Non Linear Dynamics - Methods and Tools

Yannis PAPAPHILIPPOU
Accelerator and Beam Physics group
Beams Department
CERN

CERN Accelerator School
Advanced Accelerator Physics Course 2019
Konferencecenter Metalskolen Jørlunde, Slangerup, Denmark
9-21 June 2017
Bibliography

- Books on non-linear dynamical systems

- Books on beam dynamics

- Lectures on non-linear beam dynamics
  - A. Chao, Advanced topics in Accelerator Physics, USPAS, 2000.
Content of lecture I

- Non-linear effects and their impact
- From linear to non-linear or from matrices to maps
- Introduction to TPSA and differential algebra
- Reminder of Lagrangian and Hamiltonian formalism, canonical transformation, and symplecticity
- The relativistic Hamiltonian for E/M fields
Non-linear effects

- Non-linear magnets, such as chromaticity sextupoles (especially in low emittance rings), octupoles,…

- Magnet imperfections and misalignments

- Insertion devices (wigglers, undulators) for synchrotron radiation storage rings

- Magnet fringe fields (especially in high-intensity machines)

- Power supply ripple

- Ground motion (for e+/e-)

- Electron (Ion) cloud

- Beam-beam effect (for colliders)

- Space-charge effect (especially in high-intensity machines)
Non-linear effects affect performance

- Non-linear magnets, such as chromaticity sextupoles (especially in low emittance rings), octupoles,…
- Magnet imperfections and misalignments
- Insertion devices (wigglers, undulators) for synchrotron radiation storage rings
- Magnet fringe fields (especially in high-intensity machines)
- Power supply ripple
- Ground motion (for e+/e-)
- Electron (Ion) cloud
- Beam-beam effect (for colliders)
- Space-charge effect (especially in high-intensity machines)

Performance impact

- Reduced injection efficiency
- Particle losses causing
  - Reduced intensity and beam lifetime
  - Radio-activation (equipment maintenance and lifetime)
  - Super-conducting magnet quench
  - Reduced machine availability
- Emittance increase
- Reduced number of bunches, increased crossing angle, affecting luminosity (for colliders)
- Allow to damp instabilities (see V. Kornilov lecture on “Landau damping”)
- Can be used for beam extraction
...but also cost

- Non-linear magnets, such as chromaticity sextupoles (especially in low emittance rings), octupoles,…
- Magnet imperfections and misalignments
- Insertion devices (wigglers, undulators) for synchrotron radiation storage rings
- Magnet fringe fields (especially in high-intensity machines)
- Power supply ripple
- Ground motion (for e+/e-)
- Electron (Ion) cloud
- Beam-beam effect (for colliders)
- Space-charge effect (especially in high-intensity machines)

- Performance impact
  - Reduced injection efficiency
  - Particle losses causing
    - Reduced intensity and beam lifetime
    - Radio-activation (equipment maintenance and lifetime)
    - Super-conducting magnet quench
    - Reduced machine availability
  - Emittance increase
  - Reduced number of bunches, increased crossing angle, affecting luminosity (for colliders)
  - Allow to damp instabilities (see V. Kornilov lecture on “Landau damping”)
  - Can be used for beam extraction

- Cost issues
  - Magnet field quality, alignment tolerances
  - Number of magnet corrector, power convertor families and specifications
  - Design of collimation system
From linear to non-linear or from matrices to maps
Linear system in beam dynamics

- Linear (uncoupled) transverse particle motion is described by Hill’s equation

\[ x'' + K_x(s) \cdot x = 0 \]

- Linear equations with s-dependent coefficients (harmonic oscillator with time dependent frequency)
Linear system in beam dynamics

- Linear (uncoupled) transverse particle motion is described by Hill’s equation

\[ x'' + K_x(s) \cdot x = 0 \]

- Linear equations with s-dependent coefficients (harmonic oscillator with time dependent frequency)

- In a ring (or in transport line with symmetries), coefficients are periodic \( K_x(s) = K_x(s + C) \)

- Not straightforward to derive closed analytical solutions for the whole accelerator...
Linear system in beam dynamics

- Linear (uncoupled) transverse particle motion is described by Hill’s equation

\[ x'' + K_x(s) \cdot x = 0 \]

- Linear equations with \( s \)-dependent coefficients (harmonic oscillator with time dependent frequency)

- In a ring (or in transport line with symmetries), coefficients are periodic \( K_x(s) = K_x(s + C) \)

- Not straightforward to derive closed analytical solutions for the whole accelerator...

- ...but do we really care, in particular for a system composed by discrete building blocks?
Consider $K(s) = k_0 = \text{constant}$

$$u'' + k_0 \, u = 0$$

Equations of harmonic oscillator with solution

$$u(s) = C(s) \, u(0) + S(s) \, u'(0)$$

$$u'(s) = C'(s)u(0) + S'(s)u'(0)$$
### Harmonic Oscillator

- Consider $K(s) = k_0 = \text{constant}$
  
  $$u'' + k_0 \ u = 0$$

- Equations of harmonic oscillator with solution
  
  $$u(s) = C(s) \ u(0) + S(s) \ u'(0)$$
  $$u'(s) = C'(s) u(0) + S'(s) u'(0)$$

  with

  $$C(s) = \cos(\sqrt{k_0} s) \quad \text{and} \quad S(s) = \frac{1}{\sqrt{k_0}} \sin(\sqrt{k_0} s) \quad \text{for} \ k_0 > 0$$

  $$C(s) = \cosh(\sqrt{|k_0|} s) \quad \text{and} \quad S(s) = \frac{1}{\sqrt{|k_0|}} \sinh(\sqrt{|k_0|} s) \quad \text{for} \ k_0 < 0$$

- Note that the solution can be written in matrix form

  $$\begin{pmatrix} u(s) \\ u'(s) \end{pmatrix} = \begin{pmatrix} C(s) & S(s) \\ C'(s) & S'(s) \end{pmatrix} \begin{pmatrix} u(0) \\ u'(0) \end{pmatrix}$$
Matrix formalism

- General **transfer matrix** from $s_0$ to $s$

\[
\begin{pmatrix}
  u \\
  u' \\
\end{pmatrix}_s = \mathcal{M}(s|s_0) \begin{pmatrix}
  u \\
  u' \\
\end{pmatrix}_{s_0} = \begin{pmatrix}
  C(s|s_0) & S(s|s_0) \\
  C'(s|s_0) & S'(s|s_0) \\
\end{pmatrix} \begin{pmatrix}
  u \\
  u' \\
\end{pmatrix}_{s_0}
\]

- Note that $\det(\mathcal{M}(s|s_0)) = C(s|s_0)S'(s|s_0) - S(s|s_0)C'(s|s_0) = 1$, which is always true for conservative systems.
Matrix formalism

- **General transfer matrix** from $s_0$ to $s$
  \[
  \begin{pmatrix}
  u \\
  u'
  \end{pmatrix}_{s} = M(s|s_0) \begin{pmatrix}
  u \\
  u'
  \end{pmatrix}_{s_0} = \begin{pmatrix}
  C(s|s_0) & S(s|s_0) \\
  C'(s|s_0) & S'(s|s_0)
  \end{pmatrix} \begin{pmatrix}
  u \\
  u'
  \end{pmatrix}_{s_0}
  \]

- Note that $\det(M(s|s_0)) = C(s|s_0)S'(s|s_0) - S(s|s_0)C'(s|s_0) = 1$
  which is always true for conservative systems

- Any line can be built by a series of matrix multiplications
  $M(s_n|s_0) = M(s_n|s_{n-1}) \cdots M(s_3|s_2) \cdot M(s_2|s_1) \cdot M(s_1|s_0)$

- For a full ring, the matrix multiplication will provide the full transfer matrix for 1-turn
  \[
  M_C = \begin{pmatrix}
  \cos \mu + \alpha \sin \mu & \beta \sin \mu \\
  -\gamma \sin \mu & \cos \mu - \alpha \sin \mu
  \end{pmatrix}
  \]
Make a coordinate transformation so that we get a simpler form of the matrix, i.e. **ellipses** are transformed to circles (simple rotation)

\[ M = \mathcal{A} \circ R \circ \mathcal{A}^{-1} \quad \text{or:} \quad R = \mathcal{A}^{-1} \circ M \circ \mathcal{A} \]

Using linear algebra, the solution is

\[
\mathcal{A} = \begin{pmatrix}
\frac{\sqrt{\beta(s_0)}}{\sqrt{\beta(s_0)}} & 0 \\
\frac{\alpha(s_0)}{\sqrt{\beta(s_0)}} & 1
\end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix}
\cos(\mu_x) & \sin(\mu_x) \\
-\sin(\mu_x) & \cos(\mu_x)
\end{pmatrix}
\]

This transformation can be extended to a non-linear system (see later)
Non-linear motion

- Nonlinear elements can be represented by generalized polynomials
  \[ x'' K_x(s)x = \sum_{i,j} a_{ij}(s) x^i y^j \]

- For example, general magnetic fields can be represented by the multi-pole expansion
  \[ B_y + iB_x = \sum_{n=1}^{\infty} (b_n - ia_n)(x + iy)^{n-1} \]

- Equations of motion in the horizontal plane become
  \[ x'' + K_x(s)x = -\frac{B_y(x, y, s)}{p} \]

- Closed solution does not exist, in principle!
A generalization of the matrix (which can only describe linear systems), is a map, which transforms a system from some initial to some final coordinates.

Analyzing the map, will give useful information about the behavior of the system.

There are different ways to build the map:
- Taylor (Power) maps
- Lie transformations (see Lecture II)
- Truncated Power Series Algebra (TPSA), can generate maps from straight-forward tracking (see E. Forest lecture)

Preservation of symplecticity is important.
Building a non-linear map

- For a thin quadrupole the equivalent map can be written

\[
\begin{pmatrix}
   x \\
   x' \\
   y \\
   y'
\end{pmatrix}_{s_2} = \begin{pmatrix}
   x \\
   x' \\
   y \\
   y'
\end{pmatrix}_{s_1} + \begin{pmatrix}
   0 \\
   k_1 \cdot x_{s_1} \\
   0 \\
   k_1 \cdot y_{s_1}
\end{pmatrix}
\]

or through the matrix \( M \), as

\[
\tilde{z}(s_2) = M \cdot \tilde{z}(s_1).
\]

- For a thin sextupole, we can right the coordinate transformation as

\[
\begin{pmatrix}
   x \\
   x' \\
   y \\
   y'
\end{pmatrix}_{s_2} = \begin{pmatrix}
   x \\
   x' \\
   y \\
   y'
\end{pmatrix}_{s_1} + \begin{pmatrix}
   0 \\
   \frac{1}{2} k_2 \cdot (x_{s_1}^2 - y_{s_1}^2) \\
   0 \\
   k_2 \cdot (x_{s_1} \cdot y_{s_1})
\end{pmatrix}
\]

or \( \tilde{z}(s_2) = M \circ \tilde{z}(s_1) \) where now \( M \) is a non-linear map.
A general representation for the map for the horizontal position can be

\[ x_{new} = \left( R_{11} \cdot x + R_{12} \cdot x' + R_{21} \cdot y + R_{22} \cdot y' \right) + \left( T_{111} \cdot x^2 + T_{112} \cdot x x' + T_{122} \cdot x'^2 + T_{113} \cdot x y + T_{114} \cdot x y' \right) + \ldots. \]

or, in a more compact form up to 3\textsuperscript{rd} order, for \( j = 1, \ldots, 6 \)

\[ z_{j}^{new} = \sum_{k=1}^{6} R_{jk} z_k + \sum_{k=1}^{6} \sum_{l=1}^{6} T_{jkl} z_k z_l + \sum_{k=1}^{6} \sum_{l=1}^{6} \sum_{m=1}^{6} U_{jklm} z_k z_l z_m \]
For a sextupole in one plane the representation is written as

\[
\begin{pmatrix}
  x \\
  x'
\end{pmatrix}_{new} =
\begin{pmatrix}
  R_{11} & R_{12} & T_{111} & T_{112} & T_{122} \\
  R_{21} & R_{22} & T_{211} & T_{212} & T_{222}
\end{pmatrix}
\begin{pmatrix}
  x \\
  x' \\
  x^2 \\
  xx' \\
  x'^2
\end{pmatrix}
\]

or in general for a sextupole of length \( L \) and strength \( \kappa_2 \)

\[
x_2 = x_1 + Lx_1' - k_2 \left( \frac{L^2}{4} (x_1^2 - y_1^2) + \frac{L^3}{12} (x_1 x_1' - y_1 y_1') + \frac{L^4}{24} (x_1'^2 - y_1'^2) \right)
\]

\[
x_2' = x_1' - k_2 \left( \frac{L}{2} (x_1^2 - y_1^2) + \frac{L^2}{4} (x_1 x_1' - y_1 y_1') + \frac{L^3}{6} (x_1'^2 - y_1'^2) \right)
\]

\[
y_2 = y_1 + Ly_1' + k_2 \left( \frac{L^2}{4} x_1 y_1 + \frac{L^3}{12} (x_1 y_1' + y_1 x_1') + \frac{L^4}{24} (x_1' y_1') \right)
\]

\[
y_2' = y_1' + k_2 \left( \frac{L}{2} x_1 y_1 + \frac{L^2}{4} (x_1 y_1' + y_1 x_1') + \frac{L^3}{6} (x_1' y_1') \right)
\]

But what about symplecticity?

Need to introduce Hamiltonian formalism
Introduction to Truncated Power Series Algebra (TPSA)
Let’s consider a tracked particle at position $\alpha$ and a small deviation $\Delta x$. The Taylor series around this position is

$$f(a + \Delta x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!} \Delta x^n.$$ 

$$= f(a) + \frac{f'(a)}{1!} \Delta x^1 + \frac{f''(a)}{2!} \Delta x^2 + \frac{f'''(a)}{3!} \Delta x^3 + \cdots$$
Let’s consider a tracked particle at position $a$ and a small deviation $\Delta x$. The Taylor series around this position is

$$f(a + \Delta x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!} \Delta x^n$$

$$= f(a) + \frac{f'(a)}{1!} \Delta x^1 + \frac{f''(a)}{2!} \Delta x^2 + \frac{f'''(a)}{3!} \Delta x^3 + \cdots$$

By truncating we have

$$f(a + \Delta x) = f(a) + \sum_{n=1}^{m} \frac{f^{(n)}(a)}{n!} \Delta x^n$$

and the function $f(x)$ can be represented by the vector

$$(f(a), f'(a), f''(a), \ldots, f^{(m)}(a))$$

This vector is a Truncated Power Series Algebra

We need the derivatives $f^{(n)}(a)$ of $f(x)$ at $a$ with

$$f'(a) = \frac{f(a + \epsilon) - f(a)}{\epsilon}$$

which is numerically non-trivial (small divisors, accuracy for higher orders, ...).
The basic idea is the **automatic differentiation** of results produced by a tracking code to provide the **coefficients** of a Taylor series.
The basic idea is the **automatic differentiation** of results produced by a tracking code to provide the **coefficients** of a Taylor series.

Consider a pair of real numbers \((q_0, q_1)\) and define **operations** on a pair like

\[
(q_0, q_1) + (r_0, r_1) = (q_0 + r_0, q_1 + r_1)
\]

\[
c \cdot (q_0, q_1) = (c \cdot q_0, c \cdot q_1)
\]

\[
(q_0, q_1) \cdot (r_0, r_1) = (q_0 \cdot r_0, q_0 \cdot r_1 + q_1 \cdot r_0)
\]
The basic idea is the **automatic differentiation** of results produced by a tracking code to provide the **coefficients** of a Taylor series.

Consider a pair of real numbers \((q_0, q_1)\) and define **operations** on a pair like

\[(q_0, q_1) + (r_0, r_1) = (q_0 + r_0, q_1 + r_1)\]
\[c \cdot (q_0, q_1) = (c \cdot q_0, c \cdot q_1)\]
\[(q_0, q_1) \cdot (r_0, r_1) = (q_0 \cdot r_0, q_0 \cdot r_1 + q_1 \cdot r_0)\]

and some ordering

\[(q_0, q_1) < (r_0, r_1) \quad \text{if} \quad q_0 < r_0 \quad \text{or} \quad (q_0 = r_0 \quad \text{and} \quad q_1 < r_1)\]
\[(q_0, q_1) > (r_0, r_1) \quad \text{if} \quad q_0 > r_0 \quad \text{or} \quad (q_0 = r_0 \quad \text{and} \quad q_1 > r_1)\]

implying strange relations of the form

\[(0, 0) < (0, 1) < (r, 0), \quad \forall \ r > 0\]
\[(0, 1) \cdot (0, 1) = (0, 0) \rightarrow (0, 1) = \sqrt{(0, 0)}\]
We define the **differential unit** \( \epsilon \equiv (0, 1) \), which is located between 0 and any real number (infinitesimally small).

As \((q_0, 0)\) is just a real number, we can define a **real** part and a **differential** part

\[
q_0 = R(q_0, q_1) \quad \text{and} \quad q_1 = D(q_0, q_1)
\]
We define the **differential unit** \( \epsilon \equiv (0, 1) \), which is located between 0 and any real number (infinitesimally small).

As \((q_0, 0)\) is just a real number, we can define a **real** part and a **differential** part

\[
q_0 = \mathcal{R}(q_0, q_1) \quad \text{and} \quad q_1 = \mathcal{D}(q_0, q_1)
\]

Using the previous rules we can show

\[
(1, 0) \cdot (q_0, q_1) = (q_0, q_1)
\]

\[
(q_0, q_1)^{-1} = \left( \frac{1}{q_0}, -\frac{q_1}{q_0^2} \right)
\]

A **function** acting on a pair is \( f(x) = \mathcal{R}[f(x, q_1)], \quad \forall q_1 \)

The **differential** is

\[
\mathcal{D}[f(x + \epsilon)] = \mathcal{D}[f((x, 0) + (0, 1))] = \mathcal{D}[f(x, 1)] = f'(x)
\]
Consider the function \( f(x) = x^2 + \frac{1}{x} \) with the derivative \( f'(x) = 2x - \frac{1}{x^2} \). For \( x = 2 \), we obtain \( (f(2), f'(2)) = \left( \frac{9}{2}, \frac{15}{4} \right) \).
Consider the function \( f(x) = x^2 + \frac{1}{x} \) with the derivative \( f'(x) = 2x - \frac{1}{x^2} \). For \( x = 2 \), we obtain \( (f(2), f'(2)) = \left( \frac{9}{2}, \frac{15}{4} \right) \).

Let’s use differential algebra, by substituting \( x \to (x, 1) = (2, 1) \) to the function and use the rules

\[
f[(2, 1)] = (2, 1)^2 + (2, 1)^{-1} \]

\[
= (4, 4) + \left( \frac{1}{2}, -\frac{1}{4} \right) \]

\[
= \left( \frac{9}{2}, \frac{15}{4} \right) = (f(2), f'(2))
\]

We computed exactly the derivative, only by using algebra!
Higher orders

- The operation can be extended to derivatives of order $N$ by considering that the pair becomes

$$ (q_0, 1) \rightarrow (q_0, 1, 0, 0, \ldots, 0) \quad \text{with} \quad \epsilon = (0, 1, 0, 0, \ldots, 0) $$

- We can extend the operations as

$$ (q_0, q_1, q_2, \ldots, q_N) + (r_0, r_1, r_2, \ldots, r_N) = (q_0 + r_0, q_1 + r_1, q_2 + r_2, \ldots, q_N + r_N) $$

$$ c \cdot (q_0, q_1, q_2, \ldots, q_N) = (c \cdot q_0, c \cdot q_1, c \cdot q_2, \ldots, c \cdot q_N) $$

$$ (q_0, q_1, q_2, \ldots, q_N) \cdot (r_0, r_1, r_2, \ldots, r_N) = (s_0, s_1, s_2, \ldots, s_N) $$

with

$$ s_i = \sum_{k=0}^{i} \frac{i!}{k! (i-k)!} q_k r_{i-k} $$

- For example

$$ (x, 0, 0, 0, \ldots, 0)^n = (x^n, 0, 0, 0, \ldots, 0) $$

$$ (0, 1, 0, 0, \ldots, 0)^n = (0, 0, 0, \ldots, n!, \ldots, 0) $$

$$ (x, 1, 0, 0, \ldots, 0)^2 = (x^2, 2x, 2, 0, \ldots, 0) $$

$$ (x, 1, 0, 0, \ldots, 0)^3 = (x^3, 3x^2, 6x, 6, 0, \ldots, 0) $$
The operation can be extended **more variables**

\[ x = (a, 1, 0, 0, 0...) \quad \epsilon_x = (0, 1, 0, 0, 0, \ldots) \]

\[ p_x = (b, 0, 1, 0, 0...) \quad \epsilon_{p_x} = (0, 0, 1, 0, 0, \ldots) \]

With some modified **multiplication rules**

\[(q_{00}, q_{10}, q_{01}, q_{20}, \ldots) \cdot (r_{00}, r_{10}, r_{01}, r_{20}, \ldots) = (s_{00}, s_{10}, s_{01}, s_{20}, \ldots)\]

with \[ s_{mn} = \sum_{k=0}^{m} \sum_{l=0}^{n} q_{kl} \cdot r_{m-k,n-l} \frac{m! \cdot n!}{k! \cdot (m-k)! \cdot l! \cdot (n-l)!} \]

providing \[ f(x, p_x) = \left( f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial p_x}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial p_x}, \ldots \right) \]

Using the formalism above, a **truncated Taylor map** with the desired accuracy and to **any order**, directly from **tracking data** (see more details in E. Forest lectures)
Reminder of Lagrangian and Hamiltonian formalism
Lagrangian formalism

- Describe motion of particles in $q_n$ coordinates ($n$ degrees of freedom) from time $t_1$ to time $t_2$
- It can be achieved by the Lagrangian function $L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$ with $(q_1, \ldots, q_n)$ the generalized coordinates and $(\dot{q}_1, \ldots, \dot{q}_n)$ the generalized velocities
Lagrangian formalism

- Describe motion of particles in $q_n$ coordinates ($n$ degrees of freedom) from time $t_1$ to time $t_2$

- It can be achieved by the Lagrangian function $L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$ with $(q_1, \ldots, q_n)$ the generalized coordinates and $(\dot{q}_1, \ldots, \dot{q}_n)$ the generalized velocities

- The Lagrangian is defined as $L = T - V$, i.e. difference between kinetic and potential energy

- The integral $W = \int L(q_i, \dot{q}_i, t) dt$ defines the action

- Hamilton’s principle: system evolves so as the action becomes extremum (principle of stationary action)
By using Hamilton’s principle, i.e. $\delta W = 0$, over some time interval $t_1$ and $t_2$ for two stationary points $\delta q(t_1) = \delta q(t_2) = 0$ (see appendix), the following differential equations for each degree of freedom are obtained, the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$$

In other words, by knowing the form of the Lagrangian, the equations of motion can be derived.
For a simple force law contained in a potential function, governing motion among interacting particles, the Lagrangian is (or as Landau-Lifshitz put it “experience has shown that…”)

\[ L = T - V = \sum_{i=1}^{n} \frac{1}{2} m_i q_i^2 - V(q_1, \ldots, q_n) \]

For velocity independent potentials, Lagrange equations become

\[ m_i \ddot{q}_i = - \frac{\partial V}{\partial q_i}, \]

i.e. Newton’s equations.
Some disadvantages of the Lagrangian formalism:

- No uniqueness: different Lagrangians can lead to same equations
- Physical significance not straightforward (even its basic form given more by “experience” and the fact that it actually works that way!)

Lagrangian function provides in general $n$ second order differential equations (coordinate space)

We already observed the advantage to move to a system of $2n$ first order differential equations, which are more straightforward to solve (phase space)

These equations can be derived by the Hamiltonian of the system
The Hamiltonian of the system is defined as the Legendre transformation of the Lagrangian

\[ H(q, \dot{q}, t) = \sum_i \dot{q}_i p_i - L(q, \dot{q}, t) \]

where the generalised momenta are \( p_i = \frac{\partial L}{\partial \dot{q}_i} \).

The generalised velocities can be expressed as a function of the generalised momenta if the previous equation is invertible, and thereby define the Hamiltonian of the system.

Example: consider

\[ L(q, \dot{q}) = \frac{1}{2} \sum_i m_i \dot{q}_i^2 - V(q_1, \ldots, q_n) \]

From this, the momentum can be determined as \( p_i = \frac{\partial L}{\partial \dot{q}_i} = m_i \dot{q}_i \)

which can be trivially inverted to provide the Hamiltonian

\[ H(q, p) = \sum_i \frac{p_i^2}{2m_i} + V(q_1, \ldots, q_n) \]
Hamilton’s equations

- The equations of motion can be derived from the Hamiltonian following the same variational principle as for the Lagrangian ("least" action) but also by simply taking the differential of the Hamiltonian (see appendix).

\[
\begin{align*}
\dot{q}_i &= \frac{\partial H}{\partial p_i}, \\
\dot{p}_i &= -\frac{\partial H}{\partial q}, \\
\frac{\partial L}{\partial t} &= -\frac{\partial H}{\partial t}
\end{align*}
\]

- These are indeed $2n + 2$ equations describing the motion in the “extended” phase space $(q_i, \ldots, q_n, p_1, \ldots, p_n, t, -H)$
Properties of Hamiltonian flow

- The variables \((q_i, \ldots, q_n, p_1, \ldots, p_n, t, -H)\) are called **canonically conjugate** (or canonical) and define the evolution of the system in **phase space**
- These variables have the special property that they preserve volume in phase space, i.e. satisfy the well-known **Liouville’s theorem**
- The variables used in the Lagrangian do not necessarily have this property
- Hamilton’s equations can be written in **vector form**
  \[
  \dot{z} = J \cdot \nabla H(z) \quad \text{with} \quad z = (q_i, \ldots, q_n, p_1, \ldots, p_n)
  \]
  and \(\nabla = (\partial q_i, \ldots, \partial q_n, \partial p_1, \ldots, \partial p_n)\)
- The \(2n \times 2n\) matrix \(J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\) is called the **symplectic matrix**
Nonlinear dynamics, CERN Accelerator School, June 2019

Poisson brackets

- Crucial step in study of Hamiltonian systems is identification of **integrals of motion**
- Consider a time dependent function of phase space. Its time evolution is given by

\[
\frac{d}{dt} f(p, q, t) = \sum_{i=1}^{n} \left( \frac{dq_i}{dt} \frac{\partial f}{\partial q_i} + \frac{dp_i}{dt} \frac{\partial f}{\partial p_i} \right) + \frac{\partial f}{\partial t}
\]

\[
= \sum_{i=1}^{n} \left( \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i} \right) + \frac{\partial f}{\partial t} = [H, f] + \frac{\partial f}{\partial t}
\]

where \([H, f]\) is the **Poisson bracket** of \(f\) with \(H\)

- If a quantity is explicitly **time-independent** and its Poisson bracket with the Hamiltonian vanishes (i.e. commutes with the \(H\)), it is a **constant** (or **integral**) of motion (as an **autonomous** Hamiltonian itself)
Poisson brackets’ properties

The Poisson brackets between two functions of a set of canonical variables can be defined by the differential operator

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

From this definition, and for any three given functions, the following properties can be shown:

- Bilinearity:
  \[\[a f + b g, h\] = a [f, h] + b [g, h], a, b \in \mathbb{R}\]

- Anticommutativity:
  \[[f, g] = -[g, f]\]

- Jacobi’s identity:
  \[[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0\]

- Leibniz’s rule:
  \[[f, gh] = [f, g]h + g[f, h]\]

Poisson brackets operation satisfies a Lie algebra.
Canonical transformations
Find a function for transforming the Hamiltonian from variable \((q, p)\) to \((Q, P)\) so system becomes simpler to study.

This transformation should be canonical (or symplectic), so that the Hamiltonian properties of the system are preserved.

These “mixed variable” generating functions are derived by

\[
F_1(q, Q) : p_i = \frac{\partial F_1}{\partial q_i}, \quad P_i = -\frac{\partial F_1}{\partial Q_i} \quad F_3(Q, p) : q_i = -\frac{\partial F_3}{\partial p_i}, \quad P_i = -\frac{\partial F_3}{\partial Q_i}
\]

\[
F_2(q, P) : p_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i} \quad F_4(p, P) : q_i = -\frac{\partial F_4}{\partial p_i}, \quad Q_i = \frac{\partial F_4}{\partial P_i}
\]

A general non-autonomous Hamiltonian is transformed to

\[
H(Q, P, t) = H(q, p, t) + \frac{\partial F_j}{\partial t}, \quad j = 1, 2, 3, 4
\]

One generating function can be constructed by the other through Legendre transformations, e.g.

\[
F_2(q, P) = F_1(q, Q) - Q \cdot P, \quad F_3(Q, p) = F_1(q, Q) - q \cdot p, \quad \ldots
\]

with the inner product define as \(q \cdot p = \sum_i q_i p_i\)
A fundamental property of canonical transformations is the 
**preservation of phase space volume**

This volume preservation in phase space can be represented 
in the old and new variables as

\[
\int \prod_{i=1}^{n} dp_i dq_i = \int \prod_{i=1}^{n} dP_i dQ_i
\]

The volume element in old and new variables are related 
through the Jacobian

\[
\prod_{i=1}^{n} dp_i dq_i = \frac{\partial (P_1, \ldots, P_n, Q_1, \ldots, Q_n)}{\partial (p_1, \ldots, p_n, q_1, \ldots, q_n)} \prod_{i=1}^{n} dP_i dQ_i
\]

These two relationships imply that the Jacobian of a 
canonical transformation should have determinant equal to 1

\[
\frac{\partial (P_1, \ldots, P_n, Q_1, \ldots, Q_n)}{\partial (p_1, \ldots, p_n, q_1, \ldots, q_n)} \frac{\partial (p_1, \ldots, p_n, q_1, \ldots, q_n)}{\partial (P_1, \ldots, P_n, Q_1, \ldots, Q_n)} = 1
\]
The transformation $Q = -p$, $P = q$, which interchanges conjugate variables is area preserving, as the Jacobian is

$$\frac{\partial (P,Q)}{\partial (p,q)} = \begin{vmatrix} \frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q} \end{vmatrix} = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = 1$$
Examples of transformations

- The transformation \( Q = -p , \quad P = q \), which **interchanges conjugate variables** is area preserving, as the Jacobian is

\[
\frac{\partial(P,Q)}{\partial(p,q)} = \begin{vmatrix}
\frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\
\frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q}
\end{vmatrix} = \begin{vmatrix}
0 & -1 \\
1 & 0
\end{vmatrix} = 1
\]

- On the other hand, the transformation from **Cartesian to polar** coordinates \( q = P \cos Q , \quad p = P \sin Q \) is not, since

\[
\frac{\partial(q,p)}{\partial(Q,P)} = \begin{vmatrix}
- P \sin Q & P \cos Q \\
\cos Q & \sin Q
\end{vmatrix} = -P
\]
Examples of transformations

- The transformation \( Q = -p \), \( P = q \), which interchanges conjugate variables is area preserving, as the Jacobian is

\[
\frac{\partial (P,Q)}{\partial (p,q)} = \begin{vmatrix}
\frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\
\frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q}
\end{vmatrix} = \begin{vmatrix}
0 & -1 \\
1 & 0
\end{vmatrix} = 1
\]

- On the other hand, the transformation from Cartesian to polar coordinates \( q = P \cos Q \), \( p = P \sin Q \) is not, since

\[
\frac{\partial (q,p)}{\partial (Q,P)} = \begin{vmatrix}
-P \sin Q & P \cos Q \\
\cos Q & \sin Q
\end{vmatrix} = -P
\]

- There are actually “polar” coordinates that are canonical, given by \( q = -\sqrt{2P} \cos Q \), \( p = \sqrt{2P} \sin Q \) for which

\[
\frac{\partial (q,p)}{\partial (Q,P)} = \begin{vmatrix}
\sqrt{2P} \sin Q & \sqrt{2P} \cos Q \\
-\frac{\cos Q}{\sqrt{2P}} & \frac{\sin Q}{\sqrt{2P}}
\end{vmatrix} = 1
\]
The Relativistic Hamiltonian for electromagnetic fields
Neglecting self fields and radiation, motion can be described by a “single-particle” Hamiltonian

$$H(x, p, t) = c \sqrt{(p - \frac{e}{c} A(x, t))^2 + m^2 c^2 + e\Phi(x, t)}$$

- $x = (x, y, z)$ Cartesian positions
- $p = (p_x, p_y, p_z)$ conjugate momenta
- $A = (A_x, A_y, A_z)$ magnetic vector potential
- $\Phi$ electric scalar potential

The ordinary kinetic momentum vector is written

$$P = \gamma m v = p - \frac{e}{c} A$$

with $v$ the velocity vector and $\gamma = (1 - v^2/c^2)^{-1/2}$ the relativistic factor
- It is generally a 3 degrees of freedom one plus time (i.e. 4 degrees of freedom)
- The Hamiltonian represents the total energy
  \[ H \equiv E = \gamma mc^2 + e\Phi \]
- The total kinetic momentum is
  \[ P = \left( \frac{H^2}{c^2} - m^2 c^2 \right)^{1/2} \]
- Using Hamilton's equations
  \[ (\dot{x}, \dot{p}) = \left[(x, p), H\right] \]

it can be shown that motion is governed by **Lorentz equations**
Making a series of canonical transformations and approximations (see appendix)

- From Cartesian to Frenet-Serret (rotating) coordinate system (bending in the horizontal plane)
- Changing the independent variable from time to the path length $s$
- Electric field set to zero, as longitudinal (synchrotron) motion is much slower than transverse (betatron) one
- Consider static and transverse magnetic fields
- Rescale the momentum and move the origin to the periodic orbit
- For the ultra-relativistic limit $\beta_0 \to 1$, $\frac{1}{\beta_0^2 \gamma^2} \to 0$
  the Hamiltonian becomes

$$\mathcal{H}(x, y, l, p_x, p_y, \delta) = (1 + \delta) - e\hat{A}_s - \left(1 + \frac{x}{\rho(s)}\right) \sqrt{(1 + \delta)^2 - p_x^2 - p_y^2}$$

with $\frac{P_t - P_0}{P_0} \equiv \delta$
Note that the Hamiltonian is non-linear even in the absence of any field component (i.e. for a drift)!

Last approximation: transverse momenta (rescaled to the reference momentum) are considered to be much smaller than 1, i.e. the square root can be expanded. Considering also the large machine approximation \( x << \rho \), (dropping cubic terms), the Hamiltonian is simplified to

\[
\mathcal{H} = \frac{p_x^2 + p_y^2}{2(1 + \delta)} - \frac{x(1 + \delta)}{\rho(s)} - e\hat{A}_s
\]

This expansion may not be a good idea, especially for low energy, small size rings.
Considering the general expression of the longitudinal component of the vector potential is (see appendix)

- In curvilinear coordinates (curved elements)
  \[ A_s = (1 + \frac{x}{\rho(s)})B_0 \Re \sum_{n=0}^{\infty} \frac{b_n + ia_n}{n+1} (x + iy)^{n+1} \]

- In Cartesian coordinates
  \[ A_s = B_0 \Re \sum_{n=0}^{\infty} \frac{b_n + ia_n}{n+1} (x + iy)^{n+1} \]

with the multipole coefficients being written as

\[ a_n = \frac{1}{B_0 n!} \frac{\partial^n B_x}{\partial x^n} \bigg|_{x=y=0} \quad \text{and} \quad b_n = \frac{1}{B_0 n!} \frac{\partial^n B_y}{\partial x^n} \bigg|_{x=y=0} \]

The general non-linear Hamiltonian can be written as

\[ \mathcal{H}(x, y, p_x, p_y, s) = \mathcal{H}_0(x, y, p_x, p_y, s) + \sum_{k_x, k_y} h_{k_x, k_y}(s) x^{k_x} y^{k_y} \]

with the periodic functions

\[ h_{k_x, k_y}(s) = h_{k_x, k_y}(s + C) \]
Magnetic element Hamiltonians

- **Dipole:**
  \[ H = \frac{x\delta}{\rho} + \frac{x^2}{2\rho^2} + \frac{p_x^2 + p_y^2}{2(1 + \delta)} \]

- **Quadrupole:**
  \[ H = \frac{1}{2} k_1 (x^2 - y^2) + \frac{p_x^2 + p_y^2}{2(1 + \delta)} \]

- **Sextupole:**
  \[ H = \frac{1}{3} k_2 (x^3 - 3xy^2) + \frac{p_x^2 + p_y^2}{2(1 + \delta)} \]

- **Octupole:**
  \[ H = \frac{1}{4} k_3 (x^4 - 6x^2y^2 + y^4) + \frac{p_x^2 + p_y^2}{2(1 + \delta)} \]
Assume a simple case of linear transverse magnetic fields,

\[ B_x = b_1(s)y \]
\[ B_y = -b_0(s) + b_1(s)x \]

- main bending field
- normalized quadrupole gradient
- magnetic rigidity

\[
-B_0 \equiv b_0(s) = \frac{P_0 c}{e \rho(s)} \quad [T]
\]

\[
K(s) = b_1(s) \frac{e}{c P_0} = \frac{b_1(s)}{B \rho} \quad [1/m^2]
\]

\[
B \rho = \frac{P_0 c}{e} \quad [T \cdot m]
\]
Assume a simple case of linear transverse magnetic fields,

\begin{align*}
B_x &= b_1(s)y \\
B_y &= -b_0(s) + b_1(s)x
\end{align*}

- **main bending field** \( -B_0 \equiv b_0(s) = \frac{P_0c}{e\rho(s)} \) [T]
- **normalized quadrupole gradient** \( K(s) = b_1(s) \frac{e}{cP_0} = \frac{b_1(s)}{B\rho} \) [1/m²]
- **magnetic rigidity** \( B\rho = \frac{P_0c}{e} \) [T · m]

The vector potential has only a longitudinal component which in curvilinear coordinates is

\begin{align*}
B_x &= -\frac{1}{1+\frac{x}{\rho(s)}} \frac{\partial A_s}{\partial y}, \quad B_y = \frac{1}{1+\frac{x}{\rho(s)}} \frac{\partial A_s}{\partial x}
\end{align*}

The previous expressions can be integrated to give

\begin{align*}
A_s(x, y, s) &= \frac{P_0c}{e} \left[ -\frac{x}{\rho(s)} - \left( \frac{1}{\rho(s)^2} + K(s) \right) \frac{x^2}{2} + K(s) \frac{y^2}{2} \right] = P_0c \hat{A}_s(x, y, s)
\end{align*}
The integrable Hamiltonian

The Hamiltonian for linear fields can be finally written as

\[ \mathcal{H} = \frac{p_x^2 + p_y^2}{2(1+\delta)} - \frac{x\delta}{\rho(s)} + \frac{x^2}{2\rho(s)^2} + \frac{K(s)}{2} (x^2 - y^2) \]

Hamilton’s equation are

\[ \frac{dx}{ds} = \frac{p_x}{1+\delta}, \quad \frac{dp_x}{ds} = \frac{\delta}{\rho(s)} - \left( \frac{1}{\rho^2(s)} + K(s) \right) x \]

\[ \frac{dy}{ds} = \frac{p_y}{1+\delta}, \quad \frac{dp_y}{ds} = K(s) y \]

and they can be written as two second order uncoupled differential equations, i.e. Hill’s equations

\[ x'' + \frac{1}{1+\delta} \left( \frac{1}{\rho(s)^2} + K(s) \right) x = \frac{\delta}{\rho(s)} \]

\[ y'' - \frac{1}{1+\delta} K(s) y = 0 \]

with the usual solution for \( \delta = 0 \) and \( u = x, y \)

\[ u(s) = \sqrt{\epsilon / \beta(s)} \cos(\psi(s) + \psi_0) \]

\[ u'(s) = \sqrt{\frac{\epsilon}{\beta(s)}} (\sin(\psi(s) + \psi_0) + \alpha(s) \cos(\psi(s) + \psi_0)) \]
Action-angle variables

- There is a canonical transformation to some optimal set of variables which can simplify the phase-space motion.
- This set of variables are the action-angle variables.
- The action vector is defined as the integral \( J = \int p \, dq \) over closed paths in phase space.
- An integrable Hamiltonian is written as a function of only the actions, i.e. \( H_0 = H_0(J) \). Hamilton’s equations give
  \[
  \dot{\phi}_i = \frac{\partial H_0(J)}{\partial J_i} = \omega_i(J) \Rightarrow \phi_i = \omega_i(J)t + \phi_{i0}
  \]
  \[
  \dot{J}_i = -\frac{\partial H_0(J)}{\partial \phi_i} = 0 \Rightarrow J_i = \text{const.}
  \]
  i.e. the actions are integrals of motion and the angles are evolving linearly with time, with constant frequencies which depend on the actions.
- The actions define the surface of an invariant torus, topologically equivalent to the product of \( n \) circles.
Considering on-momentum motion, the Hamiltonian can be written as

$$\mathcal{H} = \frac{p_x^2 + p_y^2}{2} + \frac{K_x(s)x^2 - K_y(s)y^2}{2}$$

The generating function from the original to action angle variables is

$$F_1(x, y, \phi_x, \phi_y; s) = -\frac{x^2}{2\beta_x(s)} [\tan \phi_x(s) + a_x(s)] - \frac{y^2}{2\beta_y(s)} [\tan \phi_y(s) + a_y(s)]$$
Accelerator Hamiltonian in action-angle variables

Considering on-momentum motion, the Hamiltonian can be written as

$$\mathcal{H} = \frac{p_x^2 + p_y^2}{2} + \frac{K_x(s)x^2 - K_y(s)y^2}{2}$$

The generating function from the original to action angle variables is

$$F_1(x, y, \phi_x, \phi_y; s) = -\frac{x^2}{2\beta_x(s)} \left[ \tan \phi_x(s) + a_x(s) \right] - \frac{y^2}{2\beta_y(s)} \left[ \tan \phi_y(s) + a_y(s) \right]$$

The old variables with respect to actions and angles are

$$u(s) = \sqrt{2\beta_u(s)} J_u \cos \phi_u(s), \quad p_u(s) = -\sqrt{\frac{2J_u}{\beta_u(s)}} \left( \sin \phi_u(s) + \alpha_u(s) \cos \phi_u(s) \right)$$

and the Hamiltonian takes the form

$$\mathcal{H}_0(J_x, J_y, s) = \frac{J_x}{\beta_x(s)} + \frac{J_y}{\beta_y(s)}$$

The “time” (longitudinal position) dependence can be eliminated by the transformation to normalized coordinate

$$\begin{pmatrix} u \\ u' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\beta}} & 0 \\ \frac{\alpha}{\sqrt{\beta}} & \sqrt{\beta} \end{pmatrix} \begin{pmatrix} u \\ u' \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} U \\ U' \end{pmatrix} = \sqrt{2J} \begin{pmatrix} \cos(\nu \phi) \\ \sin(\nu \phi) \end{pmatrix}$$

with \( \nu = \frac{1}{2\pi} \int \frac{du}{\beta(s)} \).
Content of lecture II

- Canonical perturbation theory and its limitations
- Lie formalism for building maps
- Symplectic integration
- Normal forms for non-linear systems
Canonical perturbation theory
Canonical perturbation theory

- Consider a general Hamiltonian with \( n \) degrees of freedom

\[
H(J, \varphi, \theta) = H_0(J) + \epsilon H_1(J, \varphi, \theta) + O(\epsilon^2)
\]

where the non-integrable part \( H_1(J, \varphi, \theta) \) is \( 2\pi \)-periodic on the angles \( \varphi \) and the “time” \( \theta \)

- Provided that \( \epsilon \) is sufficiently small, tori should still exist but they are distorted

- We seek a canonical transformation that could “straighten up” the tori, i.e. it could transform the non-integrable part of the Hamiltonian (at first order in \( \epsilon \)) to a function only of some new actions \( \bar{H}(\bar{J}) \) plus higher orders in \( \epsilon \)

- This can be performed by a mixed variable close to identity generating function

\[
S(\bar{J}, \varphi, \theta) = \bar{J} \cdot \varphi + \epsilon S_1(\bar{J}, \varphi, \theta) + O(\epsilon^2)
\]

for transforming old variables to new ones \( (\bar{J}, \bar{\varphi}) \)

- In principle, this procedure can be carried to arbitrary powers of the perturbation
By the canonical transformation equations (slide 19), the old action and new angle can be also represented by a power series in $\epsilon$

$$J = \bar{J} + \epsilon \frac{\partial S_1(\bar{J}, \varphi, \theta)}{\partial \varphi} + \mathcal{O}(\epsilon^2) \quad J = \bar{J} + \epsilon \frac{\partial S_1(\bar{J}, \bar{\varphi}, \theta)}{\partial \bar{\varphi}} + \mathcal{O}(\epsilon^2)$$

$$\bar{\varphi} = \varphi + \epsilon \frac{\partial S_1(\bar{J}, \varphi, \theta)}{\partial \bar{J}} + \mathcal{O}(\epsilon^2) \quad \text{or} \quad \varphi = \bar{\varphi} - \epsilon \frac{\partial S_1(\bar{J}, \bar{\varphi}, \theta)}{\partial \bar{J}} + \mathcal{O}(\epsilon^2)$$
By the canonical transformation equations (slide 19), the old action and new angle can be also represented by a power series in $\epsilon$

$$J = \bar{J} + \epsilon \frac{\partial S_1(\bar{J}, \varphi, \theta)}{\partial \varphi} + O(\epsilon^2) \quad \bar{J} = J + \epsilon \frac{\partial S_1(\bar{J}, \bar{\varphi}, \theta)}{\partial \bar{\varphi}} + O(\epsilon^2)$$

$$\bar{\varphi} = \varphi + \epsilon \frac{\partial S_1(\bar{J}, \varphi, \theta)}{\partial \bar{J}} + O(\epsilon^2) \quad \text{or} \quad \varphi = \bar{\varphi} - \epsilon \frac{\partial S_1(\bar{J}, \bar{\varphi}, \theta)}{\partial \bar{J}} + O(\epsilon^2)$$

The previous equations expressing the old as a function of the new variables assume that there is possibility to invert the equation on the left, so that $S_1(\bar{J}, \bar{\varphi}, \theta)$ becomes a function of the new variables.

The new Hamiltonian is then

$$\bar{H}(\bar{J}, \bar{\varphi}, \theta) = H(J(\bar{J}, \bar{\varphi}), \varphi(\bar{J}, \bar{\varphi}), \theta) + \epsilon \frac{\partial S_1(\bar{J}, \bar{\varphi}, \theta)}{\partial \theta} + O(\epsilon^2)$$

The second term is appearing because of the "time" dependence through $\theta$.
The question is what is the form of the \textit{generating function} that eliminates the angle dependence.

The procedure is cumbersome (see appendix for details), but here is the final result,

\[ S(\bar{\mathbf{J}}, \bar{\varphi}) = \bar{\mathbf{J}} \cdot \bar{\varphi} + \epsilon i \sum_{k \neq 0} \frac{H_{1k}(\bar{\mathbf{J}})}{k \cdot \omega(\bar{\mathbf{J}}) + p} e^{i(k \cdot \bar{\varphi} + p \theta)} + \mathcal{O}(\epsilon^2) \]

with the frequency vector \( \omega(\bar{\mathbf{J}}) = \frac{\partial H_0(\bar{\mathbf{J}})}{\partial \bar{\mathbf{J}}} \)

and the integers \( k, p \neq 0 \)

If the denominator vanishes, i.e. for the \textit{resonance condition} \( k \cdot \omega(\bar{\mathbf{J}}) + p = 0 \), the Fourier series coefficients (\textit{driving terms}) become \textit{infinite}.

It actually implies that even at \textit{first order} in the perturbation parameter and in the vicinity of a resonance, it is \textit{impossible} to construct a \textit{generating function} for seeking some \textit{approximate integrals of motion}. 
In principle, the technique works for arbitrary order, but the disentangling of variables becomes difficult even to 2nd order!!!

The solution was given in the late 60s by introducing the Lie transforms (e.g. see Deprit 1969), which are algorithmic for constructing generating functions and were adapted to beam dynamics by Dragt and Finn (1976).

On the other hand, the problem of small denominators due to resonances is not just a mathematical one. The inability to construct solutions close to a resonance has to do with the unpredictable nature of motion and the onset of chaos.

KAM theory (see appendix) developed the mathematical framework into which local solutions could be constructed provided some general conditions on the size of the perturbation and the distance of the system from resonances are satisfied.

Very difficult though to apply directly this theorem to realistic physical systems, such as a particle accelerator.
General resonance conditions

- The general resonance conditions is \( n_x \nu_x + n_y \nu_y = p \)
  with order \( n_x + n_y \)
- For all the polynomial field terms of a 2\( m \)-pole, the excited resonances (at first order) satisfy the condition \( n_x + n_y = m \)
  but there are also sub-resonances for which \( n_x + n_y < m \)
- For normal (erect) multi-poles, the resonances (at first order) are \((n_x, n_y) = (m, 0), (m - 2, \pm 2), \ldots\)
  whereas for skew multi-poles \((n_x, n_y) = (m - 1, \pm 1), (m - 3, \pm 3), \ldots\)
- If perturbation is large, all resonances can be potentially excited
- The resonance conditions form lines in frequency space and fill it up as the order grows (the rational numbers form a dense set inside the real numbers), but Fourier amplitudes should also decrease
If lattice is made out of $N$ identical cells, and the perturbation follows the same periodicity, resulting in a reduction of the resonance conditions to the ones satisfying

$$n_x \nu_x + n_y \nu_y = jN$$

These are called **systematic** resonances.

Practically, any (linear) lattice perturbation breaks super-periodicity and any **random** resonance can be excited.

Careful choice of the working point is necessary.
Lie Transformations
Symplectic maps

- Consider two sets of canonical variables $\mathbf{z}, \mathbf{\bar{z}}$ which may be even considered as the evolution of the system between two points in phase space.

- A transformation from the one to the other set can be constructed through a map $\mathcal{M} : \mathbf{z} \mapsto \mathbf{\bar{z}}$.

- The Jacobian matrix of the map $\mathcal{M} = \mathcal{M}(\mathbf{z}, t)$ is composed by the elements $M_{ij} \equiv \frac{\partial \bar{z}_i}{\partial z_j}$.

- The map is symplectic if $\mathcal{M}^T J \mathcal{M} = J$ where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

- It can be shown that $\det(M) = 1$.

- It can be shown that the variables defined through a symplectic map $[\bar{z}_i, \bar{z}_j] = [z_i, z_j] = J_{ij}$ which is a known relation satisfied by canonical variables.

- In other words, symplectic maps preserve Poisson brackets.
Are Taylor maps symplectic?

- To test the **symplecticity** of Taylor maps, we have to construct the Jacobian matrix with elements \( M_{ij} \equiv \frac{\partial \bar{z}_i}{\partial z_j} \).

- The “thick” sextupole **Taylor map**, is written

\[
\begin{align*}
x_2 &= x_1 + Lx'_1 - k_2 \left( \frac{L^2}{4} (x^2_1 - y^2_1) + \frac{L^3}{12} (x_1 x'_1 - y_1 y'_1) + \frac{L^4}{24} (x'_1 - y'_1)^2 \right) \\
x'_2 &= x'_1 - k_2 \left( \frac{L^2}{2} (x^2_1 - y^2_1) + \frac{L^4}{4} (x_1 x'_1 - y_1 y'_1) + \frac{L^3}{6} (x'_1 - y'_1)^2 \right) \\
y_2 &= y_1 + Ly'_1 + k_2 \left( \frac{L^2}{4} x_1 y_1 + \frac{L^3}{12} (x_1 y'_1 + y_1 x'_1) + \frac{L^4}{24} (x'_1 y'_1) \right) \\
y'_2 &= y'_1 + k_2 \left( \frac{L^2}{2} x_1 y_1 + \frac{L^2}{4} (x_1 y'_1 + y_1 x'_1) + \frac{L^3}{6} (x'_1 y'_1) \right)
\end{align*}
\]

- All the coefficients of the Jacobian depend on initial conditions, e.g.

\[
\frac{\partial y_2}{\partial y_1} = 1 + k_2 \left( \frac{L^2}{4} x_1 + \frac{L^3}{12} x'_1 \right)
\]

and unless appropriately chosen they cannot satisfy \( \det(M) = 1 \).

- In general, Taylor maps are **not-symplectic!**
The Poisson bracket properties satisfy what is mathematically called a \textbf{Lie} algebra

They can be represented by (Lie) operators of the form:

\[ f : g = [f, g] \quad \text{and} \quad f : 2g = [f, [f, g]] \quad \text{etc.} \]
The Poisson bracket properties satisfy what is mathematically called a Lie algebra.

They can be represented by (Lie) operators of the form:

\[ f : g = [f, g] \quad \text{and} \quad f : 2g = [f, [f, g]] \quad \text{etc.} \]

For a Hamiltonian system \( H(z, t) \) there is a formal solution of the equations of motion

\[
\frac{dz}{dt} = [H, z] =: H : z
\]

written as

\[
z(t) = \sum_{k=0}^{\infty} \frac{t^k H : k}{k!} z_0 = e^{t : H} : z_0
\]

with a symplectic map

\[ M = e^{H : k = 0} \]
The Poisson bracket properties satisfy what is mathematically called a **Lie** algebra.

They can be represented by (Lie) operators of the form:

\[ f : g = [f, g] \quad \text{and} \quad f : g^2 = [f, [f, g]] \quad \text{etc.} \]

For a Hamiltonian system \( H(z, t) \), there is a **formal solution** of the equations of motion:

\[
\frac{dz}{dt} = [H, z] = :H : z
\]

written as \( z(t) = \sum_{k=0}^{\infty} \frac{t^k : H : k}{k!} z_0 = e^{t : H :} z_0 \) with a symplectic map \( M = e^{t : H :} \).

The 1-turn accelerator map can be represented by the composition of the maps of each element:

\[
M = e^{f_2 :} e^{f_3 :} e^{f_4 :} \ldots \quad \text{where} \quad f_i \quad \text{(called the generator) is the Hamiltonian for each element, a polynomial of degree} \quad m \quad \text{in the variables} \quad z_1, \ldots, z_n
\]
<table>
<thead>
<tr>
<th>Element</th>
<th>Map</th>
<th>Lie Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drift space</td>
<td>$x = x_0 + Lp_0$</td>
<td>$\exp(: - \frac{1}{2}Lp^2:)$</td>
</tr>
<tr>
<td></td>
<td>$p = p_0$</td>
<td></td>
</tr>
<tr>
<td>Thin-lens Quadrupole</td>
<td>$x = x_0$</td>
<td>$\exp(: - \frac{1}{2f}x^2:)$</td>
</tr>
<tr>
<td></td>
<td>$p = p_0 - \frac{1}{f}x_0$</td>
<td></td>
</tr>
<tr>
<td>Thin-lens Multipole</td>
<td>$x = x_0$</td>
<td>$\exp(: \lambda x^n:)$</td>
</tr>
<tr>
<td></td>
<td>$p = p_0 + \lambda n x^{n-1}$</td>
<td></td>
</tr>
<tr>
<td>Thin-lens kick</td>
<td>$x = x_0$</td>
<td>$\exp(: \int_0^x f(x')dx':)$</td>
</tr>
<tr>
<td></td>
<td>$p = p_0 + f(x)$</td>
<td></td>
</tr>
<tr>
<td>Thick focusing quad</td>
<td>$x = x_0 \cos kL + \frac{p_0}{k} \sin kL$</td>
<td>$\exp:[ - \frac{1}{2}L(k^2x^2 + p^2):]$</td>
</tr>
<tr>
<td></td>
<td>$p = -kx_0 \sin kL + p_0 \cos kL$</td>
<td></td>
</tr>
<tr>
<td>Thick defocusing quad</td>
<td>$x = x_0 \cosh kL + \frac{p_0}{k} \sinh kL$</td>
<td>$\exp:[\frac{1}{2}L(k^2x^2 - p^2):]$</td>
</tr>
<tr>
<td></td>
<td>$p = kx_0 \sinh kL + p_0 \cosh kL$</td>
<td></td>
</tr>
<tr>
<td>Coordinate shift</td>
<td>$x = x_0 - b$</td>
<td>$\exp( : ax + bp:')$</td>
</tr>
<tr>
<td></td>
<td>$p = p_0 + a$</td>
<td></td>
</tr>
<tr>
<td>Coordinate rotation</td>
<td>$x = x_0 \cos \mu + p_0 \sin \mu$</td>
<td>$\exp:[ - \frac{1}{2}\mu(x^2 + p^2):]$</td>
</tr>
<tr>
<td></td>
<td>$p = -x_0 \sin \mu + p_0 \cos \mu$</td>
<td></td>
</tr>
<tr>
<td>Scale change</td>
<td>$x = e^{-\lambda}x_0$</td>
<td>$\exp( : \lambda xp:')$</td>
</tr>
<tr>
<td></td>
<td>$p = e^\lambda p_0$</td>
<td></td>
</tr>
</tbody>
</table>
Formulas for Lie operators

\[ :a: = 0, \quad e^{a} : = 1 \]
\[ :f:a = 0, \quad e^{f} : a = a \]
\[ :f:f = 0, \quad e^{f} : f = f \]
\[ \{ :f:, :g: \} = : [f, g] : \]
\[ e^{f} : g(X) = g(e^{f} : X) \]
\[ e^{f} : G(:g:)e^{-f} : = G(:e^{f} : g:) \]
Consider the 1D quadrupole Hamiltonian

\[ H = \frac{1}{2} (k_1 x^2 + p^2) \]

For a quadrupole of length \( L \), the map is written as

\[ e^{L/2} : (k_1 x^2 + p^2) : \]
Consider the 1D quadrupole Hamiltonian

\[ H = \frac{1}{2} (k_1 x^2 + p^2) \]

For a quadrupole of length \( L \), the map is written as

\[ e^{\frac{L}{2} : (k_1 x^2 + p^2) :} \]

Its application to the transverse variables is

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2) :} x = \sum_{n=0}^{\infty} \left( \frac{(-k_1 L^2)^n}{(2n)!} x + L \frac{(-k_1 L^2)^n}{(2n + 1)!} p \right) \]

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2) :} p = \sum_{n=0}^{\infty} \left( \frac{(-k_1 L^2)^n}{(2n)!} p - \sqrt{k_1} \frac{(-k_1 L^2)^n}{(2n + 1)!} p \right) \]
Map for quadrupole

Consider the 1D quadrupole Hamiltonian

\[ H = \frac{1}{2} (k_1 x^2 + p^2) \]

For a quadrupole of length \( L \), the map is written as

\[ e^{\frac{L}{2} : (k_1 x^2 + p^2)} : \]

Its application to the transverse variables is

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2)} : x = \sum_{n=0}^{\infty} \left( \frac{(-k_1 L^2)^n}{(2n)!} x + L \frac{(-k_1 L^2)^n}{(2n + 1)!} p \right) \]

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2)} : p = \sum_{n=0}^{\infty} \left( \frac{(-k_1 L^2)^n}{(2n)!} p - \sqrt{k_1} \frac{(-k_1 L^2)^n}{(2n + 1)!} p \right) \]

This finally provides the usual quadrupole matrix

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2)} : x = \cos(\sqrt{k_1} L)x + \frac{1}{\sqrt{k_1}} \sin(\sqrt{k_1} L)p \]

\[ e^{-\frac{L}{2} : (k_1 x^2 + p^2)} : p = -\sqrt{k_1} \sin(\sqrt{k_1} L)x + \cos(\sqrt{k_1} L)p \]
Map for general monomial

- Consider a monomial in the positions and momenta $x^n p^m$
- The map is written as $e^{\alpha x^n p^m}$:
- Its application to the transverse variables is
  - For $n \neq m$
    \[
    e^{\alpha x^n p^m} : x = x \left[ 1 + \alpha(n - m)x^{n-1}p^{m-1} \right] \frac{m}{m-n} \\
    e^{\alpha x^n p^m} : p = p \left[ 1 + \alpha(n - m)x^{n-1}p^{m-1} \right] \frac{n}{n-m}
    \]
  - For $n = m$
    \[
    e^{\alpha x^n p^n} : x = xe^{-\alpha nx^{n-1} p^{n-1}} \\
    e^{\alpha x^n p^n} : p = pe^{\alpha nx^{n-1} p^{n-1}}
    \]
Map Concatenation

For combining together the different maps, the **Campbell-Baker-Hausdorff** formula can be used. It states that for $t_1, t_2$ sufficiently small, and $A, B$ real matrices, there is a real matrix $C$ for which

$$e^{t_1 A} e^{t_2 B} = e^C$$

For map composition through Lie operators, this is translated to

$$e^{h} = e^{f} e^{g}$$

with

$$h = f + g + \frac{1}{2} : f : g + \frac{1}{12} : f^2 : g + \frac{1}{12} : g^2 : f + \frac{1}{24} : f : g : f - \frac{1}{720} : g^4 f - \frac{1}{720} : f^4 g + \ldots$$

or

$$h = f + g + \frac{1}{2} [f, g] + \frac{1}{12} [f, [f, g]] + \frac{1}{12} [g, [g, f]] + \frac{1}{24} [f, [g, [g, f]]] - \frac{1}{720} [g, [g, [g, f]]] - \frac{1}{720} [f, [f, [f, g]]] + \ldots$$

i.e. a series of Poisson bracket operations.

Note that the **full map** is by “construction” symplectic.

By truncating the **map** to a certain order, symplecticity is lost.
The Campbell-Baker-Hausdorff formula for Lie maps has another useful form, depending if the summation is done over one or the other function

\[ e^{[f]} e^{g} = e^{g} + \left( \frac{e^{g} - 1}{e^{g} - 1} f \right) + \mathcal{O}(f^2) \]

or

\[ e^{[f]} e^{g} = e^{f} + \left( \frac{f}{1 - e^{-f} g} \right) + \mathcal{O}(g^2) \]
Example: The $-I$ transformer

- Consider two identical sextupoles separated by a beam line represented by a map $\mathcal{R}$.

- The **sextupole map** can be represented at **second order** as:

$$S_2 = e^{-\frac{1}{2} L_s : H_d : e} - L_s : H_s : e - \frac{1}{2} L_s : H_d : e$$

with the **sextupole effective Hamiltonian** $H_s = \frac{1}{6} k_2 (x^3 - 3xy^2)$ and $H_d$ the **drift Hamiltonian**.

\[\begin{align*}
S & \quad \mathcal{R} \quad S \\
S & \quad \approx \quad S_2
\end{align*}\]
Consider two identical sextupoles separated by a beam line represented by a map $\mathcal{R}$.

The sextupole map can be represented at second order as

$$S_2 = e^{-\frac{1}{2}L_s:H_d:}e^{-L_s:H_s:}e^{-\frac{1}{2}L_s:H_d:}$$

with the sextupole effective Hamiltonian $H_s = \frac{1}{6}k_2(x^3 - 3xy^2)$ and $H_d$ the drift Hamiltonian.

The total map can be approximated at 2$^{nd}$ order by

$$\mathcal{M} = SRS \approx S_2 R S_2 = e^{-\frac{1}{2}L_s:H_d:}e^{-L_s:H_s:}\mathcal{R}e^{-L_s:H_s:}e^{-\frac{1}{2}L_s:H_d:}$$

with the map $\mathcal{R} = e^{-\frac{1}{2}L_s:H_d:}\mathcal{R}e^{-\frac{1}{2}L_s:H_d:}$.

$$S \approx S_2$$
Example: The $-I$ transformer

- Inserting the identity $\bar{R}\bar{R}^{-1} = I$, we have
  \[ \mathcal{M} \approx e^{-\frac{1}{2} L_s : H_d : \bar{R}\bar{R}^{-1} e^{-L_s : H_s : \bar{R}} e^{-L_s : H_s : e^{-\frac{1}{2} L_s : H_d :}} \]

- The similarity transformation can be used
  \[ \bar{R}^{-1} e^{-L_s : H_s : \bar{R}} = e^{-L_s : \bar{R}^{-1} H_s :} \]

- The map is then rewritten as
  \[ \mathcal{M} \approx e^{-\frac{1}{2} L_s : H_d : \bar{R} e^{-L_s : \bar{R}^{-1} H_s :} e^{-L_s : H_s : e^{-\frac{1}{2} L_s : H_d :}}} \]
Example: The \(-I\) transformer

- Inserting the identity \(\mathcal{R}\mathcal{R}^{-1} = \mathcal{I}\), we have

\[ \mathcal{M} \approx e^{-\frac{1}{2} L_s : H_d : \mathcal{R}\mathcal{R}^{-1}} e^{-L_s : H_s : \mathcal{R}} e^{-L_s : H_s : e^{-\frac{1}{2} L_s : H_d :}} \]

- The similarity transformation can be used

\[ \mathcal{R}^{-1} e^{-L_s : H_s : \mathcal{R}} = e^{-L_s : \mathcal{R}^{-1} H_s :} \]

- The map is then rewritten as

\[ \mathcal{M} \approx e^{-\frac{1}{2} L_s : H_d : \mathcal{R}} e^{-L_s : \mathcal{R}^{-1} H_s : \mathcal{R}} e^{-L_s : H_s : \mathcal{R}} e^{-\frac{1}{2} L_s : H_d :} \]

- If the map \(\mathcal{R}\) is chosen such that \(-\mathcal{R}^{-1} H_s = H_s\) or \(\mathcal{R} H_s = -H_s\) so that

\[ e^{-L_s : \mathcal{R}^{-1} H_s :} e^{-L_s : H_s :} = e L_s : H_s : e^{-L_s : H_s :} = \mathcal{I} \]

- In that way, the sextupole non-linearity is getting

  **eliminated** in the final map

\[ \mathcal{M} \approx e^{-\frac{1}{2} L_s : H_d : \mathcal{R}} e^{-\frac{1}{2} L_s : H_d :} = e^{-L_s : H_d : \mathcal{R}} e^{-L_s : H_d :} \]
Example: The $-I$ transformer

- Inspecting the form of $H_s$ (odd in $x$ and even in $y$), this can be achieved if the map is such that

$$\bar{\mathbf{R}} x = -x, \quad \bar{\mathbf{R}} p_x = -p_x, \quad \bar{\mathbf{R}} y = \pm y, \quad \bar{\mathbf{R}} p_y = \pm p_y$$

- In matrix form this can be written as

$$\bar{\mathbf{R}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & \pm 1 \end{pmatrix} = \begin{pmatrix} \cos \mu_x + a_x \sin \mu_x & b_x \sin \mu_x & 0 & 0 \\ -c_x \sin \mu_x & \cos \mu_x - a_x \sin \mu_x & 0 & 0 \\ 0 & 0 & \cos \mu_y + a_y \sin \mu_y & b_y \sin \mu_y \\ 0 & 0 & -c_y \sin \mu_y & \cos \mu_y - a_y \sin \mu_y \end{pmatrix}$$

- The horizontal part of the matrix is $-I_2$ and the vertical part is $\pm I_2$, which is obtained for phase advances

$$\mu_x = (2n_x + 1)\pi, \quad \mu_y = n_y \pi$$

- This is why this beam line is called a $-I$-transformer
Symplectic integration
Why symplecticity is important

- **Symplecticity** guarantees that the transformations in phase space are **area preserving**

- To understand what deviation from symplecticity produces consider the simple case of the **quadrupole** with the general matrix written as

  \[
  \mathcal{M}_Q = \begin{pmatrix}
  \cos(\sqrt{k}L) & \frac{1}{\sqrt{k}} \sin(\sqrt{k}L) \\
  -\sqrt{k} \sin(\sqrt{k}L) & \cos(\sqrt{k}L)
  \end{pmatrix}
  \]

- Take the Taylor expansion for small lengths, up to first order

  \[
  \mathcal{M}_Q = \begin{pmatrix}
  1 & L \\
  -kL & 1
  \end{pmatrix} + O(L^2)
  \]

- This is indeed not symplectic as the determinant of the matrix is equal to \( 1 + kL^2 \), i.e. there is a deviation from symplecticity at 2\(^{\text{nd}}\) order in the quadrupole length
The iterated non-symplectic matrix does not provide the well-known elliptic trajectory in phase space.

Although the trajectory is very close to the original one, it **spirals outwards towards infinity**.
Symplecticity can be restored by adding “artificially” a correcting term to the matrix to become

$$\mathcal{M}_Q = \begin{pmatrix} 1 & L \\ -kL & 1-kL^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -kL & 1 \end{pmatrix} \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix}$$

In fact, the matrix now can be decomposed as a drift with a thin quadrupole at the end.

This representation, although not exact, produces an ellipse in phase space.
The same approach can be continued to 2\textsuperscript{nd} order of the Taylor map, by adding a 3\textsuperscript{rd} order correction.

\[ \mathcal{M}_Q = \begin{pmatrix} 1 - \frac{1}{2} kL^2 & L - \frac{1}{4} kL^3 \\ -kL & 1 - \frac{1}{2} kL^2 \end{pmatrix} = \begin{pmatrix} 1 & L/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -kL & 1 \end{pmatrix} \begin{pmatrix} 1 & L/2 \\ 0 & 1 \end{pmatrix} \]

The matrix now can be decomposed as two half drifts with a thin kick at the center.

This representation is even more exact as the error now is at 3\textsuperscript{rd} order in the length.
The idea is to distribute three kicks with different strengths so as to get a final map which is more accurate than the previous ones.

For the quadrupole, one can write

$$M_Q = \begin{pmatrix} 1 & d_1 L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_1 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_2 L/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_2 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_3 L/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_3 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_4 L/2 \\ 0 & 1 \end{pmatrix}$$

which imposes $\sum d_i = \sum c_i = 1$.

A symmetry condition of this form can be added $d_1 = d_4$, $d_2 = d_3$, $c_1 = c_3$. 
The idea is to distribute **three kicks with different strengths** so as to get a final map which is more accurate than the previous ones.

For the quadrupole, one can write

\[
\mathcal{M}_Q = \begin{pmatrix} 1 & d_1 L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_1 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_2 L/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_2 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_3 L/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c_3 kL & 1 \end{pmatrix} \begin{pmatrix} 1 & d_4 L/2 \\ 0 & 1 \end{pmatrix}
\]

which imposes \( \sum d_i = \sum c_i = 1 \).

A symmetry condition of this form can be added

\[d_1 = d_4 , \quad d_2 = d_3 , \quad c_1 = c_3\]

This provides the matrix \( \mathcal{M}_Q = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \) with

\[m_{11} = m_{22} = -\frac{1}{2} kL^2 + c_1 d_2 (d_1 + c_2) k^2 L^4 - d_1 d_2 c_1 c_2 k^3 L^6\]

\[m_{12} = L - \left(\frac{c_2}{4} + d_1 d_2 + 2d_1 d_2 c_1\right) kL^3 + 2d_1 d_2 c_1 (d_1 d_2 + \frac{c_2}{2}) k^2 L^5 + d_1^2 d_2^2 c_1^2 c_2 k^3 L^7\]

\[m_{21} = -kL + c_1 d_2 (1 + c_2) k^2 L^3 - d_2^2 c_1^2 c_2 k^3 L^5\]
3-kick symplectic integrator

By imposing that the determinant is 1, the following additional relations are obtained

\[ c_1 d_2 \left( d_1 + \frac{c_2}{2} \right) = \frac{1}{24} \]

\[ \frac{c_2}{4} + d_1 d_2 + 2d_1 d_2 c_1 = \frac{1}{6} \]

\[ c_1 d_2 (1 + c_2) = \frac{1}{6} \]
3-kick symplectic integrator

By imposing that the determinant is 1, the following additional relations are obtained

\[ c_1 d_2 \left( d_1 + \frac{c_2}{2} \right) = \frac{1}{24} \]

\[ \frac{c_2}{4} + d_1 d_2 + 2d_1 d_2 c_1 = \frac{1}{6} \]

\[ c_1 d_2 (1 + c_2) = \frac{1}{6} \]

Although these are 5 equations with 4 unknowns, solutions exist

\[ d_1 = d_4 = \frac{1}{2(2 - 2^{1/3})} \quad d_2 = d_3 = \frac{1 - 2^{1/3}}{2(2 - 2^{1/3})} \]

\[ c_1 = c_3 = \frac{1}{2 - 2^{1/3}} \quad c_2 = -\frac{2^{1/3}}{2 - 2^{1/3}} \]
By imposing that the determinant is 1, the following additional relations are obtained
\[ c_1 d_2 (d_1 + \frac{c_2}{2}) = \frac{1}{24} \]
\[ \frac{c_2}{4} + d_1 d_2 + 2d_1 d_2 c_1 = \frac{1}{6} \]
\[ c_1 d_2 (1 + c_2) = \frac{1}{6} \]

Although these are 5 equations with 4 unknowns, solutions exist
\[ d_1 = d_4 = \frac{1}{2(2 - 2^{1/3})} \]
\[ d_2 = d_3 = \frac{1 - 2^{1/3}}{2(2 - 2^{1/3})} \]
\[ c_1 = c_3 = \frac{1}{2 - 2^{1/3}} \]
\[ c_2 = -\frac{2^{1/3}}{2 - 2^{1/3}} \]

This is actually the famous 7 step 4th order symplectic integrator of Forest, Ruth and Yoshida (1990). It can be generalized for any non-linear element.

It imposes negative drifts...
Yoshida has proved that a **general integrator map** of order $2k$ can be used to build a **map of order** $2k + 2$

$$S_{2k+2}(t) = S_{2k}(x_1 t) \circ S_{2k}(x_0 t) \circ S_{2k}(x_1 t)$$

with

$$x_0 = \frac{-2^{\frac{1}{2k+1}}}{2 - 2^{\frac{1}{2k+1}}} \quad , \quad x_1 = \frac{1}{2 - 2^{\frac{1}{2k+1}}}$$
Higher order integrators

- Yoshida has proved that a general integrator map of order $2^k$ can be used to build a map of order $2^k + 2$
  
  $$S_{2k+2}(t) = S_{2k}(x_1 t) \circ S_{2k}(x_0 t) \circ S_{2k}(x_1 t)$$

  with  
  $$x_0 = \frac{-2^{\frac{1}{2k+1}}}{2 - 2^{\frac{1}{2k+1}}} , \quad x_1 = \frac{1}{2 - 2^{\frac{1}{2k+1}}}$$

- For example the 4th order scheme can be considered as a composition of three 2nd order ones (single kicks)
  
  $$S_4(t) = S_2(x_1 t) \circ S_2(x_0 t) \circ S_2(x_1 t)$$

  with  
  $$x_0 = \frac{-2^{\frac{1}{3}}}{2 - 2^{\frac{1}{3}}} , \quad x_1 = \frac{1}{2 - 2^{\frac{1}{3}}}$$
Yoshida has proved that a **general integrator map** of order \(2k\) can be used to built a **map of order** \(2k + 2\)
\[
S_{2k+2}(t) = S_{2k}(x_1 t) \circ S_{2k}(x_0 t) \circ S_{2k}(x_1 t)
\]
with \(x_0 = \frac{-2^{\frac{1}{2k+1}}}{2 - 2^{\frac{1}{2k+1}}}, \quad x_1 = \frac{1}{2 - 2^{\frac{1}{2k+1}}}\)

For example the 4\(^{th}\) order scheme can be considered as a **composition** of three 2\(^{nd}\) order ones (single kicks)
\[
S_4(t) = S_2(x_1 t) \circ S_2(x_0 t) \circ S_2(x_1 t)
\]
with \(x_0 = \frac{-2^{\frac{1}{3}}}{2 - 2^{\frac{1}{3}}}, \quad x_1 = \frac{1}{2 - 2^{\frac{1}{3}}}\)

A 6\(^{th}\) order integrator can be produced by three interleaved 4\(^{th}\) order ones (9 kicks)
\[
S_6(t) = S_4(x_1 t) \circ S_4(x_0 t) \circ S_4(x_1 t)
\]
with \(x_0 = \frac{-2^{\frac{1}{5}}}{2 - 2^{\frac{1}{5}}}, \quad x_1 = \frac{1}{2 - 2^{\frac{1}{5}}}\)
Modern symplectic integration schemes

- Symplectic integrators with **positive** steps for Hamiltonian systems  \( H = A + \varepsilon B \) with both \( A \) and \( B \) integrable were proposed by **McLachlan** (1995).

- **Laskar** and **Robutel** (2001) derived all orders of such integrators.

- Consider the formal solution of the Hamiltonian system written in the Lie representation

  \[
  \ddot{x}(t) = \sum_{n \geq 0} \frac{t^n}{n!} L^n_H \ddot{x}(0) = e^{tL_H} \ddot{x}(0).
  \]

- A symplectic integrator of order \( n \) from \( t \) to \( t + \tau \) consists of approximating the Lie map \( e^{\tau L_H} = e^{\tau (L_A + L_{\varepsilon B})} \) by products of \( e^{c_i \tau L_A} \) and \( e^{d_i \tau L_{\varepsilon B}} \), \( i = 1, \ldots, n \) which integrate exactly \( A \) and \( B \) over the time-spans \( c_i \tau \) and \( d_i \tau \).

- The constants \( c_i \) and \( d_i \) are chosen to reduce the error.
The SABA\(_2\) integrator is written as

\[
SABA_2 = e^{c_1 \tau L_A} e^{d_1 \tau L_B} e^{c_2 \tau L_A} e^{d_1 \tau L_B} e^{c_1 \tau L_A},
\]

with \(c_1 = \frac{1}{2} \left(1 - \frac{1}{\sqrt{3}} \right)\), \(c_2 = \frac{1}{\sqrt{3}}\), \(d_1 = \frac{1}{2}\).

When\(\{\{A, B\}, B\}\) is integrable, e.g. when \(A\) is quadratic in momenta and \(B\) depends only in positions, the accuracy of the integrator is improved by two small negative steps

\[
SABA_2 C = e^{-\tau^3 \epsilon^2 \frac{c}{2} L_{\{\{A,B\},B\}}} \left( SABA_2 \right) e^{-\tau^3 \epsilon^2 \frac{c}{2} L_{\{\{A,B\},B\}}}
\]

with \(c = (2 - \sqrt{3})/24\).

The accuracy of SABA\(_2\)C is one order of magnitude higher than the Forest-Ruth 4\(^{th}\) order scheme.

The usual “drift-kick” scheme corresponds to the 2\(^{nd}\) order inte

\[
SABA_1 = e^{\frac{\tau}{2} L_A} e^{\tau L_B} e^{\frac{\tau}{2} L_A},
\]
SABA$_2$C integrator

- C. Skokos, YP and J. Laskar, EPAC 2008

- SABA$_2$C allows symplectic integration with positive steps

- Several orders of magnitude better precision of SABA$_2$C with respect to classical YFR integrator
Normal forms
Normal forms consists of finding a canonical transformation of the 1-turn map, so that it becomes simpler to analyze.

In the linear case, the Floquet transformation is a kind a normal form as it turns ellipses into circles.
Normal forms

- Normal forms consists of finding a canonical transformation of the 1-turn map, so that it becomes simpler to analyze.
- In the linear case, the Floquet transformation is a kind a normal form as it turns ellipses into circles.
- The transformation can be written formally as:

\[
\begin{align*}
\Phi^{-1} & \quad \Phi^{-1} \\
\mathcal{M} & \quad \mathcal{N} \\
\mathcal{N} & = \Phi \circ \mathcal{M} \circ \Phi^{-1} = e^{\hbar_{eff} f}:
\end{align*}
\]
Normal forms

- Normal forms consists of finding a canonical transformation of the 1-turn map, so that it becomes simpler to analyze.

- In the linear case, the Floquet transformation is a kind of normal form as it turns ellipses into circles.

- The transformation can be written formally as

\[
\begin{align*}
\mathbf{z} & \xrightarrow{\mathcal{M}} \mathbf{z}' \\
\Phi^{-1} \downarrow & \quad \downarrow \Phi^{-1} \\
\mathbf{u} & \xrightarrow{\mathcal{N}} \mathbf{u}'
\end{align*}
\]

with the original map \( \mathcal{M} = \Phi^{-1} \circ \mathcal{N} \circ \Phi \) and its normal form

\[
\mathcal{N} = \Phi \circ \mathcal{M} \circ \Phi^{-1} = e^{i h_{\text{eff}}}
\]

- The transformation \( \Phi = e^{i F} \) is better suited in action angle variables, i.e.

\[
\zeta = e^{-i F} \mathbf{h} \quad \text{taking the system from the original action-angle } h_{x,y}^\pm = \sqrt{2J_{x,y}} e^{\mp i \phi_{x,y}} \text{ to a new set}
\]

\[
\zeta_{x,y}^\pm(N) = \sqrt{2I_{x,y}} e^{\mp i \psi_{x,y}(N)} \quad \text{with the angles being just simple rotations,}
\]

\[
\psi_{x,y}(N) = 2\pi N \nu_{x,y} + \psi_{x,y_0} \text{ and the new effective Hamiltonian depends only on the new actions}
\]
Effective Hamiltonian

The generating function can be written as a polynomial in the new actions, i.e.

\[ F_r = \sum_{jklm} f_{jklm} \zeta_x^j \zeta_x^{-k} \zeta_y^l \zeta_y^{-m} = f_{jklm} (2I_x)^{j+k/2} (2I_y)^{l+m/2} e^{-i\psi_{jklm}} \]

There are software tools that built this transformation

Once the “new” effective Hamiltonian is known, all interesting quantities can be derived

This Hamiltonian is a function only of the new actions, and to 3rd order it is obtained as

\[ h_{eff} = \nu_x I_x + \nu_y I_y \]

\[ + \frac{1}{2} \alpha_c \delta^2 + c_{x1} I_x \delta + c_{y1} I_y \delta + c_3 \delta^3 \]

\[ + c_{xx} I_x^2 + c_{xy} I_x I_y + c_{yy} I_y^2 + c_{x2} I_x \delta^2 + c_{y2} I_y \delta^2 c_4 \delta^4 \]
Effective Hamiltonian

The correction of the tunes is given by

\[ Q_x = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial I_x} = \frac{1}{2\pi} (\nu_x + 2c_{xx} I_x + c_{xy} I_y + c_{x1}\delta + c_{x2}\delta^2) \]

\[ Q_y = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial I_y} = \frac{1}{2\pi} (\nu_y + 2c_{yy} I_y + c_{xy} I_x + c_{y1}\delta + c_{y2}\delta^2) \]

The correction to the path length is

\[ \Delta s = \frac{\partial h_{\text{eff}}}{\partial \delta} = \alpha c_3 \delta + c_3 \delta^2 + 4c_4 \delta^3 + c_{x1} I_x + c_{y1} I_y + 2c_{x2} I_x \delta + 2c_{y2} I_y \delta \]

1st, 2nd and 3rd momentum compaction
Using the BCH formula, one can prove that the composition of two maps with \(g\) small can be written as (see slide 9)

\[
e^{\cdot f} e^{\cdot g} = \exp \left[ f + \left( \frac{\cdot f}{1 - e^{-\cdot f}} \right) g + \mathcal{O}(g^2) \right]
\]
Normal form for perturbation

Using the BCH formula, one can prove that the composition of two maps with \( g \) small can be written as (see slide 9)

\[
e^{f \cdot e^{-g \cdot f}} = \exp \left[ f + \left( \frac{f}{1 - e^{-f \cdot g}} \right) g + \mathcal{O}(g^2) \right]
\]

Consider a linear map (rotation) followed by a small perturbation

\[
\mathcal{M} = e^{f_2 \cdot f_3}:
\]

We are seeking a transformation such that

\[
\mathcal{N} = \Phi \mathcal{M} \Phi^{-1} = e^{F \cdot f_2 \cdot f_3} e^{-F}:
\]
Normal form for perturbation

Using the BCH formula, one can prove that the composition of two maps with $g$ small can be written as (see slide 9)

$$e^{f}e^{g} = \exp \left[ f + \left( \frac{f}{1 - e^{-f}} \right) g + \mathcal{O}(g^2) \right]$$

Consider a linear map (rotation) followed by a small perturbation $\mathcal{M} = e^{f_2}e^{f_3}$.

We are seeking for a transformation such that $\mathcal{N} = \Phi \mathcal{M} \Phi^{-1} = e^{F}e^{f_2}e^{f_3}e^{-F}$.

This can be written as

$$\mathcal{N} = e^{f_2}e^{-f_2}e^{F}e^{f_2}e^{f_3}e^{-F}$$

$$= e^{f_2}e^{(e^{-f_2} - 1)F + f_3} + \ldots$$

This will transform the new map to a rotation to leading order.

$$F = \frac{f_3}{1 - e^{-f_2}}.$$
Consider a linear map followed by an octupole
\[ \mathcal{M} = e^{-\frac{\nu}{2} : x^2 + p^2} : e^{\frac{x^4}{4}} : = e^{2 : e^{\frac{x^4}{4}}} : \]

The generating function has to be chosen such as to make the following expression simpler
\[ (e^{-2 : e^{\frac{x^4}{4}}} - 1) F + \frac{x^4}{4} \]
Example: Octupole

- Consider a linear map followed by an octupole

\[ \mathcal{M} = e^{-\frac{\nu}{2} : x^2} + p^2 \cdot e^{\frac{x^4}{4}} \]

- The generating function has to be chosen such as to make the following expression simpler

\[ (e^{-:f_2:} - 1) F + \frac{x^4}{4} \]

- The simplest expression is the one that the angles are eliminated and there is only dependence on the action
Consider a linear map followed by an octupole

\[ \mathcal{M} = e^{-\frac{\nu}{2} : x^2 + p^2 :} e^i \frac{x^4}{4} : = e^i f_2 : e^i \frac{x^4}{4} : \]

The **generating function** has to be chosen such as to make the following expression simpler

\[ (e^{-i f_2 :} - 1) F + \frac{x^4}{4} \]

The simplest expression is the one that the **angles** are **eliminated** and there is only dependence on the action

We pass to the **action angle variable** (resonance basis)

\[ h^\pm = \sqrt{2J} e^{\mp i \phi} = x \mp i p \]

The perturbation is

\[ x^4 = (h_+ + h_-)^4 = h^\pm = h_+^4 + 4h_+^3 h_- + 6h_+^2 h_-^2 + 4h_+ h_-^3 + h_-^4 \]

**Example: Octupole**

**Non-linear dynamics, CERN Accelerator School, June 2019**
The term $6h_+^2h_-^2 = 24J^2$ is independent on the angles. Thus we may choose the generating functions such that the other terms are eliminated. It takes the form

$$F = \frac{1}{16} \left( \frac{h_+^4}{1 - e^{4iv}} + \frac{4h_+^3h_-}{1 - e^{2iv}} + \frac{4h_+h_-^3}{1 - e^{2iv}} + \frac{h_-^4}{1 - e^{4iv}} \right)$$
The term \( 6h^2_+ h^2_- = 24J^2 \) is independent on the angles. Thus we may choose the generating functions such that the other terms are eliminated. It takes the form

\[
F = \frac{1}{16} \left( \frac{h^4_+}{1 - e^{4iv}} + \frac{4h^3_+ h_-}{1 - e^{2iv}} + \frac{4h_+ h^3_-}{1 - e^{2iv}} + \frac{h^4_-}{1 - e^{4iv}} \right)
\]

The map is now written as

\[
\mathcal{M} = e^{-:F:} e^{:\iota J + \frac{3}{8} J^2 + e^{:F:}}
\]

The new effective Hamiltonian is depending only on the actions and contains the tune-shift terms.
**Example: Octupole**

- The term $6h^2_+h^2_- = 24J^2$ is independent on the angles. Thus we may choose the generating functions such that the other terms are eliminated. It takes the form

$$F = \frac{1}{16} \left( \frac{h^4_+}{1 - e^{4i\nu}} + \frac{4h^3_+h_-}{1 - e^{2i\nu}} + \frac{4h_+h^3_-}{1 - e^{2i\nu}} + \frac{h^4_-}{1 - e^{4i\nu}} \right)$$

- The map is now written as

$$\mathcal{M} = e^{-iF} : e^{\nu J + \frac{3}{8} J^2} : e^{iF} :$$

- The new effective Hamiltonian is depending only on the actions and contains the tune-shift terms.

- The generator in the original variables is written as

$$F = -\frac{1}{64} \left[ -5x^4 + 3p^4 + 6x^2p^2 + 4x^3p(2\cot(\nu) + \cot(2\nu)) + 4xp^3(2\cot(\nu) - \cot(2\nu)) \right]$$

- Constant values of the generator describe the trajectories in phase space.
It is possible by constructing the one turn map to build the generating (sometimes called "distortion") function

\[ F_r \approx \sum_{jklm} f_{jklm} J_x^{\frac{j+k}{2}} J_y^{\frac{l+m}{2}} e^{-i\psi_{jklm}} \]

For any resonance \( a\nu_x + b\nu_y = c \), and setting \( \psi_{jklm} = 0 \), the associated part of the functions is

\[ F(a,b) \approx \sum_{\substack{jklm\ \ \mid\ \ j+k+l+m \leq n \\ j+k=a,\ l+m=b}} f_{jklm} J_x^{\frac{a}{2}} J_y^{\frac{b}{2}} \]
Normal forms for LHC models

- In the LHC at injection (450 GeV), beam stability is necessary over a very large number of turns ($10^7$).
- Stability is reduced from random multi-pole imperfections mainly in the super-conducting magnets.
- Area of stability (Dynamic aperture - DA) computed with particle tracking for a large number of random magnet error distributions.
- Numerical tool based on normal form analysis (GRR) permitted identification of DA reduction reason (errors in the “warm” quadrupoles).

### Table: Numerical Tool Results

<table>
<thead>
<tr>
<th>Phase</th>
<th>Type</th>
<th>DA ($\sigma$)</th>
<th>LHC Version</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>15°</td>
<td>Warm Quads switched ON</td>
<td>Average</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>8.5</td>
</tr>
<tr>
<td>15°</td>
<td>Warm Quads switched OFF</td>
<td>Average</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>9.1</td>
</tr>
<tr>
<td>45°</td>
<td>Warm Quads switched ON</td>
<td>Average</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>9.5</td>
</tr>
<tr>
<td>45°</td>
<td>Warm Quads switched OFF</td>
<td>Average</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>9.0</td>
</tr>
</tbody>
</table>

- With “warm” quad. errors
- Without “warm” quad. errors
Summary

- **Natural way** to represent motion in an accelerator is by using maps.
- **Powerful tools** to build them from straight-forward tracking (TPSA).
- **Canonical (symplectic) transformations** enable to move from variables describing a distorted phase space to something simpler (ideally circles).
- The **generating functions** passing from the old to the new variables are bounded to **diverge** in the vicinity of **resonances** (emergence of chaos, see Lectures of NLD Phenomenology).
- Calculating this generating function with **canonical perturbation theory** becomes **hopeless** for higher orders.
- **Lie transformations** of **accelerator maps** enables derivation of the generating functions in an **algorithmic way**, in principle to arbitrary order.
- For real accelerator models, we have to rely on **symplectic integration**, i.e. **particle tracking** and **methods** to analyse it (see Lectures of NLD Phenomenology).