Linear Dynamics, Lecture 1

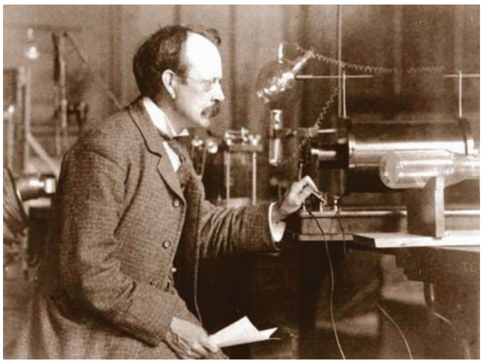
Review of Hamiltonian Mechanics

Andy Wolski

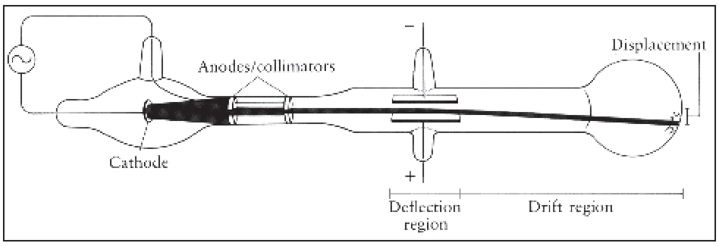
University of Liverpool, and the Cockcroft Institute, Daresbury, UK.

November, 2012

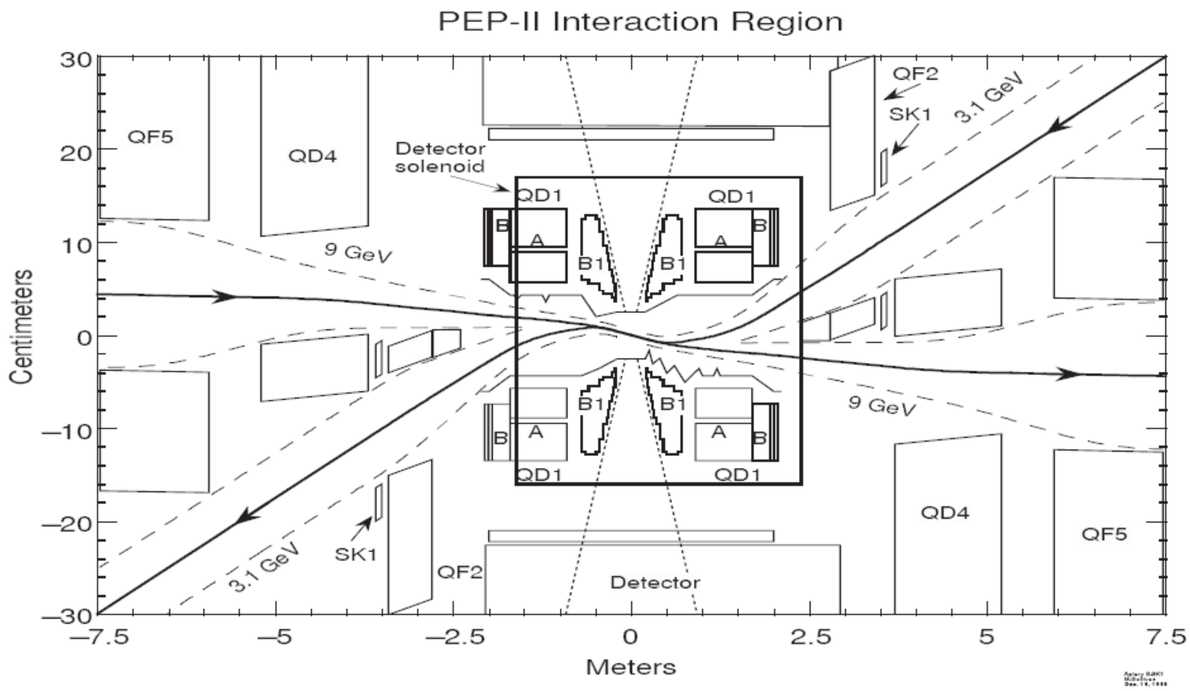
Introduction



Joseph John Thomson, 1856-1940



Early accelerators were fairly straightforward.



Modern accelerators are more sophisticated.

For modern accelerators to operate properly, the beam dynamics must be modelled and understood with very high precision. There are many effects that are important, including synchrotron radiation, interactions between the particles, interactions with the residual gas in the vacuum chamber and with the vacuum chamber itself, etc.

However, everything starts with understanding the motion of individual particles through the fields from the magnets and the RF cavities.

There are several possible approaches. We shall develop an approach starting from the fundamentals of classical mechanics. This requires more initial effort to derive the equations of motion in a form appropriate for accelerator physics; but has the benefit of providing a rigorous framework for modelling the dynamics with the precision required for modern accelerators.

Part I (Lectures 1 – 5): Dynamics of a relativistic charged particle in the electromagnetic field of an accelerator beamline.

1. Review of Hamiltonian mechanics
2. The accelerator Hamiltonian in a straight coordinate system
3. The Hamiltonian for a relativistic particle in a general electromagnetic field using accelerator coordinates
4. Dynamical maps for linear elements
5. Three loose ends: edge focusing; chromaticity; beam rigidity.

Part II (Lectures 6 – 10): Description of beam dynamics using optical lattice functions.

6. Linear optics in periodic, uncoupled beamlines
7. Including longitudinal dynamics
8. Bunches of many particles
9. Coupled optics
10. Effects of linear imperfections

There are three alternative approaches to classical mechanics: Newtonian, Lagrangian and Hamiltonian mechanics.

Formally, all these approaches are equivalent: they have the same "physical content", and any one can be derived from any of the others.

So why prefer any one over the others?

It depends on the problem you are trying to solve; the equations of motion for a given system may appear simpler in one of the approaches than in the others. As we shall see, for accelerator physics, Hamiltonian mechanics provides some great advantages.

Isaac Newton, 1643-1727

The equation of motion of a particle of mass m subject to a force $\mathbf{F}$ is:

$$\frac{d}{dt} m\dot{\mathbf{x}} = \mathbf{F}\left(\mathbf{x}, \dot{\mathbf{x}}; t\right) \tag{1}$$

where $\dot{\mathbf{x}}$ is the velocity. Note that the dot over a variable indicates the derivative with respect to time.

Consider the case of a particle of fixed mass moving in one degree of freedom, subject to a force F given by:

$$F = -m\omega^2 x \tag{2}$$

The equation of motion becomes:

$$m\frac{d}{dt}\dot{x} = -m\omega^2 x \tag{3}$$

or:

$$\frac{d^2x}{dt^2} = -\omega^2 x \tag{4}$$

This has solution:

$$x(t) = x_0 \sin(\omega t + \phi_0) \tag{5}$$

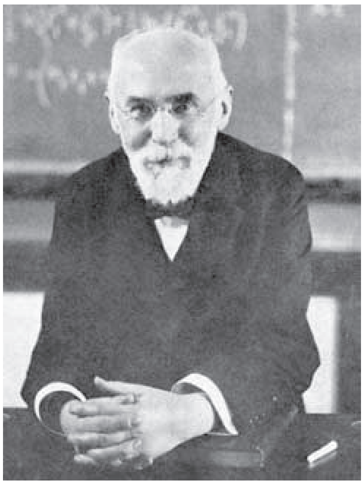
where x_0 and ϕ_0 are constants determined by the initial values of $x(t)$ and $\dot{x}(t)$.

In Newtonian mechanics, the dynamics of the system are defined by the force $\mathbf{F}$, which in general is a function of position $\mathbf{x}$, velocity $\dot{\mathbf{x}}$ and time t.

Given the function $\mathbf{F}$, we derive the equations of motion, which we must then solve to give the explicit dependence of the position $\mathbf{x}$ (and the velocity $\dot{\mathbf{x}}$) on the *independent parameter* t.

"Physics" consists of writing down the form of the function $\mathbf{F}$ for a given system.

The Lorentz Force



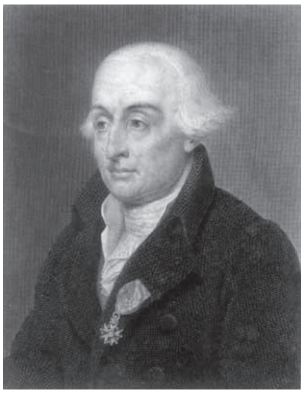
Hendrik Lorentz, 1853-1928

Of particular interest in accelerator physics is the Lorentz force for a particle of charge q moving in a region with an electric field $\mathbf{E}(\mathbf{x})$ and magnetic field $\mathbf{B}(\mathbf{x})$:

$$\mathbf{F} = q\left(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}\right) \tag{6}$$

The functions $\mathbf{E}$ and $\mathbf{B}$ must satisfy a further set of equations (Maxwell's equations), to be dealt with later.

Lagrangian Mechanics



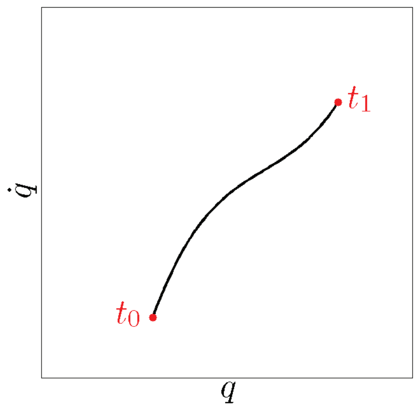
Joseph-Louis Lagrange, 1736-1813

Given a function $L(\mathbf{q}, \dot{\mathbf{q}}; t)$ (called the Lagrangian), the equations of motion for a dynamical system are given by:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \tag{7}$$

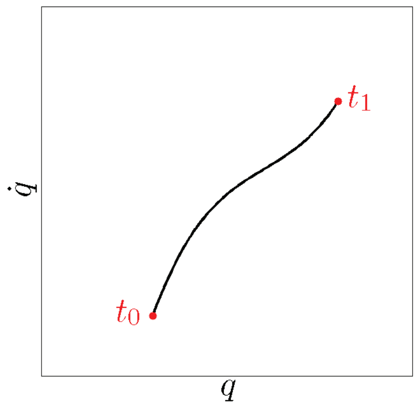
where q_i are the components of $\mathbf{q}$.

Equations (7) are known as the "Euler-Lagrange Equations."



Consider the path traced by a dynamical system on a plot of $\dot{q}$ vs q. We can evaluate the integral S of the Lagrangian L along the line:

$$S = \int_{t_0}^{t_1} L dt \tag{8}$$



It can be shown that the Euler-Lagrange equations (7) define a path for which the *action* S is a minimum, i.e.:

$$\delta S = \delta \left[\int_{t_0}^{t_1} L dt \right] = 0 \tag{9}$$

where the operator δ gives the change with respect to a change in path.

The variables q_i can be any convenient set of parameters that describe the state of the system. The coordinates in Euclidean space are an obvious example, $(\mathbf{q} = \mathbf{x})$ but not the only (or necessarily best) choice.

The question is, how do we write down the function L that contains the dynamics of the system? This question is equivalent to "How do we write down the force $\mathbf{F}$?" in Newtonian mechanics.

It turns out that in many cases the Lagrangian is given by:

$$L = T - V \tag{10}$$

where T is the kinetic energy of the system, and V is the potential energy.

Consider a particle moving in one degree of freedom, with kinetic energy T given by:

$$T = \frac{1}{2}m\dot{x}^2 \tag{11}$$

and potential energy V given by:

$$V = \frac{1}{2}m\omega^2 x^2 \tag{12}$$

The Lagrangian is then:

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2 \tag{13}$$

Inserting the Lagrangian (13) into the Euler-Lagrange equations (7), we find the equation of motion:

$$-\omega^2 m x - \frac{d}{dt}(m\dot{x}) = 0 \tag{14}$$

or:

$$\frac{d^2x}{dt^2} = -\omega^2 x \tag{15}$$

Note that in using the Euler-Lagrange Equations to derive the equation of motion from the Lagrangian, we treated the coordinates $\mathbf{x}$ and the velocities $\dot{\mathbf{x}}$ as independent of one another. Of course, they are related through differentiation – $\dot{\mathbf{x}}$ is the rate of change of $\mathbf{x}$ – but for applying the Euler-Lagrange Equations, we ignore this fact.

Lagrangian mechanics allows us to write down the equation of motion using any convenient parameters. This sometimes simplifies the problem compared to a treatment based on Newtonian mechanics.

There is a third way...

Hamiltonian Mechanics

William Rowan Hamilton, 1805-1865

Given a function $H(\mathbf{x}, \mathbf{p}; t)$ (called the Hamiltonian), the equations of motion for a dynamical system are given by Hamilton's equations:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \tag{16}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} \tag{17}$$

The Hamiltonian plays the same role in Hamiltonian mechanics as does the force in Newtonian mechanics, and the Lagrangian in Lagrangian mechanics: it defines the dynamics of the system. "Physics" consists of writing down a Hamiltonian.

We need to be careful about the meaning of the *conjugate momentum* $\mathbf{p}$: in simple cases, it is equivalent to the *mechanical momentum* $m\dot{\mathbf{x}}$ – but this is not always the case!

Formally, given a Lagrangian L, the conjugate momentum and the Hamiltonian can be derived as follows:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{18}$$

$$H = \sum_i \dot{q}_i p_i - L \tag{19}$$

The *Lagrangian* is given in terms of the coordinates and velocities. Equation (18) defines the conjugate momenta that we can use instead of the velocities. The *Hamiltonian* defined by equation (19) should be expressed purely in terms of the coordinates and conjugate momenta: the velocities should not appear in the Hamiltonian. Given a Hamiltonian, the equations of motion are Hamilton's equations (16) and (17):

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}$$

Note that in n degrees of freedom, the Euler-Lagrange equations give us n second-order differential equations, while Hamilton's equations give us $2n$ first-order differential equations. Representing the dynamics using first-order equations has certain advantages concerning linear methods, stability analysis, etc.

Consider the Lagrangian that we looked at before:

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2 \tag{20}$$

The conjugate momentum (18) is:

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \tag{21}$$

Note that as usual, we treat x and $\dot{x}$ as independent of one another. Also note that in this case, the conjugate momentum p_x is equal to the mechanical momentum $m\dot{x}$.

The Hamiltonian is:

$$H = \dot{x}p_x - L = m\dot{x}^2 - L \tag{22}$$

or:

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{23}$$

Moving from Lagrangian to Hamiltonian mechanics essentially involves making a change of variables from $\dot{\mathbf{x}}$ to $\mathbf{p}$. The Hamiltonian should always be written in terms of the conjugate momentum $\mathbf{p}$ rather than the velocity $\dot{\mathbf{x}}$.

In Lagrangian mechanics, the "state" of a system at any time is defined by specifying values for the coordinates $\mathbf{x}$ (or more generally $\mathbf{q}$) and the velocity $\dot{\mathbf{x}}$ (or $\dot{\mathbf{q}}$).

In Hamiltonian mechanics, the "state" of a system at any time is defined by specifying values for the coordinates $\mathbf{x}$ (or more generally $\mathbf{q}$) and the momentum $\mathbf{p}$.

It follows from Hamilton's equations that the Hamiltonian itself is conserved if the independent ("time-like") variable does not appear explicitly in the Hamiltonian. This can be shown as follows:

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p_x}\frac{dp_x}{dt} + \frac{\partial H}{\partial t} \tag{24}$$

Using Hamilton's equations, we have:

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{\partial H}{\partial p_x} - \frac{\partial H}{\partial p_x}\frac{\partial H}{\partial x} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{25}$$

If the Hamiltonian does not depend explicitly on t, then the Hamiltonian is conserved:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = 0 \tag{26}$$

Given the Hamiltonian (23):

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{27}$$

and Hamilton's equations, (16) and (17):

$$\frac{dx}{dt} = \frac{\partial H}{\partial p_x} \tag{28}$$

$$\frac{dp_x}{dt} = -\frac{\partial H}{\partial x} \tag{29}$$

The equations of motion are:

$$\frac{dx}{dt} = \frac{p_x}{m} \tag{30}$$

$$\frac{dp_x}{dt} = -m\omega^2 x \tag{31}$$

Note that equation (31) is equivalent to Newton's equation (1).

When deriving the equations of motion for the system from Hamilton's equations, we treat $\dot{\mathbf{x}}$ and $\mathbf{p}$ as independent of one another, even though we have a formal relationship between them.

In our simple example, the Hamiltonian (23) was:

$$H = \frac{p_x^2}{2m} + \frac{1}{2} m \omega^2 x^2 \tag{32}$$

which can be written:

$$H = T + V \tag{33}$$

for kinetic energy T and potential energy V. It appears that (at least in this case) the Hamiltonian is the "total energy" of the system, expressed in terms of the coordinates and conjugate momentum. This is a clue for writing down the Hamiltonian in more complicated systems.

Consider the Lagrangian:

$$L = \frac{1}{2} m \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} - q\phi + q\mathbf{A} \cdot \dot{\mathbf{x}} \tag{34}$$

This describes a non-relativistic particle with two components to its potential energy: one a straightforward scalar function $\phi(\mathbf{x})$ of position, and the other a function of the vector field $\mathbf{A}(\mathbf{x})$ and proportional to the particle's velocity $\dot{\mathbf{x}}$.

The conjugate momentum is:

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + qA_i \tag{35}$$

Note that in this case, the conjugate momentum $\mathbf{p}$ is *not* equal to the mechanical momentum $m\dot{\mathbf{x}}$.

The Hamiltonian is:

$$H = \mathbf{p} \cdot \dot{\mathbf{x}} - L = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \tag{36}$$

A Further Example: Dynamics in an Electromagnetic Field

After some working (see Appendix A), we find that the equation of motion (17) from the Hamiltonian (36) is (92):

$$\frac{d}{dt}(\mathbf{p} - q\mathbf{A}) = q(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}) \tag{37}$$

or:

$$\frac{d}{dt} m\dot{\mathbf{x}} = q(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}) \tag{38}$$

where the fields $\mathbf{E}$ and $\mathbf{B}$ are derived from the potentials:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{39}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{40}$$

Equation (38) is just Newton's equation (1) with the Lorentz force (6). Note that this was derived for non-relativistic particles: later we will need to derive a relativistic equation of motion.

Hamiltonian Mechanics: Some Further Remarks

Hamiltonian mechanics introduces three important and related concepts:

- canonical variables
- symplecticity
- canonical transformations

We shall briefly discuss each of these concepts in the rest of this lecture.

One of the benefits of using Hamiltonian mechanics is that it provides a highly structured framework for transforming between coordinate systems. This is important in accelerator physics, where the variables used are not always the most obvious choice.

In Hamiltonian mechanics, the *canonical variables* consist of a set of coordinates and their conjugate momenta defined by equation (18). In general:

> *The word canonical is used to indicate a particular choice from of a number of possible conventions. This convention allows a mathematical object or class of objects to be uniquely identified or standardized.*

Wolfram Mathworld, mathworld.wolfram.com

A $2n \times 2n$ *symplectic matrix* M is one that satisfies:

$$M^{\mathsf{T}} \cdot S \cdot M = S \tag{41}$$

where S is a $2n \times 2n$ matrix with block diagonals:

$$S_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{42}$$

Note that the matrix S has the properties:

$$S^{\mathsf{T}} = -S \tag{43}$$

and:

$$S^2 = -I \tag{44}$$

where I is the $2n \times 2n$ identity matrix.

... OK, but what has this got to do with Hamiltonian mechanics?

We write Hamilton's equations (16) and (17) in the form:

$$\frac{d}{dt}\vec{x} = S \cdot \nabla_{\vec{x}} H \tag{45}$$

where:

$$\vec{x} = \begin{pmatrix} x \\ p_x \\ y \\ p_y \\ \vdots \end{pmatrix} \tag{46}$$

and:

$$\nabla_{\vec{x}} = \begin{pmatrix} \partial_x \\ \partial_{p_x} \\ \partial_y \\ \partial_{p_y} \\ \vdots \end{pmatrix} \tag{47}$$

A Hamiltonian that is second-order in the dynamical variables leads to equations of motion that are linear. For a general second-order Hamiltonian:

$$\nabla_{\vec{x}} H = J \cdot \vec{x} \tag{48}$$

where J is a symmetric matrix, $J^{\mathsf{T}} = J$. Hamilton's equations can then be written:

$$\frac{d}{dt}\vec{x} = S \cdot J \cdot \vec{x} \tag{49}$$

The solution to (49) is given by:

$$\vec{x}(t) = M(t) \cdot \vec{x}(0) \tag{50}$$

where the matrix $M(t)$ is given by:

$$M(t) = \exp\left(tS \cdot J\right) \tag{51}$$

$M(t)$ is sometimes called the *transfer matrix*.

Now, note that J is symmetric and S is antisymmetric:

$$J^{\mathsf{T}} = J \qquad S^{\mathsf{T}} = -S \tag{52}$$

Also, it can be shown (see Appendix B) that:

$$S \cdot \exp(tS \cdot J) = \exp(tJ \cdot S) \cdot S \tag{53}$$

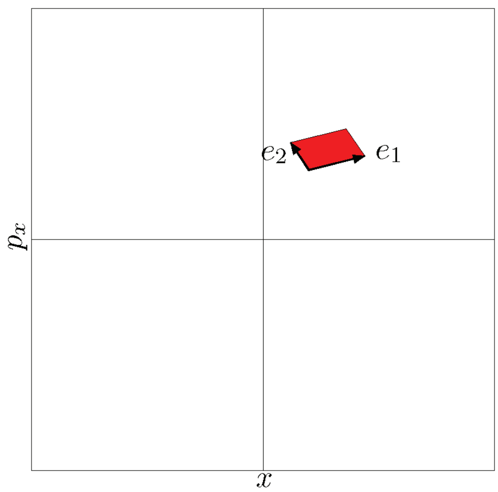
Using equations (52) and (53), we can write:

$$M^{\mathsf{T}}(t) \cdot S \cdot M(t) = \exp(-tJ \cdot S) \cdot S \cdot \exp(tS \cdot J) \tag{54}$$
$$= \exp(-tJ \cdot S) \cdot \exp(tJ \cdot S) \cdot S \tag{55}$$
$$= S \tag{56}$$

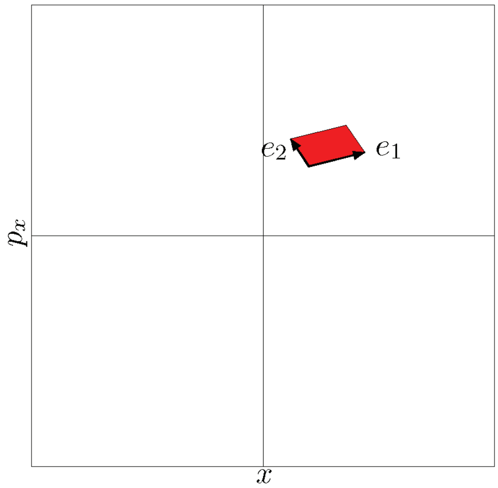
Therefore, from (41), $M(t)$ is symplectic.

Conclusion: for a linear system whose dynamics can be described by a Hamiltonian, the transfer matrix is symplectic.

... OK, but what does symplecticity mean physically?

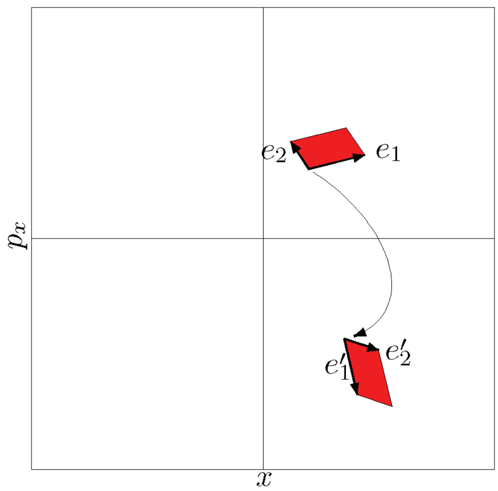


Consider an area of *phase space* (a plot of the conjugate momentum vs the corresponding coordinate) defined by vectors $\vec{e}_1$ and $\vec{e}_2$.



The area of the phase space element is:

$$A = |\vec{e}_1 \times \vec{e}_2| = \vec{e}_1^{\mathsf{T}} \cdot S \cdot \vec{e}_2 \qquad (57)$$



If the system evolves over some period of time, t:

$$\vec{e}_1 \rightarrow \vec{e}'_1 = M(t) \cdot \vec{e}_1 \quad \text{and} \quad \vec{e}_2 \rightarrow \vec{e}'_2 = M(t) \cdot \vec{e}_2 \qquad (58)$$

The area of the new section of phase space is:

$$A' = \vec{e}_1^{\prime\top} \cdot S \cdot \vec{e}_2^{\,\prime} \quad (59)$$
$$= \vec{e}_1^{\top} \cdot M^{\top}(t) \cdot S \cdot M(t) \cdot \vec{e}_2 \quad (60)$$

But for a Hamiltonian system, $M(t)$ is symplectic:

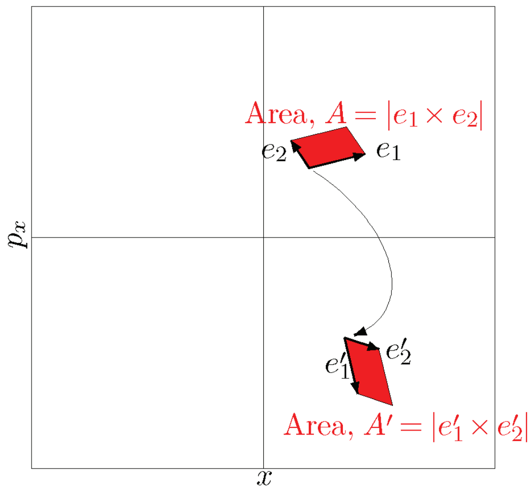
$$M^{\top}(t) \cdot S \cdot M(t) = S \quad (61)$$

It follows that the area of the new section of phase space is:

$$A' = \vec{e}_1^{\top} \cdot S \cdot \vec{e}_2 \quad (62)$$
$$= A \quad (63)$$

In other words...

Symplecticity and Hamiltonian Mechanics



...$A' = A$: the area of the element of phase space is conserved during the motion of the system.

Joseph Liouville, 1809-1882

We have shown that the area of phase space “elements” is conserved during the motion of any system whose dynamics are linear, and can be described by a Hamiltonian.

In fact, it can be shown that areas of phase space elements are conserved for all Hamiltonian systems, even when the dynamics are nonlinear. This important result is known as “Liouville’s Theorem”.

The conservation of area in phase space is an important property of Hamiltonian systems, which are also known as “conservative systems”. Not all dynamical systems are conservative. The presence of dissipative forces, such as friction, leads to a shrinkage of phase space area.

In accelerator physics, the phase space area occupied by a bunch of particles is an important quantity, and is known as the “emittance”. In an accelerator, synchrotron radiation can be analogous to friction, and can reduce the emittance of a bunch of particles. However, in many cases, it is a good approximation to neglect non-conservative forces on particles in an accelerator; in this approximation, Liouville’s theorem tells us that the emittance of a bunch of particles is conserved as the particles move through an accelerator.

One of the great advantages of Hamiltonian mechanics, is that it provides a rigorous framework for making changes of variables, and for writing down the equations of motion in the new variables. In Hamiltonian mechanics, the process of changing from one set of (canonical) variables to another is known as a "canonical transformation".

Here, we do not go through a rigorous treatment, but simply quote some useful results, and give some examples.

In general, we wish to transform from a set of "old" canonical variables $(\mathbf{q}, \mathbf{p})$ to a "new" set $(\mathbf{Q}, \mathbf{P})$. We wish to find expressions for the new variables in terms of the old variables; the requirement that the new variables be canonical is a constraint on the expressions that are allowed.

Canonical transformations may be found by means of "generating functions". A generating function is a function of some combination of old and new canonical variables, and (optionally) the independent variable, t. For example, consider:

$$F_2 = F_2(\mathbf{q}, \mathbf{P}, t) \tag{64}$$

F_2 is a function of the old coordinates, the new momenta, and the time. The old momenta are expressed as:

$$p_i = \frac{\partial F_2}{\partial q_i} \tag{65}$$

and the new coordinates are expressed as:

$$Q_i = \frac{\partial F_2}{\partial P_i} \tag{66}$$

The new Hamiltonian is given by:

$$K = H + \frac{\partial F_2}{\partial t} \tag{67}$$

Let's consider a concrete example. Consider the Hamiltonian for a certain nonlinear oscillator:

$$H = \frac{p^2}{2(1+q^2)^2} + \frac{1}{2}\left(q + \frac{1}{3}q^3\right)^2 \qquad (68)$$

If you like, you can write down the equations of motion using Hamilton's equations (16) and (17), and then try to solve them. I prefer a simpler method. Write down the generating function:

$$F_2(q, P) = \left(q + \frac{1}{3}q^3\right) P \qquad (69)$$

In this case, the generating function F_2 is independent of time, t. We can use $F_2(q, P)$ to generate new canonical variables (Q, P), and a new Hamiltonian. If we've chosen the generating function correctly, the equations of motion in the new variables will be simpler to solve than the equations of motion in the old variables. Let's try...

Using the above equations (65) and (66), we find relations between the old variables (q, p) and the new variables (Q, P):

$$p = \frac{\partial F_2}{\partial q} = \left(1 + q^2\right) P \qquad (70)$$

$$Q = \frac{\partial F_2}{\partial P} = q + \frac{1}{3}q^3 \qquad (71)$$

In other words:

$$P = \frac{p}{1+q^2} \qquad (72)$$

$$Q = q + \frac{1}{3}q^3 \qquad (73)$$

and in terms of the new variables, the Hamiltonian (68) becomes:

$$K = \frac{1}{2}P^2 + \frac{1}{2}Q^2 \qquad (74)$$

Now the solution is easy. We recognise the Hamiltonian for a simple harmonic oscillator, so the equations of motion can be solved to give:

$$Q = Q_0 \sin(t + \phi_0) \tag{75}$$

where the constants Q_0 and ϕ_0 are set by the initial conditions. In terms of the old coordinate:

$$q + \frac{1}{3}q^3 = Q_0 \sin(t + \phi_0) \tag{76}$$

This is an algebraic equation, which we can solve for q. But since this is a course on *linear* dynamics, we won't take this example any further.

Finally, we note that there are *four* "standard" mixed-variable generating functions that can be used to construct canonical transformations. The equations are as follows.

Generating function of the first kind:

$$F_1 = F_1(\mathbf{q}, \mathbf{Q}, t), \;\; p_i = \frac{\partial F_1}{\partial q_i}, \;\; P_i = -\frac{\partial F_1}{\partial Q_i}, \;\; K = H + \frac{\partial F_1}{\partial t} \tag{77}$$

Generating function of the second kind:

$$F_2 = F_2(\mathbf{q}, \mathbf{P}, t), \;\; p_i = \frac{\partial F_2}{\partial q_i}, \;\; Q_i = \frac{\partial F_2}{\partial P_i}, \;\; K = H + \frac{\partial F_2}{\partial t} \tag{78}$$

Generating function of the third kind:

$$F_3 = F_3(\mathbf{p}, \mathbf{Q}, t), \;\; q_i = -\frac{\partial F_3}{\partial p_i}, \;\; P_i = -\frac{\partial F_3}{\partial Q_i}, \;\; K = H + \frac{\partial F_3}{\partial t} \tag{79}$$

Generating function of the fourth kind:

$$F_4 = F_4(\mathbf{p}, \mathbf{P}, t), \;\; q_i = -\frac{\partial F_4}{\partial p_i}, \;\; Q_i = \frac{\partial F_4}{\partial P_i}, \;\; K = H + \frac{\partial F_4}{\partial t} \tag{80}$$

Certain classical dynamical systems can be described using Hamilton's equations:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p_x}, \qquad \frac{dp_x}{dt} = -\frac{\partial H}{\partial x} \tag{81}$$

where (x, p_x) are *dynamical variables* (coordinate x and conjugate or canonical momentum p_x); t is the *independent variable*; the Hamiltonian H is a function (of the dynamical variables) that defines the dynamics of the system.

Hamiltonian systems are *symplectic*: areas in *phase space* are conserved as the system evolves (Liouville's theorem).

We can change from one set of dynamical variables to another using a formally defined *canonical transformation*. The new variables defined by a canonical transformation also obey Hamilton's equations, for an appropriate Hamiltonian.

The Hamiltonian is:

$$H = \mathbf{p} \cdot \dot{\mathbf{x}} - L = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \tag{82}$$

The first equation of motion (16) is easy enough:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} = \frac{p_i - qA_i}{m} \tag{83}$$

This reduces to:

$$m\dot{x}_i = p_i - qA_i \tag{84}$$

which we knew already from the definition of the conjugate momentum (35).

The second equation of motion (17) needs some work...

Hamilton's second equation (17) with the Hamiltonian (36) gives for the x component of the momentum:

$$\frac{dp_x}{dt} = -\frac{\partial H}{\partial x} \tag{85}$$

$$= q\left(\dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x}\right) - q\frac{\partial \phi}{\partial x} \tag{86}$$

Now, if we define a vector field:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{87}$$

we notice that:

$$[\dot{\mathbf{x}} \times \mathbf{B}]_x = \dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x} - \dot{x}\frac{\partial A_x}{\partial x} - \dot{y}\frac{\partial A_x}{\partial y} - \dot{z}\frac{\partial A_x}{\partial z}$$

$$= \left(\dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x}\right) - \frac{dA_x}{dt} + \frac{\partial A_x}{\partial t} \tag{88}$$

It follows that we can write:

$$\frac{dp_x}{dt} = q\,[\dot{\mathbf{x}} \times \mathbf{B}]_x + q\frac{dA_x}{dt} - q\frac{\partial A_x}{\partial t} - q\frac{\partial \phi}{\partial x} \tag{89}$$

If we now define a vector field:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{90}$$

then we can write:

$$\frac{d}{dt}(p_x - qA_x) = q\,[\dot{\mathbf{x}} \times \mathbf{B}]_x + qE_x \tag{91}$$

Finally, considering all vector components, we have:

$$\frac{d}{dt}(\mathbf{p} - q\mathbf{A}) = q\,(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}) \tag{92}$$

Given the Hamiltonian (36):

$$H = \mathbf{p} \cdot \dot{\mathbf{x}} - L = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} + q\phi \qquad (93)$$

with the conjugate momentum (35):

$$\mathbf{p} = m\dot{\mathbf{x}} + q\mathbf{A} \qquad (94)$$

the equation of motion from (17) is (92):

$$\frac{d}{dt}(\mathbf{p} - q\mathbf{A}) = q\left(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}\right) \qquad (95)$$

or:

$$\frac{d}{dt} m\dot{\mathbf{x}} = q\left(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}\right) \qquad (96)$$

This is just Newton's equation (1) with the Lorentz force (6). Note that this was derived for non-relativistic particles: later we will need to derive a relativistic equation of motion.

The matrix exponential exp(A) can be defined, as for the exponential of a number, by the series:

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \qquad (97)$$

where A^n is the *matrix power* of A to the power n. Using (97) we write:

$$S \cdot \exp\left(tS \cdot J\right) = S \cdot \left(1 + tS \cdot J + \frac{1}{2}t^2 S \cdot J \cdot S \cdot J \cdots\right) \qquad (98)$$

Since $S^2 = -I$, where I is the identity matrix, we can post-multiply the right hand side of (98) by $I = -S^2$:

$$S \cdot \exp\left(tS \cdot J\right) = -S \cdot \left(1 + tS \cdot J + \frac{1}{2}t^2 S \cdot J \cdot S \cdot J \cdots\right) S^2 \qquad (99)$$

Applying the initial factor $-S$ and one final factor of S to each term in the summation gives:

$$\begin{aligned} S \cdot \exp\left(tS \cdot J\right) &= \left(1 + tJ \cdot S + \frac{1}{2}t^2 J \cdot S \cdot J \cdot S \cdots\right) S \\ &= \exp\left(tJ \cdot S\right) \cdot S \end{aligned} \qquad (100)$$