Lagrangian and Hamiltonian Dynamics

Volker Perlick

(Lancaster University)

Lecture 1

The Passage from Newtonian to Lagrangian Dynamics

(Cockcroft Institute, 22 February 2010)

Subjects covered

Lecture 1: The Passage from Newtonian to Lagrangian Dynamics

Lecture 2: Discussion of Lagrangian Dynamics and Passage to Hamiltonian Dynamics

Lecture 3: Discussion of Hamiltonian Dynamics and Phase Flows

Lecture 4: Canonical Transformations

Lecture 5: Hamilton-Jacobi Theory

Lecture 6: Stability and Linearisation

The most elementary form to describe the dynamics of particles is in terms of Newton's equations (Newtonian dynamics). For many applications, e.g.

- **accelerator physics**
- **celestial mechanics**
- **systems with constraints**
- **passage to quantum mechanics**
- **...**

Newton's equations are not appropriate.

It is much more useful to use Lagrangian or Hamiltonian dynamics.

In the first lecture we start out from Newtonian dynamics and perform the passage to Lagrangian dynamics.

First consider *one* particle with mass m and trajectory $\vec{x}(t)$.

Assume that the force $\vec{F}(\vec{x}, \vec{v}, t)$ onto the particle is known. Then Newton's equation

$$m\,\ddot{\vec{x}}(t) = \vec{F}\big(\vec{x}(t), \dot{\vec{x}}(t), t\big)$$

gives a second-order differential equation for the trajectory.

Here and in the following, $(\,)^{\cdot} = \frac{d}{dt}$.

To every initial conditions

$$\vec{x}(0) = \vec{x}_0,\ \dot{\vec{x}}(0) = \vec{v}_0$$

there is a unique solution $\vec{x}(t)$.

This follows from the existence and uniqueness theorem for ordinary differential equations.

In Cartesian coordinates:

$$\vec{x}(t) \;=\; \sum_{i=1}^{3} x^i(t)\,\vec{e}_i\;, \qquad \dot{\vec{x}}(t) \;=\; \sum_{i=1}^{3} \dot{x}^i(t)\,\vec{e}_i\;,$$

$$\ddot{\vec{x}}(t) \;=\; \sum_{i=1}^{3} \ddot{x}^i(t)\,\vec{e}_i\;, \qquad \vec{F}(\vec{x},\vec{v},t) \;=\; \sum_{i=1}^{3} F^i(\vec{x},\vec{v},t)\,\vec{e}_i\;.$$

Newton's equations of motion read

$$m\,\ddot{x}^i(t) \;=\; F^i\big(\,\vec{x}(t),\dot{\vec{x}}(t),t\,\big)\;, \qquad i=1,2,3\,.$$

These equations are *not* covariant, i.e, they do not preserve their form if we change to curvilinear coordinates. (For proof see next page.) Curvilinear coordinates are useful because they can be adapted

- **to the symmetry of the situation (in accelerator physics, e.g., choose desired path of particle as coordinate line);**
- **to constraints (not very relevant to accelerator physics).**

Transformation from Cartesian coordinates (x^1, x^2, x^3) to new coordinates (x'^1, x'^2, x'^3):

$$x^i \;=\; x^i(x'^1, x'^2, x'^3) \;: \qquad \dot{x}^i(t) \;=\; \sum_{j=1}^{3} \frac{\partial x^i}{\partial x'^j}\big(x'^1(t), x'^2(t), x'^3(t)\big)\, \dot{x}'^j(t)\,.$$

$$\dot{\vec{x}}(t) \;=\; \sum_{i=1}^{3} \dot{x}^i(t)\, \vec{e}_i \;=\; \sum_{j=1}^{3} \dot{x}'^j(t) \underbrace{\sum_{i=1}^{3} \frac{\partial x^i}{\partial x'^j}\big(x'^1(t), x'^2(t), x'^3(t)\big)\, \vec{e}_i}_{\vec{e}\,'_j\left(x'^1(t), x'^2(t), x'^3(t)\right)} \;.$$

The $\vec{e}\,'_j$ are constant only if the new coordinates are rectilinear.

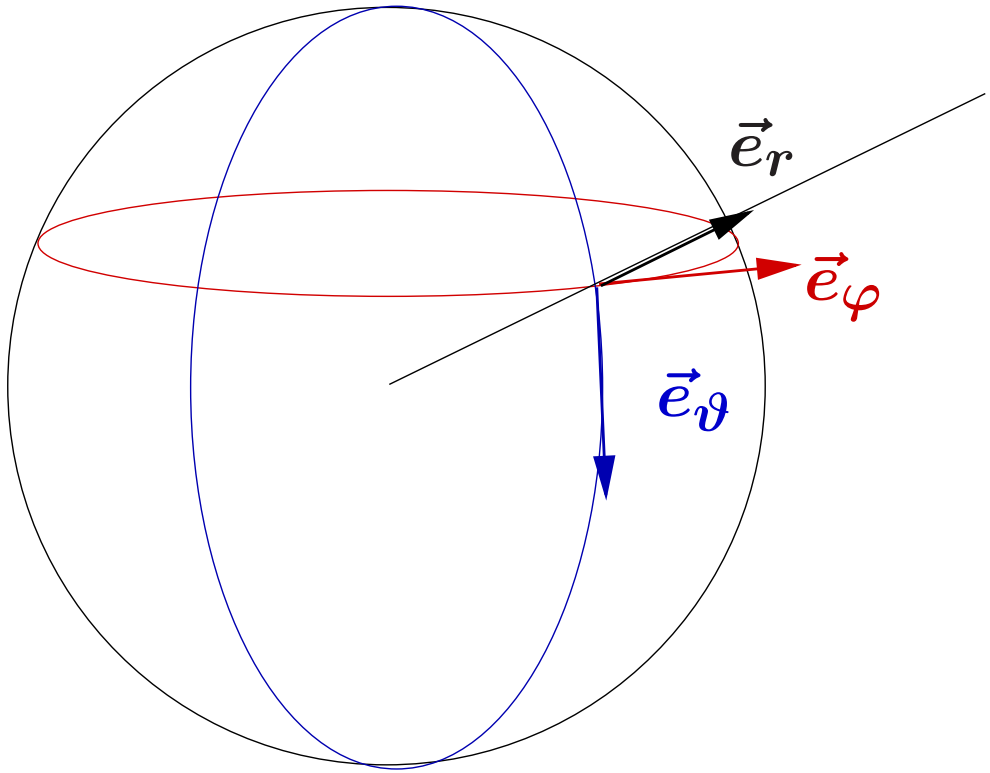
$$\ddot{\vec{x}}(t) \;=\; \sum_{i=1}^{3} \ddot{x}^i(t)\, \vec{e}_i \;=\; \sum_{j=1}^{3} \ddot{x}'^j(t)\, \vec{e}\,'_j\big(x'^1(t), x'^2(t), x'^3(t)\,\big)$$

$$+ \sum_{j=1}^{3} \sum_{i=1}^{3} \sum_{k=1}^{3} \frac{\partial^2 x^i}{\partial x'^j x'^k}\big(x'^1(t), x'^2(t), x'^3(t)\,\big)\, \dot{x}'^j(t)\, \dot{x}'^k(t)\, \vec{e}_i \;.$$

Thus, $\ddot{x}^i = 0 \;\Rightarrow\; \ddot{x}'^i \;=\; 0$ only if the x'^i are rectilinear coordinates.

Example: Newton's force-free equation in spherical polar coordinates $(x'^1, x'^2, x'^3) = (r, \vartheta, \varphi)$,

$$x^1 = r\sin\vartheta\cos\varphi\,,$$
$$x^2 = r\sin\vartheta\sin\varphi\,,$$
$$x^3 = r\cos\vartheta\,.$$



Then $\ddot{x}^i = 0$ **is equivalent to**

$$\ddot{r} - r\sin^2\vartheta\,\dot{\varphi}^2 - r\,\dot{\vartheta}^2 = 0\,,$$
$$\ddot{\vartheta} + \frac{2}{r}\dot{r}\,\dot{\vartheta} - \sin\vartheta\cos\vartheta\,\dot{\varphi}^2 = 0\,,$$
$$\ddot{\varphi} + \frac{2}{r}\dot{r}\,\dot{\varphi} + 2\cot\vartheta\,\dot{\vartheta}\,\dot{\varphi}^2 = 0\,.$$

We want to reformulate Newton's equation in a way that is covariant, i.e., we seek a formulation that takes the same form in any coordinates.

This will be the *Lagrangian formulation*. We do this first for some special cases before discussing the general framework of Lagrangian dynamics.

(a) Particle in a potential

Assume that the force is of the form

$$\vec{F}(\vec{x}) \;=\; -\,\boldsymbol{\nabla} V(\vec{x})$$

with some scalar function $V(\vec{x})$. Then Newton's equation takes the form

$$m\,\ddot{\vec{x}}(t) \;=\; -\,\boldsymbol{\nabla} V\big(\vec{x}(t)\big)\,.$$

(Example: Particle with charge q in an electrostatic field $\vec{E}(\vec{x}) \;=\; -\,\boldsymbol{\nabla}\,\phi(\vec{x})$, where $V(\vec{x}) \;=\; q\,\phi(\vec{x})$.)

Kinetic energy: $T(\dot{\vec{x}}) = \frac{m}{2}\,|\dot{\vec{x}}|^2$

Potential energy: $V(\vec{x})$

Total energy $T+V$ is preserved along trajectory:

$$\frac{d}{dt}\Big(\,T(\dot{\vec{x}}(t)) \,+\, V(\vec{x}(t))\,\Big) \,=\, \Big(\,m\,\ddot{\vec{x}}(t) \,+\, \boldsymbol{\nabla} V(\vec{x}(t))\,\Big)\cdot\dot{\vec{x}}(t) \,=\, 0\,.$$

Introduce *Lagrange function* $L = T - V$:

$$L\big(x^1,x^2,x^3,\dot{x}^1,\dot{x}^2,\dot{x}^3\big) \,=\, \frac{m}{2}\Big(\,\big(\dot{x}^1\big)^2 \,+\, \big(\dot{x}^2\big)^2 \,+\, \big(\dot{x}^3\big)^2\,\Big) \,-\, V\big(\,x^1\,,\,x^2\,,\,x^3\,\big)\,.$$

Then we have $\dfrac{\partial L}{\partial \dot{x}^i} \,=\, m\,\dot{x}^i\,,\qquad \dfrac{\partial L}{\partial x^i} \,=\, -\,\dfrac{\partial V}{\partial x^i}\,.$

Hence, Newton's equations $m\,\ddot{x}^i + \dfrac{\partial V}{\partial x^i} = 0$ **are equivalent to the *Euler-Lagrange equations***

$$\frac{d}{dt}\,\frac{\partial L}{\partial \dot{x}^i} \,-\, \frac{\partial L}{\partial x^i} \,=\, 0\,. \quad \text{(EL)}$$

The Euler-Lagrange equations are covariant, i.e., they preserve their form under arbitrary coordinate transformations.

Of course, the Lagrange function looks different if expressed in other coordinates. What covariance means is the following:

Assume that L satisfies (EL). Make a coordinate transformation

$$x^i = x^i\big(x'^1, x'^2, x'^3\big) \,.$$

Define the new Lagrange function L' via

$$L'\big(x'^1, x'^2, x'^3, \dot{x}'^1, \dot{x}'^2, \dot{x}'^3\big) = L\big(x^1, x^2, x^3, \dot{x}^1, \dot{x}^2, \dot{x}^3\big) \,.$$

Then L' satisfies the primed Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\partial L'}{\partial \dot{x}'^i} - \frac{\partial L'}{\partial x'^i} = 0 \,.$$

We now prove that the Euler-Lagrange equations are, indeed, covariant in this sense.

Assume that (EL) holds, $\displaystyle \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = 0\,.$

Make coordinate transformation $x^i = x^i(x'^1, x'^2, x'^3)\,.$

$$\dot{x}^i = \sum_{j=1}^{3} \frac{\partial x^i}{\partial x'^j}\,\dot{x}'^j\,, \qquad \frac{\partial \dot{x}^i}{\partial x'^k} = \sum_{j=1}^{3} \frac{\partial^2 x^i}{\partial x'^j x'^k}\,\dot{x}'^j\,, \qquad \frac{\partial \dot{x}^i}{\partial \dot{x}'^k} = \frac{\partial x^i}{\partial x'^k}\,.$$

$$\frac{d}{dt}\frac{\partial L'}{\partial \dot{x}'^k} = \frac{d}{dt}\Big(\sum_{i=1}^{3} \frac{\partial L}{\partial \dot{x}^i}\frac{\partial \dot{x}^i}{\partial \dot{x}'^k}\Big) = \frac{d}{dt}\Big(\sum_{i=1}^{3} \frac{\partial L}{\partial \dot{x}^i}\frac{\partial x^i}{\partial x'^k}\Big) =$$

$$= \sum_{i=1}^{3} \Big(\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i}\Big)\frac{\partial x^i}{\partial x'^k} + \sum_{i=1}^{3} \frac{\partial L}{\partial \dot{x}^i}\frac{d}{dt}\Big(\frac{\partial x^i}{\partial x'^k}\Big) =$$

$$= \sum_{i=1}^{3} \Big(\frac{\partial L}{\partial x^i}\Big)\frac{\partial x^i}{\partial x'^k} + \sum_{i=1}^{3} \frac{\partial L}{\partial \dot{x}^i}\sum_{j=1}^{3} \Big(\frac{\partial^2 x^i}{\partial x'^k \partial x'^j}\,\dot{x}'^j\Big) =$$

$$= \sum_{i=1}^{3} \Big(\frac{\partial L}{\partial x^i}\frac{\partial x^i}{\partial x'^k} + \frac{\partial L}{\partial \dot{x}^i}\frac{\partial \dot{x}^i}{\partial x'^k}\Big) = \frac{\partial L'}{\partial x'^k}\,.$$

Example: In spherical polar coordinates

$$x^1 \,=\, r\sin\vartheta\cos\varphi\,, \qquad x^2 \,=\, r\sin\vartheta\sin\varphi\,, \qquad x^3 \,=\, r\cos\vartheta\,,$$

the Lagrange function

$$L(\vec{x}^1, x^2, x^3, \dot{x}^1, \dot{x}^2, \dot{x}^3) \,=\, \frac{m}{2}\Big(\,(\dot{x}^1)^2 \,+\, (\dot{x}^2)^2 \,+\, (\dot{x}^3)^2\,\Big) \,-\, V(x^1, x^2, x^3)\,.$$

takes the form

$$L'(r, \vartheta, \varphi, \dot{r}, \dot{\vartheta}, \dot{\varphi}) \,=\, \frac{m}{2}\Big(\,\dot{r}^2 \,+\, r^2\big(\,\dot{\vartheta}^2 \,+\, \sin^2\vartheta\,\dot{\varphi}^2\,\big)\,\Big) \,-\, V'(r, \vartheta, \varphi)\,.$$

Remark: If a force field $\vec{F}(\vec{x})$ has non-vanishing curl, $\boldsymbol{\nabla}\times\vec{F}\neq\vec{0}$, it cannot be written as $\vec{F} \,=\, -\boldsymbol{\nabla}V$. It is then impossible to bring the equation of motion $m\ddot{\vec{x}} \,=\, \vec{F}$ into the form of the Euler-Lagrange equations.

(b) Charged particle in electromagnetic field (non-relativistic)

On a particle with charge q, an electromagnetic field $\vec{E}(\vec{x},t)$, $\vec{B}(\vec{x},t)$ exerts the *Lorentz force*. For non-relativistic motion (i.e., $|\dot{\vec{x}}| << c$), the Lorentz force equation reads

$$m\,\ddot{\vec{x}}(t) \;=\; q\left(\,\vec{E}\big(\vec{x}(t),t\big) \;+\; \dot{\vec{x}}(t)\times\vec{B}\big(\vec{x}(t),t\big)\,\right).$$

$\vec{E}$ and $\vec{B}$ have to satisfy the Maxwell equations, in particular

$$\nabla\cdot\vec{B} \;=\; 0\,, \qquad\qquad \nabla\times\vec{E} \;+\; \frac{\partial}{\partial t}\vec{B} \;=\; \vec{0}\,.$$

Owing to the first equation, $\vec{B}$ can be written as the curl of a vector potential $\vec{A}$,

$$\vec{B} \;=\; \nabla\times\vec{A}\,.$$

Owing to the second equation, $\vec{E}+\partial\vec{A}/\partial t$ can then be written as the gradient of a scalar potential ϕ,

$$\vec{E} \;=\; -\,\nabla\phi \;-\; \frac{\partial}{\partial t}\vec{A}\,.$$

We define the Lagrange function in this situation by

$$L \,=\, \frac{m}{2}\,\big|\,\dot{\vec{x}}\,\big|^2 \,-\, q\,\phi \,+\, q\,\vec{A}\cdot\dot{\vec{x}}\,.$$

Then the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} \,-\, \frac{\partial L}{\partial x^i} \,=\, 0$$

take the form

$$\frac{d}{dt}\big(\,m\,\dot{\vec{x}} \,+ q\,\vec{A}\,\big) \,+\, q\nabla\phi \,-\, q\,\nabla\big(\vec{A}\cdot\dot{\vec{x}}\big) \,=\, 0\,,$$

$$m\,\ddot{\vec{x}} \,=\, \underbrace{-\,q\,\nabla\phi \,-\, q\,\frac{\partial}{\partial t}\vec{A}}_{q\,\vec{E}} \,+\, \underbrace{q\Big(\,\nabla(\vec{A}\cdot\dot{\vec{x}}) \,-\, (\dot{\vec{x}}\cdot\nabla)\vec{A}\,\Big)}_{q\,\dot{\vec{x}}\times(\nabla\times\vec{A})\,=\,q\,\dot{\vec{x}}\times\vec{B}}\,.$$

Thus, the Euler-Lagrange equations are equivalent to the Lorentz force equation.

Owing to the covariance of the Euler-Lagrange equations, L can be rewritten in arbitrary coordinates if this is desired.

In addition, the freedom of changing the potentials by *gauge transformations*

$$\vec{A} \longmapsto \vec{A} + \boldsymbol{\nabla} f\,, \qquad \phi \longmapsto \phi - \frac{\partial}{\partial t} f$$

allows to change the Lagrangian according to

$$L \longmapsto L + q\frac{\partial}{\partial t} f + q\,\boldsymbol{\nabla} f \cdot \dot{\vec{x}} = L + \frac{d}{dt}(qf)$$

with an arbitrary function $f(\vec{x}, t)$.

Such a change of the Lagrangian leaves the Euler-Lagrange equations unaltered.

(c) Charged particle in electromagnetic field (relativistic)

For relativistic motion, the Lorentz force equation reads

$$\frac{d}{dt}\left(\frac{m\,\dot{\vec{x}}(t)}{\sqrt{1-\frac{|\,\dot{\vec{x}}(t)\,|^2}{c^2}}}\right) = q\left(\vec{E}(\vec{x}(t),t) + \dot{\vec{x}}(t)\times\vec{B}(\vec{x}(t),t)\right) \quad \text{(LF)}$$

where m is the particle's rest mass. The difference to the non-relativistic case is in the square-root on the left-hand side.

We define the Lagrange function

$$L = -m\,c^2\sqrt{1-\frac{|\,\dot{\vec{x}}(t)\,|^2}{c^2}} - q\,\phi + q\,\vec{A}\cdot\dot{\vec{x}}$$

which for small velocities, $\sqrt{1-\frac{|\,\dot{\vec{x}}(t)\,|^2}{c^2}} \approx 1-\frac{1}{2}\frac{|\dot{\vec{x}}(t)|^2}{c^2}$, reproduces the Lagrange function for the non-relativistic case up to an irrelevant additive constant.

With the relativistic Lagrange function, the Euler-Lagrange equations yield

$$\frac{d}{dt}\Big(\frac{m\,\dot{\vec{x}}}{\sqrt{1-\frac{|\dot{\vec{x}}|^2}{c^2}}}+q\,\vec{A}\Big)+q\nabla\phi-q\,\nabla(\vec{A}\cdot\dot{\vec{x}})=0\,,$$

$$\frac{d}{dt}\Big(\frac{m\,\dot{\vec{x}}}{\sqrt{1-\frac{|\dot{\vec{x}}|^2}{c^2}}}\Big)=-q\,\nabla\phi-q\frac{\partial}{\partial t}\vec{A}+q\Big(\nabla(\vec{A}\cdot\dot{\vec{x}})-(\dot{\vec{x}}\cdot\nabla)\vec{A}\Big)\,.$$

The right-hand side is the same as for the non-relativistic case, so

$$\frac{d}{dt}\Big(\frac{m\,\dot{\vec{x}}}{\sqrt{1-\frac{|\dot{\vec{x}}|^2}{c^2}}}\Big)=q\Big(\vec{E}+\dot{\vec{x}}\times\vec{B}\Big)\,,$$

i.e., the Euler-Lagrange equations of our Lagrange function are, indeed, equivalent to the relativistic Lorentz force equation (LF).

(d) Damped harmonic oscillator

Sometimes it is said that dissipative systems (i.e., systems with friction or other damping mechanisms) cannot be put into Lagrangian form.

This is true only as long as one wants to have a *time-independent* Lagrange function.

We consider the damped harmonic oscillator and show that it admits a (time-dependent) Lagrange function.

Newton's equation for a damped harmonic oscillator reads

$$m\,\ddot{x} \;=\; -\,k\,x \;-\; \lambda\,\dot{x}\;,$$

where the constant k describes the restoring force and the constant λ describes the damping.

For the Lagrangian

$$L(x, \dot{x}, t) \;=\; e^{\frac{\lambda t}{m}} \left(\frac{m}{2}\, \dot{x}^2 \;-\; \frac{k}{2}\, x^2 \right)$$

the Euler-Lagrange equation

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \;-\; \frac{\partial L}{\partial x} \;=\; 0$$

reads

$$\frac{d}{dt} \left(e^{\frac{\lambda t}{m}}\, m\, \dot{x} \right) \;+\; e^{\frac{\lambda t}{m}}\, k\, x \;=\; 0\,,$$

$$e^{\frac{\lambda t}{m}} \left(m\, \ddot{x} \;+\; \lambda\, \dot{\vec{x}} \;+\; k\, x \right) \;=\; 0\,,$$

so it is equivalent to Newton's equation.

This example demonstrates that sometimes it is necessary to work with an explicitly time-dependent Lagrange function even if the force is time-independent.

(e) Interacting particles

For N particles wthout constraints, the Lagrange function depends on $6N$ coordinates (and possibly on time): $3N$ position coordinates and $3N$ velocity coordinates.

For *non-interacting* particles, the total Lagrange function is of the form

$$L = \sum_{I=1}^{N} L_I$$

where I labels the particles and L_I depends on position and velocity coordinates of the Ith particle only.

In this case the $3N$ Euler-Lagrange equations for L decompose into the Euler-Lagrange equations for the L_I, i.e., the motion can be studied for each particle separately.

For *interacting* particles, the Lagrange function (if it exists) contains interaction terms, i.e., summands which depend on position and velocity coordinates of two or more particles. For pair-interaction

$$L = \sum_{I=1}^{N} \Big(L_I + \sum_{\substack{J=1 \\ J>I}}^{N} L_{IJ} \Big)$$

where L_{IJ} depends on position and velocity coordinates of the Ith and the Jth particle. Here is an example.

For two particles with (Newtonian) gravitational interaction, the Lagrange function is

$$L = \underbrace{\frac{m_1}{2} |\, \dot{\vec{x}}_1 \,|^2}_{L_1} + \underbrace{\frac{m_2}{2} |\, \dot{\vec{x}}_2 \,|^2}_{L_2} + \underbrace{\frac{G\, m_1\, m_2}{|\, \vec{x}_2 - \vec{x}_1 \,|}}_{L_{12}} .$$

(Coulomb interaction for slowly moving charged particles is analogous.)

The six Euler-Lagrange equations are

$$\cancel{m_1}\left(\ddot{\vec{x}}_1 - \frac{G\,m_2\left(\vec{x}_2 - \vec{x}_1\right)}{|\,\vec{x}_2 - \vec{x}_1\,|^3}\right) = \vec{0}\,,$$

$$\cancel{m_2}\left(\ddot{\vec{x}}_2 - \frac{G\,m_1\left(\vec{x}_1 - \vec{x}_2\right)}{|\,\vec{x}_1 - \vec{x}_2\,|^3}\right) = \vec{0}\,.$$

Summary: In many relevant (but not in all) cases, the equations of motion can be written in the form of the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = 0\,, \quad i = 1, \ldots, n\,.$$

The dynamics is then completely coded in the Lagrange function

$L(x^1 \ldots x^n, \dot{x}^1, \ldots, \dot{x}^n, t)$.

For N particles without constraints, $n = 3N$.

The Euler-Lagrange equations are covariant with respect to arbitrary coordinate transformations.

Lagrangian and Hamiltonian Dynamics

Volker Perlick

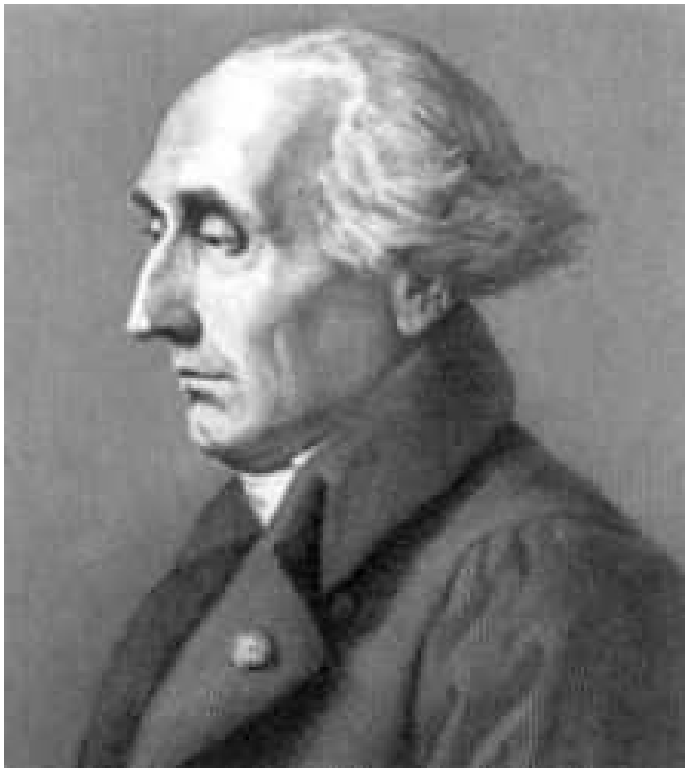
(Lancaster University)

Lecture 2

Discussion of Lagrangian Dynamics and Passage to Hamiltonian Dynamics

(Cockcroft Institute, 1 March 2010)

The inventor of Lagrangian dynamics:



Joseph-Louis Lagrange (1736, Torino – 1813, Paris)

Director of Mathematics at Berlin Academy (1766 to 1787)

“Mécanique analytique” (published 1788 in Paris)

Recall:

In the Lagrangian formulation, all information on the dynamics of a system is coded in the Lagrange function (or "Lagrangian")

$$\boldsymbol{L(x^1, \ldots, x^n, \dot{x}^1, \ldots \dot{x}^n, t)} \; .$$

n is the number of degrees of freedom.

For N particles without constraints, $n = 3N$.

Not all dynamical systems admit a Lagrangian formulation, but many important ones do. In particular, we have seen that the motion of charged particles in an electromagnetic field, both non-relativistically and relativistically, can be put into Lagrangian form.

If L is known, the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = 0 , \quad i = 1, \ldots, n$$

determine the dynamics of the system. More precisely, the Euler-Lagrange equations give a system of second-order ordinary differential equations for $\big(x^1(t), \ldots, x^n(t) \big)$,

$$\sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{x}^i \dot{x}^j} \ddot{x}^j + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{x}^i x^j} \dot{x}^j - \frac{\partial L}{\partial x^i} = 0 , \quad i = 1, \ldots, n .$$

This admits a unique solution, if initial conditions for x^i and $\dot{x}^i$ are given, provided that the system can be solved for the $\ddot{x}^j$. A necessary and sufficient condition for this is

$$\det\Big(\frac{\partial^2 L}{\partial \dot{x}^i \dot{x}^j} \Big) \neq 0 .$$

If this is true, L is called "regular".

In any of the examples considered in Lecture 1, the Lagrange function was regular. (However, non-regular Lagrangians play an inportant role in other branches of physics, in particluar in the Lagrange formalism for fields. P. A. M. Dirac has developed a special formalism for dealing with them.)

Recall that the Euler-Lagrange equations are covariant, i.e, they preserve their form under arbitrary coordinate transformations

$$\left(x^1, \ldots, x^n\right) \mapsto \left(x'^1, \ldots, x'^n\right).$$

It is often recommendable to use coordinates adapted to the symmetries of the system.

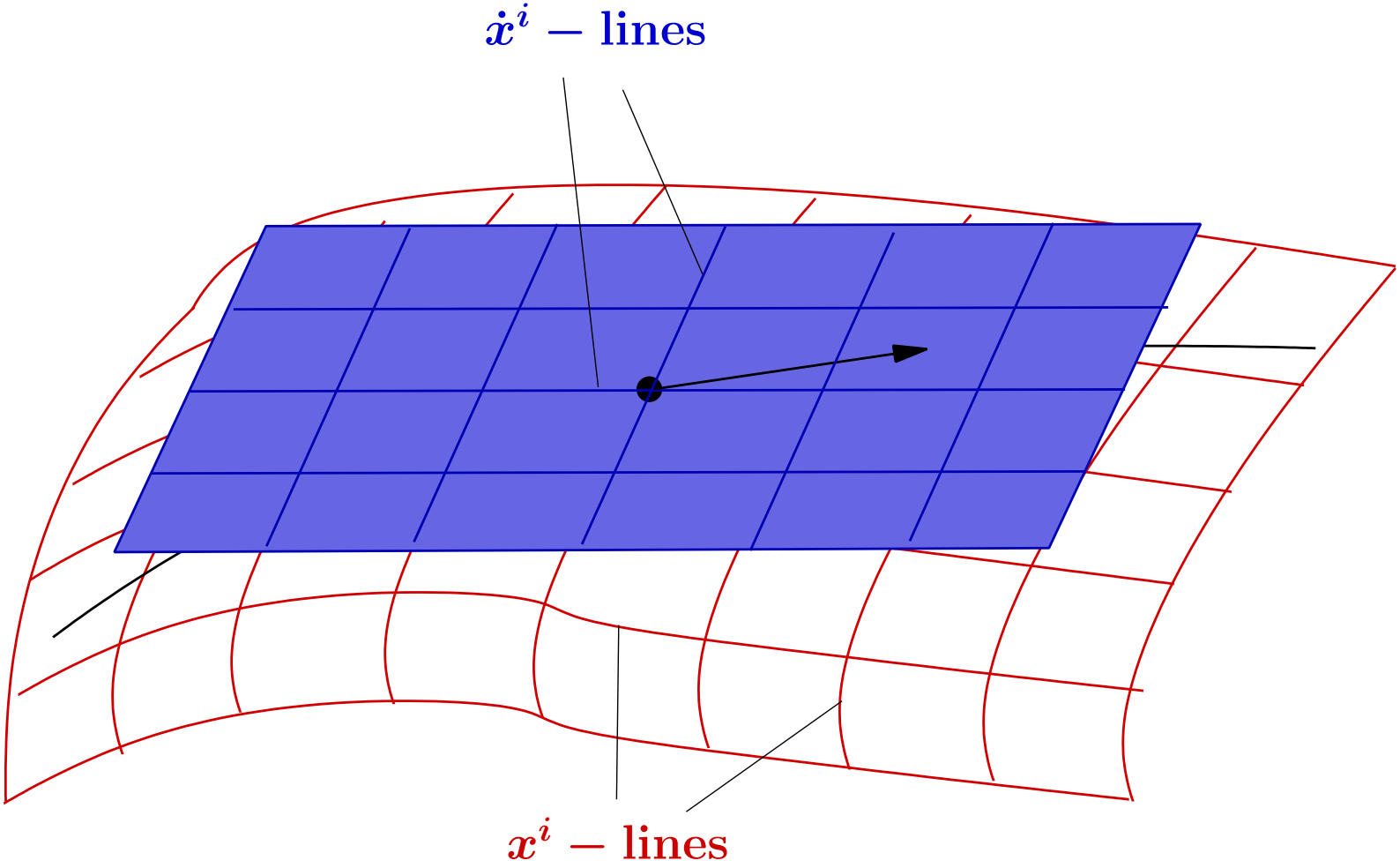
The space coordinatised by the $\left(x^1, \ldots, x^n\right)$ is called the "configuration space". The trajectories of the dynamical system are curves in the configuration space.

The sum $\left(x_1^1, \ldots, x_1^n\right) + \left(x_2^1, \ldots, x_2^n\right)$ of two points in configuration space has no coordinate-independent meaning. It is invariant only under linear coordinate transformations. However, we want to be free to make non-linear coordinate transformations.

Thus, in general the configuration space is not a vector space. It is what in Differential Geometry is called an n-dimensional "manifold". Therefore, instead of "configuration space" one often says "configuration manifold".

At any fixed point $\left(x_1^1, \ldots, x_1^n\right)$ of the configuration space, the set of possible n-tuples $\left(\dot{x}^1, \ldots, \dot{x}^n\right)$ corresponds to the set of possible tangent vectors of curves through that point. In other words, the $\left(\dot{x}^1, \ldots, \dot{x}^n\right)$ coordinatise the tangent space to the configuration manifold at the chosen point.

The union of all tangent spaces, at all points in the configuration manifold is called the "tangent bundle". In the language of Differential Geometry, the Lagrangian $L(x^1, \ldots, x^n, \dot{x}^1, \ldots, \ldots x^n, t)$ is a time-dependent function on the tangent bundle of the configuration space.



In the Lagrange formalism, there is a particularly simple relation between symmetries and conservation laws.

Assume that the Lagrange function is independent of one of the coordinates, say

$$\frac{\partial L}{\partial x^1} = 0 \, .$$

x^1 is then called a "cyclic coordinate" and the Euler-Lagrange equations say that $\frac{\partial L}{\partial \dot{x}^1}$ is a constant of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^1} = 0 \, .$$

Thus, integrating the equations of motion is greatly simplified if we can choose the coordinates such that L is independent of some of the coordinates.

Example: Consider a Lagrange function for one particle in spherical polar coordinates, $L(r, \vartheta, \varphi, \dot{r}, \dot{\vartheta}, \dot{\varphi})$. If L is independent of φ, we get

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}} = 0 \; .$$

The constant of motion $\partial L/\partial \dot{\varphi}$ is the z-component of the angular momentum. If L is independent of φ and ϑ, this argument applies to arbitrarily rotated axes, i.e., all three components of the angular momentum must be conserved.

The case that L is independent of one of the velocity coordinates, say $\partial L/\partial \dot{x}^1 = 0$, only occurs for systems with constraints. x^1 is then called a "Lagrange multiplier".

There is a more general connection between symmetries and conservation laws, for transformations not necessarily along x^i lines or $\dot{x}^i$-lines. We will discuss it in the Hamltonian formalism.

The Euler-Lagrange equations are equivalent to a variational principle, known as the "Principle of Stationary Action" or "Hamilton's Principle". (The latter name is a bit confusing: Hamilton's Principle belongs into Lagrangian dynamics, not into Hamiltonian dynamics!)

The "action functional"

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

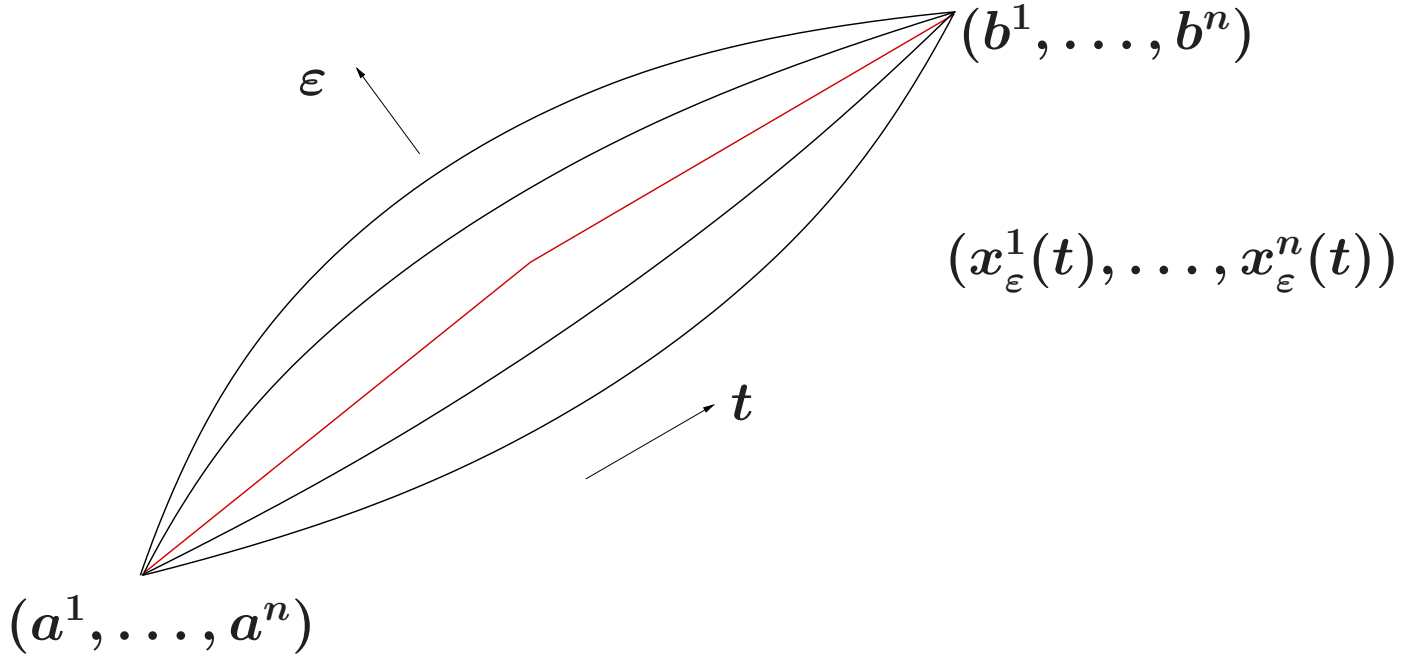
assigns to each curve in the configuration space a number. The Principle of Stationary Action says that the solution curves of the Euler-Lagrange equations are those curves that make the action functional stationary, i. e., a minimum, a maximum, or a saddle.

More precisely: Fix a curve in configuration space and consider a "variation", i.e., a one-parameter family of neighbouring curves $\left(x^1_\varepsilon(t),\ldots,x^n_\varepsilon(t)\right)$, all defined on the same time interval $[t_1,t_2]$ with the same end points,

$$\left(x^1_\varepsilon(t_1),\ldots,x^n_\varepsilon(t_1)\right) \;=\; \left(a^1,\ldots,a^n\right)\,,$$

$$\left(x^1_\varepsilon(t_2),\ldots,x^n_\varepsilon(t_2)\right) \;=\; \left(b^1,\ldots,b^n\right)\,,$$

such that $\varepsilon = 0$ gives the original curve.



Claim: The original curve is a solution of the Euler-Lagrange equations if and only if $\partial_\varepsilon S\,\big|_{\varepsilon=0} = 0$ for all possible variations.

Proof:

$$\partial_\varepsilon S = \partial_\varepsilon \int_{t_1}^{t_2} L\big(x_\varepsilon^1(t), \ldots, x_\varepsilon^n(t), \dot{x}_\varepsilon^1(t), \ldots, \dot{x}_\varepsilon^n(t), t \big)\, dt =$$

$$= \int_{t_1}^{t_2} \Big(\sum_{i=1}^{n} \frac{\partial L}{\partial x_\varepsilon^i}\, \partial_\varepsilon x_\varepsilon^i + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_\varepsilon^i}\, \partial_\varepsilon \dot{x}_\varepsilon^i \Big)\, dt =$$

$$= \int_{t_1}^{t_2} \Big(\sum_{i=1}^{n} \frac{\partial L}{\partial x_\varepsilon^i}\, \partial_\varepsilon x_\varepsilon^i + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_\varepsilon^i}\, (\partial_\varepsilon x_\varepsilon^i)^{\cdot} \Big)\, dt =$$

$$= \int_{t_1}^{t_2} \sum_{i=1}^{n} \frac{\partial L}{\partial x_\varepsilon^i}\, \partial_\varepsilon x_\varepsilon^i\, dt + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_\varepsilon^i}\, \partial_\varepsilon x_\varepsilon^i \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \sum_{i=1}^{n} \Big(\frac{\partial L}{\partial \dot{x}_\varepsilon^i}\Big)^{\cdot}\, \partial_\varepsilon x_\varepsilon^i\, dt\,.$$

The middle term vanishes because $\partial_\varepsilon x_\varepsilon^i(t_1) = \partial_\varepsilon x_\varepsilon^i(t_2) = 0$ as all curves have the same end points.

So we get

$$\partial_\varepsilon S \;=\; \int_{t_1}^{t_2} \sum_{i=1}^{n} \Big(\frac{\partial L}{\partial x^i_\varepsilon} \;-\; (\frac{\partial L}{\partial \dot{x}^i_\varepsilon})^{\cdot} \Big)\, \partial_\varepsilon x^i_\varepsilon \, dt \;.$$

Setting $\varepsilon = 0$,

$$\partial_\varepsilon S\big|_{\varepsilon=0} \;=\; \int_{t_1}^{t_2} \sum_{i=1}^{n} \Big(\frac{\partial L}{\partial x^i} \;-\; (\frac{\partial L}{\partial \dot{x}^i})^{\cdot} \Big)\, \partial_\varepsilon x^i_\varepsilon \Big|_{\varepsilon=0} \, dt \;.$$

The right-hand side is zero, for all possible variations, if and only if the bracket vanishes, i.e., if and only if the Euler-Lagrange equations hold. □

Remark: For many cases of interest, the solution is actually a minimum. However, saddles also occur frequently. Therefore, the older name "Principle of Minimal Action" is not justified.

If the Lagrange function is changed by a total time derivative,

$$L \mapsto L + \frac{d}{dt} f \,,$$

where f is an arbitrary function of x^i, $\dot{x}^i$ and t, the action functional $S = \int_{t_1}^{t_2} L\, dt$ changes according to

$$S \mapsto S + f\big(x^1(t), \ldots, x^n(t), \dot{x}^1(t), \ldots, \dot{x}^n(t), t\big)\Big|_{t_1}^{t_2} .$$

For variations with fixed end points, $S \mapsto S +$ constant .

This proves that adding a total time derivative to the Lagrangian does not change the solutions to the Euler-Lagrange equations. We have already met an example of this kind: For a charged particle in an electromagnetic field, a gauge transformation of the electromagnetic potential adds a total time derivative to the Lagrangian.

We now perform the passage to the Hamiltonian formulation. In a sense, the Hamiltonian formalism will turn out to be even more powerful than the Lagrangian formalism. In the Lagrangian formalism the equations of motion are covariant with respect to arbitrary transformations of the x^i. The Hamiltonian formalism puts the equations of motion into a form that is covariant even with respect to a larger group of transformations, called "canonical transformations".

Given a Lagrangian, we define the "canonical momenta"

$$p_i = \frac{\partial L}{\partial \dot{x}^i} \quad i = 1, \ldots, n \ .$$

This set of equations can be solved for the $\dot{x}^i$ if and only if the Lagrangian is regular,

$$\det\Big(\frac{\partial^2 L}{\partial \dot{x}^i \dot{x}^j}\Big) \neq 0 \ .$$

If this regularity condition is satisfied, we can define the "Hamilton function" (or "Hamiltonian" for short) as

$$H\big(x^1,\ldots,x^n,p_1,\ldots,p_n,t\big) \;=\; \sum_{i=1}^{n} p_i\,\dot{x}^i \;-\; L\big(x^1,\ldots,x^n,\dot{x}^1,\ldots,\dot{x}^n,t\big)\,.$$

On the right-hand side the $\dot{x}^i$ have to be expressed in terms of the x^j and p_j which, by our regularity assumption, is possible.

The transformation from L to H , at fixed $x^1,\ldots,x^n,t$, is called a "Legendre transformation".

The total differential of the Hamiltonian is

$$dH \;=\; \cancel{\sum_{i=1}^{n} p_i\,d\dot{x}^i} \;+\; \sum_{i=1}^{n} \dot{x}^i\,dp_i \;-\; \sum_{i=1}^{n} \frac{\partial L}{\partial x^i}\,dx^i \;-\; \cancel{\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}^i}\,d\dot{x}^i} \;-\; \frac{\partial L}{\partial t}\,dt\,.$$

We compare with the identity

$$dH \;=\; \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\,dp_i \;+\; \sum_{i=1}^{n} \frac{\partial H}{\partial x^i}\,dx^i \;+\; \frac{\partial H}{\partial t}\,dt\,.$$

We find

$$\frac{\partial H}{\partial p_i} = \dot{x}^i \,, \qquad \frac{\partial H}{\partial x^i} = -\frac{\partial L}{\partial x^i} \,, \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \,.$$

Thus, for regular Lagrangians, the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} = \dot{p}_i$$

are equivalent to "Hamilton's equations"

$$\frac{\partial H}{\partial p_i} = \dot{x}^i \,, \qquad \frac{\partial H}{\partial x^i} = -\dot{p}_i \,, .$$

The Euler-Lagrange equations are second-order ordinary differential equations for $x^1(t), \ldots, x^n(t)$.

Hamilton's equations are first-order ordinary differential equations for $x^1(t), \ldots, x^n(t), p_1(t), \ldots, p_n(t)$.

We now determine the Hamiltonian for some examples.

(a) Particle in a potential

Recall the Lagrangian for a particle in a potential $V(\vec{x})$,

$$L(\vec{x},\dot{\vec{x}}) \;=\; T(\,\dot{\vec{x}}\,) \;-\; V(\,\vec{x}\,) \;=\; \frac{m}{2}\,|\,\dot{\vec{x}}\,|^2 \;-\; V(\,\vec{x}\,)\,.$$

In Cartesian coordinates, the canonical momenta coincide with the usual mechanical momentum coordinates,

$$p_i \;=\; \frac{\partial L}{\partial \dot{x}^i} \;=\; m\,\dot{x}^i\,.$$

Solving for the velocities gives, in vector notation,

$$\dot{\vec{x}} \;=\; \frac{1}{m}\,\vec{p}\,.$$

Thus, the Hamiltonian is

$$H(\vec{x},\vec{p}) \;=\; \vec{p}\cdot\dot{\vec{x}} \;-\; L(\vec{x},\dot{\vec{x}}) \;=\; \vec{p}\cdot\dot{\vec{x}} \;-\; \frac{m}{2}\,|\,\dot{\vec{x}}\,|^2 \;+\; V(\,\vec{x}\,) \;=$$

$$=\; \frac{1}{m}\,|\,\vec{p}\,|^2 \;-\; \frac{m}{2}\,\frac{1}{m^2}\,|\,\vec{p}\,|^2 \;+\; V(\,\vec{x}\,) \;=\; \frac{1}{2m}\,|\,\vec{p}\,|^2 \;+\; V(\,\vec{x}\,)\,.$$

As

$$\frac{1}{2m}\left|\,\vec{p}\,\right|^2 \;=\; \frac{m}{2}\left|\,\dot{x}\,\right|^2$$

is the kinetic energy, the Hamiltonian is equal to the total energy, $H = T + V$, expressed in terms of position and momentum coordinates.

(b) Charged particle in electromagnetic field (non-relativistic)

For a charged particle in an electromagnetic field, the non-relativistic Lagrangian was

$$L \;=\; \frac{m}{2}\left|\,\dot{\vec{x}}\,\right|^2 \;-\, q\,\phi \;+\; q\,\vec{A}\cdot\dot{\vec{x}}$$

where ϕ and $\vec{A}$ are the electromagnetic potentials. In Cartesian coordinates, the canonical momenta are

$$p_i \;=\; \frac{\partial L}{\partial \dot{x}^i} \;=\; m\,\dot{x}^i \;+\; q\,A_i\,.$$

Solving for the velocities gives, in vector notation,

$$\dot{\vec{x}} \;=\; \frac{1}{m}\,(\,\vec{p}\,-\,q\,\vec{A}\,)\,.$$

The Hamiltonian reads

$$H(\vec{x},\vec{p}) \;=\; \vec{p}\cdot\dot{\vec{x}} \;-\; L(\vec{x},\dot{\vec{x}}) \;=$$

$$=\vec{p}\cdot\frac{1}{m}\,(\vec{p}-q\vec{A}) \;-\; \frac{m}{2}\,\frac{1}{m^2}\,\big|\,\vec{p}\,-\,q\,\vec{A}\,\big|^2 \;+\; q\,\phi \;-\; q\,\vec{A}\cdot\frac{1}{m}\,(\vec{p}-q\vec{A}) \;=$$

$$=\;\frac{1}{2m}\,\big|\,\vec{p}\,-\,q\,\vec{A}\,\big|^2 \;+\; q\,\phi\,.$$

The first term is, again, the kinetic energy.

(c) Charged particle in electromagnetic field (relativistic)

For relativistic motion, the Lagrangian was

$$L \;=\; -\,m\,c^2\,\sqrt{1\;-\;\frac{|\,\dot{\vec{x}}(t)\,|^2}{c^2}}\;\;-\;q\,\phi\;+q\,\vec{A}\cdot\dot{\vec{x}}$$

Now the canonical momenta, in Cartesian coordinates, are

$$p_i \;=\; \frac{\partial L}{\partial \dot{x}^i} \;=\; \frac{m\,\dot{x}^i}{\sqrt{1\;-\;\frac{|\,\dot{\vec{x}}\,|^2}{c^2}}}\;+\;q\,A_i\,.$$

To solve this for the velocities, we use vector notation,

$$\vec{p}\;-\;q\,\vec{A}\;=\;\frac{m\,\vec{v}}{\sqrt{1\;-\;\frac{|\,\vec{v}\,|^2}{c^2}}}\,.$$

Squaring both sides and solving for $|\,\vec{v}\,|^2$ yields

$$|\,\vec{v}\,|^2\;=\;\frac{c^2\,\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}{c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}\,.$$

This allows to express the velocities in terms of the momenta,

$$\vec{v} \;=\; \frac{c\,(\,\vec{p}\;-\;q\,\vec{A}\,)}{\sqrt{\,c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}}\,.$$

Then the Hamiltonian takes the form

$$H(\vec{x},\vec{p}) \;=\; \vec{p}\cdot\vec{v}\;-\;L(\vec{x},\vec{v}) \;=$$

$$=\; \frac{c\,\vec{p}\cdot\,(\,\vec{p}\;-\;q\,\vec{A}\,)}{\sqrt{\,c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}} \;+\; m\,c^2\,\frac{c\,m}{\sqrt{\,c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}} +$$

$$+\,q\,\phi\;-\;\frac{q\,c\,\vec{A}\cdot\,(\,\vec{p}\;-\;q\,\vec{A}\,)}{\sqrt{\,c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}}\,. \;=$$

$$=\; c\,\sqrt{\,c^2\,m^2\;+\;\left|\,\vec{p}\;-\;q\,\vec{A}\,\right|^2}\;+\;q\,\phi\;.$$

(d) Damped harmonic oscillator

Recall the Lagrangian for the damped harmonic oscillator,

$$L(x, \dot{x}, t) \;=\; e^{\frac{\lambda t}{m}} \left(\frac{m}{2}\, \dot{x}^2 \;-\; \frac{k}{2}\, x^2 \right).$$

The canonical momentum is

$$p \;=\; e^{\frac{\lambda t}{m}}\, m\, \dot{x}\ ,$$

and the Hamiltonian

$$H(x, p, t) \;=\; p\, \dot{x} \;- L(x, \dot{x}, t) \;=$$

$$= \; e^{-\frac{\lambda t}{m}}\, \frac{1}{m}\, p^2 \;-\; e^{\frac{\lambda t}{m}}\, \frac{m}{2}\, e^{-\frac{2\lambda t}{m}}\, \frac{1}{m^2}\, p^2 \;+\; e^{\frac{\lambda t}{m}}\, \frac{k}{2}\, x^2 \;=$$

$$= \; e^{-\frac{\lambda t}{m}}\, \frac{1}{2m}\, p^2 \;+\; e^{\frac{\lambda t}{m}}\, \frac{k}{2}\, x^2\ .$$

In the next lecture we discuss some general features of the Hamiltonian formalism.

Lagrangian and Hamiltonian Dynamics

Volker Perlick

(Lancaster University)

Lecture 3

Discussion of Hamiltonian Dynamics and Phase Flows

(Cockcroft Institute, 1 March 2010)

The inventor of Hamiltonian dynamics:

Sir William Rowan Hamilton (1805, Dublin – 1865, Dublin)

Professor of Astronomy at Trinity College (1827 – 1865)

"Theory of Systems of Rays" (1827)

"On a General Method in Dynamics" (1834/35)

Recall: In the Hamiltonian formulation, all information on the dynamics of a system is coded in the Hamiltonian function

$$H(x^1, \ldots, x^n, p_1, \ldots p_n, t) \, .$$

For N particles without constraints, $n = 3N$.

The passage from the Lagrangian to the Hamiltonian description is possible if the Lagrangian is regular,

$$\det\Big(\frac{\partial^2 L}{\partial \dot{x}^i \dot{x}^j}\Big) \neq 0 \, .$$

Then the set of equations

$$p_i = \frac{\partial L}{\partial \dot{x}^i} \, , \quad i = 1, \ldots, n$$

can be solved for the $\dot{x}^i$ and the Hamiltonian can be defined as the Legendre transform of the Lagrangian,

$$H\big(x^1, \ldots, x^n, p_1, \ldots, p_n, t\big) = \sum_{i=1}^{n} p_i \, \dot{x}^i - L\big(x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n, t\big) \, .$$

If H is known, Hamilton's equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,,$$

determine the dynamics.

This is a first-order system of ordinary differential equations for curves $x^1(t), \ldots, x^n(t), p_1(t), \ldots, p_n(t)$. For every set of initial values $x^i(t_0)$, $p_j(t_0)$ there is a unique solution.

Along a solution of Hamilton's equation,

$$\frac{dH}{dt} = \sum_{i=1}^{n} \Big(\frac{\partial H}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} \Big) + \frac{\partial H}{\partial t} =$$

$$= \sum_{i=1}^{n} \Big(-\frac{dp_i}{dt}\frac{dx^i}{dt} + \frac{dx^i}{dt}\frac{dp_i}{dt} \Big) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}\,.$$

Thus, if H does not depend on time explicitly, H is a constant of motion.

Hamilton's equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}$$

are more symmetric than the Euler-Lagrange equations. E.g., they preserve their form under the transformation

$$x'^i = p_i, \qquad p'_i = -x^i,$$

where, of course, the transformed Hamiltonian is given as

$$H'(x'^1, \ldots, x'^n, p'_1, \ldots, p'_n) = H(x^1, \ldots, x^n, p_1, \ldots, p_n).$$

Transformations that leave the form of Hamilton's equations unchanged are called "canonical transformations" and will be discussed in Lecture 4 in detail. In general, a canonical transformation mixes the x^i with the p_i.

Recall: In the Lagrange formalism we were allowed to make only transformations of the x^i alone ("point transformations"),

$$x^i \;=\; x^i(x'^1, x'^2, x'^3)\,,$$

with the transformation of the $\dot{x}^i$ determined by differentiation along curves,

$$\dot{x}^i \;=\; \sum_{j=1}^{n} \frac{\partial x^i}{\partial x'^j}\,\dot{x}'^j\,, \qquad \frac{\partial \dot{x}^i}{\partial \dot{x}'^k} \;=\; \frac{\partial x^i}{\partial x'^k}\,.$$

Under such point transformations, the p_i transform as

$$p_i \;=\; \frac{\partial L}{\partial \dot{x}^i} \;=\; \sum_{k=1}^{n} \frac{\partial L'}{\partial \dot{x}'^k}\frac{\partial \dot{x}'^k}{\partial \dot{x}^i} \;=\; \sum_{k=1}^{n} \frac{\partial L'}{\partial \dot{x}'^k}\frac{\partial x'^k}{\partial x^i} \;=\; \sum_{k=1}^{n} p'_k\,\frac{\partial x'^k}{\partial x^i}\,.$$

We see that the transformation matrix of the p_i is the inverse of the transformation matrix of the $\dot{x}^i$, hence

$$\sum_{i=1}^{n} p_i\,\dot{x}^i = \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} p'_k\,\frac{\partial x'^k}{\partial x^i}\,\frac{\partial x^i}{\partial x'^j}\,\dot{x}'^j = \sum_{k=1}^{n}\sum_{j=1}^{n} p'_k\,\delta^k_j\,\dot{x}'^j = \sum_{j=1}^{n} p'_j\,\dot{x}'^j\,.$$

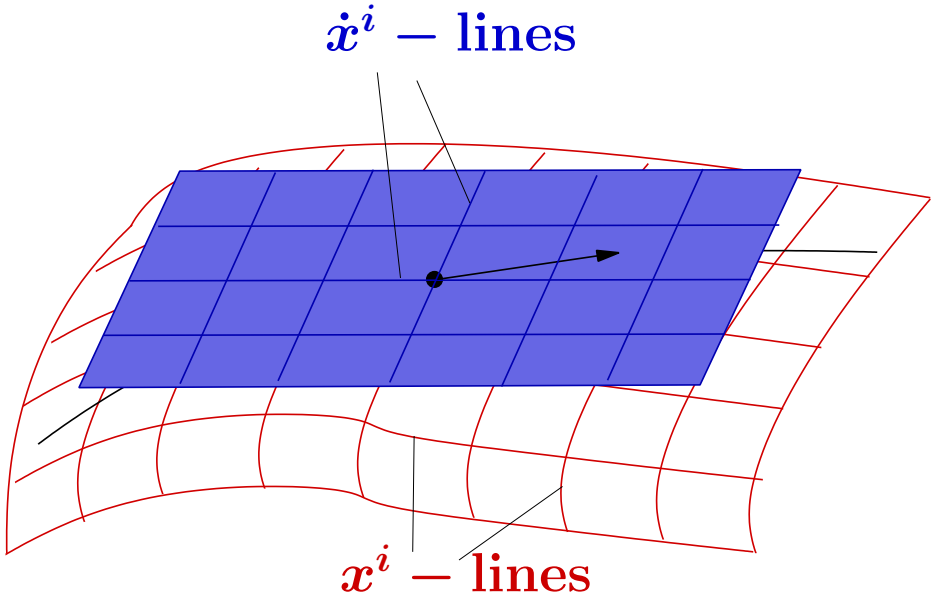
The $\dot{x}^i$ are coordinates of tangent vectors to curves in the configuration manifold, the p_i are coordinates of covectors (duals to tangent vectors).

Usual convention: Use upper indices for vectors and lower indices for covectors. Then summing over an upper and a lower index always gives a coordinate-independent expression.

The union of all cotangent spaces is called the "cotangent bundle" to the configuration manifold. The Hamiltonian is a (possibly time-dependent) function on the cotangent bundle.

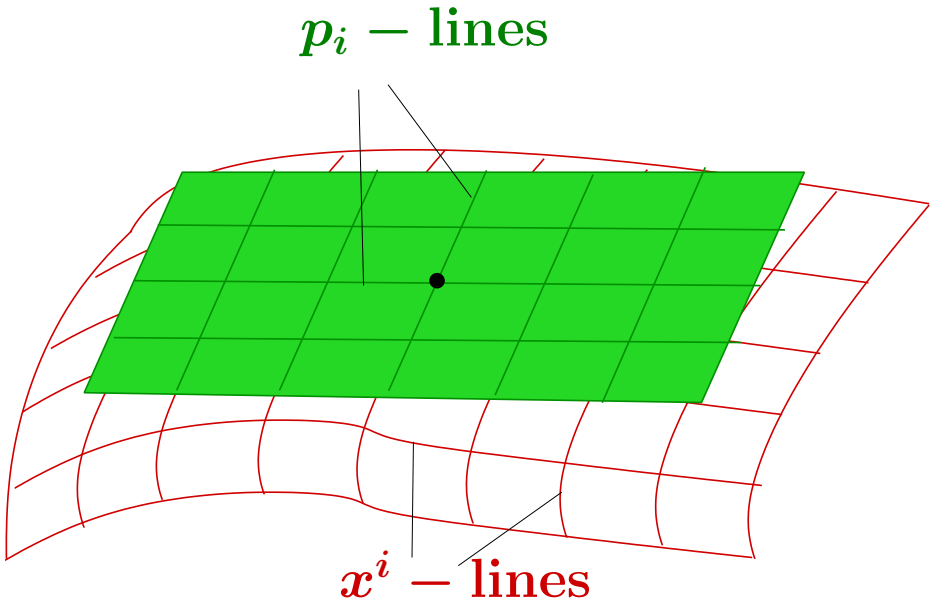
In the physical literature, the cotangent bundle of the configuration manifold is called the "phase space".

tangent bundle



Lagrangian L

cotangent bundle



Hamiltonian H

If a Hamiltonian is given that does not explicitly depend on time, through each point of phase space there is exactly one solution curve to Hamilton's equations. This is similar to the flow lines of a stationary flow. One speaks of the "phase flow" determined by the Hamiltonian. A picture of the phase flow is called a "phase portrait".

For systems with one degree of freedom, the shape of the flow lines is determined by the fact that along them

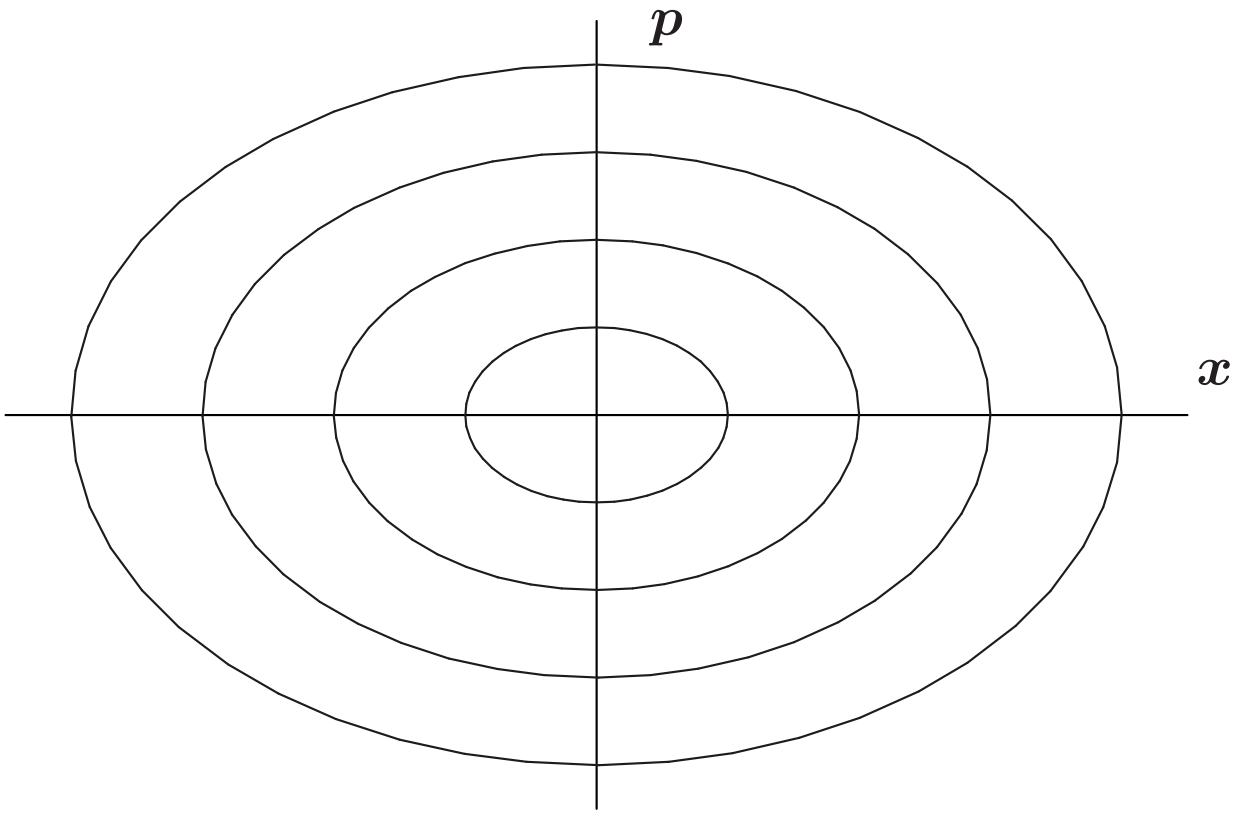
$$H = \text{constant} .$$

For systems with two degrees of freedom, the phase space is already four-dimensional, i.e., one can depict only projections of the phase flow into lower-dimensional spaces.

Example: For the undamped harmonic oscillator, the flow lines are the curves where the Hamiltonian

$$H(x,p) = \frac{p^2}{2m} + \frac{k\,x^2}{2}$$

is constant. Thus, the flow lines are ellipses.



For the damped harmonic oscillator,

$$m\,\ddot{x} + \lambda\,\dot{x} + k\,x = 0\,,$$

the Hamiltonian depends explicitly on time,

$$H(x,p,t) = e^{-\frac{\lambda\,t}{m}}\,\frac{p^2}{2\,m} + e^{\frac{\lambda\,t}{m}}\,\frac{k}{2}\,x^2\,.$$

Hence, the phase flow is non-stationary and the Hamiltonian is not preserved along the flow lines,

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \neq 0\,.$$

As we do not have a constant of motion, we have to solve the equations of motion to determine the phase space trajectories. As the equation of the damped harmonic oscillator is a linear ordinary differential equation, this can be done with an exponential ansatz.

In the case of undercritical damping,

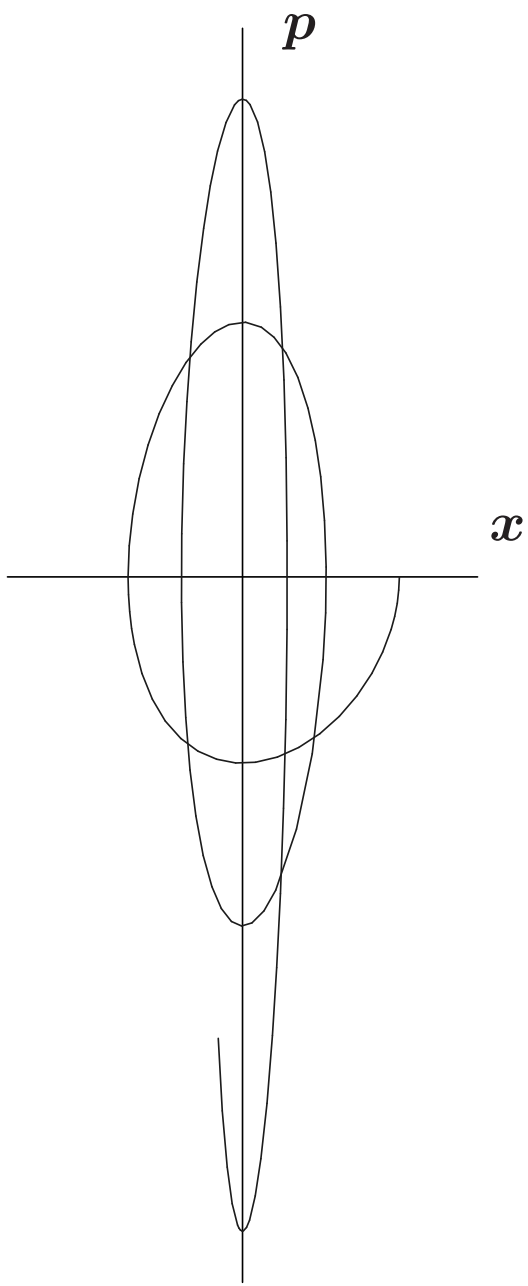
$$0 \leq \frac{k}{m} - \frac{\lambda^2}{4\,m^2} = \omega^2\,,$$

the general solution to Hamilton's equations is

$$x(t) = e^{-\frac{\lambda t}{2m}} \left(x(0)\Big(\cos(\omega t) + \frac{\lambda}{2m\omega}\sin(\omega t)\Big) + \frac{p(0)}{m\omega}\sin(\omega t) \right),$$

$$p(t) = e^{\frac{\lambda t}{2m}} \left(-x(0) \left(m\omega + \frac{\lambda^2}{4m\omega}\right)\sin(\omega t) + p(0) \left(\cos(\omega t) - \frac{\lambda}{2m\omega}\sin(\omega t)\right) \right).$$

In the following picture $\big(x(t), p(t)\big)$ is plotted for a particular initial condition $\big(x(0), p(0)\big)$.


p
x

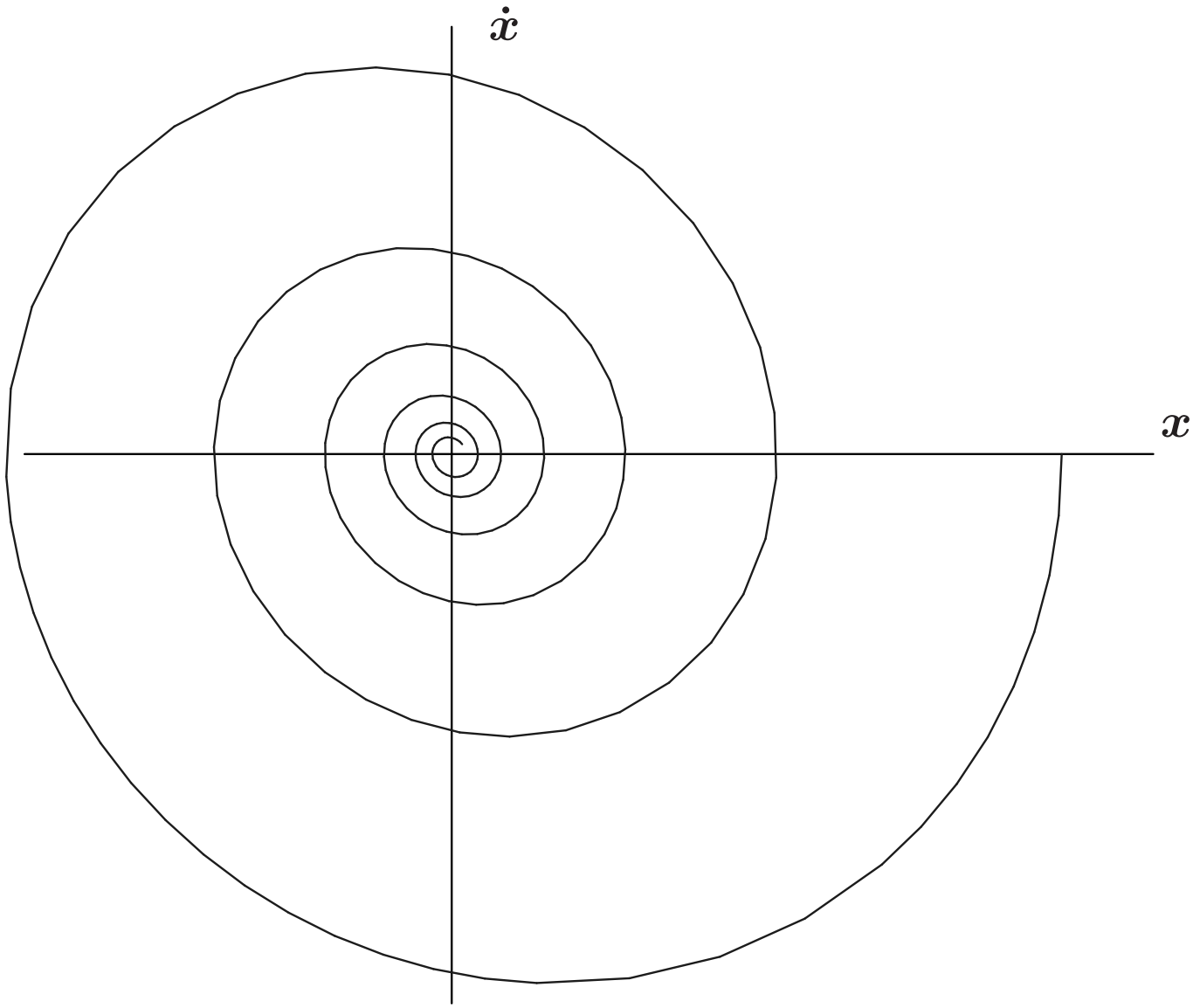
You might be surprised that $p(t)$ becomes arbitrarily large for $t \to \infty$. Note, however, that $\dot{x}(t) \to 0$ for $t \to \infty$, because p and $\dot{x}$ are related by

$$p(t) \;=\; m\,\dot{x}(t)\,e^{\frac{\lambda t}{m}}\,,$$

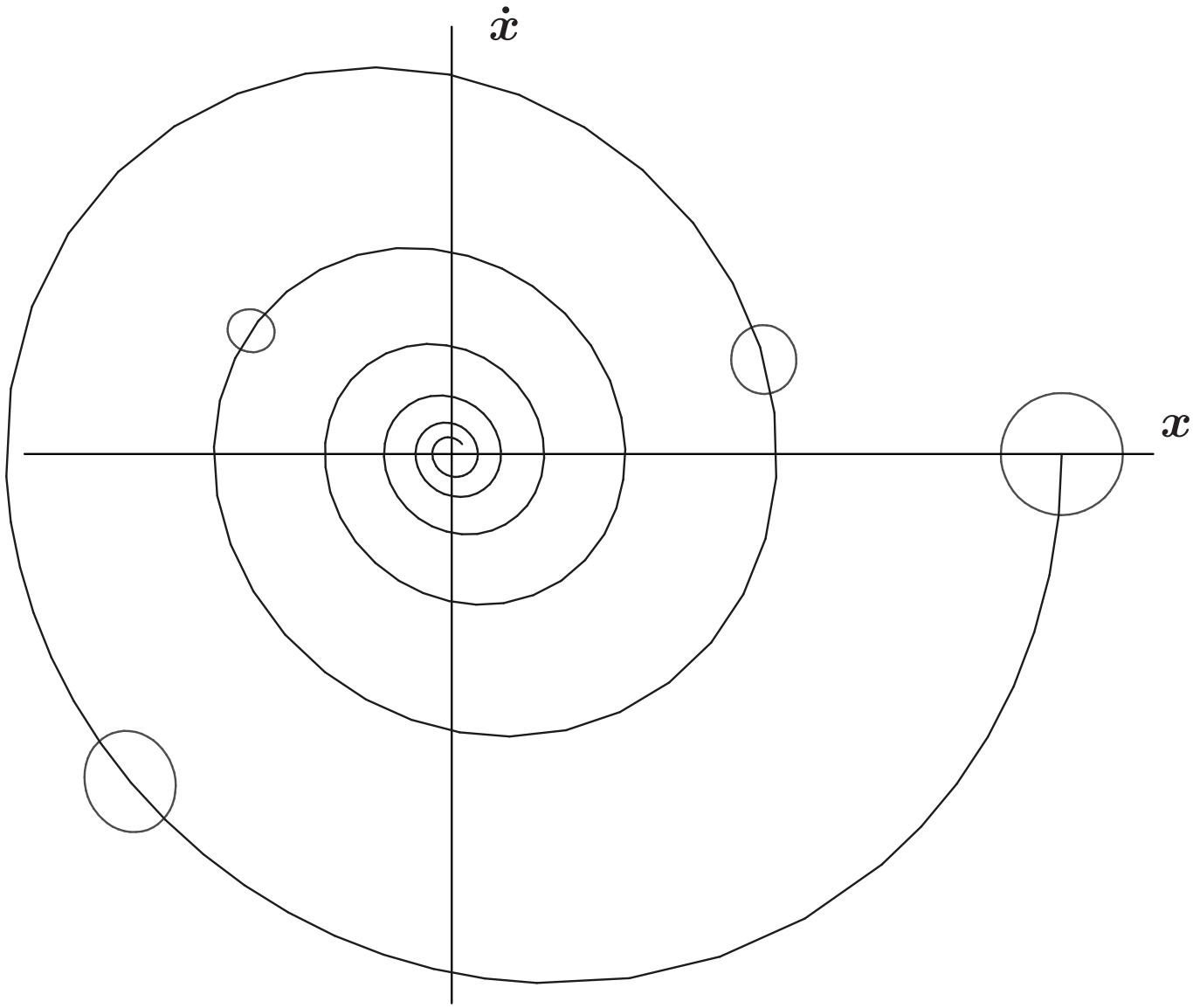
so

$$m\dot{x}(t) = e^{-\frac{\lambda t}{2m}}\left(-x(0)\Big(m\omega + \frac{\lambda^2}{4m\omega}\Big)\sin(\omega t) + p(0)\Big(\cos(\omega t) - \frac{\lambda}{2m\omega}\sin(\omega t)\Big)\right).$$

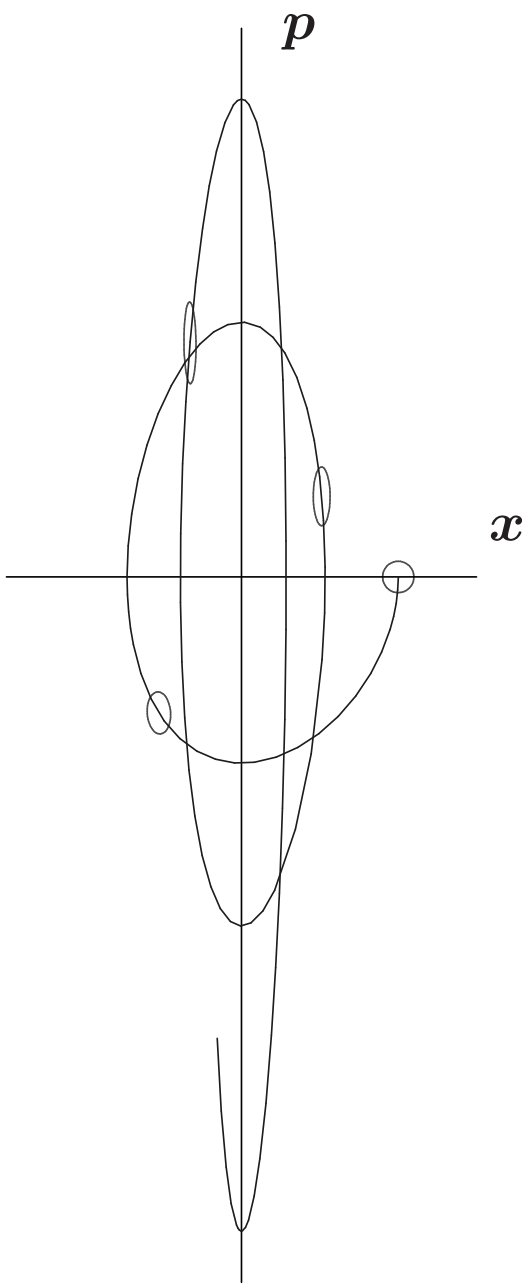
Thus, a plot in the $(x, \dot{x})-$space instead of the $(x, p)-$space looks completely different, see next page.


$\dot{x}$
x

In $(x, \dot{x})-$space, the flow of the damped harmonic oscillator shrinks the area. This is characteristic of dissipative systems.



In the $(x, p)-$space, however, the flow of the damped harmonic oscillator shrinks the area in the $x-$dimension and blows it up in the $p-$dimension.



We will now prove that in $(x, p)-$space the flow of the damped harmonic oscillator preserves the phase space area, i.e., that it behaves like an incompressible fluid.

The solution to Hamilton's equations can be written in matrix form,

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} \;=\; M(t) \begin{pmatrix} x(0) \\ p(0) \end{pmatrix}$$

where

$$M(t) \;=\; \begin{pmatrix} e^{-\frac{\lambda t}{2m}}\big(\cos(\omega t) + \frac{\lambda}{2m\omega}\sin(\omega t)\big) & e^{-\frac{\lambda t}{2m}}\frac{1}{m\omega}\sin(\omega t) \\ -e^{\frac{\lambda t}{2m}}\big(m\omega + \frac{\lambda^2}{4m\omega}\big)\sin(\omega t) & e^{\frac{\lambda t}{2m}}\big(\cos(\omega t) - \frac{\lambda}{2m\omega}\sin(\omega t)\big) \end{pmatrix} .$$

Thus, the flow is a one-parameter family of linear maps $M(t)$. Since

$$\det\big(M(t)\big) \;=\; \cos^2(\omega t) \;-\; \frac{\lambda^2 \sin^2(\omega t)}{4m^2\omega^2} \;+\; \sin^2(\omega t) \;+\; \frac{\lambda^2 \sin^2(\omega t)}{4m^2\omega^2} \;=\; 1\,,$$

the flow preserves, indeed, the area of the phase space volume.

This is the special case of a general result, known as Liouville's Theorem: The flow of a (possibly time-dependent) Hamiltonian in $2n-$dimensional phase space preserves the $2n-$dimensional phase space volume. We will prove it in the next lecture when we discuss canonical transformations.

We will then also verify that the phase space volume has a coordinate-independent meaning. This is another important advantage of the Hamilton over the Lagrange formalism. In the tangent bundle the volume depends on the chosen coordinates.

Lagrangian and Hamiltonian Dynamics

Volker Perlick

(Lancaster University)

Lecture 4

Canonical Transformations

(Cockcroft Institute, 15 March 2010)

In this lecture we will discuss the group of transformations that leave Hamilton's equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,,$$

invariant. We begin by introducing a useful notation.

For any two phase space functions $f(x^1,\ldots,x^n,p_1,\ldots,p_n,t)$ and $g(x^1,\ldots,x^n,p_1,\ldots,p_n,t)$ define the "Poisson bracket"

$$\{f,g\} = \sum_{i=1}^{n}\Big(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i}\Big)\,.$$

We compare this expression with the total time derivative of f along a solution to Hamilton's equations,

$$\frac{df}{dt} = \sum_{i=1}^{n}\Big(\frac{\partial f}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}\Big) + \frac{\partial f}{\partial t} =$$

$$= \sum_{i=1}^{n}\Big(\frac{\partial f}{\partial x^i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial x^i}\Big) + \frac{\partial f}{\partial t}\,.$$

This shows that

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} .$$

Thus, if f does not depend on time explicitly, its total time derivative along solutions of Hamilton's equations is given by the Poisson bracket with the Hamiltonian.

In particular: If f does not depend on time explicitly, we have

$$\frac{df}{dt} = 0 \quad \Longleftrightarrow \quad \{f, H\} = 0 .$$

i.e., f is a constant of motion if and only if its Poisson bracket with H vanishes.

The Poisson bracket

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial x^i} \right)$$

satisfies the following calculation rules.

- Linearity: $\{f, a\,g + b\,h\} = a\,\{f, g\} + b\,\{f, h\}$.
- Antisymmetry: $\{f, g\} = -\{g, f\}$.
- Jacobi identity: $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$.

These are the defining properties of a Lie Algebra.

In addition, the Poisson bracket satisfies the derivation rule

$$\{f, g\,h\} = g\,\{f, h\} + h\,\{f, g\} .$$

We are now ready to define canonical transformations.

Definition: A transformation

$$x'^i = x'^i(x^1, \ldots, x^n, p_1, \ldots, p_n) \,,$$

$$p'_i = p'_i(x^1, \ldots, x^n, p_1, \ldots, p_n) \,,$$

is called "canonical" if

$$\{x'^i, x'^j\} = 0 \,, \qquad \{p'_i, p'_j\} = 0 \,, \qquad \{x'^i, p'_j\} = \delta^i_j \,.$$

In the literature one can find several alternative definitions of canonical transformations, all of which turn out to be equivalent to the one chosen here. One of these alternative characterisations is the following.

Claim: A transformation is canonical if and only if it leaves the Poisson bracket of any two phase space functions invariant,

$$\{f, g\} = \{f', g'\}' .$$

Here

$$f'(x'^1, \ldots, x'^n, p'_1, \ldots, p'_n, t) = f(x^1, \ldots, x^n, p_1, \ldots, p_n, t) ,$$

$$g'(x'^1, \ldots, x'^n, p'_1, \ldots, p'_n, t) = g(x^1, \ldots, x^n, p_1, \ldots, p_n, t) ,$$

$$\{f', g'\}' = \sum_{i=1}^{n} \left(\frac{\partial f'}{\partial x'^i} \frac{\partial g'}{\partial p'_i} - \frac{\partial f'}{\partial p'_i} \frac{\partial g'}{\partial x'^i} \right) .$$

Proof: For arbitrary transformations, we find

$$\{f,g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} \right) =$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \left(\left(\frac{\partial f'}{\partial x'^j}\frac{\partial x'^j}{\partial x^i} + \frac{\partial f'}{\partial p'_j}\frac{\partial p'_j}{\partial x^i} \right) \left(\frac{\partial g'}{\partial x'^k}\frac{\partial x'^k}{\partial p_i} + \frac{\partial g'}{\partial p'_k}\frac{\partial p'_k}{\partial p_i} \right) \right.$$

$$\left. - \left(\frac{\partial f'}{\partial x'^j}\frac{\partial x'^j}{\partial p_i} + \frac{\partial f'}{\partial p'_j}\frac{\partial p'_j}{\partial p_i} \right) \left(\frac{\partial g'}{\partial x'^k}\frac{\partial x'^k}{\partial x^i} + \frac{\partial g'}{\partial p'_k}\frac{\partial p'_k}{\partial x^i} \right) \right) =$$

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \left(\frac{\partial f'}{\partial x'^j}\frac{\partial g'}{\partial x'^k}\{x'^j, x'^k\} + \frac{\partial f'}{\partial p'_j}\frac{\partial g'}{\partial p'_k}\{p'_j, p'_k\} \right.$$

$$\left. + \frac{\partial f'}{\partial x'^j}\frac{\partial g'}{\partial p'_k}\{x'^j, p'_k\} + \frac{\partial f'}{\partial p'_j}\frac{\partial g'}{\partial x'^k}\{p'_j, x'^k\} \right) .$$

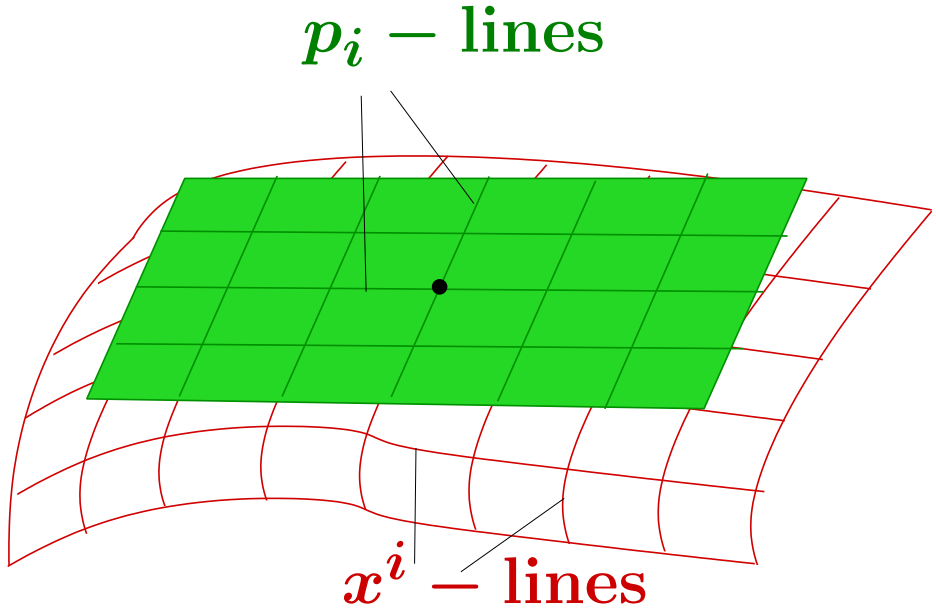
The right-hand side is equal to $\{f',g'\}'$ for all functions f' and g' if and only if the transformation is canonical. □

Note: Here one has to be careful about the notation. For $f = x^i$, the function f' is NOT equal to x'^i; it is rather x^i expressed in terms of the x'^i and the p'_i, i.e., $x'^i \neq x^{i\prime}$.

Examples:

- Every point transformation is a canonical transformation.
- The transformation $x'^i = p_i$, $p'_i = -x^i$ is canonical.

Our old picture remains true only under point transformations, not under arbitrary canonical transformations.



The fact that canonical transformations are of crucial relevance for the Hamiltonian formalism is in the following result.

Claim: A transformation is canonical if and only if it preserves the form of Hamilton's equations for all Hamiltonian functions, i.e., if and only if

$$\dot{x}^i = \frac{\partial H}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,,$$

implies

$$\dot{x}'^i = \frac{\partial H'}{\partial p'_i}\,, \qquad \dot{p}'_i = -\frac{\partial H'}{\partial x'^i}\,,$$

where

$$H'(x'^1, \ldots, x'^n, p'_1, \ldots, p'_n, t) = H(x^1, \ldots, x^n, p_1, \ldots, p_n, t)\,.$$

Proof: Assume that Hamilton's equations hold. Then

$$\dot{p}'_i = \sum_{k=1}^{n} \left(\frac{\partial p'_i}{\partial x^k} \dot{x}^k + \frac{\partial p'_i}{\partial p_k} \dot{p}_k \right) = \sum_{k=1}^{n} \left(\frac{\partial p'_i}{\partial x^k} \frac{\partial H}{\partial p_k} - \frac{\partial p'_i}{\partial p_k} \frac{\partial H}{\partial x^k} \right) =$$

$$\sum_{k=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial p'_i}{\partial x^k} \left(\frac{\partial H'}{\partial x'^j} \frac{\partial x'^j}{\partial p_k} + \frac{\partial H'}{\partial p'_j} \frac{\partial p'_j}{\partial p_k} \right) - \frac{\partial p'_i}{\partial p_k} \left(\frac{\partial H'}{\partial x'^j} \frac{\partial x'^j}{\partial x^k} + \frac{\partial H'}{\partial p'_j} \frac{\partial p'_j}{\partial x^k} \right) \right).$$

After rearranging, we find

$$\dot{p}'_i = \sum_{j=1}^{n} \left(\frac{\partial H'}{\partial x'^j} \{p'_i, x'^j\} + \frac{\partial H'}{\partial p'_j} \{p'_i, p'_j\} \right) .$$

An analogous calculation leads to

$$\dot{x}'^i = \sum_{j=1}^{n} \left(\frac{\partial H'}{\partial p'_j} \{x'^i, p'_j\} + \frac{\partial H'}{\partial x'^j} \{x'^i, x'^j\} \right) .$$

Thus, the primed Hamilton equations hold, for any H', if and only if the transformation is canonical. □

There are actually two different aspects to the relation between Hamilton's equations and canonical transformations.

First, as we have seen, canonical transformatiosn leave Hamilton's equations invariant.

Second, the phase flow defined by the Hamiltonian can be viewed as a one-parameter family of canonical transformations.

This can be formulated in the following way.

Claim: Consider any Hamiltonian $H\big(x^1,\ldots,x^n,p_1,\ldots,p_n,t\big)$. Fix a time t_0 and a time interval ε. Then the transformation

$$\Big(x^1(t_0),\ldots,x^n(t_0),p_1(t_0),\ldots,p_n(t_0)\Big)$$

$$\longmapsto\ \Big(x^1(t_0+\varepsilon),\ldots,x^n(t_0+\varepsilon),p_1(t_0+\varepsilon),\ldots,p_n(t_0+\varepsilon)\Big)$$

is canonical, where the $x^i(t)$ and $p_j(t)$ satisfy Hamilton's equations.

Proof: It suffices to give the proof for infinitesimally small ε because then the general result follows by integration. As

$$x^i(t_0+\varepsilon) \;=\; x^i(t_0) \;+\; \varepsilon\,\dot{x}^i(t_0) \;=$$

$$=\; x^i(t_0) \;+\; \varepsilon\,\frac{\partial H}{\partial p_i}\Big(x^1(t_0),\ldots,x^n(t_0),p_1(t_0),\ldots,p_n(t_0),t_0\Big)\;,$$

$$p_i(t_0+\varepsilon) \;=\; p_i(t_0) \;+\; \varepsilon\,\dot{p}^i(t_0) \;=$$

$$=\; p_i(t_0) \;-\; \varepsilon\,\frac{\partial H}{\partial x^i}\Big(x^1(t_0),\ldots,x^n(t_0),p_1(t_0),\ldots,p_n(t_0),t_0\Big)\;,$$

we have to prove that the transformation

$$x'^i \;=\; x^i \;+\; \varepsilon\;\frac{\partial H}{\partial p_i}\Big(x^1,\ldots,x^n,p_1,\ldots,p_n,t_0\Big)\;,$$

$$p'_i \;=\; p_i \;-\; \varepsilon\;\frac{\partial H}{\partial x^i}\Big(x^1,\ldots,x^n,p_1,\ldots,p_n,t_0\Big)\;,$$

is canonical.

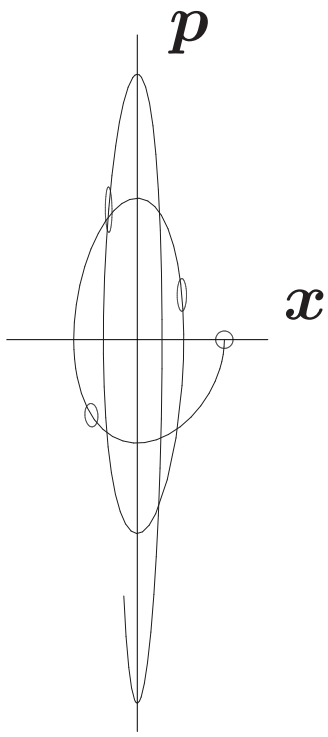
This is true since

$$\{x'^i, x'^j\} = \{x^i, x^j\} + \varepsilon \left\{ x^i , \frac{\partial H}{\partial p_j} \right\} + \varepsilon \left\{ \frac{\partial H}{\partial p_i} , x^j \right\} =$$

$$= 0 + \varepsilon \frac{\partial^2 H}{\partial p_i p_j} - \varepsilon \frac{\partial^2 H}{\partial p_j p_i} = 0 ,$$

and, analogously, $\{p'_i, p'_j\} = 0$ and $\{x'^i, p'_j\} = \delta^i_j$. □

Thus, shifting the points along the solution curves to Hamilton's equations always gives a canonical transformation. Recall our example of the damped harmonic oscillator.



We will now relate canonical transformations to what is called the symplectic structure of phase space. To that end, we rewrite the Poisson bracket in matrix notation,

$$\{f,g\} = \left(\frac{\partial f}{\partial x^1} \cdots \frac{\partial f}{\partial x^n} \frac{\partial f}{\partial p_1} \cdots \frac{\partial f}{\partial p_n}\right) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial g}{\partial x^1} \\ \vdots \\ \frac{\partial g}{\partial x^n} \\ \frac{\partial g}{\partial p_1} \\ \vdots \\ \frac{\partial g}{\partial p_n} \end{pmatrix} .$$

The antisymmetric $2n \times 2n$ matrix

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

is called the "symplectic matrix" in $2n$ dimensions.

The Poisson bracket reflects the "symplectic structure" of phase space.

"Symplectic" in greek has the same meaning as "complex" in latin.

The name was introduced by Herrman Weyl in the 1920s.

It refers to the fact that the $2n \times 2n$ matrix J is "composed" of four $n \times n$ blocks, corresponding to the phase space being coordinatised by $x^1, \ldots, x^n$ and $p_1, \ldots, p_n$.

Canonical coordinates can be characterised by the fact that they preserve the symplectic structure in a certain sense.

To work this out, we write the chain rule in matrix form.

$$\left(\frac{\partial f}{\partial x^1} \cdots \frac{\partial f}{\partial x^n} \; \frac{\partial f}{\partial p_1} \cdots \frac{\partial f}{\partial p_n} \right) =$$

$$= \left(\frac{\partial f'}{\partial x'^1} \cdots \frac{\partial f'}{\partial x'^n} \; \frac{\partial f'}{\partial p'_1} \cdots \frac{\partial f'}{\partial p'_n} \right) \underbrace{\begin{pmatrix} \left(\frac{\partial x'^j}{\partial x^k}\right) & \left(\frac{\partial x'^j}{\partial p_k}\right) \\ \left(\frac{\partial p'_j}{\partial x^k}\right) & \left(\frac{\partial p'_j}{\partial p_k}\right) \end{pmatrix}}_{S}$$

and

$$\begin{pmatrix} \frac{\partial g}{\partial x^1} \\ \vdots \\ \frac{\partial g}{\partial x^n} \\ \frac{\partial g}{\partial p_1} \\ \vdots \\ \frac{\partial g}{\partial p_n} \end{pmatrix} = \underbrace{\begin{pmatrix} \left(\frac{\partial x'^j}{\partial x^k}\right) & \left(\frac{\partial p'_j}{\partial x^k}\right) \\ \left(\frac{\partial x'^j}{\partial p_k}\right) & \left(\frac{\partial p'_j}{\partial p_k}\right) \end{pmatrix}}_{S^T} \begin{pmatrix} \frac{\partial g'}{\partial x'^1} \\ \vdots \\ \frac{\partial g'}{\partial x'^n} \\ \frac{\partial g'}{\partial p'_1} \\ \vdots \\ \frac{\partial g'}{\partial p'_n} \end{pmatrix} .$$

The unprimed and primed Poisson brackets are

$$\{f,g\} = \left(\frac{\partial f'}{\partial x'^1} \cdots \frac{\partial f'}{\partial x'^n}\, \frac{\partial f'}{\partial p'_1} \cdots \frac{\partial f'}{\partial p'_n}\right) S\, J\, S^T \begin{pmatrix} \frac{\partial g'}{\partial x'^1} \\ \vdots \\ \frac{\partial g'}{\partial x'^n} \\ \frac{\partial g'}{\partial p'_1} \\ \vdots \\ \frac{\partial g'}{\partial p'_n} \end{pmatrix},$$

$$\{f',g'\}' = \left(\frac{\partial f'}{\partial x'^1} \cdots \frac{\partial f'}{\partial x'^n}\, \frac{\partial f'}{\partial p'_1} \cdots \frac{\partial f'}{\partial p'_n}\right) J \begin{pmatrix} \frac{\partial g'}{\partial x'^1} \\ \vdots \\ \frac{\partial g'}{\partial x'^n} \\ \frac{\partial g'}{\partial p'_1} \\ \vdots \\ \frac{\partial g'}{\partial p'_n} \end{pmatrix}.$$

Thus, the transformation is canonical if and only if its Jacobi matrix leaves the symplectic matrix invariant.

$$\boldsymbol{S}\,\boldsymbol{J}\,\boldsymbol{S}^{T} \;=\; \boldsymbol{J}\,;$$

For this reason, canonical transformations are also called "symplectic transformations".

From the determinant theorem we find for a canonical transformation

$$\det(\boldsymbol{S})\,\det(\boldsymbol{J})\,\det(\boldsymbol{S}^{T}) \;=\; \det(\boldsymbol{J})\;.$$

As $\det(\boldsymbol{S}^{T}) = \det(\boldsymbol{S})$ and $\det(\boldsymbol{J}) = 1$, we have found that a canonical transformation has to satisfy

$$\det(\boldsymbol{S}) \;=\; \pm 1\;.$$

As the volume element transforms with the Jacobi determinant S,

$$dx'^1 \cdots dx'^n \, dp'_1 \cdots dp'_n = \det(S) \, dx^1 \cdots dx^n \, dp_1 \cdots dp_n \,,$$

we have proven that a canonial transformation leaves the phase space volume invariant (up to sign; a negative sign corresponds to a change of orientation from right-handed to left-handed.)

In particular, we have proven Liouville's theorem:

The phase flow of any Hamiltonian is volume-preserving, i.e., it behaves like an incompressible fluid.

Lagrangian and Hamiltonian Dynamics

Volker Perlick

(Lancaster University)

Lecture 5

Hamilton-Jacobi Theory

(Cockcroft Institute, 15 March 2010)

In this lecture we will discuss Hamilton-Jacobi theory which is about a particular method for solving Hamilton's equations.

Recall: Hamilton's equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}\,, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}\,,$$

preserve their form under canonical transformations

$$x'^i = x'^i(x^1, \ldots, x^n, p_1, \ldots, p_n)\,,$$

$$p'_i = p'_i(x^1, \ldots, x^n, p_1, \ldots, p_n)\,,$$

$$\{x'^i, x'^j\} = 0\,, \qquad \{p'_i, p'_j\} = 0\,, \qquad \{x'^i, p'_j\} = \delta^i_j\,,$$

where the Poisson brackets are defined by

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i} \right).$$

The basic idea of Hamilton-Jacobi theory is to find a canonical transformation such that in the new coordinates the Hamiltonian depends only on the p'_i, and not on the x'^i. Then Hamilton's equations read

$$\dot{x}'^i \;=\; \frac{\partial H'}{\partial p'_i}\;, \qquad \dot{p}'_i \;=\; 0\;.$$

The second set of equations says that the p'_i are constants of motion,

$$p'_i(t) \;=\; p'_i(0)\;.$$

Thus the first set, whose right-hand side depends only on the p'_i, can be integrated,

$$x'^i(t) \;=\; \frac{\partial H'}{\partial p'_i}\Big(p'_1(0),\ldots,p'_n(0)\Big)\,t \;+\; x'^i(0)\;.$$

So the problem has been shifted from solving Hamilton's equations to finding the desired canonical transformation.

To work this out, we proceed in three steps.

- **We discuss some properties of constants of motion in the Hamiltonian formalism.**

- **We consider Hamiltonian dynamical systems which admit n independent constants of motion such that the Poisson bracket of any pair of them vanishes. This is a necessary condition such that they can serve as momentum coordinates. Systems which admit such constants of motion are called "completely integrable".**

- **We investigate how to actually find a canonical transformation such that the new momenta are constants of motion. This will be done in terms of a "generating function", and the partial differential equation this generating function has to satisfy is known as the Hamilton-Jacobi equation.**

Throughout we consider a time-independent Hamiltonian

$$H(x^1, \ldots, x^n, p_1, \ldots, p_n)$$

on a $2n-$dimensional phase space.

We begin by discussing some properties of constants of motion. Recall that for any $f(x^1, \ldots, x^n, p_1, \ldots, p_n)$ that does not depend explicitly on time, we have

$$\frac{df}{dt} = \{f, H\} ,$$

i.e., f is a constant of motion if and only if its Poisson bracket with H vanishes. What does $\{f, H\} = 0$ mean?

f determines a "flow" on phase space, given by solving Hamilton's equations with f instead of H,

$$\frac{dx^i}{d\varepsilon} = \frac{\partial f}{\partial p_i} , \qquad \frac{dp_i}{d\varepsilon} = -\frac{\partial f}{\partial x^i} .$$

We write ε instead of t because the curve parameter has not the physical meaning of time if f is not the Hamiltonian of the system.

Now the differentiation of H along any such flow line is

$$\frac{dH}{d\varepsilon} = \sum_{i=1}^{n} \Big(\frac{\partial H}{\partial p_i} \frac{dp_i}{d\varepsilon} + \frac{\partial H}{\partial x^i} \frac{dx^i}{d\varepsilon} \Big) =$$

$$= \sum_{i=1}^{n} \Big(- \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial x^i} + \frac{\partial H}{\partial x^i} \frac{\partial f}{\partial p_i} \Big) = - \{ f, H \} \, .$$

Thus, if $\{f, H\} = 0$, the flow determined by f leaves H invariant. This shows that every constant of motion is associated with a one-parameter group of symmetry transformations. This is the general relation between symmetries and constants of motion in the Hamiltonian formalism.

We will now define the notion of "complete integrability". Again, we restrict to systems with a time-independent Hamiltonian. (Time-dependent Hamiltonians can be treated in a space-time formalism, adding t as an additional configuration space coordinate.)

Definition: Hamilton's equations with a time-independent Hamiltonian $H(x^1, \ldots, x^n, p_1, \ldots, p_n)$ are called "completely integrable" if there are n functions

$$f_1(x^1, \ldots, x^n, p_1, \ldots, p_n)\,, \ \ldots\ , f_n(x^1, \ldots, x^n, p_1, \ldots, p_n)$$

which

(a) are constants of motion, i.e. $\{H, f_i\} = 0\,,$

(b) have vanishing Poisson brackets, i.e. $\{f_i, f_j\} = 0\,,$

(c) have linearly independent differentials $df_1, \ldots, df_n\,.$

Note that (b) and (c) are necessary conditions for the existence of a canonical transformation such that $p'_i = f_i$ for $i = 1, \ldots, n$.

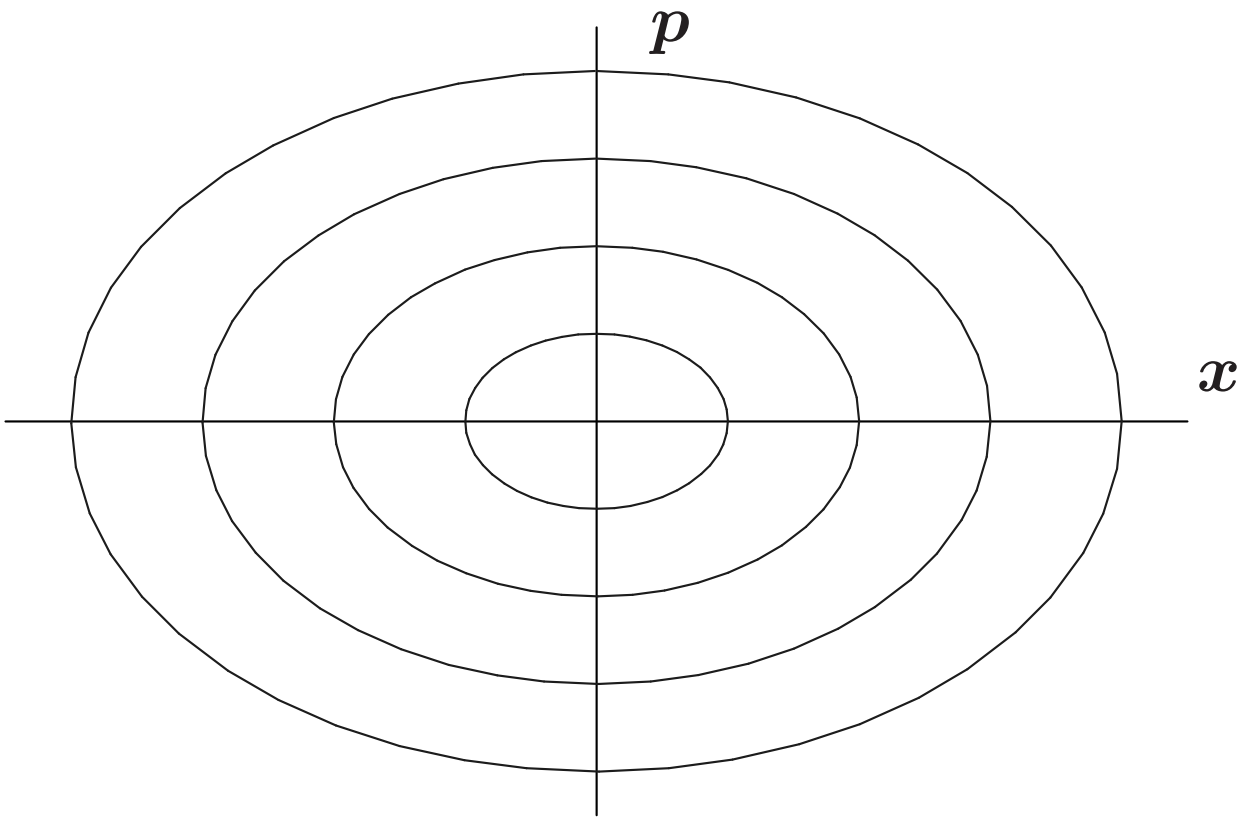
One can always choose $f_1 = H$. One just has to omit points where $dH = 0$ ("equilibrium points" or "fixed points of the Hamiltonian flow"). There (c) cannot hold with $f_1 = H$.

Complete integrability is a global notion. One wants to have the f_i on all of phase space (perhaps with some exceptional points, where the differentials become linearly dependent, omitted), not just on local neighborhoods. On sufficiently small neighborhoods of a non-equilibrium point, the desired f_i always exist.

Instead of "completely integrable" one often says "exactly integrable" or just "integrable". The latter name is misleading. Hamilton's equations are always integrable in the sense that there is a unique solution curve to every initial condition. However, they are in general not completely integrable.

Every system which is simple enough such that the solution can be explicitly written down is, indeed, completely intregrable, but many physically interesting systems are not completely integrable. Here are some examples.

Example 1: Every Hamiltonian system with one degree of freedom, $n = 1$, is completely integrable, with $f_1 = H$. One has to remove the equilibrium points. For the undamped harmonic oscillator, there is one equilibrium point at $x = 0, p = 0$.



Example 2: For a particle in a rotationally symmetric potential, Hamilton's equations are completely integrable. The Hamiltonian is given, in Cartesian coordinates, as

$$H(x^1, x^2, x^3, p_1, p_2, p_3) \;=\; \frac{p_1^2 + p_2^2 + p_3^2}{2m} \;+\; V(r)\;,$$

$$r \;=\; \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}\,.$$

The f_1, f_2, f_3 can be found in the following way: Consider the components of the angular momentum

$$L_1 = x^2 p_3 - x^3 p_2\,, \quad L_2 = x^3 p_1 - x^1 p_3\,, \quad L_3 = x^1 p_2 - x^2 p_1\,.$$

By direct computation of the Poisson brackets, one verifies that

$$\{L_i, H\} \;=\; 0\,,$$

i.e., the L_i are constants of motion.

However, we cannot choose $f_i = L_i$ for $i = 1, 2, 3$ because

$$\{L_1, L_2\} = L_3\,, \quad \{L_2, L_3\} = L_1\,, \quad \{L_3, L_1\} = L_2\,,$$

whereas we want to have $\{f_i, f_j\} = 0$. Fortunately, a quick calculation shows that

$$\{|\vec{L}|^2, L_3\} = 0\,,$$

where

$$|\vec{L}|^2 = L_1^2 + L_2^2 + L_3^2\,,$$

so we can choose

$$f_1 = H\,, \quad f_2 = |\vec{L}|^2\,, \quad f_3 = L_3\,.$$

The differentials df_1, df_2, df_3 are linearly dependent at points where $\vec{p}$ is a multiple of $\vec{x}$.

Example 3: Hamilton's equations for a relativistic particle in a constant magnetic field is completely integrable. In Lecture 2 we have seen that, if the magnetic field is $\vec{B} = \nabla \times \vec{A}$, the Hamiltonian in Cartesian coordinates reads

$$H(x^1, x^2, x^3, p_1, p2, p3) \;=\; c\,\sqrt{c^2\,m^2 \;+\; \left|\,\vec{p} \;-\; q\,\vec{A}\,\right|^2}\,.$$

For a constant magnetic field we can choose the vector potential $\vec{A}$ as

$$\vec{A} \;=\; \begin{pmatrix} -B_0\,x^2/2 \\ B_0\,x^1/2 \\ 0 \end{pmatrix}\;, \qquad \nabla \times \vec{A} \;=\; \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}\;.$$

The desired constants of motion are

$$f_1 \;=\; H\,, \quad f_2 \;=\; p_3\,, \quad f_3 \;=\; L_3 \;=\; x^1 p_2 - x^2 p_1\,.$$

The differentials df_1, df_2, df_3 are linearly dependent at points where $p_1 - qA_1 = p_2 - qA_2 = 0$.

Example 4: The Three-Body Problem is not completely integrable. For three particles, moving under the influence of their mutual Newtonian gravitational attraction, it is impossible to find nine constants of motion with the desired properties.

By definition, completely integrable systems admit n independent constants of motion with vanishing Poisson brackets. We will now investigate how to find a canonical transformation that makes these constants of motion into the new momenta. To that end we introduce the notion of a "generating function".

We restrict to the four standard types of generating functions which are used in practice. Not all canonical transformations can be generated by one of them.

Claim:

Any function $F_1(x^1,\ldots,x^n,x'^1,\ldots,x'^n)$ with $\det\left(\frac{\partial^2 F_1}{\partial x^i x'^j}\right)=0$ generates a canonical transformation

$$p_i = \frac{\partial F_1}{\partial x^i}\,, \qquad p'_i = -\frac{\partial F_1}{\partial x'^i}\,.$$

Any function $F_2(x^1,\ldots,x^n,p'_1,\ldots,p'_n)$ with $\det\left(\frac{\partial^2 F_2}{\partial x^i p'_j}\right)=0$ generates a canonical transformation

$$p_i = \frac{\partial F_2}{\partial x^i}\,, \qquad x'_i = \frac{\partial F_2}{\partial p'_i}\,.$$

Any function $F_3(p_1,\ldots,p_n,x'^1,\ldots,x'^n)$ with $\det\left(\frac{\partial^2 F_3}{\partial p_i x'^j}\right)=0$ generates a canonical transformation

$$x^i = -\frac{\partial F_3}{\partial p_i}\,, \qquad p'_i = -\frac{\partial F_3}{\partial x'^i}\,.$$

Any function $F_4(p_1,\ldots,p_n,p'_1,\ldots,p'_n)$ with $\det\left(\frac{\partial^2 F_4}{\partial p_i p'_j}\right)=0$ generates a canonical transformation

$$x^i = -\frac{\partial F_4}{\partial p_i}\,, \qquad x'^i = \frac{\partial F_4}{\partial p'_i}\,.$$

F_1, F_2, F_3, F_4 are called "generating functions of 1st, 2nd, 3rd, or 4th kind", respectively.

Proof: We give the proof for generating functions of the 2nd kind which are most frequently used. (The proof for 1st, 3rd and 4th kind is analogous.) Let $F_2\big(x^1, \ldots, x^n, p'_1, \ldots, p'_n\big)$, be given. The condition $\det\left(\frac{\partial^2 F_2}{\partial x^i p'_j}\right) = 0$ guarantees that the set of equations

$$p_i \;=\; \frac{\partial F_2}{\partial x^i}\,, \qquad x'^i \;=\; \frac{\partial F_2}{\partial p'_i}\,,$$

can be solved for the primed coordinates. We have to prove that this transformation is canonical. The differential of F_2 is

$$dF_2 \;=\; \sum_{i=1}^{n} \frac{\partial F_2}{\partial x^i}\, dx^i \;+\; \sum_{j=1}^{n} \frac{\partial F_2}{\partial p'_j}\, dp'_j \;=$$

$$= \sum_{i=1}^{n} p_i\, dx^i \;+\; \sum_{j=1}^{n}\sum_{i=1}^{n} x'^j \left(\frac{\partial p'_j}{\partial x^i}\, dx^i \;+\; \frac{\partial p'_j}{\partial p_i}\, dp_i\right).$$

$$dF_2 \;=\; \sum_{i=1}^{n} \underbrace{\Big(p_i \;+\; \sum_{j=1}^{n} x'^j \,\frac{\partial p'_j}{\partial x^i}\Big)}_{A_i}\, dx^i \;+\; \sum_{i=1}^{n} \underbrace{\sum_{j=1}^{n} x'^j \,\frac{\partial p'_j}{\partial p_i}}_{B^i}\, dp_i \,.$$

As dF_2 is a total differential, we must have

$$0 \;=\; \frac{\partial A_i}{\partial x^k} \,-\, \frac{\partial A_k}{\partial x^i}\,, \quad 0 \;=\; \frac{\partial B^i}{\partial p_k} \,-\, \frac{\partial B^k}{\partial p_i}\,, \quad 0 \;=\; \frac{\partial A_k}{\partial p_i} \,-\, \frac{\partial B^i}{\partial x^k}$$

$$0 \;=\; \sum_{j=1}^{n} \Big(\frac{\partial x'^j}{\partial x^k}\frac{\partial p'_j}{\partial x^i} \;+\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial x^i \partial x^k}} \;-\; \frac{\partial x'^j}{\partial x^i}\frac{\partial p'_j}{\partial x^k} \;-\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial x^k \partial x^i}}\Big)$$

$$0 \;=\; \sum_{j=1}^{n} \Big(\frac{\partial x'^j}{\partial p_k}\frac{\partial p'_j}{\partial p_i} \;+\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial p_i \partial p_k}} \;-\; \frac{\partial x'^j}{\partial p_i}\frac{\partial p'_j}{\partial p_k} \;-\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial p_k \partial p_i}}\Big)$$

$$0 \;=\; \delta^i_k \;+\; \sum_{j=1}^{n} \Big(\frac{\partial x'^j}{\partial p_i}\frac{\partial p'_j}{\partial x^k} \;+\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial x^k \partial p_i}} \;-\; \frac{\partial x'^j}{\partial x^k}\frac{\partial p'_j}{\partial p_i} \;-\; \cancel{x'^j \,\frac{\partial^2 p'_j}{\partial p_i \partial x^k}}\Big)$$

These three sets of equations can be comprised in matrix form,

$$S^T\, J\, S \;=\; J\;,$$

where S is the Jacobi matrix of the transformation and J is the symplectic matrix,

$$S \;=\; \begin{pmatrix} \left(\frac{\partial x'^j}{\partial x^k}\right) & \left(\frac{\partial x'^j}{\partial p_k}\right) \\ \\ \left(\frac{\partial p'_j}{\partial x^k}\right) & \left(\frac{\partial p'_j}{\partial p_k}\right) \end{pmatrix}\;, \qquad J \;=\; \begin{pmatrix} 0 & 1 \\ \\ -1 & 0 \end{pmatrix}\;.$$

This equation is equivalent to

$$S\, J\, \left(\, S^T\, J\, S\,\right) S^{-1}\, J \;=\; S\, J\, J\, S^{-1}\, J\;,$$

and, as $J\, J \;=\; -1$, to

$$S\, J\, S^T \;=\; J\;.$$

This proves that the transformation is canonical ($=$symplectic), recall Lecture 4. □

The goal of Hamilton-Jacobi theory is to find a canonical transformation

$$p_i \;=\; \frac{\partial W}{\partial x^i}\;, \qquad x'^i \;=\; \frac{\partial W}{\partial p'_i}\;,$$

where $W\big(x^1,\ldots,x^n,p'_1,\ldots,p'_n\big)$ is a generating function of the second kind, such that the new Hamiltonian is independent of the x'^i. Then the p'_i are constants of motion,

$$p'_i(t) \;=\; \alpha_i\;,$$

and the Hamiltonian is a constant of motion,

$$H'(p'_1,\ldots,p'_n) \;=\; E\;.$$

The equation $H\big(x^1,\ldots,x^n,p_1,\ldots,p_n\big) \;=\; H'(p'_1,\ldots,p'_n)$ becomes a non-linear first order partial differential equation for W,

$$H\Big(x^1,\ldots,x^n,\frac{\partial W}{\partial x^1},\ldots,\frac{\partial W}{\partial x^n}\Big) \;=\; E\;.$$

This is the (time-independent) "Hamilton-Jacobi equation".

A "complete integral" of the Hamilton-Jacobi equation is a solution $W\left(x^1, \ldots, x^n, \alpha_1, \ldots, \alpha_n\right)$ that depends on n parameters $\alpha_1, \ldots, \alpha_n$ such that the derivatives

$$\frac{\partial W}{\partial \alpha_1}\ , \ \cdots\ , \ \frac{\partial W}{\partial \alpha_n}$$

are linearly independent. E is then a function of the $\alpha_1, \ldots, \alpha_n$. (If one chooses the Hamiltonian as one of the constants of motion, say $p'_1 = H'$, one has $E = \alpha_1$.) The general solution to Hamilton's equations is then

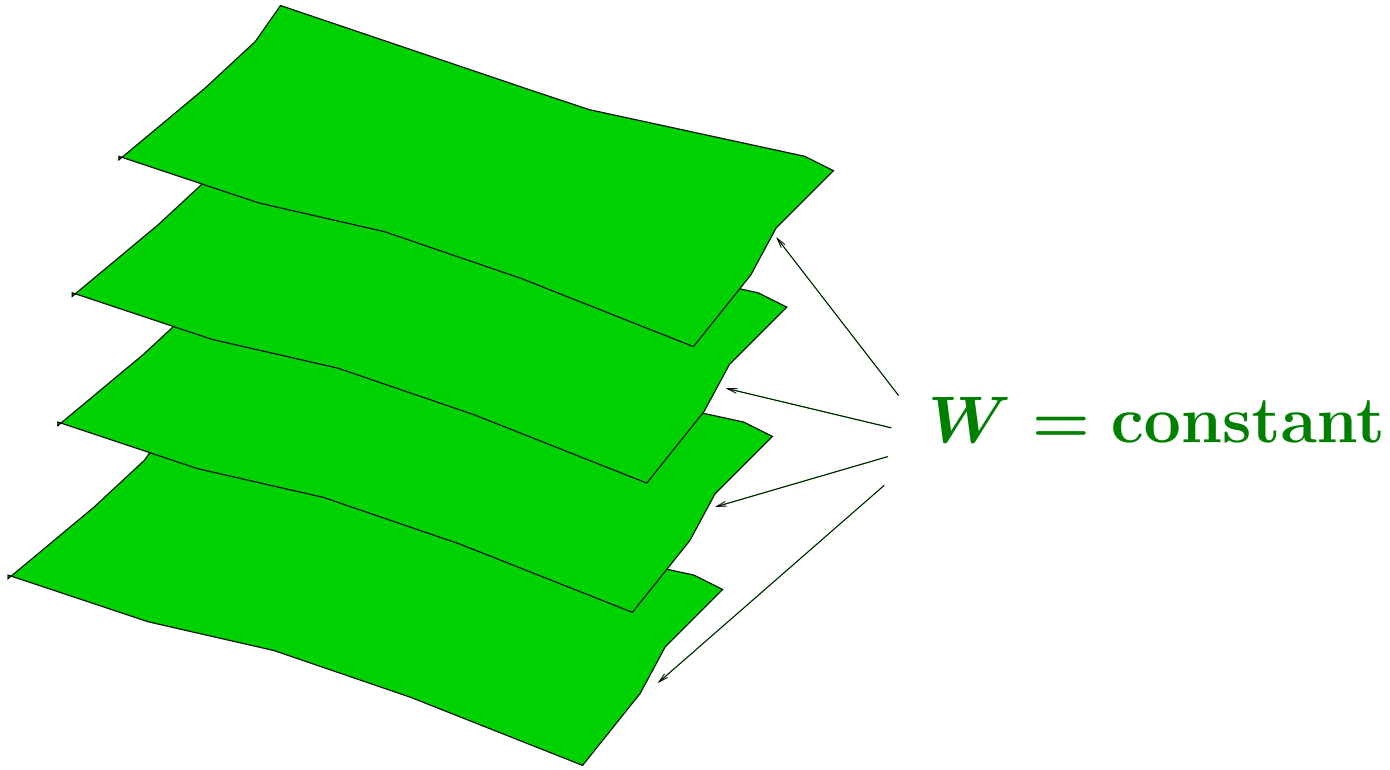
$$p'_i(t) \ = \ \alpha_i \ , \qquad x'^i(t) \ = \ \frac{\partial E}{\partial \alpha_i}\, t \ + \ \beta^i$$

which can be adapted to arbitrary initial conditions.

For simple examples (harmonic oscillator, particle in homogeneous field etc.), the Hamilton-Jacobi method is of no advantage. However, for many difficult problems it is the only method which allows to find the solution or, at least, to investigate its qualitative properties.

If we fix the values of the parameters $\alpha_1, \ldots, \alpha_n$, a solution to the Hamilton-Jacobi equation foliates the $x-$space into "phase surfaces" or "wave fronts"

$$W\left(x^1, \ldots, x^n, \alpha_1, \ldots, \alpha_n\right) \;=\; \text{constant}\,.$$



Example: We solve the Hamilton-Jacobi equation for the two-dimensional harmonic oscillator,

$$H(x^1, x^2, p_1, p_2) \;=\; \frac{p_1^2 \,+\, p_2^2}{2\,m} \,+\, \frac{k_1 (x^1)^2}{2} \,+\, \frac{k_2 (x^2)^2}{2}\,.$$

The Hamilton-Jacobi equation reads

$$\frac{1}{2\,m}\left(\frac{\partial W}{\partial x^1}\right)^2 \,+\, \frac{1}{2\,m}\left(\frac{\partial W}{\partial x^2}\right)^2 \,+\, \frac{k_1 (x^1)^2}{2} \,+\, \frac{k_2 (x^2)^2}{2} \;=\; E\,.$$

We make a separation ansatz,

$$W \;=\; W_1 \,+\, W_2\,, \qquad \frac{\partial W_1}{\partial x^2} \;=\; \frac{\partial W_2}{\partial x^1} \;=\; 0\,.$$

The Hamilton-Jacobi equation takes the form

$$\underbrace{\frac{1}{2\,m}\left(\frac{\partial W_1}{\partial x^1}\right)^2 \,+\, \frac{k_1 (x^1)^2}{2}}_{\alpha_1} \,+\, \underbrace{\frac{1}{2\,m}\left(\frac{\partial W_2}{\partial x^2}\right)^2 \,+\, \frac{k_2 (x^2)^2}{2}}_{\alpha_2} \;=\; E\,.$$

From this equation we read that α_1 and α_2 depend neither on x^1 nor on x^2.

The equations

$$\left(\frac{\partial W_1}{\partial x^1}\right)^2 = 2m\left(\alpha_1 - \frac{k_1}{2}(x^1)^2\right),\ W_1 = \int \sqrt{2m\left(\alpha_1 - \frac{k_1}{2}(x^1)^2\right)}\, dx^1 ,$$

$$\left(\frac{\partial W_2}{\partial x^2}\right)^2 = 2m\left(\alpha_2 - \frac{k_2}{2}(x^2)^2\right),\ W_2 = \int \sqrt{2m\left(\alpha_2 - \frac{k_2}{2}(x^2)^2\right)}\, dx^2 ,$$

yield the desired complete integral $W = W_1 + W_2$, with

$$E = \alpha_1 + \alpha_2 .$$

In the primed coordinates, integration of Hamilton's equations is now trivial,

$$x'^i(t) = \frac{\partial E}{\partial \alpha_i} t + \beta_i = t + \beta_i , \qquad i = 1, 2 .$$

To find the trajectories in the unprimed coordinates, we use

$$x'^i = \frac{\partial W}{\partial \alpha_i} , \qquad i = 1, 2 .$$

This results in

$$t + \beta_i = \frac{\partial}{\partial \alpha_i} \int \sqrt{2\,m\,(\alpha_i - \frac{k_i}{2}\,(x^i)^2\,)}\; dx^i =$$

$$= \int \frac{m\,dx^i}{\sqrt{2\,m\,(\alpha_i - \frac{k_i}{2}\,(x^i)^2\,)}} = \sqrt{\frac{m}{k_i}}\,\arcsin\Big(\frac{\sqrt{k_i}\,x^i}{\sqrt{2\,\alpha_i}}\Big)\;.$$

Solving for x^i yields

$$x^i(t) = \sqrt{\frac{2\,\alpha_i}{k_i}}\,\sin\Big(\sqrt{\frac{k_i}{m}}\,(t+\beta_i)\Big)\;.$$

Lagrangian and Hamiltonian Dynamics

Volker Perlick

(Lancaster University)

Lecture 6

Stability and Linearisation

(Cockcroft Institute, 22 March 2010)

In this lecture we will discuss the notions of stability and linearisation. Both notions are related to the question of how a dynamical system behaves in the neighborhood of a fixed point.

We will consider two different notions of stability:

- **Lyapunov stability**
- **asymptotic stability**

These notions can be defined for time-independent (= autonomous) first-order system of ordinary differential equations,

$$\frac{d}{dt} z^A(t) \;=\; F_A\big(\, z^1(t)\,,\,\ldots\,,\,z^\ell(t)\,\big)\,, \qquad A \;=\; 1,\ldots,\ell\,.$$

This includes all time-independent Hamiltonian systems, see next page.

Hamilton's equations can be written in the form

$$\frac{d}{dt}\begin{pmatrix} z^1 \\ \vdots \\ z^{2n} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ & \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial z^1} \\ \vdots \\ \frac{\partial H}{\partial z^{2n}} \end{pmatrix} ,$$

where $\ell = 2n$ and

$$z^1 = x^1 \, , \, \ldots \, , \, z^n = x^n \, , \, z^{n+1} = p_1 \, , \, \ldots \, , \, z^{2n} = p_n \, .$$

If the Hamiltonian is time-independent, this system of equations is, indeed, of the desired form

$$\frac{d}{dt} z^A \, = \, F_A\big(\, z^1 \, , \, \ldots \, , \, z^\ell \, \big) \, , \qquad A \, = \, 1, \ldots , \ell \, .$$

The inventor of stability theory:

Aleksandr Mikhailovich Lyapunov (1857, Yaroslavl – 1918, Odessa)

Professor of Applied Mathematics, St. Petersburg (1902 – 1917)

“General problem of the stability of motion” (1892)

We now introduce the stability notions. Assume that our sytem

$$\frac{d}{dt} z^A(t) \;=\; F_A\big(\, z^1(t)\,,\,\ldots\,,\,z^\ell(t)\,\big)\,, \qquad A \;=\; 1,\ldots,\ell$$

has an equilibrium point (= fixed point) at $\zeta^1,\ldots,\zeta^\ell$.

By definition, this means that

$$F_A\big(\,\zeta^1\,,\,\ldots\,,\,\zeta^\ell\,\big) \;=\; 0\,, \qquad A \;=\; 1\,\ldots,\ell\,,$$

so the solution with initial condition $z^A(0) = \zeta^A$ is the constant curve $z^A(t) = \zeta^A$.

The following definition says under what condition an equilibrium point is called Lyapunov stable or asymptotically stable.

Definition: The equilibrium point $\zeta^1, \ldots, \zeta^\ell$ is said to be

- **"Lyapunov stable" if for every $\varepsilon > 0$ there is a $\delta > 0$ such that**

$$\sqrt{\sum_{A=1}^{\ell} \left(z^A(0) - \zeta^A\right)^2} < \delta \implies \sqrt{\sum_{A=1}^{\ell} \left(z^A(t) - \zeta^A\right)^2} < \varepsilon \text{ for all } t > 0 ;$$

- **"attractive" if there is a $\delta > 0$ such that**

$$\sqrt{\sum_{A=1}^{\ell} \left(z^A(0) - \zeta^A\right)^2} < \delta \implies z^A(t) \to \zeta^A \text{ for } t \to \infty ;$$

- **"asymptotically stable" if it is Lyapunov stable and attractive.**

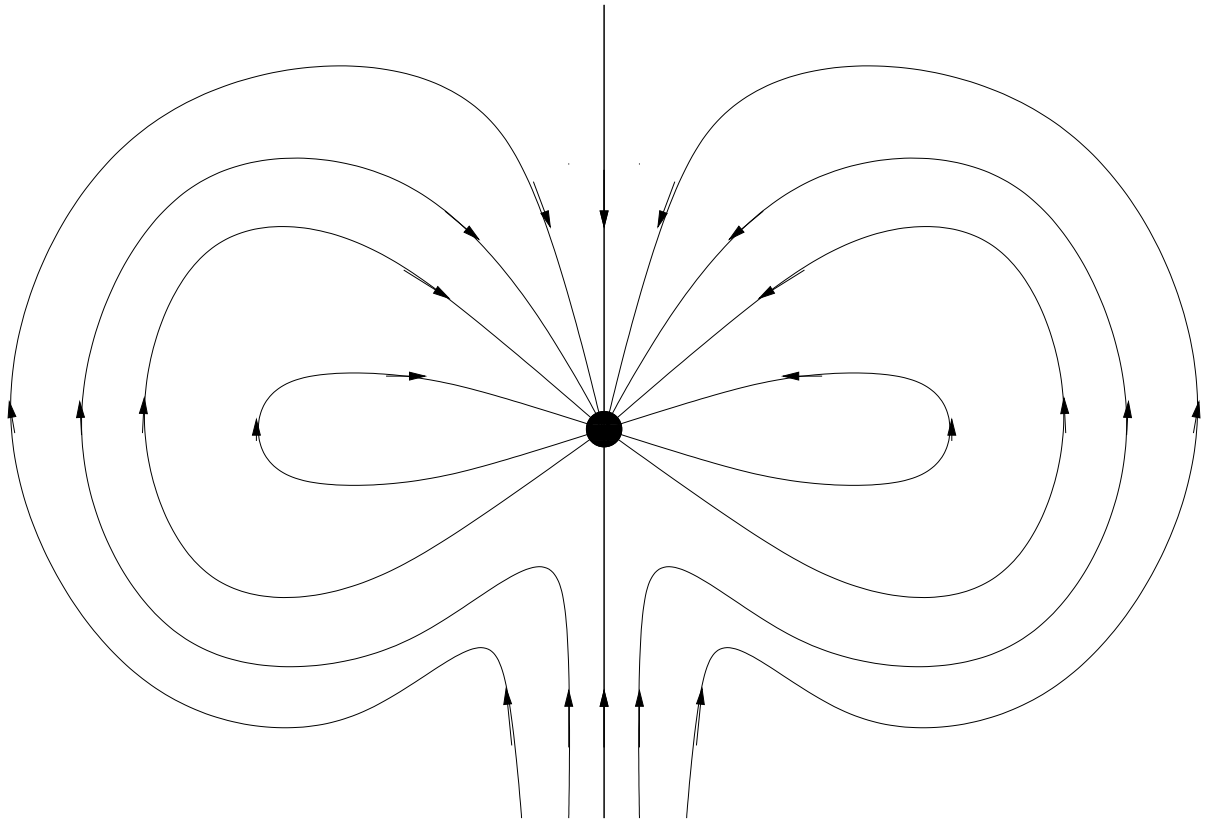
Roughly speaking:

An equilibrium point is Lyapunov stable if solutions stay arbitrarily close to it if they start sufficiently close to it.

An equilibrium point is attractive if solutions converge upon it if they start sufficiently close to it.

An equilibrium point is asymptotically stable if solutions converge upon it if they start sufficiently close to it and stay close to it during this convergence process.

Note that, in general, an equilibrium point may be attractive without being Lyapunov stable, see example next page.



This picture shows the solution curves (flow lines) of a two-dimensional system, $\ell = 2$. The equilibrium point (heavy dot) is attractive but not Lyapunov stable.

A sufficient (but not necessary) condition for stability is the existence of a Lyapunov function:

Theorem: Let $\Phi(z^1, \ldots, , z^\ell)$ be a differentiable function, defined in a neighborhood of an equilibrium point $\zeta^1, \ldots, \zeta^\ell$, such that

(a)

$$\Phi(z^1, \ldots, z^\ell) \begin{cases} = 0 & \text{if } (z^1, \ldots, z^\ell) = (\zeta^1 \ldots, \zeta^\ell) , \\ > 0 & \text{if } (z^1, \ldots, z^\ell) \neq (\zeta^1 \ldots, \zeta^\ell) , \end{cases}$$

(b)

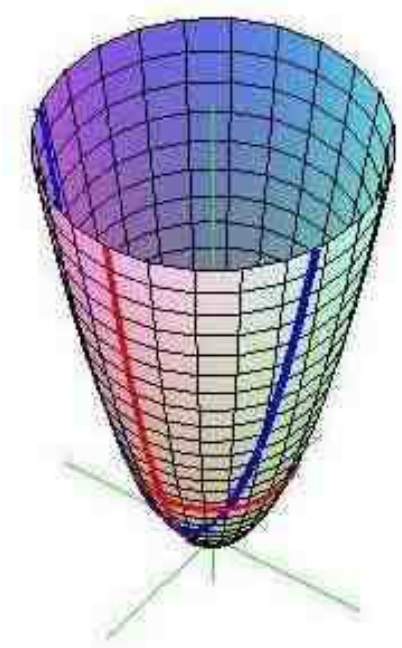
$$\sum_{A=1}^{\ell} \frac{\partial \Phi(z^1, \ldots, z^\ell)}{\partial z^A} F_A(z^1, \ldots, z^\ell) \leq 0$$

$$\text{for all } (z^1, \ldots, z^\ell) \neq (\zeta^1 \ldots, \zeta^\ell) .$$

Then $(\zeta^1, \ldots, \zeta^\ell)$ is Lyapunov stable. If the inequality in (b) is strict, $(\zeta^1, \ldots, \zeta^\ell)$ is asymptotically stable. Φ is called a "Lyapunov function" in the first case and a "strict Lyapunov function" in the second case.

This result is easy to understand: Condition (a) implies that Φ has a minimum at the equilibrium point. Condition (b) requires that Φ is non-increasing along solution curves,

$$\frac{d\Phi}{dt} = \sum_{A=1}^{\ell} \frac{\partial \Phi}{\partial z^A} \frac{dz^A}{dt} = \sum_{A=1}^{\ell} \frac{\partial \Phi}{\partial z^A} F_A \leq 0 \,.$$



Graph of a Lyapunov function

Example: Consider the damped harmonic oscillator

$$m\,\frac{d^2x}{dt^2} + \lambda\,\frac{dx}{dt} + k\,x = 0\,.$$

In first-order form, with

$$z^1 = x, \quad z^2 = \frac{dx}{dt}\,,$$

the equation of motion reads

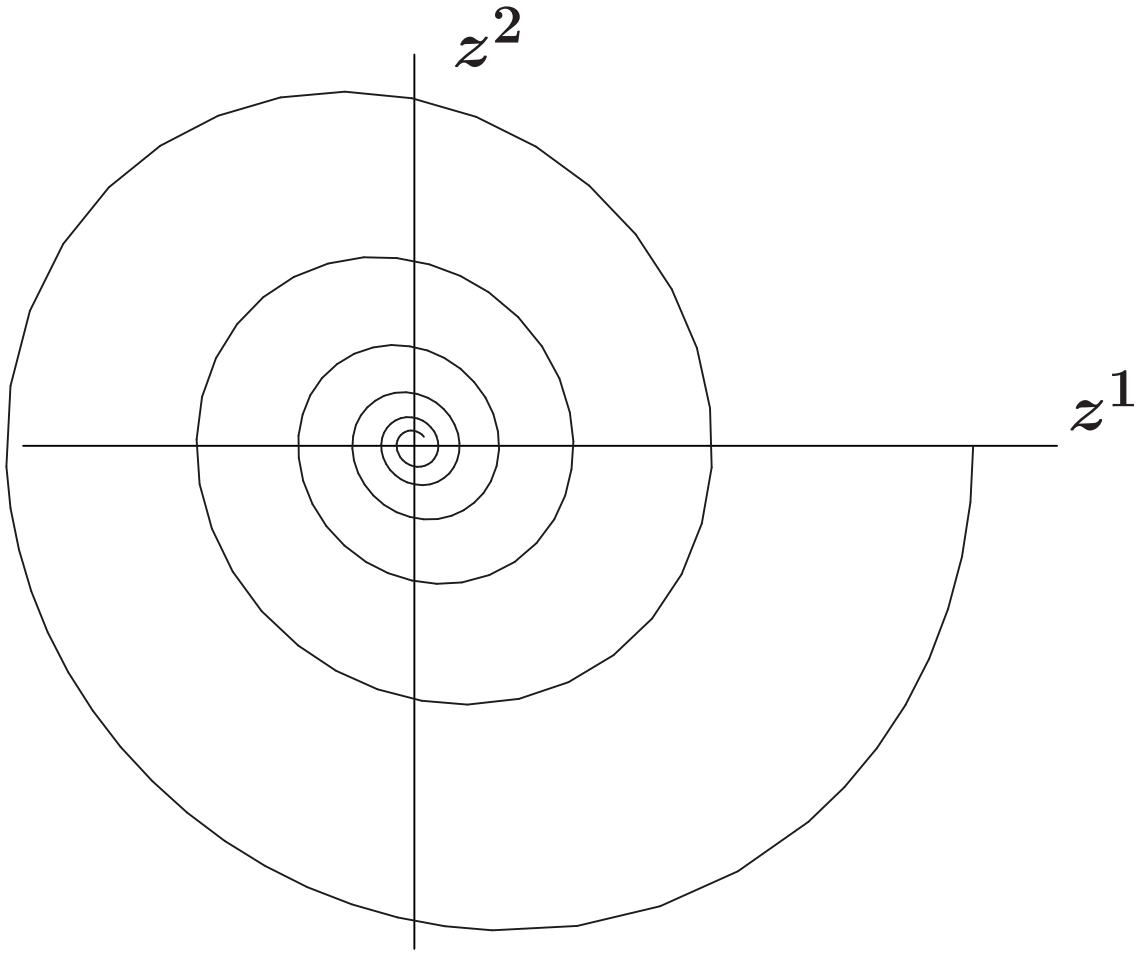
$$\frac{dz^1}{dt} = z^2\,, \qquad \frac{dz^2}{dt} = -\frac{k}{m}\,z^1 - \frac{\lambda}{m}\,z^2\,.$$

For non-vanishing, undercritical damping, $0 < \lambda < 2\sqrt{m\,k}$, the function

$$\Phi\big(z^1, z^2\big) = \frac{m}{2}\Big(z^2 + \frac{\alpha\,\lambda}{m}\,z^1\Big)^2 + \frac{k}{2}\,(z^1)^2$$

is a strict Lyapunov function, for any $0 < \alpha < 1$.

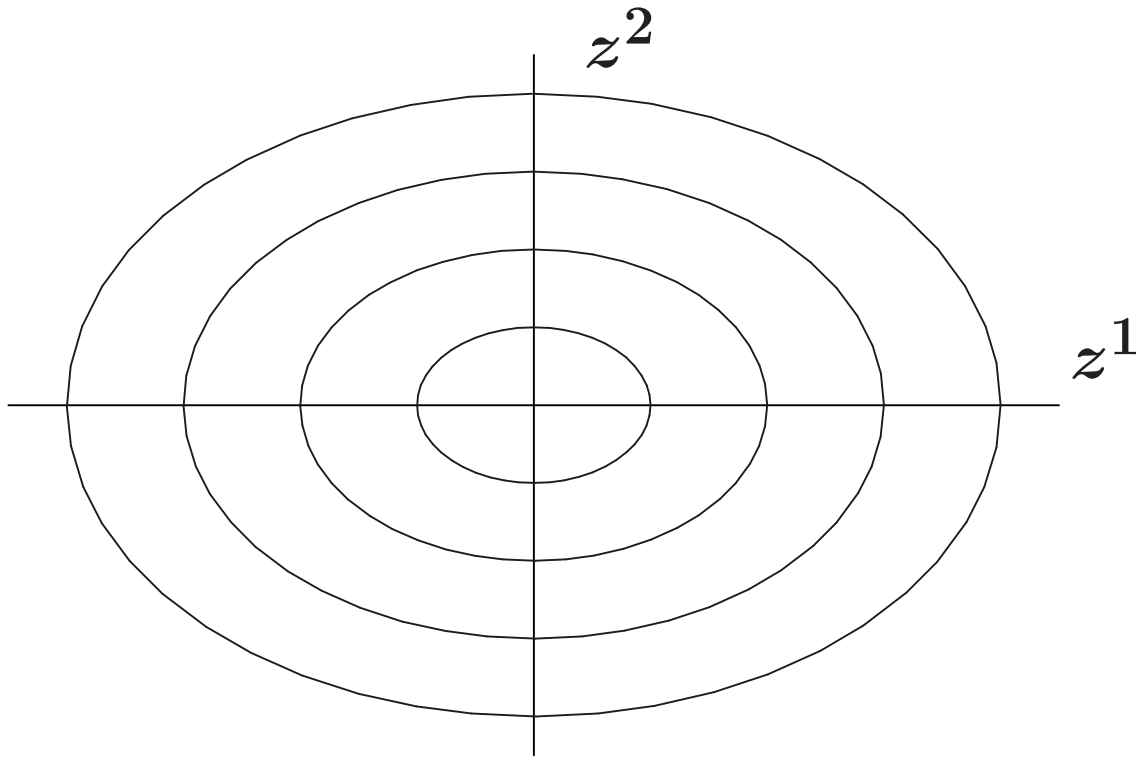
This demonstrates that the equilibrium is asymptotically stable.



For the undamped case, $\lambda = 0$, the mechanical energy

$$\Phi(z^1, z^2) = \frac{m}{2}(z^2)^2 + \frac{k}{2}(z^1)^2$$

is a Lyapunov function. This demonstrates that the equilibrium is Lyapunov stable. However, it is not asymptotically stable, so it is impossible to find a strict Lyapunov function.



We now turn to linear systems,

$$\frac{d}{dt}\begin{pmatrix} z^1(t) \\ \vdots \\ z^\ell(t) \end{pmatrix} = Q \begin{pmatrix} z^1(t) \\ \vdots \\ z^\ell(t) \end{pmatrix} ,$$

where Q is a constant $\ell \times \ell-$matrix. The general solution is

$$\begin{pmatrix} z^1(t) \\ \vdots \\ z^\ell(t) \end{pmatrix} = e^{tQ} \begin{pmatrix} z^1(0) \\ \vdots \\ z^\ell(0) \end{pmatrix} ,$$

where

$$e^{tQ} = \sum_{N=0}^{\infty} \frac{t^N}{N!} Q^N .$$

In this case $\left(z^1 , \ldots , z^\ell \right) = (0, \ldots, 0)$ is an equilibrium point. Its stability can be investigated with the methods of linear algebra (matrix calculation).

If we can find an $\ell \times \ell-$matrix M such that

(a) M is symmetric and has strictly positive eigenvalues,

(b) $Q^T M + MQ$ has non-positive eigenvalues,

then the function

$$\Phi(z^1, \ldots, z^\ell) = (z^1 \cdots z^\ell)\, M \begin{pmatrix} z^1 \\ \vdots \\ z^\ell \end{pmatrix}$$

is a Lyapunov function, i.e., the equilibrium is Lyapunov stable.

If, instead of condition (b), we have the stronger condition that $Q^T M + MQ$ has strictly negative eigenvalues, the Lyapunov function is strict and the equilibrium is asymptotically stable.

A Hamiltonian system is linear if and only if the Hamiltonian is a quadratic form (plus an irrelevant constant),

$$H(\, z^1 \,, \ldots \,, \, z^{2n} \,) \;=\; \sum_{A=1}^{2n} \sum_{B=1}^{2n} \; H_{AB} \, z^A \, z^B \;\;,$$

with constant coefficients H_{AB}, where

$$z^1 = x^1 \;,\; \ldots \;,\; z^n = x^n \;,\; z^{n+1} = p_1 \;,\; \ldots \;,\; z^{2n} = p_n \;\;.$$

If a non-linear Hamiltonian system is given, it can be "linearised" in the following way:

Choose canonical coordinates such that

$$(\, x^1 \,, \ldots \,, \, x^n \,, \, p_1 \,, \ldots \,, \, p_n \,) \;=\; (\, 0 \,, \ldots \,, \, 0 \,)$$

is an equilibrium point.

Then Taylor-expand the Hamiltonian around this point up to second order in all coordinates.

The zeroth order term is irrelevant and can be dropped; first order terms do not appear because the zero point is an equilibrium point. So the resultant Hamiltonian is a quadratic form and gives linear equations of motion. It is called the "linearised Hamiltonian".

Solving Hamilton's equations with the linearised Hamiltonian gives a good approximation to the original solution curves as long as the latter stay close to the equilibrium point.

If the equilibrium point is Lyapunov stable, this is true for all time if the initial conditions are chosen sufficiently close to the equilibrium point.

If the equilibrium point is not Lyapunov stable, the approximation may become very bad for large times, even if the initial condition was chosen very close to the equilibrium point.

Example 1: The "Duffing oscillator" is an anharmonic oscillator with Hamiltonian

$$H(x,p) \;=\; \frac{p^2}{2m} + \frac{k\,x^2}{2} + \frac{\mu\,x^4}{4}\,.$$

$x = p = 0$ is an equilibrium point. The linearised Hamiltonian gives the harmonic oscillator,

$$H^{\rm lin}(x,p) \;=\; \frac{p^2}{2m} + \frac{k\,x^2}{2}\,.$$

The equilibrium point is Lyapunov stable. Thus, if the initial condition is chosen close to the equilibrium point, the linearisation is a good approximation for all time. In other words: for small elongations the Duffing oscillator behaves like the harmonic oscillator.

Example 2: A trivial but instructive example is the Hamiltonian for the free relativistic particle,

$$H \;=\; c\,\sqrt{m^2\,c^2 \;+\; p_1^2 \;+\; p_2^2 \;+\; p_3^2}\;.$$

$x^1 = x^2 = x^3 = p_1 = p_2 = p_3 = 0$ is an equilibrium point, corresponding to the particle being at rest at the origin of the coordinate system. Taylor expansion yields

$$H \;=\; m\,c^2\,\sqrt{1 \;+\; \frac{p_1^2 \;+\; p_2^2 \;+\; p_3^2}{m^2\,c^2}} \;=\; m\,c^2\,\Big(\,1 \;+\; \frac{p_1^2 \;+\; p_2^2 \;+\; p_3^2}{2\,m^2\,c^2} \;+\; \ldots\,\Big)\,.$$

After dropping the irrelevant constant term, linearisation gives the Hamiltonian for the free non-relativistic particle,

$$H^{\rm lin} \;=\; \frac{p_1^2 \;+\; p_2^2 \;+\; p_3^2}{2\,m}\,.$$

Clearly, the equilibrium is not Lyapunov stable.

In accelerator physics, one considers a charged particle in an electromagnetic field and chooses a reference trajectory.

One decomposes the motion into longitudinal and transverse dynamics.

By a series of transformations one puts the Hamiltonian into a form such that the reference trajectory becomes an (approximate) equilibrium point for the transverse dynamics.

Linearisation of the transverse dynamics around this equilibrium point gives an approximation which is good if the particle stays close to the reference trajectory.

The derivation of the linearised accelerator Hamiltonian is given in detail by Andy Wolski in his Cockcroft Lectures on "Linear Dynamics".

We conclude with some references on Lagrangian and Hamiltonian Dynamics.

Elementary textbooks:

H. Goldstein: "Classical Mechanics" Addison-Wesley (1950)

E. Sudarshan and N. Mukunda: "Classical Dynamics" Wiley (1974)

H. Iro: "A Modern Approach to Classical Mechanics" World Scientic (2002)

With emphasis on geometry:

V. Arnold: "Mathematical Methods of Classical Mechanics" Springer (1974)

On stability:

A. Lyapunov: "Stability of motion" Academic Press (1966) [Translation of a paper from 1893]