\ \delta_v p_2 \\ \delta_v q_3 \\ \delta_v p_3 \end{pmatrix}$$

and a square matrix η

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

the bilinear invariant is written as $\tilde{x}_u \eta x_\nu$ = constant, where $\tilde{x}$ means the transposed of x. Let us now make a canonical transformation to new variables. The variations in the new system are represented by primes: x'_u and x'_v. Having small variations there must be a linear relationship between the primed and unprimed quantities, expressed by a matrix M:

$$x'_u = Mx_u \quad , \quad x'_v = Mx_v.$$

The bilinear invariant in the new system equals

$$\tilde{x}'_u \eta x'_v = \tilde{x}_u \tilde{M} \eta M x_v = \tilde{x}_u \eta x_v.$$

From this follows $\tilde{M} \eta M = \eta.$

The matrix M is called symplectic. All canonical transformations show this symplectic behaviour. Thus also a transformation that shows the progress in time of the phase space for small variations from t_1 to t_2. Thus ion optical transfer matrices, as long as they use canonical conjugated variables, show this property. Mapping the phase space at t_1 to the phase space at t_2 by numerical methods must show symplecticity. This is a good check, together with conservation of phase-space volume, of the numerical results.

As $\tilde{M} \eta M = \eta,$

we have $\det \tilde{M} \eta M = \det \eta = 1,$

and therefore $\det M = 1.$

This shows that the phase volume remains constant after any canonical transformation (Liouville).

4. APPLICATIONS OF THE CANONICAL TRANSFORMATIONS, ACTION AND ANGLE VARIABLES

4.1. Removal of a first degree part in the Hamiltonian

A first degree part in the Hamiltonian often describes a situation in which a central orbit is not known beforehand. In the case of a synchrotron a change in the central orbit occurs if a small variation in the energy is present. Normally in this case the central orbit is a well known fixed orbit for a certain energy. In cyclotrons one really has to find central orbits by calculation (numerical or analytical).

Let us assume a simple Hamiltonian representing linear motion in a periodic accelerator such as a cyclotron or a synchrotron

$$H = A_1 \, x \sin n\theta + \frac{1}{2} p^2 + \frac{1}{2} \omega^2 x^2 + \ldots$$

We want to remove the first degree part of this Hamiltonian and try a transformation of the form

$$\tilde{x} = x - x_0(\theta), \qquad \tilde{p} = p - p_0(\theta).$$

The transformation is made via a generating function $\phi(p,\tilde{x},\theta)$.

Then

$$x = -\frac{\partial \phi}{\partial p} \quad \rightarrow \quad \phi = -\tilde{x}p - x_0(\theta)p + f(\tilde{x}),$$

$$\tilde{p} = -\frac{\partial \phi}{\partial \tilde{x}} \quad \rightarrow \quad \phi = -\tilde{x}p + p_0(\theta)\tilde{x} + g(p).$$

Thus a suitable ϕ is given by

$$\phi = -\tilde{x}p - x_0(\theta)p + p_0(\theta)\tilde{x},$$

and

$$\tilde{H} = H + \frac{\partial \phi}{\partial \theta} .$$

Note that $\phi = -\tilde{x}p$ results in the unit transformation. It is quite often advisable to start with this one and add a number of correction terms to it in such a way that the new variables are roughly equal to the old ones. The new Hamiltonian becomes

$$\tilde{H} = A_1\tilde{x} \sin n\theta + A_1 x_0 \sin n\theta + \frac{1}{2}\tilde{p}^2 + \tilde{p}p_0 + \frac{1}{2} p_0^2 + \frac{1}{2} \omega^2 \tilde{x}^2 + \omega^2 \tilde{x} x_0 + \frac{1}{2} \omega^2 x_0^2$$

$$- x_0'\tilde{p} + p_0'\tilde{x} + \ldots\ldots$$

with $$x_0' = \frac{dx_0}{d\theta} \quad \text{etc.}$$

The functions $x_0(\theta)$, $p_0(\theta)$ are chosen such that the first degree part vanishes.

Then $$\tilde{H} = \frac{1}{2}\tilde{p}^2 + \frac{1}{2}\omega^2\tilde{x}^2 \ .$$

The Hamiltonian is now explicitly time independent and thus constant. The problem is solved (see Section 1.5). The solution for x_0, p_0 is found by putting the coefficients of $\tilde{x}$ and $\tilde{p}$ zero:

Coefficient of $\tilde{x}$: $A_1 \sin n\theta + \omega^2 x_0 + p_0' = 0.$

Coefficient of $\tilde{p}$: $p_0 - x_0' = 0.$

Thus $$x_0'' + \omega^2 x_0 = A_1 \sin n\theta.$$

Substituting $x_0 = A_0 \sin n\theta + B_0 \cos n\theta$, one finds in this simple case:

$$B_0 = 0 \ , \quad A_0 = \frac{A_1}{n^2 - \omega^2} \ ,$$

and

$$x_0 = \frac{A_1}{n^2 - \omega^2} \sin n\theta, \qquad p_0 = \frac{nA_1 \cos n\theta}{n^2 - \omega^2} \ .$$

One must be careful about the influence of higher-degree terms in H, which also may have θ-dependent coefficients. Furthermore second-order constant (A^2) coefficients may arise in the equations yielding x_0 and p_0. These may be important if the circumference of an orbit has to be investigated as they directly have an influence on synchronism.

4.2. Removal of an oscillating part in the second degree of the Hamiltonian.

An oscillating part in the second degree arises when alternating focusing and defocusing elements are present along the particle orbit in a circular accelerator. The Hamiltonian may have a form like

$$H = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 \ (1 + A_n \sin n\theta) \ x^2.$$

Normally A_n is small and n large.

We try a transformation generated by $\phi = \phi(\tilde{x},p,\theta)$ with

$$\frac{\partial\phi}{\partial\theta}(\tilde{x},p,\theta) = -\frac{1}{2}\omega^2 A_n \sin n\theta\, \tilde{x}^2 .$$

At first sight this eliminates the oscillating term as long as $\tilde{x} = x$(note that $\bar{H} = H + \frac{\partial\phi}{\partial\theta}$). Therefore ϕ is now given by

$$\phi = -\tilde{x}p + \frac{1}{2}\frac{\omega^2}{n} A_n \cos n\theta\, \tilde{x}^2 .$$

Furthermore $$x = -\frac{\partial\phi}{\partial p} = \tilde{x} \quad \text{and} \quad \tilde{p} = -\frac{\partial\phi}{\partial\tilde{x}} = p - \frac{\omega^2}{n} A_n \cos n\theta\, \tilde{x} .$$

The result of the substitution of $\tilde{x}$ and $\tilde{p}$ in the Hamiltonian shows a cross term $\tilde{p}\,\tilde{x}$ with a periodic coefficient. Therefore try as a next step

$$\phi = -\tilde{x}p + F_1\tilde{x}p + F_2\tilde{x}^2 + F_3p^2 ,$$

with F_1, F_2, F_3 periodic functions in θ (period n). Some lengthy calculations follow, but it will prove to be possible to choose F_1, F_2, F_3 such that all first-order oscillating terms vanish.

A method used often to tackle the problem is the transformation to action and angle variables. The conjugated variables p,x are rapidly changing over a large range. Therefore new variables which give less variation will make it easier to proceed.

4.3. Action and angle variables

A system of one degree of freedom will be considered. For a closed curve, $S = \oint pdq$ will be constant and is equivalent to Liouvilles statement of conservation of area in the two dimensional phase space. In Fig. 11 a closed curve in phase space is given with particles on it. The enclosed area remains constant during time evolution. Suppose now that the closed line coincides with a flow line. We thus have an oscillatory motion.

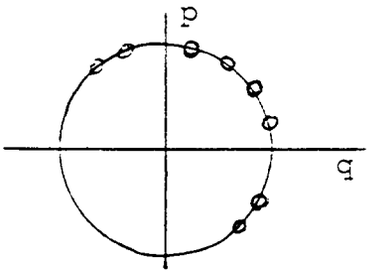

Figure 11. Particles on a closed curve in two-dimensional phase space.

Suppose now that an infinite number of particles lie on the flow line. They will all perform the oscillatory motion many times.

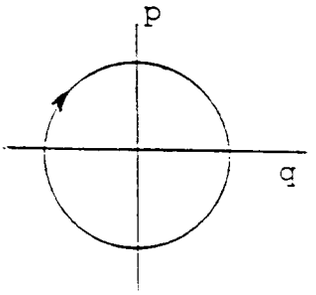

Figure 12. A flow line in phase space for an oscillatory motion.

At a certain time we start to change slowly a parameter due to which the flow line will change slowly. We follow one particle and observe it at a second time, where we stop the parameter change but where we still have an oscillatory behaviour, thus a closed flow line. The new flow line may be as given in Fig. 13.

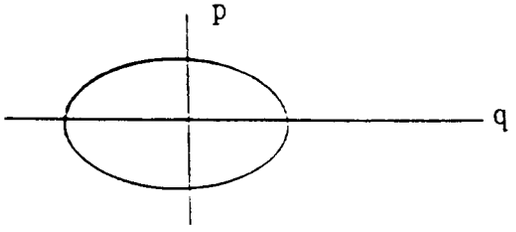

Figure 13. The flow line in phase space a long time after the situation presented in figure 12.

During the time interval many oscillations are performed by all particles. On the average they all will have experienced roughly the same variation of the slow parameter change (except for initial and final conditions). The result will be that all particles will approximately

remain on the flow line; the slower the change the more this becomes true. As the circulation integral remains constant we may state that also for each particle alone

$$J = \oint p\,dq = \text{constant}.$$

This integral is called an adiabatic constant, the variation an adiabatic variation. The quantity J is called "the action" variable.

For a harmonic oscillator the action equals

$$J = \oint p\,dq = \pi\omega A^2 = \text{constant}.$$

$$\left(H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 x^2, \qquad p = \dot{x}\ ,\ \dot{p} = -\omega^2 x\ ,\ x = A\sin\omega t,\right.$$

$$\left.\oint p\,dx = \text{Area} = \pi\omega A^2.\right)$$

The adiabatic invariance states that for a slow variation of ω with time the action J remains constant.

Thus the oscillation amplitude varies inversely proportional with the square root of ω

$$A \sim \omega^{-1/2}\ .$$

We now take J as a new canonical momentum which must be acquired after a canonical transformation. As J is a constant, the new Hamiltonian does not depend on its conjugate coordinate ϕ. Therefore

$$\bar{H} = \bar{H}(J).$$

We take $\bar{H}(J) = \omega J$ and try to find the transformation.

$$\bar{H} = H + \frac{\partial G}{\partial t} = \omega J\ , \qquad \dot{\phi} = \omega,$$

$$G = G(x,\phi), \qquad J = -\frac{\partial G}{\partial \phi}\ , \qquad p = \frac{\partial G}{\partial x}\ .$$

Thus we have to find a generating function which satisfies

$$H\left(\frac{\partial G}{\partial x}, x, t\right) + \frac{\partial G}{\partial t} = -\omega\,\frac{\partial G}{\partial \phi}\ .$$

If G is known, in fact p,x can be solved algebraically from J and ϕ, which

are a constant and a cyclic variable ($\phi = \omega t$). This procedure is the start for the Hamilton-Jacobi method.

In general for one degree of freedom we like to find a function G such that

$$H(\frac{\partial G}{\partial x}, x, t) + \frac{\partial G}{\partial t} = \bar{H}(J, \phi).$$

If we have succeeded in finding this function G, the problem is solved as $\bar{H}$ = constant and J, ϕ are algebraically related to each other.

In the following H is taken explicitly time independent and also the generating function is assumed to be time independent.

We take $$H = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 x^2 .$$

The value of J equals $$J = \oint p dx = \pi \omega A^2 ,$$

with A the oscillation amplitude. The new Hamiltonian is given by

$$H = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 x^2 = \frac{\omega J}{2\pi} = \bar{H} \quad , \quad \dot{\varphi} = \frac{\omega}{2\pi} \quad , \quad J = \text{constant} .$$

We further know $$J = - \frac{\partial G}{\partial \varphi} \quad , \quad p = \frac{\partial G}{\partial x} .$$

Therefore after substitution in the Hamiltonian:

$$- \frac{\omega}{2\pi} \frac{\partial G}{\partial \varphi} = \frac{1}{2} (\frac{\partial G}{\partial x})^2 + \frac{1}{2} \omega^2 x^2 .$$

Suppose $G = \phi(\varphi) \cdot X(x)$ (separation of variables), then

$$- \frac{\omega}{2\pi} X \phi' = \frac{1}{2} X'^2 \phi^2 + \frac{1}{2} \omega^2 x^2 ,$$

or

$$- \frac{\omega \phi'}{2\pi} = \frac{1}{2} \frac{X'^2}{X} \phi^2 + \frac{1}{2} \frac{\omega^2 x^2}{X} .$$

φ-dependent x-dependent

Take $X = \frac{1}{2} \omega^2 x^2$, $X' = \omega^2 x$, then $\frac{X'^2}{X} = 2\omega^2$.

Then $-\frac{\omega\phi'}{2\pi} = 1 + \omega^2\phi^2$ and $\phi = \frac{1}{\omega}\,\mathrm{cotg}\,2\pi\varphi$,

$$G = X\phi = \frac{1}{2}\,\omega x^2\,\mathrm{cotg}\,2\pi\varphi.$$

From this generating function the relations between the old and new variables follow:

$$p = \frac{\partial G}{\partial x} = \omega x\,\mathrm{cotg}\,2\pi\varphi, \quad J = -\frac{\partial G}{\partial \varphi} = \frac{\pi\omega x^2}{\sin^2 2\pi\varphi},$$

or

$$p = \sqrt{\frac{J}{\pi\omega}}\cos 2\pi\varphi, \qquad x = \sqrt{\frac{\omega J}{\pi}}\sin 2\pi\varphi, \quad \bar{H} = \frac{\omega J}{2\pi}.$$

Application of scaling $\bar{J} = \frac{J}{2\pi}$ and $\bar{\varphi} = 2\pi\varphi$ yields the more normal form. The scaling could be omitted by directly taking a different G ($\bar{H}$ remains the same):

$$G = \frac{1}{2}\,\omega x^2\,\mathrm{cotg}\,\varphi, \qquad \bar{H}\,(\bar{J},\bar{\varphi}) = \omega\bar{J},$$

$$p = \sqrt{2J\omega}\cos\varphi, \qquad x = \sqrt{\frac{2J}{\omega}}\sin\varphi.$$

(Remark. Try the generating function $G = \frac{1}{2}\,\omega x^2 \tan\varphi$. What is the new momentum, and the new coordinate?)

4.4. Removal of an oscillating part in the second degree of H, using action and angle variables

We return to the Hamiltonian in second degree with an oscillating part (Section 4.2):

$$H = \frac{1}{2}\,p^2 + \frac{1}{2}\,\omega^2\,(1 + A_n \sin n\theta)x^2.$$

First step: change to action and angle variables

$$x = \sqrt{\frac{2J}{\omega}}\cos\varphi, \quad p = \sqrt{2J\omega}\sin\varphi,$$

(φ = momentum, J = coordinate).

Then $$H = \omega J + \omega J\cos^2\varphi\, A_n \sin n\theta.$$

In this Hamiltonian rapidly oscillating terms are present. Therefore, remembering that φ is still rapidly changing we put $\omega = k + \Delta\omega$, with k an

integer and $\Delta\omega$ generally small. The second step involves a change in the momentum φ:

$$\bar{\varphi} = \varphi + k\theta.$$

As φ changes roughly with $\omega\theta$ we see that $\bar{\varphi}$ will change roughly with $\Delta\omega\theta$.

This transformation is generated by a function

$$G = J\bar{\varphi} - Jk\theta.$$

The new Hamiltonian becomes

$$H = \Delta\omega J + \omega J\cos^2 (\bar{\varphi} - k\theta)A_n \sin n\theta,$$

$$H = \Delta\omega J + Jf(\theta,\varphi).$$

In the third step the oscillating part is removed (in fact replaced by a new term that has second order of size in A_n and that is generally much smaller). A transformation function $G(\bar{J},\varphi,\theta)$ is applied with, in any case, two properties: first we should be somewhere around the unit transformation (therefore a part $-\bar{J}\varphi$ is needed); second we like to remove the oscillating part $Jf(\theta,\varphi)$:

$$\bar{H} = \Delta\omega J + Jf(\theta,\varphi) + \frac{\partial G}{\partial\theta}.$$

A good trial is thus $G(\bar{J},\varphi,\theta) = -\bar{J}\varphi - \bar{J}U_2(\theta,\varphi)$,

with $\frac{\partial U_2}{\partial\theta} = f.$

It then follows that $\bar{\bar{H}} = (\Delta\omega + < \frac{\partial U_2}{\partial\varphi} \cdot \frac{\partial U_2}{\partial\theta} >)\bar{J}.$

The oscillating terms are at least of order (A^2) and can be eliminated by a second step, then yielding a 4th order constant part and a 4th order oscillating part, etc. For many applications the above approximation is sufficient.

A resonance occurs if $n = 2k$, then the Hamiltonian becomes, in most important order,

$$H = \Delta\omega J + \alpha\cos 2\varphi J.$$

If $|\alpha|<|\Delta\omega|$ then there will always be stability. If $|\alpha|>|\Delta\omega|$ then one can

find an angle φ for which the total coefficient of J equals zero, driving the action variable to infinity, as H = constant.

As a general remark: rapidly oscillating terms can be averaged over a long period and then yield a zero effect. However, it may be that the above given transformation shows that small remaining constant parts may be important for a good estimation of frequency deviations. The transformation for eliminating oscillating terms can be applied also for higher degree terms in the Hamiltonian, following roughly the same lines.

4.5. Trajectories in a solenoidal lens

Quite often these trajectories are described in cylindrical coordinates. However, a term with $\frac{p_\theta}{r}$ then appears in the Hamiltonian, giving rise to problems for particles around the axis:

$$H = \frac{1}{2m} \{(\frac{p_\theta}{r} - eA_\theta)^2 + (p_r - eA_r)^2 + (p_z - eA_z)^2\}.$$

With cylinder symmetry there is θ-independency. Further an adequate vector potential can be represented by one component $A_\theta(r,z)$ with z the position along the axis. Apart from the situation for small r there is a coupling between the different dimensions. It is more profitable to use cartesian coordinates. For a homogeneous magnetic field the vector components are given by

$$A_x = -\frac{1}{2} By \quad , \qquad A_y = +\frac{1}{2} Bz \ .$$

The Hamiltonian is taken to be $K = -\pi_z$ (see Section 1.7) and the resulting square root is expanded up to the second degree. Scaled momenta are used:

$$\pi_x = \frac{p_x}{p_o} \ , \quad \pi_y = \frac{p_y}{p_o} \ , \ \pi_z = \frac{p_z}{p_o} \ .$$

Then $$K = -\pi_z = \frac{1}{2} (\pi_x + \frac{1}{2} \frac{eBy}{p_o})^2 + \frac{1}{2} (\pi_y - \frac{1}{2} \frac{eBx}{p_o})^2 \ .$$

The coupling between x and y space is solved by applying a coordinate transformation such that the new XY-system is rotating with half the cyclotron-frequency w.r.t. the old system (this new frame is sometimes called the Larmor frame):

$$x = \bar{x} \cos \varphi - \bar{y} \sin \varphi,$$
$$y = \bar{x} \sin \varphi + \bar{y} \cos \varphi,$$

with $$\frac{d\varphi}{dz} = - \frac{\omega}{v_0} = - \frac{\frac{1}{2} eB}{mv_0} = \frac{-1}{2R_c} ,$$

where B is the magnetic induction along the axis (B = B(z)), v_0 the particle velocity, and R_c the cyclotron radius in a magnetic field with induction B. The transformation is generated by

$$G = G(x, \bar{\pi}_x, y, \bar{\pi}_y) = \bar{\pi}_x x \cos \varphi + \bar{\pi}_x y \sin \varphi - \bar{\pi}_y x \sin \varphi + \bar{\pi}_y y \sin \varphi.$$

This generating function is found by first making the coordinate transformation, remarking that $\bar{x} = \frac{\partial G}{\partial \bar{\pi}_x}$, etc. A form of G follows and the new momenta are found from $\pi_x = \frac{\partial G}{\partial x}$ etc. The new Hamiltonian is

$$\bar{K} = K + \frac{\partial G}{\partial z} = \frac{1}{2} \bar{\pi}_x^2 + \frac{1}{2} \left(\frac{1}{4} \frac{e^2 B^2}{p_0^2}\right) \bar{x}^2 + \frac{1}{2} \bar{\pi}_y^2 + \frac{1}{2} \left(\frac{1}{4} \frac{e^2 B^2}{p_0^2}\right) \bar{y}^2.$$

The problem is uncoupled and therefore solved (see Section 1.5). The Hamiltonian represents two equal lenses with the same sign in a rotating frame.

In matrix notation $$\underline{x}_f = R_f^{-1} L R_i \underline{x}_i,$$

where R is the rotation matrix, L the lens matrix and, x the four vector x, x', y, y' outside magnetic field and x, π_x, y, π_y inside magnetic field.

The initial rotation can be taken equal to zero: R_i = unit matrix. Inside the magnetic field canonical momenta must be used. The relation between the canonical variables and the geometrical variables is given by

$$\bar{x} = \begin{pmatrix} x \\ \pi_x \\ y \\ \pi_y \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2R_c} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2R_c} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix} .$$

Suppose that inside the magnetic field $x'' = 0 = y''$, occurring for example in the situation of an Electron Cyclotron Resonance (ECR) source, then

$$\bar{x} = \begin{pmatrix} x \\ \frac{-y}{2R_c} \\ y \\ \frac{x}{2R_c} \end{pmatrix} .$$

If the beam cross section is circular and the field is cylindrically symmetrical, the emittance area in X and Y phase space is given by

$$\epsilon_x = \epsilon_y = \frac{\pi r^2}{2R_c} ,$$

where r is the radius of the beam cross section. Further one observes that the phase spaces are strongly coupled (one point within the X phase space has only one point in the Y phase space). Because of the cylinder symmetry rotational matrices can be skipped. The above matrix relating the divergencies with the momenta is denoted by C, the lens matrix of the solenoidal field by L; we then get

$$\underset{\text{outside B}}{x_f} = \underset{\text{inside B}}{L\,\underline{x}_i} = LC\tilde{\underline{x}}_i .$$

where $\tilde{x} = (x,x'',y,y'')$

The transformation matrix LC keeps the volume from $(x,x'y,y')$ space to $(x_f, \pi_{x_f}, y_f, \pi_{y_f})$ constant. So Liouville is 'valid' but is the transformation canonical? Therefore LC must be symplectic. It is easy to check that

$$\tilde{C}\eta C \neq \eta,$$

but $\tilde{L}\eta L = \eta$, so L is symplectic. Clearly C is not a canonical transformation (x' and y' were not canonical momenta). A 2 * 2 matrix with determinant $\Delta = 1$ is symplectic.

4.6. An evaluation of a general Hamiltonian

The Hamiltonian for particle trajectories in electric and magnetic fields is given by

$$H = \{ E_r^2 + (p_x - eA_x)^2 c^2 + (p_z - eA_z)^2 c^2 + \left(\frac{p_s}{1+\frac{x}{\rho(s)}} - eA_s\right)^2 c^2 \}^{1/2} + eV.$$

The representation is given in curvilinear coordinates in the median plane. It is important first to have an appropriate representation of the vector potential, secondly one has to realize that electric fields in cavities do not follow from a scalar potential function, third the choice of a vector potential is not free when the orbit encloses time varying magnetic fields (betratron acceleration [8]). RF acceleration should thus be treated via the vector potential. Remember that, within the Lorentz gauge,

$$E = - \nabla \cdot V - \frac{\partial A}{\partial t} .$$

In the case of stationary fields, it is often useful to find a representation of the vector potential such that

$$A_s = 0 \text{ on a reference orbit.}$$

Then around a circle (radius R) and assuming cylinder symmetry

$$A_s = - B_0 x + \frac{1}{2} \left(\frac{B_0}{\rho} - b_1\right) x^2 + \frac{1}{2} b_1 z^2 + \ldots\ldots, \quad b_1 = \frac{1}{R} \frac{dB}{dx} ,$$
$$A_x = 0, \quad A_z = 0.$$

We will now treat a solution for which there is
- no time dependency
- no acceleration
- cylinder symmetry and medium plane symmetry.

The aim is to find a separated representation for the three dimensions and only the linear motion will be considered. Four steps are involved:
1. Elimination of first degree parts

2. Scaling
3. Expansion of the square root, skipping constant terms
4. A final transformation for separating the dimensions.

4.6.1 Elimination of first degree parts.

For this a reference momentum p_0 is chosen:

$$p_s = \tilde{p}_s + p_0.$$

In the case of cylinder symmetry p_0 is taken not to be a function of s

$$p_0 = - eB_0\rho.$$

By this choice first degree parts disappear.

4.6.2 Scaling.

The scaled coordinates and momenta are

$$\bar{x} = \frac{x}{R}, \quad \bar{z} = \frac{z}{R}, \quad \bar{s} = \frac{s}{R},$$

$$\bar{p}_x = \frac{p_x}{p_0}, \quad \bar{p}_z = \frac{p_z}{p_0}, \quad \bar{p}_s = \frac{\tilde{p}_s}{p_0},$$

$$\tau = \omega_0 t, \quad \omega_0 = \frac{c}{R}\left(1 - \frac{1}{\gamma_0^2}\right)^{1/2}.$$

c is the velocity of light and $\gamma_0 = \frac{m}{m_0}$, where m_0 is the rest mass. The new Hamiltonian is scaled as

$$\bar{H} = \frac{H}{p_0 R\omega_0}.$$

4.6.3 Expansion of the square root.

After expansion of the square root up to the second degree in the variables the new Hamiltonian becomes

$$H = \frac{1}{2} p_x^2 + \frac{1}{2}(1-n)x^2 + \frac{1}{2} p_z^2 + \frac{1}{2} nz^2 + \frac{1}{2\gamma_0^2} p_s^2 + p_s - p_s x + \ldots, \quad \gamma_0 = \frac{m}{m_0}.$$

In this Hamiltonian the bars have been omitted, further n means the field

index ($n = - b_1$).

The Hamiltonian is (was already from the beginning!) independent of s. Therefore p_s = constant. For $p_s \neq 0$ a first degree part in x is still present: we then did not make the right choice for p_o. For $p_s = 0$ there is no first degree part and the Hamiltonian describes the two well known betatron oscillations as two separate modes.

What does the coupling term $p_s x$ mean? Even for $p_s = 0$ we may not skip it.

$$\dot{s} = \frac{\partial H}{\partial p_s} = - x + 1, \qquad (p_s = 0).$$

The position of the particle on the trajectory thus depends on x, the radial coordinate of the radial oscillation. Differently said: the particle position gives rise to a phase which depends on the radial coordinate.

4.6.4 The canonical tranformation.

The coupling is removed by a canonical transformation generated by

$$G = \bar{p}_x\, x + \bar{p}_z\, z + \bar{p}_s\, s - \frac{1}{1-n}\, \bar{p}_s\, \bar{p}_x.$$

The generating function in this case is time independent so that

$$\bar{H} = H.$$

The relation between old and new variables is given by

$$x = \bar{x} + \frac{\bar{p}_s}{1-n}, \qquad p_x = \bar{p}_x,$$

$$z = \bar{z}, \qquad p_z = \bar{p}_z,$$

$$s = \bar{s} + \frac{\bar{p}_x}{1-n}, \qquad p_s = \bar{p}_s.$$

The new Hamiltonian is

$$H = \frac{1}{2}\, \bar{p}_x^2 + \frac{1}{2}\,(1-n)\, \bar{x}^2 + \frac{1}{2}\, \bar{p}_z^2 + \frac{1}{2}\, n\bar{z}^{\,2} + \frac{1}{2}\, \bar{p}_s^2 \left(\frac{1}{\gamma_o^2} - \frac{1}{1-n}\right) + \bar{p}_s.$$

Now the three dimensions are uncoupled. Note the large difference in the

coefficient before $\bar{p}_s^2$ w.r.t. the coefficient of p_s^2 in the Hamiltonian H. Due to this substantial change the above transformation must always be carried out for the study of longitudinal effects [8].

The transformation shows that there is an extra contribution in x due to the fact that $p_s \neq 0$ gives a new equilibrium orbit. In s there is a contribution due to the betatron oscillations resulting in the fact that the real particle phase $\bar{s}$ differs from a measured phase by $\frac{\bar{p}_x}{1-n}$. In cyclotrons this is illustrated by the fact that the orbit centre is shifted (see Fig. 14).

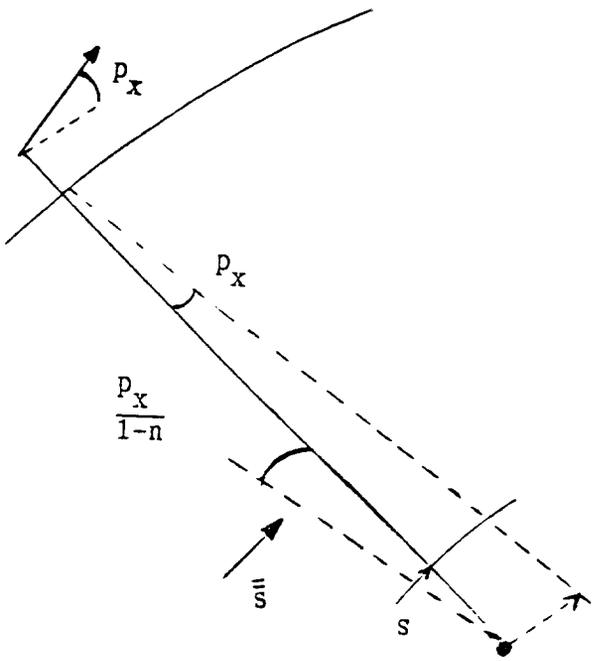

Figure 14. Shift of orbit centre due to betatron motion.

4.7. Some final remarks

In many practical cases the transformations are rather complicated and it is advisable to carry them out in steps, each step having a special purpose. However, the examples should give some guidance on how to start

and how to proceed. Sometimes it is sufficient to choose the position along the axis as a new independent variable in order to get the solution of the problem. For numerical calculations one has to take care of approximations made in the equations of motion as then these may turn out not to be canonical. It is a safer precedure to start with a proper Hamiltonian and then skip those terms that are of minor importance.

ACKNOWLEDGEMENT

The authors thank Prof. F.W. Sluijter and Prof. B.J. Verhaar for carefully reading the manuscript and for useful comments.

REFERENCES

[1] E.T. Whittaker, A treatise on the analytical dynamics of particles and rigid bodies, Cambridge University Press, 1960

[2] L.D. Landau and E.M. Lifshitz, Course of theoretical physics, Third edition, Volume 1, Pergamon Press

[3] V.I. Arnold, Mathematical methods of classical mechanics, Springer Verlag

[4] C. Lanczos, The variational principles of mechanics, Second Edition, University of Toronto Press, Toronto

[5] H.C. Corben and P. Stehle, Classical mechanics, 2nd Edition, John Wiley & Sons, Inc., New York

[6] H. Goldstein, Classical mechanics, 2nd Edition, Addison-Wesley Publishing Co.

[7] J.S. Bell, Hamiltonian mechanics, Proceedings CERN Accelerator School, September 1985, CERN 87-03, Volume I. pp. 5-40

[8] C.J.A. Corsten, Resonance and coupling effects in circular accelerators, Ph.D. Thesis, September 1982, Eindhoven University of Technology