

Selected Notes from the USPAS Winter 2008 Accelerator Physics Course

Introduction.

These notes are a sub-set of lectures presented in the Accelerator Physics course (USPAS, Santa Rosa, California, January 14-25, 2008) taught by Eduard Pozdeyev, Todd Satogata and myself.

These notes comprised of five files: 1) My lecture notes with focus on a complete Hamiltonian dynamics in accelerators without simplification of plane trajectories or absence of coupling; 2) Two lecture notes by Eduard Pozdeyev: “Bogolyubov and Metropolsky’s averaging method” and “Space charge effects in accelerators”; 3) Three lecture notes by Todd Satogata: “Transport matrices”, ‘Poisson Brackets and Lie Operators” and “Colliders and Beam-Beam Effects”; 4) Problems for home works and mid-term exam; 5) Solutions to selected problems.

A number of lectures with introductory character or been short reviews of complex subjects (such as instabilities or electron cloud) are omitted from these notes.

My notes are based on my mostly unpublished notes I wrote during period from 1976 to 1979 in Novosibirsk. Only few copies (in Russian) were distributed to my colleagues in Novosibirsk Institute of Nuclear Physics. The goal of these notes is a complete description starting from the arbitrary reference orbit, explicit expressions for 4-potential and accelerator Hamiltonian and finishing with parameterization with action and angle variables. To a large degree follow logic developed in *Theory of Cyclic Particle Accelerators* by A.A.Kolomensky and A.N.Lebedev [Kolomensky], but going beyond in a number of directions. One of unusual feature is these notes use of matrix function and Sylvester formula for calculating matrices of arbitrary elements.

Teaching this course motivated me to translate significant part of my notes into the English. I also included some introductory materials following *Classical Theory of Fields* by L.D. Landau and E.M. Lifshitz [Landau]. A large number of short notes covering various techniques are placed in the Appendices.

I never found time to turn these notes into a perfect text I would like to see, but decided not to delay any further and to release them in the present imperfect form. Neither I had time to find all necessary references. I ask readers to excuse awkward places, lack of references and, unfortunately, some remaining typos in formulae.

Copyrights for original material presented in these notes are reserved.

Vladimir N Litvinenko

September 1, 2010

Content

Page	
1. Lecture 1: Particles In Electromagnetic Fields ¹ .	3
2. Lecture 2. Accelerator Hamiltonian	19
3. Lectures 3 & 4. Linear systems, matrix functions & transport matrices	34
4. Lecture 5. Parameterization of motion, modes, amplitudes and phases...	50
5. Lecture 6. Action – phase variables, solving standard problems	56
6. Lecture 7. Synchrotron oscillations	67
7. Lecture 8. Effects of synchrotron radiation	70
Appendix A: 4-D metric of special relativity	78
Appendix B: Lorentz group	85
Appendix C: Vlasov equation for an ensemble of particles.	88
Appendix D: Recap of Lecture 2	89
Appendix E: Matrix functions and Projection operators	92
Appendix F: Inhomogeneous solution	99
Appendix G: Canonical variables – one more touch.	100
Appendix H: Fokker-Plank Equation	104
Appendix I: Standard perturbation method	109
Appendix J: Arbitrary order Hamiltonian expansion with Mathematica	110
References	116

¹ People fluent in the classical EM, theory, the relativistic, Lagrangian and Hamiltonian mechanics can skip this lecture.

Lecture 1: Particles In Electromagnetic Fields.

Fundamentals of Hamiltonian Mechanics

http://en.wikipedia.org/wiki/Hamilton_principle

1.0. Least-Action Principle and Hamiltonian Mechanics

Let us refresh our knowledge of some aspects of the **Least-Action Principle** (humorously termed the *coach potato principle*) and **Hamiltonian Mechanics**. The **Principle of Least Action** is the most general formulation of laws governing the motion (evolution) of systems of particles and fields in physics. In mechanics, it is known as the **Hamilton's Principle**, and states the following:

- 1) A mechanical system with n degrees of freedom is fully characterized by a monotonic generalized coordinate, t , the full set of n coordinates $q = \{q_1, q_2, q_3, \dots, q_n\}$ and their derivatives $\dot{q} = \{\dot{q}_1, \dot{q}_2, \dot{q}_3, \dots, \dot{q}_n\}$ that are denoted by dots above a letter. We study the dynamics of the system with respect to t . All the coordinates, $q = \{q_1, q_2, q_3, \dots, q_n\}$; $\dot{q} = \{\dot{q}_1, \dot{q}_2, \dot{q}_3, \dots, \dot{q}_n\}$ should be treated as a functions of t that itself should be treated as an independent variable.
- 2) Each mechanical system can be fully characterized by the **Action Integral**:

$$S(A,B) = \int_A^B L(q, \dot{q}, t) dt \quad (L1.1)$$

that is taken between two events A and B described by full set of coordinates $^*(q, t)$. The function under integral $L(q, \dot{q}, t)$ is called the system's **Lagrangian function**. Any system is fully described by its action integral.

After that, applying **Lagrangian mechanics involves just n second -order ordinary differential equations: $\ddot{q} = f(q, \dot{q})$.**

We can find these equations, setting variation of δS_{AB} to zero:

$$\begin{aligned} \delta S_{AB} &= \delta \left(\int_A^B L(q, \dot{q}, t) dt \right) = \int_A^B \left\{ \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right\} dt = \int_A^B \left\{ \frac{\partial L}{\partial q} \delta q dt + \frac{\partial L}{\partial \dot{q}} \delta dq \right\} = \\ & \left[\frac{\partial L}{\partial \dot{q}} \delta q \right]_A^B + \int_A^B \left\{ \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right\} \delta q dt = 0 \quad ; \quad (L1.2) \end{aligned}$$

and taking into account $\delta q(A) = \delta q(B) = 0$. Thus, we have integral of the function in the brackets, multiplied by an arbitrary function $\delta q(t)$ equals zero.

Therefore, we must conclude that the function in the brackets also equals zero and thus obtain **Lagrange's equations**:

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = 0. \quad (L1.3)$$

* For one particle, the full set of event coordinates is the time and location of the particle. The integral is taken along a particle's world line (its unique path through 4-dimensional space-time) and is a function of both the end points and the intervening trajectory.

Explicitly, this represents a set of n second-order equations

$$\frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i} = \frac{\partial L(q, \dot{q}, t)}{\partial q_i} \Leftrightarrow \sum_{j=1}^n \left(\ddot{q}_j \frac{\partial^2 L(q, \dot{q}, t)}{\partial \dot{q}_i \partial \dot{q}_j} + \dot{q}_j \frac{\partial^2 L(q, \dot{q}, t)}{\partial \dot{q}_i \partial q_j} \right) = \frac{\partial L(q, \dot{q}, t)}{\partial q_i}.$$

The partial derivative of the Lagrangian over \dot{q} is called generalized (**canonical**) momentum:

$$P^i = \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i}; \quad (L1.4)$$

and the partial derivative of the Lagrangian over q is called the generalized force:

$$f^i = \frac{\partial L(q, \dot{q}, t)}{\partial q_i} : (L1.4) \text{ can be rewritten in more familiar form: } \frac{dP^i}{dt} = f^i. \text{ Then, by a definition,$$

the energy (Hamiltonian) of the system is:

$$H = \sum_{i=1}^n P^i \dot{q}_i - L = \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L; L = \sum_{i=1}^n P^i \dot{q}_i - H. \quad (L1.5)$$

Even though the Lagrangian approach fully describes a mechanical system it has some significant limitations. It treats the coordinates and their derivatives differently, and allows only coordinate transformations $q' = q'(q, t)$. There is more powerful method, the **Hamiltonian or Canonical Method**. The Hamiltonian is considered as a function of coordinates and momenta, which are treated equally. Specifically, pairs of coordinates with their conjugate momenta (L1.4) (q_i, P_i) or (q^i, P_i) are called canonical pairs. The Hamiltonian method creates many links between classical and quantum theory wherein it becomes an operator. Before using the Hamiltonian, let us prove that it is really function of (q, P, t) : i.e., that the full differential of the Hamiltonian is

$$dH(q, P, t) = \sum_{i=1}^n \left(\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial P^i} dP^i \right) + \frac{\partial H}{\partial t} dt. \quad (L1.6)$$

Using equation (L1.5) explicitly, we can easily prove it:

$$\begin{aligned} dH &= d \sum_{i=1}^n \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - dL = \sum_{i=1}^n \left\{ \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \dot{q}_i d \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt \right\} = \\ &= \sum_{i=1}^n \left\{ \dot{q}_i d \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \right\} = \sum_{i=1}^n \left\{ \dot{q}_i dP^i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \right\} = \sum_{i=1}^n \left(\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial P^i} dP^i \right) + \frac{\partial H}{\partial t} dt. \end{aligned}$$

wherein we substitute $d(\partial L / \partial \dot{q}_i) = dP^i$ with the expression for generalized momentum. In addition to this proof, we find some ratios between the Hamiltonian and the Lagrangian:

$$\left. \frac{\partial H}{\partial q_i} \right|_{P=const} = - \left. \frac{\partial L}{\partial q_i} \right|_{\dot{q}=const}; \quad \left. \frac{\partial H}{\partial t} \right|_{P,q=const} = - \left. \frac{\partial L}{\partial t} \right|_{q,\dot{q}=const}; \quad \dot{q}_i = \frac{dq_i}{dt} = \frac{\partial H}{\partial P^i};$$

wherein we should very carefully and explicitly specify what type of partial derivative we use. For example, the Hamiltonian is function of (q, P, t) : thus, partial derivative on q must be taken with constant momentum and time. For the Lagrangian, we should keep $\dot{q}, t = const$ to partially differentiate on q .

The last ratio gives us the first Hamilton's equation, while the second one comes from Lagrange's equation (L1.5-11):

$$\dot{q}_i = \frac{dq_i}{dt} = \frac{\partial H}{\partial P^i};$$

$$\frac{dP^i}{dt} = \frac{d}{dt} \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i} = \frac{dP^i}{dt} = \frac{\partial L}{\partial q_i} \Big|_{\dot{q}=const} = - \frac{\partial H}{\partial q_i} \Big|_{P=const};$$
(L1.7)

both of which are given in compact form below in (L1.11).

Now, to state this in a formal way. The Hamiltonian or Canonical Method uses a Hamiltonian function to describe a mechanical system as a function of coordinates and momenta:

$$H = H(q, P, t)$$
(L1.8)

Then using eq. (L1.5), we can write the action integral as

$$S = \int_A^B \left(\sum_{i=1}^n P^i \frac{dq_i}{ds} - H(q, P, t) \right) dt = \int_A^B \left(\sum_{i=1}^n P^i dq_i - H(q, P, t) dt \right);$$
(L1.9)

The total variation of the integral can be separated into the variation of the end points, and the variation of the integral argument:

$$\begin{aligned} \delta \int_A^B f(x) dt &= \int_{A+\delta A}^{B+\delta B} f(x+\delta x) dt - \int_A^B f(x) dt = \int_B^{B+\delta B} f(x+\delta x) dt + \int_{A+\delta A}^A f(x+\delta x) dt + \int_A^B f(x+\delta x) dt - \int_A^B f(x) dt = \\ &= f(B)\Delta t_B - f(A)\Delta t_A + \int_A^B (f(x+\delta x) - f(x)) dt; \quad \Delta t_C = t(C + \delta C) - t(C); \quad \text{for } C = A, B. \end{aligned}$$

The first term represents the variation caused by a change of integral limits (events), while the second represents the variation of the integral between the original limits (events). The total variation of the action integral (L1.9) can be separated similarly:

$$\begin{aligned} \delta S &= \left[\sum_{i=1}^n P^i \Delta q_i - H \Delta t \right]_A^B + \int_A^B \left(\delta \sum_{i=1}^n P^i dq_i - \sum_{i=1}^n \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt \right) \right) = \\ &= \left[\sum_{i=1}^n P^i \Delta q_i - H \Delta t \right]_A^B + \sum_{i=1}^n \int_A^B \left(\delta P^i dq_i + P^i d\delta q_i - \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt \right) \right); \end{aligned}$$
(L1.10)

This equation encompasses everything: The expressions for the Hamiltonian and the momenta through the action and **Hamiltonian equations of motion**. Now we consider variation in both the coordinates and momenta that are treated equally: $\delta q; \delta P$.

To find the equation of motion we set constant events and $\delta q(A) = \delta q(B) = 0$; the first term disappears, and the minimal-action principle gives us

$$\delta S = \sum_{i=1}^n \int_A^B \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt - \delta P^i dq_i - P^i d\delta q_i \right) = 0,$$

which, after integration by parts of the last term translates into

$$\begin{aligned}\delta S &= \left[-\sum_{i=1}^n P^i \delta q_i \right]_A^B + \sum_{i=1}^n \int_A^B \left(\frac{\partial H}{\partial q_i} \delta q_i dt + \frac{\partial H}{\partial P^i} \delta P^i dt - \delta P^i dq_i + dP^i \delta q_i \right) = \\ &= \sum_{i=1}^n \int_A^B \left(\left\{ \frac{\partial H}{\partial q_i} + \frac{dP^i}{dt} \right\} \delta q_i dt + \left\{ \frac{\partial H}{\partial P^i} - \frac{dq_i}{dt} \right\} \delta P^i dt \right) = 0;\end{aligned}$$

where the variation of coordinates and momenta are considered to be independent. Therefore, both expressions in brackets must be zero at a real trajectory. This gives us the **Hamilton's equations of motion**:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial P^i}; \quad \frac{dP^i}{dt} = -\frac{\partial H}{\partial q_i}. \quad (\text{L1.11})$$

It is easy to demonstrate that these equations are exactly equivalent to the Lagrange's equation of motion. This is not surprising because they are obtained from the same principle of least action and describe the motion of the same system. Let us also look at the full derivative of the Hamiltonian:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_{i=1}^n \left(\frac{\partial H}{\partial P^i} \frac{dP^i}{dt} + \frac{\partial H}{\partial q_i} \frac{dq_i}{dt} \right) = \frac{\partial H}{\partial t} + \sum_{i=1}^n \left(-\frac{\partial H}{\partial q_i} \frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial P^i} \right) = \frac{\partial H}{\partial t}.$$

This equation means that the Hamiltonian is constant if it does not depend explicitly on t . It is an independent derivation of energy conservation for closed system. The conservation of momentum is apparent from equation (L1.11), viz., if the Hamiltonian does not depend explicitly on the coordinates, then momentum is constant. All these conservation laws result from the general theorem by **Emmy Noether** : *Any one-parameter group of diffeomorphisms operating in a phase space $((q, \dot{q}, t)$ for Lagrangian $((q, P, t)$ for Hamiltonian) and preserving the Lagrangian/Hamiltonian function equivalent to existence of the (first order) integral of motion.* (Informally, it can be stated as, for every differentiable symmetry created by local actions there is a corresponding conserved current).

Returning to the Eq. (L1.10), we now can consider motion **along real trajectories**. Here, the variation of the integral is zero and the connection between the action and the Hamiltonian variables is obtained by differentiation of the first term:

$$H = -\frac{\Delta S}{\Delta t} \Big|_{\Delta q=0} = -\frac{\partial S}{\partial t}; \quad P^i = \frac{\Delta S}{\Delta q_i} \Big|_{\Delta q_{k \neq i}=0} = \frac{\partial S}{\partial q_i}; \quad S = \int (P_i dq_i - H dt);$$

Along real
Trajectory

(L1.12)

Thus, knowing the action integral we can find the Hamiltonian and canonical (generalized) momenta from solving (L1.12) without using the Lagrangian. All conservation laws emerge naturally from (L1.10): if nothing depends on t , then H is conserved (i.e., the energy). If nothing depends on position, then the momenta are conserved: $P^i(A) = P^i(B)$. Finally, we write the Hamiltonian equations for one particle using the Cartesian frame:

$$S = \int (\vec{P} d\vec{r} - H(\vec{r}, \vec{P}, t) dt) \quad (\text{L1.13})$$

$$H(\vec{r}, \vec{P}, t) = -\frac{\partial S}{\partial t}; \quad \vec{P} = \frac{\partial S}{\partial \vec{r}};$$

$$\frac{d\vec{r}}{dt} = \frac{\partial H}{\partial \vec{P}}; \quad \frac{d\vec{P}}{dt} = -\frac{\partial H}{\partial \vec{r}}; \quad \frac{dH}{dt} = \frac{\partial H}{\partial t}$$

Hamiltonian method gives us very important tool – the general change of variables: $\{P_i, q_i\} \rightarrow \{\tilde{P}_i, \tilde{q}_i\}$, called **Canonical transformations**. From the least-action principle, two systems are equivalent if they differ by a full differential: (we assume the summation on repeating indices $i=1,2,3$, $a_i b_i \equiv \sum_i a_i b_i$; $a^\alpha b_\alpha \equiv \sum_\alpha a^\alpha b_\alpha$ and the use of co- and contra-variant vector components for the non-unity metrics tensor)

$$\delta \int P_i dq_i - H dt = 0 \propto \delta \int \tilde{P}_i d\tilde{q}_i - \tilde{H} dt = 0 \rightarrow P_i dq_i - H dt = \tilde{P}_i d\tilde{q}_i - \tilde{H} dt + dF \quad (\text{L1.14})$$

where F is the so-called generating function of the transformation. Rewriting (L1.14), reveals that $F = F(q_i, \tilde{q}_i, t)$:

$$dF = P_i dq_i - \tilde{P}_i d\tilde{q}_i + (H' - H) dt; \quad \tilde{P}_i = \frac{\partial F}{\partial \tilde{q}_i}; \quad P_i = -\frac{\partial F}{\partial q_i}; \quad H' = H + \frac{\partial F}{\partial t}. \quad (\text{L1.15})$$

In fact, generating functions on any combination of old coordinates or old momenta with new coordinates of new momenta are possible, totaling $4 = 2 \times 2$ combinations:

$$F(q, \tilde{q}, t) \Rightarrow dF = P_i dq_i - \tilde{P}_i d\tilde{q}_i + (H' - H) dt; \quad P_i = \frac{\partial F}{\partial q_i}; \quad \tilde{P}_i = -\frac{\partial F}{\partial \tilde{q}_i}; \quad H' = H + \frac{\partial F}{\partial t}.$$

$$\Phi(q, \tilde{P}, t) = F + \tilde{q}_i \tilde{P}_i \Rightarrow d\Phi = P_i dq_i + \tilde{q}_i d\tilde{P}_i + (H' - H) dt; \quad P_i = \frac{\partial \Phi}{\partial q_i}; \quad \tilde{q}_i = \frac{\partial \Phi}{\partial \tilde{P}_i}; \quad H' = H + \frac{\partial \Phi}{\partial t};$$

$$\Omega(P, \tilde{q}, t) = F - P_i q_i \Rightarrow d\Omega = -q_i dP_i - \tilde{P}_i d\tilde{q}_i + (H' - H) dt; \quad q_i = -\frac{\partial \Omega}{\partial P_i}; \quad \tilde{P}_i = -\frac{\partial \Omega}{\partial \tilde{q}_i}; \quad H' = H + \frac{\partial \Omega}{\partial t};$$

$$\Lambda(P, \tilde{P}, t) = \Phi - P_i q_i \Rightarrow d\Lambda = \tilde{q}_i d\tilde{P}_i - q_i dP_i + (H' - H) dt; \quad q_i = -\frac{\partial \Lambda}{\partial P_i}; \quad \tilde{q}_i = \frac{\partial \Lambda}{\partial \tilde{P}_i}; \quad H' = H + \frac{\partial \Lambda}{\partial t}; \quad (\text{L1.15}')$$

The most trivial canonical transformation is $\tilde{q}_i = P_i$; $\tilde{P}_i = -q_i$ with trivial generation function of

$$F(q, \tilde{q}) = q_i \tilde{q}_i \quad P_i = \frac{\partial F}{\partial q_i} = \tilde{q}_i; \quad \tilde{P}_i = -\frac{\partial F}{\partial \tilde{q}_i} = -q_i; \quad H' = H$$

Hence, this is direct proof that in the Hamiltonian method the coordinates and momenta are treated equally, and that the meaning of canonical pair (and its connection to Poisson brackets) has fundamental nature.

The most non-trivial finding from the Hamiltonian method is that the motion of a system, i.e., the evolution of coordinates and momenta also entails a Canonical transformation: $q_i(t + \tau) = \tilde{q}_i(q_i(t), P_i(t), t)$; $P_i(t + \tau) = \tilde{P}_i(q_i(t), P_i(t), t)$; with generation function being the action integral along a real trajectory (L1.12):

$$S = \int_A^{t+\tau} (P_i dq_i - H dt) - \int_A^t (P_i dq_i - H dt);$$

$$dS = P_i(t + \tau) dq_i - P_i(t) dq_i + (H_{t+\tau} - H_t) dt$$

1.1 Relativistic Mechanics

From here further: $i=0,1,2,3$ in Minkowski space with $(1,-1,-1,-1)$ metric.

see Appendix A for 4-D metric, vectors and tensors

Let's use Principle of Least Action for a relativistic particle. To determine *the action integral for a free particle* (which does not interact with the rest of the world), we must ensure that the action integral does not depend on our choice of the inertial system. Otherwise, the laws of the particle motion also will depend on the choice of the reference system, which contradicts the first principle of relativity. Therefore, the action must be invariant of Lorentz transformations and rotation in 3D space; i.e., it must depend on a 4D scalar. So far, *from Appendix A*, we know of one 4D scalar for a free particle: the interval. We can employ it as trial function for the action integral, and, by comparing the result with classical mechanics find a constant α connecting the action with the integral of the interval: $ds^2 = dx^i dx_i \equiv \sum_{i=1}^4 dx^i dx_i = (cdt)^2 - (d\vec{r})^2$

$$S = -\alpha \int_A^B ds = -\alpha \int_A^B \sqrt{(cdt)^2 - d\vec{r}^2}. \quad (\text{L1.16})$$

The minus sign before the integral reflects a natural phenomenon: the law of inertia requires a resting free particle to stay at rest in inertial system. The interval $ds = cdt$ has a maximum possible value ($cdt \geq \sqrt{(cdt)^2 - d\vec{r}^2}$) and requires for the action to be minimal, that the sign is set to be "-".

The integral (L1.16) is taken along the world line of the particle. The initial point A (event) determines the particle's start time and position, while the final point B (event) determines its final time and position. The action integral (L1.16) can be represented as integral with respect to the time:

$$S = -\alpha \int_A^B \sqrt{(cdt)^2 - d\vec{r}^2} = -\alpha c \int_A^B dt \sqrt{1 - \vec{v}^2 / c^2} = \int_A^B \mathbf{L} dt; \quad \mathbf{L} = -\alpha c \sqrt{1 - \frac{\vec{v}^2}{c^2}}; \quad \vec{v} = \frac{d\vec{r}}{dt};$$

where \mathbf{L} signifies the Lagrangian function of the mechanical system. It is important to note that while the action is an invariant of the Lorentz transformation, the Lagrangian is not. It must depend on the reference system because time depends on it. To find coefficient α , we compare the relativistic form with the known classical form by expanding \mathbf{L} by \vec{v}^2 / c^2 :

$$\mathbf{L} = -\alpha c \sqrt{1 - \frac{\vec{v}^2}{c^2}} \approx -\alpha c + \alpha \frac{\vec{v}^2}{2c}; \quad \mathbf{L}_{\text{classical}} = m \frac{\vec{v}^2}{2};$$

which confirms that α is positive and $\alpha = mc$, where m is the mass of the particle. Thus, we found the action and the Lagrangian for a relativistic particle:

$$S = -mc \int_A^B ds; \quad (\text{L1.17})$$

$$\mathbf{L} = -mc^2 \sqrt{1 - \frac{\vec{v}^2}{c^2}}; \quad (\text{L1.18})$$

The energy and momentum of the particles are defined by the standard relations eqs. (L1.4) and (L1.5):

$$\vec{p} = \frac{\partial \mathbf{L}}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = \gamma m\vec{v}; \quad (\text{L1.19})$$

$$E = \vec{p}\vec{v} - L = \gamma mc^2; \quad \gamma = 1/\sqrt{1 - \vec{v}^2/c^2} \quad (\text{L1.20})$$

with ratio between them of

$$E^2 = \vec{p}^2 c^2 + (mc^2)^2. \quad (\text{L1.21})$$

The energy of the resting particle does not go to zero as in classical mechanics but is equal to the famous Einstein value, $E = mc^2$; with the standard classical additions at low velocities ($v \ll c; p \ll mc$): $E \cong mc^2 + m \frac{\vec{v}^2}{2} \cong mc^2 + \frac{\vec{p}^2}{2m}$.

Four-momentum, conservation laws. The least-action principle gives us the equations of motion and an expression for the momentum of a system. Let us consider the total variation of an action for a single particle:

$$\begin{aligned} \delta S &= -mc \delta \int_A^B ds = -mc \delta \int_A^B \sqrt{dx^i dx_i} = -mc \left\{ \int_A^B \sqrt{d\delta x^i dx_i} + \sqrt{dx^i d\delta x_i} \right\} = \\ &= -mc \left\{ \int_A^B \frac{d\delta x^i dx_i}{2\sqrt{dx^i dx_i}} + \frac{dx^i d\delta x_i}{2\sqrt{dx^i dx_i}} \right\} = -mc \int_A^B \frac{dx^i d\delta x_i}{ds} = -mc \int_A^B u^i d\delta x_i; \end{aligned}$$

where $u^i \equiv dx^i/ds$ is 4-velocity. Integrating by parts,

$$\delta S = -mc u^i \delta x_i \Big|_A^B + mc \int_A^B \delta x_i \frac{du^i}{ds} ds; \quad (\text{L1.22})$$

we obtain the expression that can be used for all purposes. First, using the least-action principle with fixed A and B $\delta x_i(A) = \delta x_i(B) = 0$, to derive the conservation of 4-velocity for a free particle: $\frac{du^i}{ds} = 0$; $u^i = \text{const}$ or **the inertia law**.

Along a real trajectory $mc \int_A^B \delta x_i \frac{du^i}{ds} ds = 0$ the action is a function of the limits A and B (see eq.

(L1.12): $\delta S_{\text{real traj}} = (-E \delta t + \vec{P} \delta \vec{r}) \Big|_A^B$, i.e., $dS_{\text{real traj}} = -E dt + \vec{P} d\vec{r}$ is the full differential of t and \vec{r} with energy and momentum as the parameters. We note that this form of the action already is a Lorentz invariant:

$$\delta S_{\text{real traj}} = (-E \delta t + \vec{P} \delta \vec{r}) \Big|_A^B = (-P^i \delta x_i) \Big|_A^B;$$

i.e. classical Hamiltonian mechanics always encompassed a relativistic form and a metric: a

scalar δS is a 4-product of P^i and δx_i with the metric (1,-1,-1,-1). Probably one of most remarkable things in physics is that its classic approach detected the metric of 4-D space and time at least a century before Einstein and Poincaré.

To get 4-momentum, we consider a real trajectory $mc \int_A^B \delta x_i \frac{du^i}{ds} ds = 0$ and set $\delta x_i(B) = \delta x_i$:

$$p^i = -\frac{\partial S}{\partial x_i} = -\partial^i S = mc u^i = (\gamma mc, \gamma m \vec{v}) = (E/c, \vec{p}) \quad (\text{L1.23})$$

with an obvious scalar product ($u^i u_i = 1$, see Appendix A. eq. (A.42))

$$p^i p_i = E^2/c^2 - \vec{p}^2 = m^2 c^2 u^i u_i = m^2 c^2. \quad (\text{L1.24})$$

Equivalent forms of presentation are

$$p^i = (E/c, \vec{p}) \equiv m \gamma_v(c, \vec{v}) \equiv \frac{(mc, m\vec{v})}{\sqrt{1 - v^2/c^2}} \quad (\text{L1.25})$$

and, Lorentz transformation (P^i is a 4-vector, K' moves with $\vec{V} = \hat{e}_x V$):

$$E = \gamma_v(E' + c\beta_v p'_x); p_x = \gamma_v(p'_x + \beta_v E'/c); p_{y,z} = p'_{y,z}; \gamma_v = 1/\sqrt{1 - \beta_v^2}; \beta_v = V/c; \quad (\text{L1.26})$$

where subscripts are used for γ, β to define the velocity to which they are related. .

Equation (L1.24) expresses energy, velocity, and the like in terms of momenta and allows us to calculate all differentials:

$$E = c\sqrt{\vec{p}^2 + m^2 c^2}; dE = cd\sqrt{\vec{p}^2 + m^2 c^2} = \frac{d\vec{p} \cdot c\vec{p}}{\sqrt{\vec{p}^2 + m^2 c^2}} = \frac{c^2 \vec{p} \cdot d\vec{p}}{E} = \vec{v} \cdot d\vec{p}; \quad (\text{L1.27})$$

$$\vec{v} = \frac{c\vec{p}}{\sqrt{\vec{p}^2 + m^2 c^2}}; \vec{a} dt = d\vec{v} = d \frac{c\vec{p}}{\sqrt{\vec{p}^2 + m^2 c^2}} = \frac{c(d\vec{p}(\vec{p}^2 + m^2 c^2) - \vec{p}(\vec{p} d\vec{p}))}{(\sqrt{\vec{p}^2 + m^2 c^2})^3} = c \frac{d\vec{p} \cdot m^2 c^2 + [\vec{p} \times [d\vec{p} \times \vec{p}]]}{(\sqrt{\vec{p}^2 + m^2 c^2})^3}, \quad (\text{L1.28})$$

Coefficients $\gamma = E/mc^2$; $\vec{\beta} = \vec{v}/c$ differ from the above by constants, and satisfy similar relations.

The conservation laws reflect the homogeneity of space and time (see Mechanics): these natural laws do not change even if the origin of the coordinate system is shifted by δx . Then, $\delta x_i(A) = \delta x_i(B) = \delta x_i$. We can consider a closed system of particles (without continuous interaction, i.e., for most of the time they are free). Their action is sum of the individual actions, and

$$\sum_a \delta S_a = -(\sum_a m_a c u^i_a) \delta x_i|_A^B = -(\sum_a m_a c u^i_a) \delta x_i|_A^B = \{\sum_a p^i_a(A) - \sum_a p^i_a(B)\} \delta x_i = 0 \quad (\text{L1.29})$$

$$\sum_a p^i_a(A) = \sum_a p^i_a(B) = \left(\sum_a E_a/c, \sum_a \vec{p} \right) = \text{const}. \quad (\text{L1.30})$$

1.2 Particles in the 4-potential of the EM field.

The EM field propagates with the speed of light, i.e., it is a natural product of relativistic 4-D space-time; hence, the 4-potential is not an odd notion!

In contrast with the natural use of the interval for deriving the motion of the free relativistic particle, there is no clear guideline on what type of term should be added into action integral to describe a field. It is possible to consider some type of scalar function $\int A(x^i)ds^2$ to describe electromagnetic fields, but this would result in wrong equations of motion. Nevertheless, the next guess is to use a product of 4-vectors $A^i dx_i$, and surprisingly it does work, even though we do not know why? **Hence, the fact that electromagnetic fields are fully described by the 4-vector of potential $A^i = (A^0, \vec{A})$ must be considered as an experimental fact!**

Nevertheless, it looks natural that the interaction of a charge with electromagnetic field is represented by the scalar product of two 4-vectors with the $-e/c$ coefficient chosen by convention:

$$S_{\text{int}} = -\frac{e}{c} \int_A^B A^i dx_i; \quad A^i \equiv (A^0, \vec{A}) \equiv (\varphi, \vec{A}) \quad (\text{L1.31})$$

where the integral is taken along the particle's world line. A charge e and speed of the light c are moved outside the integral because they are constant; hence, **we use the conservation of the charge e and constancy of the speed of the light !**

It is essential that field is GIVEN, SINCE we are CONSIDERING a particle interacting with a given field.

Turning our attention back to the Least-Action Principle and Hamiltonian Mechanics

The standard presentation of 4-potential is

$$A^i \equiv (A^0, \vec{A}) \equiv (\varphi, \vec{A}); \quad (\text{L1.32})$$

where φ is called the scalar potential and \vec{A} is termed the vector potential of electromagnetic field.

Gauge Invariance. As we discussed earlier the action integral is not uniquely defined; we can add to it an arbitrary function of coordinates and time without changing the motion: $S' = S + f(x_i)$. This corresponds to adding the full differential of f in the integral (L1.31)

$$S' = \int_A^B \left(-mcds - \frac{e}{c} A^i dx_i + dx_i \partial^i f \right).$$

This signifies that the 4-potential is defined with sufficient flexibility to allow the addition of any 4-gradient to it (let us choose $f(x_i) = \frac{e}{c} g(x_i)$)

$$A'^i = A^i - \partial^i g(x_i) = A^i - \frac{\partial g}{\partial x_i}; \quad (\text{L1.33})$$

² You can check that this function will give the equations of motion $(mc - A) \frac{du^i}{ds} + \partial^i A = 0$.

without affecting the motion of the charge, a fact called **THE GAUGE INVARIANCE** .

WE SHOULD BE AWARE THAT THE EVOLUTION OF THE SYSTEM DOES NOT CHANGE BUT APPEARANCE OF THE EQUATION OF THE MOTION FOR THE SYSTEM COULD CHANGE. FOR EXAMPLE, AS FOLLOWS FROM (L1.33), THE CANONICAL MOMENTA WILL CHANGE:

$$P'^i = P^i - \partial^i f .$$

NEVERTHELESS, ONLY THE APPEARANCE OF THE SYSTEM IS ALTERED, NOT ITS EVOLUTION. OFTEN THE RIGHT CHOICE OF THE GAUGE INVARIANCE ALLOWS THE EQUATION OF MOTION TO BE INTEGRATED EASILY. MEASURABLE VALUES (SUCH AS FIELDS, MECHANICAL MOMENTUM) DO NOT DEPEND UPON IT. ONE MIGHT CONSIDER GAUGE INVARIANCE AS AN INCONVENIENCE, BUT, IN PRACTICE, IT PROVIDES A GREAT OPPORTUNITY TO FIND A GAUGE IN WHICH THE PROBLEM BECOMES MORE COMPREHENSIBLE AND SOLVABLE.

The action is an additive function: therefore, the action of a charge in electromagnetic field is simply the direct sum of a free particle's action and action of interaction: (remember $ds = ds^2 / ds = dx^i dx_i / ds = u^i dx_i$)

$$S = \int_A^B \left(-mc ds - \frac{e}{c} A^i dx_i \right) = \int_A^B \left(-mc u^i - \frac{e}{c} A^i \right) dx_i \quad (\text{L1.34})$$

Then the total variation of the action is

$$\begin{aligned} \delta S &= \delta \int_A^B \left(-mc ds - \frac{e}{c} A^i dx_i \right) = \int_A^B \left(-mc \frac{dx^i d\delta x_i}{ds} - \frac{e}{c} A^i d\delta x_i - \frac{e}{c} \delta A^i dx_i \right) = \\ &= - \left[\left(mc u^i + \frac{e}{c} A^i \right) \delta x_i \right]_A^B + \int_A^B \left(mc \frac{du^i}{ds} \delta x_i ds + \frac{e}{c} \delta x_i dA^i - \frac{e}{c} \delta A^i dx_i \right) = 0. \end{aligned} \quad (\text{L1.35})$$

That gives us a 4-momentum

$$P^i = - \frac{\delta S}{\delta x_i} = \left(mc u^i + \frac{e}{c} A^i \right) = \left(H/c, \vec{P} \right) = p^i + \frac{e}{c} A^i; \quad (\text{L1.36})$$

with

$$\begin{aligned} H = E &= c \left(mc u^0 + \frac{e}{c} A^0 \right) = \gamma mc^2 + e\varphi = c \sqrt{m^2 c^2 + \vec{p}^2} + e\varphi; \\ \vec{P} &= \gamma m \vec{v} + \frac{e}{c} \vec{A} = \vec{p} + \frac{e}{c} \vec{A}; \Rightarrow \vec{p} = \vec{P} - \frac{e}{c} \vec{A}. \end{aligned} \quad (\text{L1.37})$$

The Hamiltonian must be expressed in terms of generalized 3-D momentum, $\vec{P} = \vec{p} + \frac{e}{c} \vec{A}$ and it is

$$H(\vec{r}, \vec{P}, t) = c \sqrt{m^2 c^2 + \left(\vec{P} - \frac{e}{c} \vec{A} \right)^2} + e\varphi; \quad (\text{L1.38})$$

with Hamiltonian equation following from it:

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{\partial H}{\partial \vec{P}} = \frac{\vec{P}c - e\vec{A}}{\sqrt{m^2c^2 + \left(\vec{P} - \frac{e}{c}\vec{A}\right)^2}}$$

$$\frac{d\vec{P}}{dt} = \frac{d\vec{p}}{dt} + \frac{e}{c} \frac{d\vec{A}}{dt} = -\frac{\partial H}{\partial \vec{r}} = -e\vec{\nabla}\varphi - e \frac{\{(\vec{P} - \frac{e}{c}\vec{A}) \cdot \vec{\nabla}\} \vec{A}}{\sqrt{m^2c^2 + \left(\vec{P} - \frac{e}{c}\vec{A}\right)^2}} = -e\vec{\nabla}\varphi - \frac{e}{c}(\vec{v} \cdot \vec{\nabla})\vec{A};$$

From this equation we can derive (without any elegance!) the equation for mechanical momentum $\vec{p} = \gamma m \vec{v}$. We will not do it here, but rather we will use easier way to obtain the 4D equation of motion via the least-action principle. We fix A and B to get from equation (L1.35)

$$\delta S = \int_A^B \left(mc u^i \delta x_i ds + \frac{e}{c} \delta x_i dA^i - \frac{e}{c} \delta A^k dx_k \right) = \int_A^B \left(mc \frac{du^i}{ds} \delta x_i ds + \frac{e}{c} \frac{\partial A^i}{\partial x_k} \delta x_i dx_k - \frac{e}{c} \frac{\partial A^k}{\partial x_i} \delta x_i dx_k \right) =$$

$$\int_A^B \left(\frac{dp^i}{ds} + \frac{e}{c} \left[\frac{\partial A^i}{\partial x_k} - \frac{\partial A^k}{\partial x_i} \right] u_k \right) \delta x_i ds = 0. \quad (\text{L1.39})$$

As usual, the expression inside the round brackets must be set at zero to satisfy (L1.39); i.e., we have the equations of charge motion in an electromagnetic field:

$$mc \frac{du^i}{ds} \equiv \frac{dp^i}{ds} = \frac{e}{c} F^{ik} u_k; \quad (\text{L1.40})$$

wherein we introduce an anti-symmetric **electromagnetic field tensor**

$$F^{ik} = \frac{\partial A^k}{\partial x_i} - \frac{\partial A^i}{\partial x_k}. \quad (\text{L1.41})$$

Electromagnetic field tensor: The Gauge Invariance can be verified very easily:

$$F^{ik} = \frac{\partial A^k}{\partial x_i} - \frac{\partial A^i}{\partial x_k} = F^{ik} - \frac{\partial^2 g}{\partial x_i \partial x_k} + \frac{\partial^2 g}{\partial x_k \partial x_i} = F^{ik};$$

which means that the equation of motion (L1.40) is not affected by the choice of the gauge, and the **electromagnetic field tensor is defined uniquely!** Using the Landau convention, we can represent the asymmetric tensor by two 3-vectors (see Appendix A):

$$F^{ik} = (-\vec{E}, \vec{B}); F_{ik} = (\vec{E}, \vec{B});$$

$$F^{ik} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix}. \quad (\text{L1.42})$$

\vec{E} is the so-called vector of the electric field and \vec{B} is the vector of the magnetic field. Note the occurrence of the Lorentz group generator (see Appendix B) in (L1.42). The 3D expressions of the field vectors can be obtained readily:

$$E^\alpha = F^{\alpha 0} = \frac{\partial A^0}{\partial x_\alpha} - \frac{\partial A^\alpha}{\partial x_0} = -\frac{\partial \varphi}{\partial r_\alpha} - \frac{1}{c} \frac{\partial A^\alpha}{\partial t}; \quad \alpha = 1, 2, 3; \quad \vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \text{grad}\varphi; \quad (43)$$

$$B^\alpha = -\frac{1}{2} e^{\alpha\kappa\lambda} F^{\kappa\lambda} = e^{\alpha\kappa\lambda} \left(\frac{\partial A^\lambda}{\partial x_\kappa} - \frac{\partial A^\kappa}{\partial x_\lambda} \right); \quad \vec{B} = \text{curl}\vec{A}; \quad F^{\kappa\lambda} = e^{\lambda\kappa\alpha} H_\alpha. \quad (L1.44)$$

A 3D asymmetric tensor $e^{\alpha\kappa\lambda}$ and the *curl* definition are used to derive last equation and use Greek symbols for the spatial 3D components. The electric and magnetic fields are also Gauge invariant being components of Gauge invariant tensor.

(We have the first pair of Maxwell's equations without further calculation using the fact that differentiation is symmetric operator ($\partial^i \partial^k \equiv \partial^k \partial^i$):

$$e_{iklm} \partial^k F^{lm} = e_{iklm} \partial^k (\partial^l A^m - \partial^m A^l) = 2e_{iklm} (\partial^k \partial^l) A^m = 0; \quad (L1.45)$$

or explicitly:

$$\partial^k F^{lm} + \partial^l F^{mk} + \partial^m F^{kl} = 0. \quad (L1.46)$$

A simple exercise gives the 3D form of the first pair of Maxwell equations. They also can be attained using (L1.43) and (L1.44) and known 3D equivalencies: $\text{div}(\text{curl}\vec{A}) \equiv 0$; $\text{curl}(\text{grad}\varphi) \equiv 0$:

$$\begin{aligned} \vec{E} &= -\text{grad}\varphi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t}; & \text{curl}\vec{E} &= -\text{curl}(\text{grad}\varphi) - \frac{1}{c} \text{curl} \frac{\partial \vec{A}}{\partial t} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}; \\ \vec{B} &= \text{curl}\vec{A}; & \text{div}\vec{B} &= \text{div}(\text{curl}\vec{A}) \equiv 0; \end{aligned} \quad (L1.47)$$

I note that (L1.47) is the exact 3D equivalent of invariant 4D Maxwell equations (L1.45) that you may wish to verify yourself. There are 4 equations in (L1.45): $i=0,1,2,3$. The *div* is one equation and *curl* gives three (vector components) equations. Even the 3D form looks very familiar; the beauty and relativistic invariance of the 4D form makes it easy to remember and to use.

EM Fields transformation, Invariants of the EM field. The 4-potential was defined as 4-vector and it transforms as 4-vector. The electric and magnetic fields, as components of the asymmetric tensor, follow its transformation rules (See Appendix A).

$$\begin{aligned} \varphi &= \gamma(\varphi' + \beta A'_x); A_x = \gamma(A'_x + \beta\varphi'); \\ E_y &= \gamma(E'_y + \beta B'_z); E_z = \gamma(E'_z - \beta B'_y); \\ B_y &= \gamma(B'_y - \beta E'_z); B_z = \gamma(B'_z + \beta E'_y). \end{aligned} \quad (L1.48)$$

and the rest is unchanged. An important repercussion from these transformations is that the separation of the electromagnetic field in two components is an artificial one. They translate into each other when the system of observation changes and **MUST** be measured in the same units (Gaussian). The rationalized international system of units (SI) system measures them in V/m, Oe, A/m and T. Why not use also a horse power per square mile an hour, the old British thermal units as well? This makes about the same sense as using Tesla or A/m.

While the values and directions of 3D field components are frame-dependent, two 4-scalars can be build from the EM 4-tensor $F^{ik} = (-\vec{E}, \vec{B})$

$$F^{ik} F_{ik} = \text{inv}; \quad e^{iklm} F_{ik} F_{lm} = \text{inv}; \quad (L1.49)$$

which in the 3D-form appear as

$$\vec{B}^2 - \vec{E}^2 = \text{inv}; \quad (\vec{E} \cdot \vec{B}) = \text{inv}. \quad (L1.50)$$

This conveys a good sense what can and cannot be done with the 3D components of electromagnetic fields. Any reference frame can be chosen and both fields transferred in a minimal number of components limited by (L1.50). For example; 1) if $|\vec{E}| > |\vec{B}|$ in one system it is true in all systems and vice versa; and (L1.2) if fields are perpendicular in one frame, $(\vec{E} \cdot \vec{B}) = 0$, this is true in all frames. When $(\vec{E} \cdot \vec{B}) = 0$ a frame can always be found where E or B are equal to zero (locally!).

Lorentz form of equation of a charged particle's motion.

The equations of motion (L1.40) can be rewritten in the form:

$$\begin{aligned} \frac{dE}{dt} &= c \frac{dp^0}{dt} = eF^{0k}v_k = e\vec{E} \cdot \vec{v}; & v_k &= \frac{dx_k}{dt} = (c, -\vec{v}) \\ \frac{d\vec{p}}{dt} &= e \left(\hat{e}_\alpha F^{\alpha k} \frac{v_k}{c} \right) = \frac{e}{c} \left(\hat{e}_\alpha \cdot cF^{\alpha 0} - \hat{e}_\alpha \cdot F^{\alpha k} v_k \right) = e\vec{E} + \hat{e}_\alpha e^{\alpha k \lambda} B_\lambda \frac{v_k}{c} = e\vec{E} + \frac{e}{c} [\vec{v} \times \vec{B}]. \end{aligned} \quad (L1.51)$$

So, we have expressions for the generalized momentum and energy of the particle in an electromagnetic field. Generalized momentum is equal to the particle's mechanical momentum plus the vector potential scaled by e/c . The total energy of the charged particle is its mechanical energy, γmc^2 , plus its potential energy, $e\varphi$, in an electromagnetic field. The Standard Lorentz (not Hamiltonian!) equations of motion for $\vec{p} = \gamma m\vec{v}$ are

$$\frac{d\vec{p}}{dt} = e\vec{E} + \frac{e}{c} [\vec{v} \times \vec{B}]. \quad (L1.52)$$

with the force caused by the electromagnetic field (Lorentz force) comprised of two terms: the electric force, which does not depend on particle's motion, and, the magnetic force that is proportional to the vector product of particle velocity and the magnetic field, i.e., it is perpendicular to the velocity. Accordingly, the magnetic field does not change the particle's energy. We derived it in Eq. (L1.51):

$$mc^2 \frac{d\gamma}{dt} = e\vec{E} \cdot \vec{v}; \quad (L1.53)$$

Eqs. (L1.52) and (L1.53) are generalized equations. Using directly standard Lorentz equations of motion in a 3D form is a poor option. The 4D form is much better (see below) and, from all points of view, the Hamiltonian method is much more powerful!

It is worth noting that the 4D form of the charge motion (L1.40) and its matrix form is the most compact one,

$$u^i = \frac{dx^i}{ds}; \quad mc \frac{du^i}{ds} = \frac{e}{c} F^i_k u^k; \Rightarrow \frac{d}{ds} [x] = [I] \cdot [u]; \quad \frac{d}{ds} [u] = \frac{e}{mc^2} [F] \cdot [u] \quad (L1.54)$$

and, in many cases, it is very useful. We treat the \mathbf{x} , \mathbf{u} as a vectors, and $[F]$ as the 4x4 matrix. $[I]$ is just the unit 4x4 matrix It has interesting formal solution in the matrix form:

$$[u] = e^{\int \frac{e}{mc^2} [F] ds} [u_0]; \quad [x] = [x_0] + \left[\int ds e^{\int \frac{e}{mc^2} [F] ds} \right] [u_0] \quad (L1.55)$$

Its resolution is well defined when applied to the motion of a charged particle in uniform, constant EM field:

$$[u] = e^{\frac{e}{mc^2}[F](s-s_0)} [u_0]; [x] = [x_o] + \left[\int e^{\frac{e}{mc^2}[F](s-s_0)} ds \right] [u_0] \quad (L1.56)$$

THE LORENTZ GROUP OF THEORETICAL PHYSICS (SEE APPENDIX B) IS FASCINATING, AND THE FACT THAT EM FIELD TENSOR HAS THE SAME STRUCTURE AS THE GENERATOR OF LORENTZ GROUP IS NO COINCIDENCE – RATHER, IT IS INDICATION THAT PHYSICISTS HAVE PROBABLY COME VERY CLOSE TO THE ROOTS OF NATURE IN THIS SPECIFIC DIRECTION. THIS STATEMENT IS FAR FROM TRUTH FOR OTHER FUNDAMENTAL FORCES AND INTERACTIONS.

To conclude this subsection, we will take one step further from (L1.54) and write a totally linear evolution equation for a combination of 4D vectors

$$\frac{d}{ds} \begin{bmatrix} x \\ u \end{bmatrix} = [\Lambda] \cdot \begin{bmatrix} x \\ u \end{bmatrix}; [\Lambda] = \begin{bmatrix} 0 & I \\ 0 & \frac{e}{mc^2} F \end{bmatrix} \quad (L1.57)$$

where $[\Lambda]$ is an 8x8 degenerated matrix. Similarly to (L1.55) and (L1.56)

$$\begin{bmatrix} x \\ u \end{bmatrix} = e^{\int [\Lambda] ds} \cdot \begin{bmatrix} x \\ u \end{bmatrix}_o; \begin{bmatrix} x \\ u \end{bmatrix} = e^{[\Lambda](s-s_0)I} \cdot \begin{bmatrix} x \\ u \end{bmatrix}_o \text{ for } [\Lambda] = \text{const}; \quad (L1.58)$$

Power of Hamiltonian method: Phase space and invariants

As discussed, when the Hamiltonian of the system does not depend on either the coordinates or momenta, we automatically gain “for-free” an integral of motion or the invariant. There are more general invariants of motion that exist for all Hamiltonian systems – the Poincaré invariants. Let us consider a Hamiltonian system that is described by the set of coordinates, Canonical momenta, and the independent variable s . In matrix form, the Hamiltonian equations are written as

$$H = H(Q, P, s); X = \begin{bmatrix} Q_1 \\ P_1 \\ \dots \\ Q_n \\ P_n \end{bmatrix}; \frac{dX}{ds} = S \frac{\partial H}{\partial X}; S = \begin{bmatrix} 0 & 1 & \dots & 0 & 0 \\ -1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & -1 & 0 \end{bmatrix}; \quad (L1.93)$$

where S is a generator (norm) of a symplectic group of matrixes (two different but closely related types of mathematical groups). The space of coordinates and momenta is called phase space of the system with dimension $2n$. Here, consider an ensemble of particles in this phase space whose motions satisfy the Hamiltonian equations; then, their motion is completely determined by their initial position in the phase space. This means that in the Hamiltonian system the phase-space trajectories of particles, which initially were separated, will never cross! Consider one trajectory in the phase space $X_o(s)$, which satisfies the Hamiltonian equation (L1.93) and another trajectory with an infinitesimally small deviation from X_o

$$\begin{aligned}
X_1 &= X_o(s) + \delta X; \\
\frac{dX_o}{dz} &= S \frac{\partial H(X)}{\partial X} \Big|_{X=X_o}; \quad ; \quad H(X + \Delta X) = H(X) + \frac{\partial H(X)}{\partial X} \Delta X + \frac{1}{2} \frac{\partial^2 H(X)}{\partial X^2} \Delta X^2 + O(\Delta X^3); \\
\frac{d(X_o(s) + \delta X)}{ds} &= S \frac{\partial H(X)}{\partial X} \Big|_{X=X_o(s)+\delta X} \cong S \frac{\partial H(X)}{\partial X} \Big|_{X=X_o} + SH(s)\delta X + O(\delta X^2); \quad (L1.94) \\
H(s) &= \frac{1}{2} \frac{\partial^2 H(X)}{\partial X^2} \Big|_{X=X_o(s)}; \\
[H(s)]_{ij} &= \left[\frac{\partial^2 H}{\partial x_i \partial x_j} \right]_{X=X_o(s)} \Rightarrow [H(s)]_{ij} \equiv [H(s)]_{ji} \Rightarrow H^T(s) = H(s)
\end{aligned}$$

with a symmetric $2n \times 2n$ matrix $\mathbf{H}(s)$. Thus, the equation of motion for a small deviations about the known trajectory are linear but s-dependent, and can be expressed via the linear transform matrix $M(s)$:

$$\begin{aligned}
\frac{d\delta X}{dz} &= S \cdot H(s) \cdot \delta X \\
\delta X(s) &= M(s) \cdot \delta X(0); \quad (95) \\
\frac{dM}{ds} &= S \cdot H(s) \cdot M \quad \Rightarrow \frac{d\delta X}{ds} = \frac{dM(s)}{ds} \cdot \delta X(0) = S \cdot H(s) \cdot M \cdot \delta X(0) = S \cdot H(s) \cdot \delta X \#
\end{aligned}$$

The matrix $M(s)$ is symplectic, which is proven as follows:

$$\begin{aligned}
M^T(z) \cdot S \cdot M(z) &= S \quad - \text{symplectic condition}; \quad M(0) = \hat{1} \Rightarrow M^T(0) \cdot S \cdot M(0) = S; \\
\frac{d(M^T S M)}{ds} &= \frac{dM^T}{ds} S M + M^T S \frac{dM}{ds} \\
\frac{dM^T}{ds} &= (S \cdot H \cdot M)^T = M^T H^T S^T = -M^T H S \\
&\text{because } H^T = H; \quad S^T = -S; \quad S^2 = -\hat{1} \\
\frac{d(M^T S M)}{ds} &= -M^T H S S M + M^T S S H M = M^T H M - M^T H M \equiv 0 \\
&\Rightarrow M^T(s) \cdot S \cdot M(s) = \text{const} = M^T(0) \cdot S \cdot M(0) = S \quad \#
\end{aligned}$$

The symplectic condition has two asymmetric $2n \times 2n$ matrixes on both sides

$$M^T(s) \cdot S \cdot M(s) = S \quad (L1.96)$$

and imposes $n(2n-1)$ conditions on the matrix M . These conditions result in invariants of motion for the ensembles of particles, called Poincaré invariants. Accordingly, for 3-D motion, there are 15 Poincaré invariants! The most well-known one, the conservation of the phase space volume (Liouville's theorem), is a consequence of the unit determinant of the matrix M :

$$\begin{aligned}
\det[M^T(s) \cdot S \cdot M(s)] &= \det S \rightarrow \\
(\det M(s))^2 &= 1 \rightarrow \det M = \pm 1; \\
\text{but } \det M(0) &= 1 \rightarrow \det M = 1 \#
\end{aligned}
\tag{L1.97}$$

Next, we consider an infinitesimally small phase-space volume ΔV_{2n} around a known trajectory and its transformation:

$$\begin{aligned}
\Delta V_{2n}(s) &= \det \left| \frac{\partial \Delta X(s)}{\partial \Delta X(0)} \right| \Delta V_{2n}(0) = \det M(s) \cdot \Delta V_{2n}(0) \rightarrow \\
\Delta V_{2n}(s) &= \Delta V_{2n}(0) = \text{const}
\end{aligned}
\tag{L1.98}$$

This is the core for the so-called Vlasov equations describing collective effects in accelerators (see Appendix 3).

The 3-dimensional volume occupied by the particles often is termed 3-D beam emittance. The rest of the Poincaré invariants represent similar conservation laws for the sum of projections on hyper-surfaces in $2n$ -phase space. In the general case, interpreting coupled motion invariant is too involved for this introductory lecture.

But there is one nice instance wherein the Hamiltonian is decoupled, i.e., it is the direct sum of individual Hamiltonians for each canonical pair.

$$H = H(Q, P, s) = \sum_{k=1}^3 H_k(Q_k, P_k, s)
\tag{L1.99}$$

Then, the phase space is two-dimensional and the area of the space phase occupied by the beam is called beam emittance for a specific dimension – horizontal (x, Px), vertical (y, Py) or longitudinal (-t,E). All three emittances are constants (integrals) of motion.

Lecture 2. Accelerator Hamiltonian

2.1 Accelerator coordinate system.

In accelerator physics we usually study beams of particles, i.e. particles moving in approximately the same direction (a huge difference from detectors) with approximately the same momenta. It is traditional, and very useful to choose one particle in the beam as the reference particle and study its trajectory $\vec{r}_o(t)$ as natural reference. Furthermore, most accelerator equipment is bolted to the floor and, hence, can be better described by its position in space than its existence in time. This is the reason why accelerator physicists decided to use length along the reference trajectory, s , as independent coordinate instead of time:

$$s(t) = \int_{t_i}^t |d\vec{r}_o(t)| = \int_{t_i}^t |\vec{v}_o(t)| dt; \quad (L2.1)$$

$$\vec{v}_o(t) = \frac{d\vec{r}_o(t)}{dt}; \quad \gamma(t) = 1/\sqrt{1 - \vec{v}_o^2(t)/c^2}; \quad \vec{p}_o(t) = \gamma(t)m\vec{v}_o(t); \quad E_o(t) = \gamma(t)mc^2$$

It is important for independent variable to be a monotonous function (as is time), which requires that the reference particle never stops moving (except possibly at the beginning and the end of the reference trajectory).

Reference trajectories

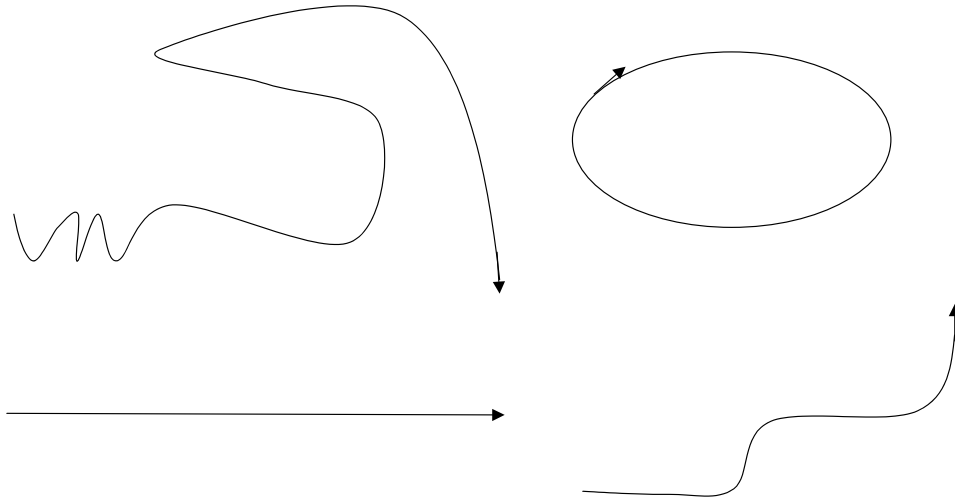


Fig. 1. Various possible reference trajectories, from a simple straight pass to a circular one, though all other possibilities.

The reference trajectory is determined by initial 4-momentum of the reference particle and the EM field along its trajectory. We should consider that trajectory is given (and from $\vec{r}_o(t)$ we also know the particle's 4-momentum in each point of trajectory) and so satisfy the equation of motion (for example eqs. (L1.52) and (L1.53)).

Usually EM fields are designed for the existence of such a trajectory (within constrains of Maxwell equation). Herein, the words reference trajectory and orbit are used interchangeably.

Inverting (L2.1) we can write the 4D trajectory at the function of s :

$$\vec{r} = \vec{r}_o(s); t = t_o(s); \vec{p} = \vec{p}_o(s), E = E_o(s). \quad (\text{L2.2})$$

with the charge to the designer of accelerator to make it real trajectory:

$$\frac{d\vec{p}_o(s)}{ds} = \frac{dt_o(s)}{ds} \left(e\vec{E}(\vec{r}_o(s), t_o(s)) + \frac{e}{c} [\vec{v}_o(s) \times \vec{B}(\vec{r}_o(s), t_o(s))] \right) \quad (\text{L2.3})$$

Starting from this point, we use following conventions: Derivatives of any function with respect to the time will be shown by appropriate number of dots, while appropriate number of symbol ' will be used to indicate derivatives with respect to s:

$$f' = \frac{df}{ds}; f'' = \frac{d^2f}{ds^2} \dots \dots f\dot{=} = \frac{df}{dt}; f\ddot{=} = \frac{d^2f}{dt^2}. \quad (\text{L2.4})$$

There is infinite variety of possible reference trajectories. The most popular ones are flat, i.e. they lie in a plane. A typical example is the circular orbit of a storage ring with a horizontal trajectory. Many of reference orbits are piece-wise combinations of trajectories lying in various planes. Still, there are 3D reference orbits by design. As the matter of fact, all real reference orbits are 3D because of the field errors in magnets, and errors in aligning these magnets.

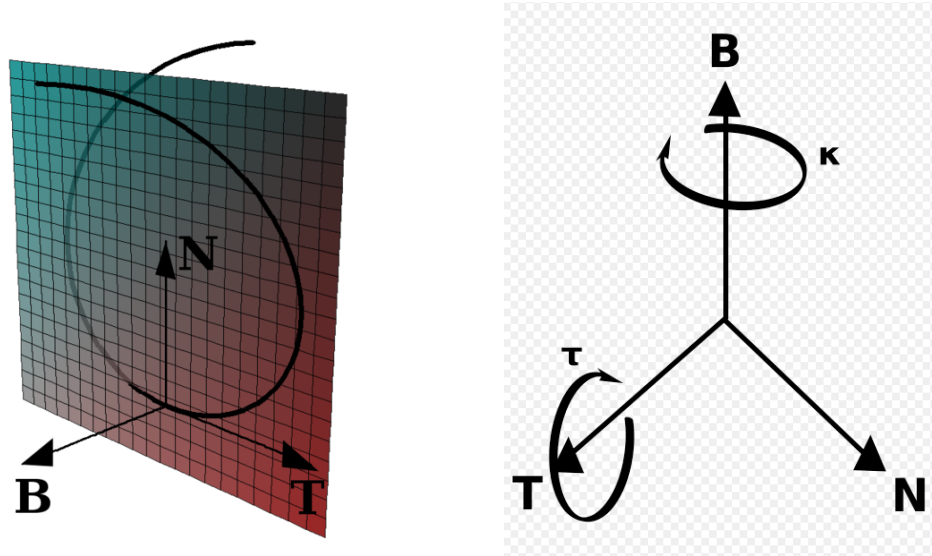


Fig. 2. Illustration of Frenet-Serret formulas and system from <http://en.wikipedia.org/wiki/Frenet-Serret>

Hence, there is no good reason not to start this discussion from general 3D reference trajectory. Fortunately two French mathematicians, Jean Frédéric Frenet and Joseph Alfred Serret, in the mid-nineteenth century developed such a coordinate system, which is described by the Frenet-Serret formulas in classical differential geometry (O.Struik, Dirk J., Lectures on Classical Differential Geometry, Addison-Wesley, Reading, Mass, 1961). The Frenet-Serret coordinate system often is called the natural coordinate system. One important feature is that it has non-diagonal metrics. Hence, we have a bit of differential geometry to spice the mix.

Figures 2 and 3 illustrate the Frenet-Serret coordinate system and define 3 orthogonal unit vectors: Normal $\hat{e}^1 = \vec{n}(s)$, tangent $\hat{e}^2 = \vec{\tau}(s)$, and normal and bi-normal $\hat{e}^3 = \vec{b}(s) = [\vec{n} \times \vec{\tau}]$:

$$(\vec{n} \cdot \vec{\tau}) = (\vec{b} \cdot \vec{n}) = (\vec{b} \cdot \vec{\tau}) = 0.$$

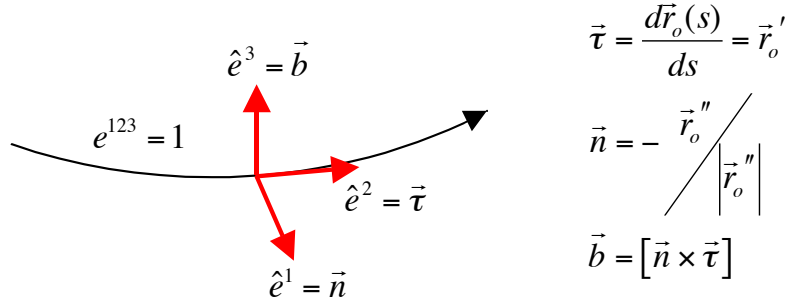


Fig. 3. Unit vectors in the Frenet-Serret coordinate system and their definitions

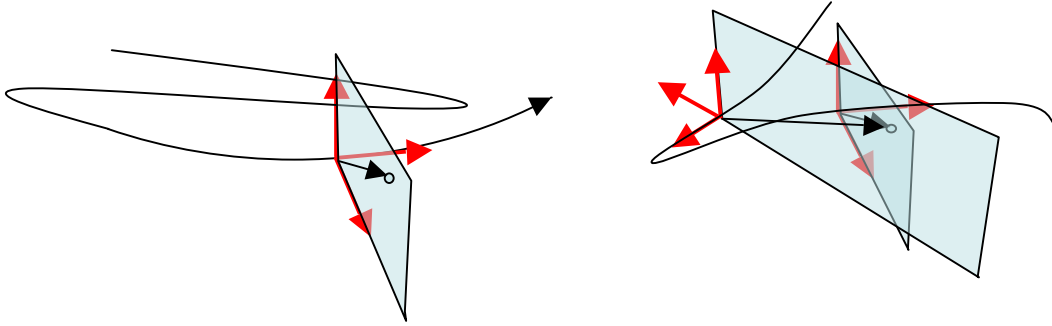


Fig. 4. Expansion of particle's position in Frenet-Serret frame.

The reference trajectory must be smooth, with finite second derivatives, etc....etc... The position of any particle located in close proximity³ to the reference trajectory can uniquely expressed as

$$\vec{r} = \vec{r}_o(s) + x \cdot \vec{n}(s) + y \cdot \vec{b}(s). \quad (\text{L2.5})$$

i.e., it is fully described by 3 contra-variant coordinates:

$$q^1 = x; \quad q^2 = s, \quad q^3 = y. \quad (\text{L2.5-1})$$

The vectors $\{\vec{n}, \vec{\tau}, \vec{b}\}$ satisfy Frenet-Serret formulae:

$$\frac{d\vec{\tau}}{ds} = K(s) \cdot \vec{n}; \quad \frac{d\vec{n}}{ds} = K(s) \cdot \vec{\tau} - \kappa(s) \cdot \vec{b}; \quad \frac{d\vec{b}}{ds} = \kappa(s) \cdot \vec{n}; \quad (\text{L2.6})$$

where

$$K(s) = 1/\rho(s) \quad (\text{L2.6-1})$$

is the curvature of the trajectory, and $\kappa(s)$ is its torsion. If the torsion is equal to zero, the trajectory remains in one plane, as designed for majority of accelerators. Curvature of trajectory

³ Proximity to the reference orbit is important for the uniqueness of the extension (L2.5): As shown on the figure above, equation (L2.6-2) may have multiple solutions if the requirement of proximity is not applied, i.e., the expansion (L2.5) may have multiple branches and mathematically become too involved.

is more common – each dipole magnet makes trajectory to curve.

As shown in Fig.4, the transverse part of the position vector $\vec{r}_\perp = x \cdot \vec{n}(s) + y \cdot \vec{b}(s)$ lies in the plane defined by the normal and by-normal unit vectors $(\vec{n}(s), \vec{b}(s))$, while s is defined from equation:

$$(\vec{r} - \vec{r}_o(s)) \cdot \vec{\tau}(s) = 0. \quad (\text{L2.6-2})$$

Now we expand the differential geometry:

$$d\vec{r} = \sum_{i=1}^3 \vec{a}_i dq^i = \vec{n}dx + \vec{b}dy + \left\{ (1+Kx)\vec{\tau} + \kappa(\vec{n}y - \vec{b}x) \right\} ds \quad (\text{L2.7})$$

with the co-variant basis of

$$\vec{a}_i = \frac{\partial \vec{r}}{\partial q^i}; \quad \vec{a}_1 = \vec{n}; \quad \vec{a}_2 = (1+Kx)\vec{\tau} + \kappa(\vec{n}y - \vec{b}x); \quad \vec{a}_3 = \vec{b}; \quad (\text{L2.8})$$

A co-variant basis vector is readily derived from the orthogonal conditions:

$$\vec{a}_i \vec{a}^j = \delta_i^j; \quad \vec{a}^1 = \vec{n} - \frac{\kappa y}{1+Kx} \vec{\tau}; \quad \vec{a}^2 = \frac{\vec{\tau}}{1+Kx} + \kappa(\vec{n}y - \vec{b}x); \quad \vec{a}^3 = \vec{b} + \frac{\kappa x}{1+Kx} \vec{\tau}; \quad (\text{L2.9})$$

The components of the co- and contra-variant metric tensors are defined as follows:

$$g_{ik} = \vec{a}_i \cdot \vec{a}_k = \begin{bmatrix} 1 & \kappa y & 0 \\ \kappa y & (1+Kx)^2 + \kappa^2(x^2 + y^2) & -\kappa x \\ 0 & -\kappa x & 1 \end{bmatrix}$$

$$g^{ik} = \vec{a}^i \cdot \vec{a}^k = \frac{1}{(1+Kx)^2} \cdot \begin{bmatrix} (1+Kx)^2 + \kappa^2 y^2 & -\kappa y & -\kappa^2 xy \\ -\kappa y & 1 & \kappa x \\ -\kappa^2 xy & \kappa x & (1+Kx)^2 + \kappa^2 x^2 \end{bmatrix} \quad (\text{L2.10})$$

$$g_o = \det[g_{ik}] = (1+Kx)^2$$

Any vector can be expanded about both co- and contra-variant bases, as well can $\{\vec{n}, \vec{\tau}, \vec{b}\}$:

$$\vec{R} \equiv R_x \vec{n} + R_s \vec{\tau} + R_y \vec{b} \equiv \sum_k R^k \vec{a}_k \equiv \sum_k R_k \vec{a}^k$$

$$R_k = \vec{R} \cdot \vec{a}_k; \quad R_1 = R_x; \quad R_2 = (1+Kx)R_s + \kappa(R_x y - R_y x); \quad R_3 = R_y; \quad (\text{L2.11})$$

$$R^k = \vec{R} \cdot \vec{a}^k; \quad R^1 = R_x - \frac{\kappa y}{1+Kx} R_s; \quad R^2 = \frac{R_s}{1+Kx} + \kappa(R_x y - R_y x); \quad R^3 = R_y + \frac{\kappa x}{1+Kx} R_s;$$

All this is trivial, and finally differential operators will look like:

$$\vec{\nabla} \varphi = \vec{a}^k \frac{\partial \varphi}{\partial q^k}; \quad \text{div} \vec{A} = (\vec{\nabla} \cdot \vec{A}) = \frac{1}{\sqrt{g_o}} \frac{\partial}{\partial q^k} (\sqrt{g_o} A^k);$$

$$\text{curl} \vec{A} = [\vec{\nabla} \times \vec{A}] = \frac{e^{ikl}}{\sqrt{g_o}} \frac{\partial A_l}{\partial q^k} \vec{a}_i; \quad \Delta \varphi = \vec{\nabla}^2 \varphi = \frac{1}{\sqrt{g_o}} \frac{\partial}{\partial q^i} \left(\sqrt{g_o} g^{ik} \frac{\partial \varphi}{\partial q^k} \right). \quad (\text{L2.12})$$

As discussed before, the Hamiltonian of a charged particle in EM field in Cartesian coordinate system is

$$H(\vec{r}, \vec{P}, t) = c \sqrt{m^2 c^2 + \left(\vec{P} - \frac{e}{c} \vec{A} \right)^2} + e\varphi, \quad (\text{L1.38, Lecture 1})$$

where the canonical momentum is $\vec{P} = \vec{p} + \frac{e}{c} \vec{A}$. Let us explore how we can make the transformation to our “curved and twisted” coordinate system. The easiest way is to apply canonical transformation with generation function

$$F(\vec{P}, q^i) = -\vec{P} \cdot \left(\vec{r}_o(s) + x \cdot \vec{n}(s) + y \cdot \vec{b}(s) \right). \quad (\text{L2.13})$$

to our new coordinates (L2.6):

$$q^1 = x; \quad q^2 = s, \quad q^3 = y. \quad (\text{L2.14-1})$$

with new momenta obtained by simple differentiation

$$P_1 = P_x; \quad P_2 = (1 + Kx)P_s + \kappa(P_x y - P_y x); \quad P_3 = P_y; \quad (\text{L2.14-2})$$

that alter the appearance of the Hamiltonian (L1.38)

$$H = c \sqrt{\left((1 + Kx)^{-2} \left(\left(P_2 - \frac{e}{c} A_2 \right) + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) \right) \right)^2 + \left(P_1 - \frac{e}{c} A_1 \right)^2 + \left(P_3 - \frac{e}{c} A_3 \right)^2 + m^2 c^2} + e\varphi \quad (\text{L2.15})$$

This is still the Hamiltonian with t as independent variable and three sets of canonical pairs $\{q^1, P_1\}, \{q^2, P_2\}, \{q^3, P_3\}$. Now, we change the independent variable to s by the easiest method, that, as always, is using the least-action principle: we consider the conjugate momentum to s , P_2 , as a function of the remaining canonical variables: $\{q^1, P_1\}, \{q^3, P_3\}, \{-t, H\}$

$$S = \int_A^B P_1 dq^1 + P_2(\dots) ds + P_3 dq^3 - H dt; \quad \delta S = 0; \quad (\text{L2.16})$$

Notably, the coordinates and time, the canonical momenta and the Hamiltonian appear in the 4-D scalar product form in the action integral.

$$P_i dx^i; \quad x^i = \{ct, x, s, y\}; \quad P_i = \{H/c, -P_1, -P_2, -P_3\}; \quad i = 0, 1, 2, 3.$$

This equivalency of the time and space is fundamental to the relativistic theory. Next is the following

$$H = H(x^i, P_1, P_2, P_3) \xrightarrow{\text{solve}} P_2 = P_2(x^i, P_1, P_2, H) \quad \text{rename} \quad P_i = -H; \quad h^* = -P_2(x^i, P_1, P_2, H) \quad (\text{L2.17})$$

$$S = \int P_1 dx + P_3 dy + P_2 dz - H dt \equiv \int P_1 dx - P_3 dy - h^* dz + P_1 dt$$

These two forms are identical, so the immediate conclusion (we did explicitly in our class)⁴ is that

$$\begin{aligned} x' &= \frac{dx}{ds} = \frac{\partial h^*}{\partial P_1}; & \frac{dP_1}{ds} &= -\frac{\partial h^*}{\partial x}; & y' &= \frac{dy}{ds} = \frac{\partial h^*}{\partial P_3}; & \frac{dP_3}{ds} &= -\frac{\partial h^*}{\partial y} \\ t' &= \frac{dt}{ds} = \frac{\partial h^*}{\partial P_t} \equiv -\frac{\partial h^*}{\partial H}; & \frac{dP_t}{ds} &= -\frac{\partial h^*}{\partial t} \rightarrow \frac{dH}{ds} = \frac{\partial h^*}{\partial t} \end{aligned} \quad (\text{L2.18})$$

$$\begin{aligned} h^* &= -(1+Kx) \sqrt{\frac{(H - e\varphi)^2}{c^2} - m^2 c^2 - \left(P_1 - \frac{e}{c} A_1\right)^2 - \left(P_3 - \frac{e}{c} A_3\right)^2} \\ &\quad + \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3\right) - \kappa y \left(P_1 - \frac{e}{c} A_1\right) \end{aligned} \quad (\text{L2.19})$$

Thus, by choosing one of coordinates as independent variable, the new Hamiltonian is nothing but its conjugate canonical momentum with a minus sign. Applying a canonical transformation that exchanges the coordinate with momentum and then employs a new coordinate (old momentum) as the independent variable it would turn the old coordinate into the new Hamiltonian. In all cases, the Hamiltonian is the function of the remaining canonical variables. This capability of the Hamiltonian systems is unique and one we can take advantage of. An important restriction is the monotonous behavior of independent variable. Otherwise, some or all of the derivatives can be infinite in the point where the independent variable stumbles (i.e., where the new time stops).

The equations (L2.18) and (L2.19) are the general form of the single-particle Hamiltonian equation in an accelerator. It undoubtedly is nonlinear (the square root signifies relativistic mechanics), and cannot be solved analytically in general. Only few specific cases allow such solutions.

The only additional option we have is to choose a gauge for the 4-potential. One good choice (my preference) is to make the vector potential equal to zero at the reference trajectory. Two other auxiliary conditions will allow us to express the components of the 4-vector potential in a

⁴ As always,

$$\delta S_{AB} = \delta \left(\int_A^B P_i dq^i - H dt \right) = \int_A^B \left(\sum_{i=1,3} \left(\delta P_i dq^i + P_i d\delta q^i + \frac{\partial P_2}{\partial q^i} \delta q^i ds + \frac{\partial P_2}{\partial P_i} \delta P_i ds \right) + \left(-\delta H dt - H d\delta t + \frac{\partial P_2}{\partial t} \delta t ds + \frac{\partial P_2}{\partial H} \delta H ds \right) \right) = 0$$

and integrating by parts $\sum_{i=1,3} P_i \delta q^i - H \delta t \Big|_A^B \equiv 0$, we obtain the direct equivalent of eqs. (L2.18,19):

$$\begin{aligned} \delta S_{AB} &= \int_A^B \left(\sum_{i=1,3} \left(\delta P_i \left(\frac{\partial P_2}{\partial P_i} ds + dq^i \right) + \delta q^i \left(\frac{\partial P_2}{\partial q^i} ds - dP_i \right) \right) + \left(\delta t \left(dH + \frac{\partial P_2}{\partial t} ds \right) + \delta H \left(\frac{\partial P_2}{\partial H} ds - dt \right) \right) \right) = 0 \\ \frac{dq^i}{ds} &= -\frac{\partial P_2}{\partial P_i}; & \frac{dt}{ds} &= \frac{\partial P_2}{\partial H} ds; & \frac{dP_i}{ds} &= +\frac{\partial P_2}{\partial q^i}; & \frac{dH}{ds} &= -\frac{\partial P_2}{\partial t} \end{aligned}$$

form of the Taylor series:

$$a) \bar{A}(s,0,0,t) = 0; \quad b) \partial_x^n A_1 \Big|_{s,0,0,t} = \partial_y^n A_3 \Big|_{s,0,0,t} = 0; \quad c) \frac{\partial A_1}{\partial x} \Big|_{s,0,0,t} + \frac{\partial A_3}{\partial y} \Big|_{s,0,0,t} = 0 \quad (\text{L2.20})$$

that can be achieved by gauge transformation

$$\begin{aligned} \bar{A} &= \tilde{A} - \bar{\nabla} f; \quad \varphi = \varphi + \frac{1}{c} \frac{\partial f}{\partial t}; \quad f = f_a = f_b + f_c \\ f_a &= \int_0^s \tilde{A}_2(s_1,0,0,t) ds_1 + \tilde{A}_1(s_1,0,0,t) \cdot x + \tilde{A}_3(s_1,0,0,t) \cdot y \\ f_b &= \sum_{n=1}^{\infty} \left(\partial_x^n \tilde{A}_1 \Big|_{s,0,0,t} \frac{x^{n+1}}{(n+1)!} + \partial_y^n \tilde{A}_3 \Big|_{s,0,0,t} \frac{y^{n+1}}{(n+1)!} \right) \\ f_c &= \frac{1}{2} \sum_{n,k=0}^{\infty} \partial_x^n \partial_y^k (\partial_y \tilde{A}_1 + \partial_x \tilde{A}_3) \frac{x^{n+1}}{(n+1)!} \frac{y^{k+1}}{(n+1)!} \end{aligned} \quad (\text{L2.21})$$

Conditions (L2.20) have following important consequences:

$$a) \partial_s^k \partial_t^l \bar{A}(s,0,0,t) = 0; \quad b) A_1(s,x,0,t) = 0; A_3(s,0,y,t) = 0; \quad c) \partial_s^k \partial_t^l \partial_x^m \partial_y^n \left(\frac{\partial A_1}{\partial x} + \frac{\partial A_3}{\partial y} \right) = 0 \quad (\text{L2.20+})$$

After a one-page-long exercise, using the first pair of Maxwell equations (eqs. (L1.47), Lect. 1) and conditions (L2.20), one can express the 4-potential in this gauge though the components of the magnetic- and electric- fields, in other words, make an unique vector potential:

$$\begin{aligned} A_1 &= \frac{1}{2} \sum_{n,k=0}^{\infty} \partial_x^k \partial_y^n B_s \Big|_{ro} \frac{x^k}{k!} \frac{y^{n+1}}{(n+1)!}; \quad A_3 = -\frac{1}{2} \sum_{n,k=0}^{\infty} \partial_x^k \partial_y^n B_s \Big|_{ro} \frac{x^{k+1}}{(k+1)!} \frac{y^n}{n!} \\ A_2 &= \sum_{n=1}^{\infty} \left\{ \partial_x^{n-1} \left((1+Kx) B_y + \kappa x B_s \right) \Big|_{ro} \frac{x^n}{n!} - \partial_y^{n-1} \left((1+Kx) B_x - \kappa y B_s \right) \Big|_{ro} \frac{y^n}{n!} \right\} + \\ &+ \frac{1}{2} \sum_{n,k=1}^{\infty} \left\{ \partial_x^{n-1} \partial_y^k \left((1+Kx) B_y + \kappa x B_s \right) \Big|_{ro} \frac{x^n}{n!} \frac{y^k}{k!} - \partial_x^n \partial_y^{k-1} \left((1+Kx) B_x - \kappa y B_s \right) \Big|_{ro} \frac{x^n}{n!} \frac{y^k}{k!} \right\}; \quad (\text{L2.22}) \\ \varphi &= \varphi_o(s,t) - \sum_{n=1}^{\infty} \partial_x^{n-1} E_x \Big|_{ro} \frac{x^n}{n!} - \sum_{n=1}^{\infty} \partial_y^{n-1} E_y \Big|_{ro} \frac{y^n}{n!} - \frac{1}{2} \sum_{n,k=1}^{\infty} \left(\partial_x^{n-1} \partial_y^k E_x \Big|_{ro} + \partial_x^n \partial_y^{k-1} E_y \Big|_{ro} \right) \frac{x^n}{n!} \frac{y^k}{k!}; \end{aligned}$$

where $f|_{ro}; (f)_{ro}$ denotes that the value of the function f is taken at the reference orbit $r_o(s)$: i.e., at $x=0; y=0$, but in an arbitrary moment of time t . We reserve the notions $f|_{ref}; (f)_{ref}$ for values taken at the reference trajectory $\vec{r} = \vec{r}_o(s)$ at the reference time $t = t_o(s)$. It is noteworthy that the value of our new Hamiltonian for the reference particle is the full particle's momentum with the minus sign:

$$h^*|_{ref} = -P_o(s) \quad (\text{L2.23})$$

We should note that $\varphi_o(s,t)$ is determined with the accuracy of an arbitrary constant, which can be eliminated by requesting $\varphi_o(s_o, t_o(s_o)) = 0$ at some point along the reference trajectory.

The coefficients in (L2.22) can be expanded further using a trivial time series

$$f(t) = f(t_o(s)) + \sum_{n=1}^{\infty} \frac{d^n f}{dt^n} \Big|_{t=t_o(s)} \frac{(t - t_o(s))^n}{n!}.$$

One important feature of this expansion that no conditions in the EM field are assumed; thus, it can be in free-space field (typical for single-particle dynamics) or a field with sources (for example, charges and currents of beam are examples). Hence, the expansion is applicable to any arbitrary accelerator problem.

2.2 An equilibrium particle and a reference trajectory.

A particle that follows the reference trajectory is called an equilibrium (or reference) one:

$$\vec{r} = \vec{r}_o(s); \quad t = t_o(s); \quad H = H_o(s) = E_o(s) + \varphi_o(s, t_o(s)), \quad (\text{L2.23})$$

with $x = 0; \quad y = 0; \quad p_x = 0; \quad p_y = 0$. This is where condition (L2.20a) $\vec{A}|_{ref} = 0$ is useful, i.e., for

$$x|_{ref} = 0; \quad y|_{ref} = 0; \quad P_1|_{ref} = p_x|_{ref} + \frac{e}{c} A_1|_{ref} \equiv 0; \quad P_3|_{ref} = p_y|_{ref} + \frac{e}{c} A_3|_{ref} \equiv 0. \quad (\text{L2.24})$$

The differential form of (L2.24)

$$\frac{dx}{ds} \Big|_{ref} = \frac{\partial h^*}{\partial P_1} \Big|_{ref} = 0; \quad \frac{dy}{ds} \Big|_{ref} = \frac{\partial h^*}{\partial P_3} \Big|_{ref} = 0; \quad \frac{dP_1}{ds} \Big|_{ref} = -\frac{\partial h^*}{\partial x} \Big|_{ref} = 0; \quad \frac{dP_3}{ds} \Big|_{ref} = -\frac{\partial h^*}{\partial y} \Big|_{ref} = 0; \quad (\text{L2.25})$$

should be combined with the expression for the Hamiltonian (L2.19). The two first equations in (L2.25) give us the already known conditions, viz., that of the zero transverse component of momentum. The following two equations are not as trivial; they set the two conditions at the reference orbit. Completing a trivial differentiation on x (where most of the terms are turned into zero at the reference orbit, except $\partial_x \varphi$ and $\partial_x A_2$) we have

$$-\frac{\partial h^*}{\partial x} \Big|_{ref} = K\sqrt{G} \Big|_{ref} - (1 + Kx)_{ref} \frac{\left(\left[\frac{eE}{c^2} \frac{\partial \varphi}{\partial x} \right] + p_x \frac{e}{c} \frac{\partial A_1}{\partial x} + p_y \frac{e}{c} \frac{\partial A_3}{\partial x} \right)_{ref}}{\sqrt{G} \Big|_{ref}} + \left[\frac{e}{c} \frac{\partial A_2}{\partial x} \right]_{ref} + \kappa (p_y)_{ref} + \kappa \left(\frac{e}{c} \frac{\partial A_1}{\partial x} y - \frac{e}{c} \frac{\partial A_3}{\partial x} x \right)_{ref} = 0$$

$$E \equiv (H - e\varphi); \quad G = \frac{E^2}{c^2} - m^2 c^2 - p_x^2 - p_y^2; \quad \sqrt{G} \Big|_{ref} = p_o \dots$$

Note: The term(s) that do not vanish at the limit are identified by the square brackets [...]

and using the above expansions, we derive the well-know equation for the curvature of the trajectory:

$$K(s) \equiv \frac{1}{\rho(s)} = -\frac{e}{p_o c} \left(B_y \Big|_{ref} + \frac{E_o}{p_o c} E_x \Big|_{ref} \right). \quad (\text{L2.26})$$

Differentiation on y is similar

$$-\frac{\partial h^*}{\partial y} \Big|_{ref} = -(1 + Ky)_{ref} \frac{\left(\left[\frac{eE}{c^2} \frac{\partial \varphi}{\partial y} \right] + p_x \frac{e}{c} \frac{\partial A_1}{\partial y} + p_y \frac{e}{c} \frac{\partial A_3}{\partial y} \right)_{ref}}{\sqrt{G} \Big|_{ref}} + \left[\frac{e}{c} \frac{\partial A_2}{\partial y} \right]_{ref} - (p_x)_{ref} + \kappa \left(\frac{e}{c} \frac{\partial A_1}{\partial y} y - \frac{e}{c} \frac{\partial A_3}{\partial y} x \right)_{ref} = 0$$

and yields

$$B_x|_{ref} = \frac{E_o}{p_o c} E_y|_{ref}, \quad (L2.27)$$

That represents only the absence of “vertical curvature”. The difference between (L2.26) and (L2.27) arises from the choice of coordinates in Frenet-Serret system: \mathbf{x} (i.e., \mathbf{q}_1) corresponds to the plane where trajectory bends.

The conditions in (L2.23) for the arrival time of the reference particle and values of its Hamiltonian are also informative, but not surprising. First, the condition on the arrival time

$$\frac{dt_o(s)}{ds} = -\frac{\partial h^*}{\partial H}|_{ref} = \frac{H - e\varphi}{c^2 \sqrt{G}}|_{ref} = \frac{H_o - e\varphi_o}{p_o c^2} \equiv \frac{E_o}{p_o c^2} = \frac{1}{v_o(s)} \quad (L2.28)$$

gives an understandable definition of velocity along trajectory: $\mathbf{v} = d\mathbf{s}/dt$, and the velocity of the reference particle $v_o = p_o c^2 / E_o$.

The condition on energy (3D Hamiltonian) gives

$$\frac{\partial h^*}{\partial t}|_{ref} = (1 + Kx)|_{ref} \frac{\left(\frac{eE}{c^2} \frac{\partial \varphi}{\partial t} + p_x \frac{e}{c} \frac{\partial A_1}{\partial t} + p_y \frac{e}{c} \frac{\partial A_3}{\partial t} \right)|_{ref}}{\sqrt{G}|_{ref}} + \left(\frac{e}{c} \frac{\partial A_2}{\partial t} \right)|_{ref} + \kappa \left(\frac{e}{c} \frac{\partial A_1}{\partial x} y - \frac{e}{c} \frac{\partial A_3}{\partial x} x \right)|_{ref} = \frac{eE_o}{p_o c^2} \frac{\partial \varphi}{\partial t}|_{ref} \quad (L2.29)s$$

$$\frac{dH_o(s)}{ds} = \frac{\partial h^*}{\partial t}|_{ref} = \frac{eE_o}{p_o c^2} \frac{\partial \varphi}{\partial t}|_{ref};$$

which can be transferred using $H = E + e\varphi$ and $d\varphi_o(s, t_o(s)) = \frac{\partial \varphi_o}{\partial s} ds + \frac{\partial \varphi_o}{\partial t} \frac{ds}{v_o(s)}$ into the energy gain of the reference particle along its trajectory:

$$\frac{dE_o(s)}{ds} = \frac{d(H_o(s) - \varphi_o(s, t_o(s)))}{ds} = -e \frac{\partial \varphi}{\partial s}|_{ref} \equiv eE_2(s, t_o(s)). \quad (L2.30)$$

As discussed before, accelerator designers face the problem of ensuring that the reference particle faithfully follows the reference trajectory. Our goal is to use the above conditions to the maximum, and, as we see below, to eliminate zero-order terms from the equations of motion. By selecting the reference trajectory as basis for our coordinate system, we set the transverse coordinates and momenta at zero at the reference orbit. Hence, two canonical pairs have a good and solid origin.

The third pair $(-t, H)$ is odd; it is not zero for the reference particle. Furthermore, it has different units. Hence, we can move step forward with a more natural Canonical pair $\{q_\tau = -ct, p_\tau = H/c\}$ - whose generating function is obvious: $\Phi(q = -t, \tilde{P} = p_\tau) = -ct \cdot p_\tau$. In this case, the analogy is complete: $q_\tau = -ct$ has the dimension of distance and is just $-\mathbf{x}_o$ in 4D space, while $p_\tau = H/c$ has the dimension of momentum and is just \mathbf{P}_o in 4D space.

We also should select variables that are zero at the reference orbit. The following pair is one of better choices:

$$\{\tau = -c(t - t_o(s)), \delta = (H - E_o(s) - e\varphi_o(s, t))/c\}, \quad (L2.31)$$

which are zero for the reference particle. Generation function is easily to come with:

$$\Phi(q, \tilde{P}, s) = \tilde{P}_1 x + \tilde{P}_3 y - (E_o(s) + c\delta)(t - t_o(s)) - e \int^t \varphi_o(s, t_1) dt_1, \quad (\text{L2.32})$$

and it produces what is desired:

$$\begin{aligned} P_1 &= \frac{\partial \Phi}{\partial x} = \tilde{P}_1; \quad P_3 = \frac{\partial \Phi}{\partial y} = \tilde{P}_3; \quad H = \frac{\partial \Phi}{\partial(-t)} = E_o + c\delta + e\varphi_o(s, t); \\ \tilde{q}_1 &= \frac{\partial \Phi}{\partial \tilde{P}_1} = x; \quad \tilde{q}_3 = \frac{\partial \Phi}{\partial \tilde{P}_3} = y; \quad \tilde{q}_\delta = \frac{\partial \Phi}{\partial \delta} = -c(t - t_o(s)) = \tau \\ \tilde{h} &= h + \frac{\partial \Phi}{\partial s} = h + \frac{E_o(s) + c\delta}{v_o(s)} + E'_o(s)\tau/c - e \int^t \varphi'_o(s, t_1) dt_1 \end{aligned} \quad (\text{L2.33})$$

The change to the Hamiltonian comprises meaningful terms as well as just a trivial function of s , $\mathbf{g}(s)$:

$$\frac{\partial \Phi}{\partial s} = \frac{c}{v_o(s)} \delta - e\varphi_{||}(s, \tau) + g(s); \quad g(s) = E_o(s)/v_o(s) - e \int^{t_o(s)} \varphi'_o(s, t_1) dt_1 \quad (\text{L2.34})$$

$$\varphi_{||}(s, \tau) =_{def} \frac{\partial}{\partial s} \int_0^{-\tau/c} (\varphi_o(s, t_o(s) + \zeta) - \varphi_o(s, t_o(s))) d\zeta \equiv - \int_0^{-\tau/c} (E_2(s, t_o(s) + \zeta) - E_2|_{ref}) d\zeta$$

where we used eq. (L2.30) as $E'_o(s) = -e \frac{\partial \varphi}{\partial s} \Big|_{ref}$. Additive $\mathbf{g}(s)$ simply can be dropped from the

Hamiltonian - it does not change equations of motion. Now the only remaining task is to express the new Hamiltonian function with an updated canonical pair (L2.33) and (L2.19):

$$\begin{aligned} \tilde{h} &= -(1 + Kx) \sqrt{p_o^2 + \frac{2E_o}{c} \left(\delta - \frac{e}{c} \varphi_\perp \right) + \left(\delta - \frac{e}{c} \varphi_\perp \right)^2 - \left(P_1 - \frac{e}{c} A_1 \right)^2 - \left(P_3 - \frac{e}{c} A_3 \right)^2} + \\ &+ \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) + \frac{c}{v_o} \delta - \frac{e}{c} \varphi_{||}(s, \tau) \end{aligned} \quad (\text{L2.35})$$

where we used following trivial expansion and definition:

$$\frac{(E_o + c\delta + e\varphi_o(s, t) - e\varphi)^2}{c^2} - m^2 c^2 = p_o^2 + \frac{2E_o}{c} \left(\delta - \frac{e}{c} \varphi_\perp \right) + \left(\delta - \frac{e}{c} \varphi_\perp \right)^2; \quad (\text{L2.36})$$

$$\varphi_\perp =_{def} \varphi(s, x, y, t) - \varphi_o(s, t) \equiv \varphi(s, x, y, t) - \varphi(s, 0, 0, t)$$

2.3 Scaling variables.

Frequently, it is useful to scale one of canonical variables. Typical scaling in accelerator physics involves dividing the canonical momenta P_1, P_3, δ by the momentum of the reference particle:

$$\pi_1 = \frac{P_1}{p_o}; \quad \pi_3 = \frac{P_3}{p_o}; \quad \pi_o = \frac{\delta}{p_o}. \quad (\text{L2.37})$$

These variables are dimensionless and also are close to $x', y', \delta E / p_o c$ for small deviations. Such scaling only is allowed in Hamiltonian mechanics when the scaling parameter is constant, i.e., is not function of s . Thus, scaling by the particle's momentum remains within the framework

of Hamiltonian mechanics only if the reference particle's momentum is constant, that is, when the longitudinal electric field is zero along the reference particle's trajectory (i.e. at moment $t=t_o(s)$). One similarly can scale the coordinates by a constant.

$$\xi_1 = \frac{x}{L}; \quad \xi_3 = \frac{y}{L}; \quad \xi_o = \frac{\tau}{L}.$$

Scaling by a constant is easy; divide the Hamiltonian by the constant and rename the variables. Hence, transforming (L2.37) with constant, called p_o , will make Hamiltonian (L2.35) into

$$\begin{aligned} \tilde{h} = & -(1+Kx) \sqrt{1 + \frac{2E_o}{p_o c} \left(\delta - \frac{e}{p_o c} \varphi_{\perp} \right) + \left(\delta - \frac{e}{p_o c} \varphi_{\perp} \right)^2 - \left(\pi_1 - \frac{e}{p_o c} A_1 \right)^2 - \left(\pi_3 - \frac{e}{p_o c} A_3 \right)^2} + \\ & + \frac{e}{p_o c} A_2 + \kappa x \left(\pi_3 - \frac{e}{p_o c} A_3 \right) - \kappa y \left(\pi_1 - \frac{e}{p_o c} A_1 \right) + \frac{c}{v_o} \delta - \frac{e}{p_o c} \varphi_{\parallel}(s, \tau) \end{aligned} \quad (\text{L2.35 @})$$

constant energy

Usage of this Hamiltonian is very popular for storage rings or transport channels, wherein the energy of the particles remains constant in time. It should not be employed for particles undergoing an acceleration.

2.4 Expanding the Hamiltonian.

Expanding the Hamiltonian (L2.35) is a nominal tool in accelerator physics that allows the separation of the effects of various orders and sometimes the use of perturbation-theory approaches. Having completed the process of creating canonical variables, which are zero for the reference particle, the next step is to assume (which is true for operational accelerators) that the relative deviations of momenta are small

$$\left| \frac{P_1}{p_o} \right| \ll 1; \quad \left| \frac{P_3}{p_o} \right| \ll 1; \quad \left| \frac{\delta}{p_o} \right| \ll 1;$$

and that the EM fields are sufficiently smooth around the reference trajectory to allow expansion in terms of x ; y ; τ . We will consider⁵ all six variables to be of the same order (of infinitesimality, α). We call the order of expansion to be the maximum total power in a product that is any combination of $x, y, \tau, P_1, P_2, \delta$. Unless there is a good reason not to do so, we truncate the series using this rule.

The general expansion of Hamiltonian (L2.35) can be accomplished via the already derived expansion for 4-potential and the well-known expansion of the square root:

$$\begin{aligned} \tilde{h} = & -(1+Kx) p_o \sqrt{1 + \frac{2E_o}{p_o c} \left(\frac{\delta}{p_o} - \frac{e}{p_o c} \varphi_{\perp} \right) + \left(\frac{\delta}{p_o} - \frac{e}{p_o c} \varphi_{\perp} \right)^2 - \left(\frac{P_1}{p_o} - \frac{e}{p_o c} A_1 \right)^2 - \left(\frac{P_3}{p_o} - \frac{e}{p_o c} A_3 \right)^2} + \\ & + \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) + \frac{c}{v_o} \delta - \frac{e}{c} \varphi_{\parallel}(s, \tau) \end{aligned} \quad ;(\text{L2.38})$$

$$\sqrt{1+g} = 1 + \sum_{n=1}^{\infty} (-1)^{n-1} \frac{g^n (2n-3)!!}{2^n n!} = 1 + \frac{y}{2} + O(y^2).$$

⁵ Sometimes, one can keep explicit the time dependence of fields and expand only the rest of the variables. One such case is an approximate, and useful, description of synchrotron oscillation.

Let us discuss a few general issues first. If we expand (L2.37) into an finite series

$$\tilde{h} = -p_o(s) + \sum_{\nu=1}^N \sum_{\substack{i+j+k+ \\ l+m+n=\nu}} C_{ijklm} x^i y^j \tau^k P_1^l P_3^m \delta^n + O(\alpha^{N+1}) = -p_o + \sum_{\nu=1}^N \sum_{\substack{p_i=\nu \\ i=1,6}} C_{\nu\{p_i\}} \prod_{i=1}^6 x_i^{p_i} + O(\alpha^{N+1}); \quad (\text{L2.39})$$

where we introduce our phase-space vector

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \equiv \begin{bmatrix} x \\ P_1 \\ y \\ P_2 \\ \tau \\ \delta \end{bmatrix}; \quad (\text{L2.40})$$

which we will use in its 2n-dimensional phase-space appetence

$$X^T = [q^1 \ P_1 \ \dots \ \dots \ q^n \ P_n] = [x_1 \ x_2 \ \dots \ \dots \ x_{2n-1} \ x_{2n}], \quad (\text{L2.41})$$

where T stands for “transposed”. Using this notion, Hamiltonian equations can be written as one

$$\frac{dX}{ds} = S \cdot \frac{\partial H}{\partial X} \quad \Leftrightarrow \quad \frac{dx_i}{ds} = S_{ij} \cdot \frac{\partial H}{\partial x_j} \equiv \sum_{j=1}^{2n} S_{ij} \cdot \frac{\partial H}{\partial x_j} \quad (\text{L2.42})$$

wherein we introduce matrix S – a generator of the symplectic group (see further). The matrix, S , is asymmetric, with $S_{2m-1,2m} = 1 = -S_{2m,2m-1}$, $m = 1, \dots, n$, and other elements are zero.

In matrix form S has n diagonal blocks with a 2×2 matrix σ , and the rest is the field of zeros:

$$S = \begin{bmatrix} \sigma & 0 & \dots & 0 \\ 0 & \sigma & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \sigma \end{bmatrix}_{2n \times 2n}; \quad \sigma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (\text{L2.43})$$

This is one of the very important objects in Hamiltonian mechanics.

Using expansion (L2.39), we would dispose of $p_o(s)$ as valueless, and look initially at the first-order terms in the expansion:

$$\tilde{h} = \sum_{i=1,6} C_{1i} x_i + O(\alpha^2); \quad (\text{L2.39-0})$$

That should give us the zero- order terms in the equation of motion:

$$\frac{dx_k}{ds} = \sum_{j=1}^{2n} S_{kj} \cdot C_{1j} + O(\alpha)$$

By design of our variables (section 2.2) $\left. \frac{dx_k}{ds} \right|_{ref} = 0$, i.e. all zero order terms are equal to zero.

Hence, the expansion of our Hamiltonian does not have first-order terms:

$$h = \sum_{\nu=2}^N \sum_{\substack{p_i=\nu \\ i=1,6}} C_{\nu\{p_i\}} \prod_{i=1}^6 x_i^{p_i} + O(\alpha^{N+1}); \quad (\text{L2.44})$$

Next, we look at the second-order term that is of foremost importance in accelerator physics. We can write it as a quadratic form in matrix notations:

$$h = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 h_{ij} x_i x_j + O(\alpha^3) \equiv \frac{1}{2} X^T \cdot H \cdot X + O(\alpha^3); \quad (\text{L2.45})$$

Matrix H can be chosen to be symmetric by a simple $H = (\tilde{H}^T + H)/2$ and noting the obvious $Z^T (\tilde{H}^T - H) Z \equiv 0$, $\forall Z$ equivalence of the asymmetric matrix. Applying (L2.42) with $\partial_{z_i} (\sum h_{jk} x_j x_k) = (h_{jk} + h_{kj}) x_i = 2h_{jk} x_i$ the linear part of the equations of motion is obtained:

$$\frac{dX}{ds} = D \cdot X + O(\alpha^2); \quad D = S \cdot H \quad (\text{L2.45})$$

We finish this section with the explicit form of the first non-trivial term in the expansion of (L2.38):

$$\begin{aligned} \tilde{h} = & \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_3 - yP_1) + \\ & \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x \delta + g_y y \delta + F_x x \tau + F_y y \tau \end{aligned} \quad ; \quad (\text{L2.46})$$

with

$$\begin{aligned} \frac{F}{p_o} = & \left[-K \cdot \frac{e}{p_o c} B_y - \frac{e}{p_o c} \frac{\partial B_y}{\partial x} + \left(\frac{e B_s}{2p_o c} \right)^2 \right] - \frac{e}{p_o v_o} \frac{\partial E_x}{\partial x} - 2K \frac{e E_x}{p_o v_o} + \left(\frac{m e E_x}{p_o^2} \right)^2; \\ \frac{G}{p_o} = & \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial y} + \left(\frac{e B_s}{2p_o c} \right)^2 \right] - \frac{e}{p_o v_o} \frac{\partial E_y}{\partial y} + \left(\frac{m e E_y}{p_o^2} \right)^2; \quad ; \quad (\text{L2.47}) \\ \frac{2N}{p_o} = & \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial x} - \frac{e}{p_o c} \frac{\partial B_y}{\partial y} \right] - K \cdot \frac{e}{p_o c} B_x - \frac{e}{p_o v_o} \left(\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} \right) - 2K \frac{e E_y}{p_o v_o} + \left(\frac{m e E_z}{p_o^2} \right) \left(\frac{m e E_x}{p_o^2} \right) \\ L = & \kappa + \frac{e}{2p_o c} B_s; \quad \frac{U}{p_o} = \frac{e}{p_o c^2} \frac{\partial E_s}{\partial t}; \quad g_x = \frac{(mc)^2 \cdot e E_x}{p_o^3} - K \frac{c}{v_o}; \quad g_y = \frac{(mc)^2 \cdot e E_y}{p_o^3}; \\ F_x = & \frac{e}{c} \frac{\partial B_y}{\partial ct} + \frac{e}{v_o} \frac{\partial E_x}{\partial ct}; \quad F_y = -\frac{e}{c} \frac{\partial B_x}{\partial ct} + \frac{e}{v_o} \frac{\partial E_y}{\partial ct}. \end{aligned}$$

If momentum p_o is constant, we can use (L2.37) and rewrite Hamiltonian of the linearized motion (L2.46) as

$$\begin{aligned} \tilde{h}_n = & \frac{\pi_1^2 + \pi_3^2}{2} + f \frac{x^2}{2} + n \cdot xy + g \frac{y^2}{2} + L(x\pi_3 - y\pi_1) + \\ & \frac{\pi_o^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + u \frac{\tau^2}{2} + g_x x \pi_o + g_y y \pi_o + f_x x \tau + f_y y \tau \end{aligned} \quad ; \quad (\text{L2.46-n})$$

with

$$f = \frac{F}{p_o}; \quad n = \frac{N}{p_o}; \quad g = \frac{G}{p_o}; \quad u = \frac{U}{p_o}; \quad f_x = \frac{F_x}{p_o}; \quad f_y = \frac{F_y}{p_o}; \quad (\text{L2.47-n})$$

Note that

$$x' = \frac{\partial h_n}{\partial \pi_1} = \pi_1 - Ly; \quad y' = \frac{\partial h_n}{\partial \pi_3} = \pi_3 + Lx; \quad ; \quad (\text{L2.48})$$

i.e. as soon as $L=0$, we can use traditional x' and y' as reduced momenta.

For flat reference orbit - $\kappa = 0$, in the absence of transverse coupling ($L=0$, $N=0$) and transverse electric fields, DC magnetic fields it becomes a much simpler and much more familiar form:

$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + G \frac{y^2}{2} + \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x \delta$$

or

$$\tilde{h}_n = \frac{x'^2 + y'^2}{2p_o} + f \frac{x^2}{2} + g \frac{y^2}{2} + \frac{\pi_o^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + u \frac{\tau^2}{2} + g_x x \delta$$
(L2.49)

with

$$f = \left[-K \cdot \frac{e}{p_o c} B_y - \frac{e}{p_o c} \frac{\partial B_y}{\partial x} \right]; \quad g = \frac{e}{p_o c} \frac{\partial B_x}{\partial y}; \quad u = \frac{e}{p_o c^2} \frac{\partial E_s}{\partial t}; \quad g_x = -K \frac{c}{v_o}.$$
(L2.50)

Finally, see an Appendix where Mathematica tool allowing expansion of the accelerator to an arbitrary order is presented.

Lecture 3. Equations of motion, matrix functions & transport matrices

First, we review what we already discussed: Accelerator Hamiltonian (see Appendix D).

Important (self-evident?) features are

$$\mathbf{S}^T = -\mathbf{S}; \quad \mathbf{S}^2 = -\mathbf{I}; \quad \mathbf{S}^T \mathbf{S} = \mathbf{I}; \quad \mathbf{S}^{-1} = -\mathbf{S} \quad (\text{L3-0})$$

3.1 Linear equations of motion

Thus, we finished by concluding that the first not-trivial term in the accelerator Hamiltonian expansion is a quadratic term of canonical momenta and coordinates. This Hamiltonian can be written in the matrix form (letting n be a dimension of the Hamiltonian system with n canonical pairs $\{q_i, P_i\}$)

$$H = \frac{1}{2} \sum_{i=1}^{2n} \sum_{j=1}^{2n} h_{ij}(s) x_i x_j \equiv \frac{1}{2} X^T \cdot \mathbf{H}(s) \cdot X; \quad (\text{L3-1})$$

$$X^T = [q^1 \quad P_1 \quad \dots \quad \dots \quad q^n \quad P_n] = [x_1 \quad x_2 \quad \dots \quad \dots \quad x_{2n-1} \quad x_{2n}],$$

with the self-evident feature that a symmetric matrix can be chosen

$$\mathbf{H}^T = \mathbf{H} \quad (\text{L3-2})$$

(to be exact, a quadratic form with any asymmetric matrix has zero value). The equations of motion are just a set of $2n$ linear ordinary differential equations with s -dependent coefficients:

$$\frac{dX}{ds} = \mathbf{D}(s) \cdot X; \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{H}(s). \quad (\text{L3-3})$$

One important feature of this system is that

$$\text{Trace}[\mathbf{D}] = 0, \quad (\text{L3-4})$$

(the trivial proof is based on $\text{Trace}[\mathbf{AB}] = \text{Trace}[\mathbf{BA}]$; $\text{Trace}[\mathbf{A}^T] = \text{Trace}[\mathbf{A}]$ and $(\mathbf{SH})^T = -(\mathbf{HS})$). i.e., the Wronskian determinant of the system (<http://en.wikipedia.org/wiki/Wronskian>) is equal to one. The famous Liouville theorem comes from well-known operator formula $\frac{d \det[\mathbf{W}(s)]}{ds} = \text{Trace}[\mathbf{D}]$; we do not need it here because we will have an easier method of proof.

The solution of any system of first-order linear differential equations can be expressed through its $2n$ initial conditions X_o at azimuth s_o

$$X(s_o) = X_o, \quad (\text{L3-5})$$

through the transport matrix $\mathbf{M}(s_o/s)$:

$$X(s) = \mathbf{M}(s_o|s) \cdot X_o. \quad (\text{L3-6})$$

There are two simple proofs of this theorem. The first is an elegant one: Let us consider the matrix differential equation

$$\mathbf{M}' \equiv \frac{d\mathbf{M}}{ds} = \mathbf{D}(s) \cdot \mathbf{M}; \quad (\text{L3-7})$$

with a unit matrix as its initial condition at azimuth s_o

$$\mathbf{M}(s_o) = \mathbf{I}. \quad (\text{L3-8})$$

Such solution exists⁶ and then we readily see that

$$X(s) = \mathbf{M}(s) \cdot X_o. \quad (\text{L3-7-1})$$

satisfies eq.(L3-3):

$$\frac{dX}{ds} = \frac{d\mathbf{M}(s)}{ds} \cdot X_o = \mathbf{D}(s) \cdot \mathbf{M}(s) \cdot X_o \equiv \mathbf{D}(s) \cdot X \#.$$

A more traditional approach to the same solution is to use the facts that a) there exists a solution of equation (L3-3) with arbitrary initial conditions (less-trivial statement); and, b) any linear combination of the solutions also is a solution of eq. (L3-3) (very trivial one). Considering a set of solutions of eq.(L3-3) $M_k(s)$, $k=1, \dots, 2n$, with initial conditions at azimuth s_o , then

$$M_1(s_o) = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix}; M_2(s_o) = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \\ 0 \end{bmatrix}; \dots; M_{2n}(s_o) = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix}; \quad \frac{dM_k(s)}{ds} = \mathbf{D}(s) \cdot M_k(s); \quad (\text{L3-9})$$

and their linear combination

$$X(s) = \sum_{k=1}^{2n} x_{ko} \cdot M_k(s), \quad (\text{L3-10})$$

which satisfies the initial condition (L3-5)

$$X(s_o) = \sum_{k=1}^{2n} x_{ko} \cdot M_k(s_o) = \begin{bmatrix} x_{1,0} \\ x_{2,0} \\ \dots \\ x_{2n-1,0} \\ x_{2n,0} \end{bmatrix} = X_o. \quad (\text{L3-11})$$

Now, we recognize that our solution (L3-10) is nothing other than the transport matrix eq. (L3-7-1) with matrix $\mathbf{M}(s)$ being a simple combination of $2n$ columns $M_k(s)$:

$$\mathbf{M}(s) = [M_1(s), M_2(s), \dots, M_{2n}(s)].$$

Eq. (L3-9) then makes it equivalent to eqs. (L3-7) and (L3-8). Finally, we use notion $\mathbf{M}(s_o|s)$ to clearly demonstrate that $\mathbf{M}(s_o) = \mathbf{I}$ at azimuth s_o

In differential calculus, the solution is defined as

$$\mathbf{M}(s_o|s) = \exp \left[\int_{s_o}^s \mathbf{D}(s) ds \right] = \lim_{N \rightarrow \infty} \prod_{k=1}^N (\mathbf{I} + \mathbf{D}(s_k)) \Delta s; \quad (\text{L3-12})$$

$$\Delta s = (s - s_o)/N; s_k \in \{s_o + (k-1) \cdot \Delta s, s_o + k \cdot \Delta s\}$$

The fact that the transport matrix for a linear Hamiltonian system has unit determinant (i.e., the absence of dissipation!)

⁶ Mathematically, it is nothing else but $M(s) = \lim_{N \rightarrow \infty} \prod_{k=1}^N (\mathbf{I} + \mathbf{D}(s_k)) \Delta s; \Delta s = (s - s_o)/N; s_k = s_o + k \cdot \Delta s.$

$$\det \mathbf{M} = \exp \left[\int_{s_0}^s \text{Trace}(\mathbf{D}(s)) ds \right] = 1. \quad (\text{L3-13})$$

is the first indicator of the advantages that follow. Let us consider the invariants of motion characteristic of linear Hamiltonian systems, i.e., invariants of the symplectic phase space⁷. Starting from the bilinear form of two independent solutions of eq. (L3-3), $X_1(s)$ and $X_2(s)$, (it is obvious that $X^T \mathbf{S} X = 0$) we show that

$$X_2^T(s) \cdot \mathbf{S} \cdot X_1(s) = X_2^T(s_0) \cdot \mathbf{S} \cdot X_1(s_0) = \text{inv}. \quad (\text{L3-14})$$

The proof is straightforward

$$\frac{d}{ds} (X_2^T \cdot \mathbf{S} \cdot X_1) = X_2^{T'} \cdot \mathbf{S} \cdot X_1 + X_2^T \cdot \mathbf{S}' \cdot X_1 + X_2^T \cdot \mathbf{S} \cdot X_1' = X_2^T \cdot ((\mathbf{SD})^T \mathbf{S} + \mathbf{SSD}) \cdot X_1 = 0.$$

Proving that transport matrices for Hamiltonian system are symplectic is very similar:

$$\mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M} = \mathbf{S}. \quad (\text{L3-15})$$

Beginning from the simple fact that the unit matrix is symplectic: $\mathbf{I}^T \cdot \mathbf{S} \cdot \mathbf{I} = \mathbf{S}$, i.e. $\mathbf{M}(s_0|s_0)$ is symplectic, and following with the proof that $\mathbf{M}^T(s_0|s) \cdot \mathbf{S} \cdot \mathbf{M}(s_0|s) = \mathbf{M}^T(s_0|s_0) \cdot \mathbf{S} \cdot \mathbf{M}(s_0|s_0) = \mathbf{S}$:

$$\frac{d}{ds} (\mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M}) = \mathbf{M}^{T'} \cdot \mathbf{S} \cdot \mathbf{M} + \mathbf{M}^T \cdot \mathbf{S}' \cdot \mathbf{M} + \mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M}' = \mathbf{M}^T \cdot ((\mathbf{SD})^T \mathbf{S} + \mathbf{SSD}) \cdot \mathbf{M} = \mathbf{0} \#$$

Symplectic square matrices of dimensions $2n \times 2n$, which include unit matrix \mathbf{I} , create a symplectic group, where the product of symplectic matrices also is a symplectic matrix.⁸ The symplectic condition (L3-15) is very powerful and should not be underappreciated. Before going further, we should ask ourselves several questions: How can the inverse matrix of \mathbf{M} be found? Are there invariants of motion to hold-on to? Can something specific be said about a real accelerator wherein there are small but all-important perturbations beyond the linear equation of motions?

As you probably surmised, the Hamiltonian method yield many answers, and is why it is so vital to research.

We can count them: The general transport matrix \mathbf{M} (solution of $\mathbf{M}' = \mathbf{D}(s) \cdot \mathbf{M}$ with arbitrary \mathbf{D}) has $(2n)^2$ independent elements. Because the symplectic condition $\mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M} - \mathbf{S} = \mathbf{0}$ represents an asymmetric matrix with n-diagonal elements equivalently being zeros, and the conditions above and below the diagonal are identical – then only the $n(2n-1)$ condition remains and only the $n(2n+1)$ elements are independent. For $n=1$ (1D) there is only one condition, for $n=2$ there are 6 conditions, and $n=3$ (3D) there are 15 conditions. Are these facts of any use in furthering this exploration?

First, symplecticity makes the matrix determinant to be unit⁹:

⁷ Phase space is defined as the $2n$ -dimensional space of canonical variable $\{q_i, p_i\}$, that is, the space where this Hamiltonian system evolves.

⁸ Group G is defined as a set of elements, with a definition of a product of any two elements of the group; $P = A \cdot B \in G$; $A, B \in G$. The product must satisfy the associative law: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$; there is an unit element in the group $I \in G$; $I \cdot A = A \cdot I = A$; $\forall A \in G$; and inverse elements: $\forall A \in G$; $\exists B$ (called A^{-1}) $\in G$: $A^{-1} A = A A^{-1} = I$.

⁹ Look at a simple $n=1$ case with 2×2 matrices to verify that the symplectic product is reduced to determine

$$\mathbf{M}_{2 \times 2} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}; \mathbf{S}_{2 \times 2} = \sigma; \Rightarrow \mathbf{M}^T \cdot \sigma \cdot \mathbf{M} = \det \mathbf{M} \cdot \sigma \quad (\text{Note-4})$$

$$\det[\mathbf{M}^T(s) \cdot \mathbf{S} \cdot \mathbf{M}(s)] = \det \mathbf{S} \rightarrow (\det \mathbf{M}(s))^2 = 1 \rightarrow \det \mathbf{M} = \pm 1; \quad \det \mathbf{M}(0) = 1 \rightarrow \det \mathbf{M} = 1 \quad \#$$

i.e., it preserves the 2n-D phase space volume occupied by the ensemble of particles (system):

$$\int \prod_{i=1}^n dq_i dP^i = inv \quad (\text{L3-16})$$

The other invariants preserved by symplectic transformations are called Poincaré invariants and are the sum of projections onto the appropriate over-manifold in two, four... (2n-2) dimensions:

$$\sum_{i=1}^n \iint dq_i dP^i = inv; \quad \sum_{i \neq j} \iint \iint dq_i dP^i dq_j dP^j = inv \dots \quad (\text{L3-17})$$

For example, matrix \mathbf{M} can be represented as n^2 combinations of 2x2 matrices M_{ij} :

$$\mathbf{M} = \begin{bmatrix} M_{11} & \dots & M_{1n} \\ \dots & \dots & \dots \\ M_{n1} & \dots & M_{nn} \end{bmatrix} \quad (\text{L3-18})$$

Using equation (Note-4), we easily demonstrate the requirement for the symplectic condition (L3-15) is that the sum of determinants in each row of these 2x2 matrices is equal to one; the same is true for the columns:

$$\sum_{i=1}^n \det[M_{ij}] = \sum_{j=1}^n \det[M_{ij}] = 1 \quad (\text{L3-19})$$

with a specific prediction for decoupled matrices, which are block diagonal:

$$\mathbf{M} = \begin{bmatrix} M_{11} & 0 \dots & 0 \\ 0 & \dots & 0 \\ 0 & 0 \dots & M_{nn} \end{bmatrix}; \quad \det[M_{ii}] = 1. \quad (\text{L3-20})$$

Other trivial and useful features:

$$\mathbf{M} = [C_1 \quad C_2 \quad \dots \quad C_{2n-1} \quad C_{2n}] \Rightarrow C_{2k+1}^T S C_{2k} = -C_{2k}^T S C_{2k+1} = 1, \text{ others are } 0 \quad (\text{L3-18-1})$$

$$\mathbf{M} = \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_{2n-1} \\ L_{2n} \end{bmatrix} \Rightarrow -L_{2k+1}^T S L_{2k} = L_{2k}^T S L_{2k+1} = 1, \text{ others are } 0 \quad (\text{L3-18-2})$$

We could go further, but we will stop here by showing the most incredible feature of symplectic matrices, viz., that it is easy to find their inverse (recall there is no general rule for inverting a 2n x 2n matrix!) Thus, multiplying eq. (L3-15) from left by $-\mathbf{S}$ we get

$$-\mathbf{S} \cdot \mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M} = \mathbf{I} \Rightarrow \mathbf{M}^{-1} = -\mathbf{S} \cdot \mathbf{M}^T \cdot \mathbf{S} \quad (\text{L3-21})$$

As an easy exercise for 2x2 symplectic (i.e. with unit determinant – see note below) matrices, you can show that $\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (L3-21) gives $\mathbf{M}^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. It is a much less trivial task to invert

6x6 matrix; hence, the power of symplecticity allows us enact many theoretical manipulations that otherwise would be impossible. Obviously, and easy to prove, transposed symplectic and inverse symplectic matrices also are also symplectic:

$$\mathbf{M}^{-1T} \cdot \mathbf{S} \cdot \mathbf{M}^{-1} = \mathbf{S}; \quad \mathbf{M} \cdot \mathbf{S} \cdot \mathbf{M}^T = \mathbf{S}. \quad (\text{L3-21-1})$$

3.2 Calculating matrices

Next, we focus on the question of how matrices are calculated. Equation (L3-12) imparts the general idea that they can be integrated piece-wise wherein the coefficients in the Hamiltonian expansion do not change significantly. In practice, accelerators are built from elements, which, to a certain extent, offers such conditions. The typical elements in high-energy accelerators contributing to the linear part of the equations are the drift (free space, vacuum), the dipoles with and without transverse gradient, the quadrupoles (both normal and SQ), and the RF cavities for acceleration and bunching. Except for the last, the typical elements are magnetic and DC (or varying very slowly compared with the passing or turn-around time for particles). Electric elements are rare, so simplifying the Hamiltonian to some degree:

$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_3 - yP_1) + \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x \delta; \quad (\text{L2-46})$$

$$\frac{F}{p_o} = \left[K^2 - \frac{e}{p_o c} \frac{\partial B_y}{\partial x} + \frac{1}{2} \left(\frac{e B_s}{p_o c} \right)^2 \right] \frac{G}{p_o} = \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial y} + \frac{1}{2} \left(\frac{e B_s}{p_o c} \right)^2 \right]; \quad \frac{2N}{p_o} = \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial x} - \frac{e}{p_o c} \frac{\partial B_y}{\partial y} \right]; \quad (\text{L2-47})$$

$$\frac{L}{p_o} = \kappa + \frac{e}{2p_o c} B_s; \quad \frac{U}{p_o} = \frac{e}{p_o c^2} \frac{\partial E_s}{\partial t}; \quad g_x = -K \frac{c}{v_o};$$

In general, many elements of accelerators are designed to keep constant the coefficients of Hamiltonian expansion, with exception of edges. Here we will consider simple rigid-edge elements, where the magnetic field ends abruptly (compared with the wavelength of betatron oscillations).

Hence, initially we will explore a general way of calculating matrices, and then consider few examples. When the matrices \mathbf{D} are piece-wise constant and the \mathbf{D} from different elements do not commute, we can write

$$\mathbf{M}(s_o|s) = \prod_i \mathbf{M}(s_{i-1}|s_i); \quad \mathbf{M}(s_{i-1}|s) = \prod_{\text{elements}} \exp[\mathbf{D}_i(s - s_{i-1})] \quad (\text{L3-22})$$

The definition of the matrix exponent is very simple

$$\exp[\mathbf{A}] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}; \quad \exp[\mathbf{D} \cdot s] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{D}^k s^k}{k!} \quad (\text{L3-23})$$

According to the general theorem of Hamilton-Kelly, the matrix is a root of its characteristic equation:

$$d(\lambda) = \det[\mathbf{D} - \lambda \mathbf{I}]; \quad d(\lambda_k) = 0 \quad (\text{L3-24})$$

$$d(\mathbf{D}) \equiv 0 \quad (\text{L3-25})$$

i.e., a root of a polynomial of order $\leq 2n$. There is a theorem in theory of polynomials (rather easy to prove) that any polynomial $p_1(\mathbf{x})$ of power n can be expressed via any polynomial $p_2(\mathbf{x})$

of power $m < n$ as

$$p_1(x) = p_2(x) \cdot d(x) + r(x)$$

where $r(x)$ is a polynomial of power less than m . Accordingly, series (L3-23) can be always truncated to

$$\exp[\mathbf{D}] = I + \sum_{k=1}^{2n-1} c_k \mathbf{D}^k, \quad (\text{L3-26})$$

with the remaining daunting task of finding coefficients c_k ! There are two ways of doing this; one is a general, and the other is case specific, but an easy one. Starting from a specific case when the matrix \mathbf{D} is nilpotent ($m < 2n+1$), i.e.,

$$\mathbf{D}^m = 0.$$

In this case, $\mathbf{D}^{m+j} = 0$ the truncation is trivial:

$$\exp[D] = I + \sum_{k=1}^{m-1} \frac{D^k}{k!}. \quad (\text{L3-27})$$

We lucky to have such a beautiful case in hand – a drift, where all fields are zero and $K=0$ and $\kappa=0$:

$$\tilde{h} = \frac{\pi_1^2 + \pi_3^2}{2} + \frac{\pi_\delta^2}{2} \cdot \frac{m^2 c^2}{p_o^2}; \quad \mathbf{D} = \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{bmatrix}; D = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; D_2 = \begin{bmatrix} 0 & \frac{m^2 c^2}{p_o^2} \\ 0 & 0 \end{bmatrix}; \quad (\text{L3-28})$$

where it is easy to check: $D^2 = 0$. Hence, the 6x6 matrix of drift with length l will be

$$\mathbf{M}_{drift} = \exp[\mathbf{D} \cdot l] = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{D}^k l^k}{k!} = \mathbf{I} + \mathbf{D} \cdot l = \begin{bmatrix} M_t & 0 & 0 \\ 0 & M_t & 0 \\ 0 & 0 & M_\tau \end{bmatrix}; M_t = \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix}; M_\tau = \begin{bmatrix} 1 & l/(\beta_o \gamma_o)^2 \\ 0 & 1 \end{bmatrix}; \quad (\text{L3-29})$$

The general evaluation of the matrix exponent in (L3-22) is straightforward using the eigen values of the D-matrix:

$$\det[\mathbf{D} - \lambda \cdot \mathbf{I}] = \det[\mathbf{SH} - \lambda \cdot \mathbf{I}] = 0 \quad (\text{L3-30})$$

When the eigen values are all different ($2n$ numerically different eigen values, $\lambda_i = \lambda_j \Rightarrow i = j$, no degeneration, i.e., \mathbf{D} can be diagonalized),

$$\mathbf{D} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{-1}; \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ & & \dots & 0 \\ 0 & 0 & 0 & \lambda_{2n} \end{pmatrix};$$

we can use Sylvester's formula that is correct for any analytical $f(\mathbf{D})$, http://en.wikipedia.org/wiki/Sylvester's_formula for evaluating (L3-22):

$$\exp[\mathbf{D}s] = \sum_{k=1}^{2n} e^{\lambda_k s} \prod_{j \neq k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \quad (\text{L3-31})$$

Another easy case is when \mathbf{D} can be diagonalized, even though the number of different eigen values is $m < 2n$ (there is degeneration, i.e. some eigen values have multiplicity > 1). We can use again simple Sylvester's formula (L3-31) again, which just has fewer elements (m instead of $2n$):

$$\exp[\mathbf{D}s] = \sum_{k=1}^m e^{\lambda_k s} \prod_{\lambda_j \neq \lambda_k} \frac{\mathbf{D} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \quad (\text{L3-32})$$

Furthermore, in most general case when matrix \mathbf{D} cannot be diagonalized (i.e. there is degeneracy, some of eigen values have multiplicity, and \mathbf{D} can be only reduced to a Jordan form) we can still write a specific from (generalization of Sylvester's formula):

$$\exp[\mathbf{D}s] = \sum_{k=1}^m \left[e^{\lambda_k s} \prod_{i \neq k} \left\{ \frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \sum_{j=0}^{n_k-1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\}^{n_i} \sum_{p=0}^{n_k-1} \frac{s^p}{p!} (\mathbf{D} - \lambda_k \mathbf{I})^p \right] \quad (\text{L3-33})$$

where $n_k < 2n$ is so called height of the eigen value λ_k . Details of the definitions and as well the proof of Sevester's formulae are given in Appendix E. It is also shown there that n_k can be replaced in (L3-33) by any number $n_k > n_k$ – it will add only term, which are zeros, but can make (L3-33) look more uniform. One of the logical choices will be $n_k = \max\{n_k\}$. The other natural choice will be $n_k = 2n+1-m$, especially if computer does it for you. Eq. (L3-33) is a bit uglier than (L3-31), but still can be used with some elegance.

Eigen values split into pairs with the opposite sign because it is a Hamiltonian system:

$$\begin{aligned} \det[\mathbf{S}\mathbf{H} - \lambda \cdot \mathbf{I}] &= \det[\mathbf{S}\mathbf{H} - \lambda \cdot \mathbf{I}]^T = \det[-\mathbf{H}\mathbf{S} - \lambda \cdot \mathbf{I}] = \\ &(-1)^{2n} \det[\mathbf{H}\mathbf{S} + \lambda \cdot \mathbf{I}] = \det(\mathbf{S}^{-1}[\mathbf{H}\mathbf{S} + \lambda \cdot \mathbf{I}]\mathbf{S}) = \det[\mathbf{S}\mathbf{H} + \lambda \cdot \mathbf{I}] \end{aligned} \quad (\text{L3-35})$$

First, it makes finding eigen values a easier problem, because characteristic equation is bi-quadratic:

$$\det[\mathbf{D} - \lambda \mathbf{I}] = \prod_i (\lambda_i - \lambda)(-\lambda_i - \lambda) = \prod_i (\lambda^2 - \lambda_i^2) = 0. \quad (\text{L3-35-1})$$

For accelerator elements it is of paramount importance, 1D case is reduces to trivial (L3-37), 2D case is reduced to solution of quadratic equation and 3D case (6D phase space) required to solve cubic equation. For analytical work it gives analytical expressions – compare it with attempt to write analytical formula for roots of a generic polynomial of 6-order? It simply does not exist! Thus, we have an extra gift for accelerator physics – the roots can be written and studied!

It is also allow us to simplify (L3-31) into

$$\begin{aligned} \exp[\mathbf{D}s] &= \left\{ \sum_{k=1}^n e^{\lambda_k s} \frac{\mathbf{D} + \lambda_k \mathbf{I}}{2\lambda_k} \prod_{j \neq k} \left(\frac{\mathbf{D}^2 - \lambda_j^2 \mathbf{I}}{\lambda_k^2 - \lambda_j^2} \right) - e^{-\lambda_k s} \frac{\mathbf{D} - \lambda_k \mathbf{I}}{2\lambda_k} \prod_{j \neq k} \left(\frac{\mathbf{D}^2 - \lambda_j^2 \mathbf{I}}{\lambda_k^2 - \lambda_j^2} \right) \right\} \\ \exp[\mathbf{D}s] &= \sum_{k=1}^n \left(\frac{e^{\lambda_k s} + e^{-\lambda_k s}}{2} \mathbf{I} + \frac{e^{\lambda_k s} - e^{-\lambda_k s}}{2\lambda_k} \mathbf{D} \right) \prod_{j \neq k} \left(\frac{\mathbf{D}^2 - \lambda_j^2 \mathbf{I}}{\lambda_k^2 - \lambda_j^2} \right) \end{aligned} \quad (\text{L3-36})$$

where index k goes only through n pairs of $\{\lambda_k, -\lambda_k\}$. While (L3-36) does not look simpler, it really makes it easier (4 times less calculations) when we do it by hands... For example we can look at 1D case. First, we can easily see that

$$\lambda_1 = -\lambda_2 = \lambda; \quad \lambda^2 = -\det[\mathbf{D}] \quad (\text{L3-37})$$

Thus, it is non-degenerated case only when $\det[\mathbf{D}] \neq 0$. (L3-31) give us a simple two-piece expression :

$$\exp[\mathbf{D}s] = e^{\lambda s} \frac{\mathbf{D} - \lambda \mathbf{I}}{2\lambda} - e^{-\lambda s} \frac{\mathbf{D} + \lambda \mathbf{I}}{2\lambda} \quad (\text{L3-38})$$

while (L3-36) bring it home right away:

$$\begin{aligned}\exp[\mathbf{D}s] &= \mathbf{I} \cdot \frac{e^{\lambda s} + e^{-\lambda s}}{2} + \mathbf{D} \frac{e^{\lambda s} - e^{-\lambda s}}{2\lambda}; \\ \exp[\mathbf{D}s] &= \mathbf{I} \cdot \cosh|\lambda|s + \frac{\mathbf{D} \sinh|\lambda|s}{|\lambda|}; \quad \det[\mathbf{D}] < 0; \quad |\lambda| = \sqrt{-\det[\mathbf{D}]} \\ \exp[\mathbf{D}s] &= \mathbf{I} \cdot \cos|\lambda|s + \frac{\mathbf{D} \sin|\lambda|s}{|\lambda|}; \quad \det[\mathbf{D}] > 0; \quad |\lambda| = \sqrt{\det[\mathbf{D}]}\end{aligned}\tag{L3-39}$$

The case $\det[\mathbf{D}] = 0$ means in this case that \mathbf{D} is nilpotent: eqs (L3-24-25) look like follows

$$\det \mathbf{D} = 0 \Rightarrow \lambda_1 = -\lambda_2 = 0; \quad d(\lambda) = \det[\mathbf{D} - \lambda \mathbf{I}] = (\lambda_1 - \lambda)(-\lambda_1 - \lambda) = \lambda^2 \Rightarrow \mathbf{D}^2 = 0\tag{L3-40}$$

hence

$$\exp[\mathbf{D}s] = \mathbf{I} + \mathbf{D}s; \quad \det[\mathbf{D}] = 0;\tag{L3-41}$$

Naturally, (L3-41) is result of full-blown degenerated case – eq. (L3-33), but it also can be obtained as a limit case of (L3-39) when $|\lambda| \rightarrow 0$.

The value of this approach to matrix calculation is that we do not need to memorize all the different ways of deriving the matrices of various elements in accelerators, and ways of solving a myriad of systems of 2, 4, 6... linear differential equations. Just a smart “coach potato principle” allover again....

The elements of 6x6, 4x4, or 2x2 accelerator matrixes (often called \mathbf{R} or \mathbf{T}) are numerated as R_{ij} , where i is the line number and j is the column number. For example, R_{56} will signify an increment in τ (-arrival time by c) caused by the particle’s energy change, δ . Let’s look at most trivial case of decoupled transverse motion.

Most accelerators have a flat orbit ($\kappa=0$), avoid longitudinal fields ($B_s=0$), and do not have the SQ-quadrupole ($N=0$). Let us examine a magnetic element (no RF field) and a field in vacuum, where

$$\vec{\nabla} \times \vec{B} = 0 \Rightarrow \frac{\partial B_y}{\partial x} = \frac{\partial B_x}{\partial y}.$$

This renders the one-liner Hamiltonian: (the momenta are normalized)

$$\tilde{h} = \left(\frac{\pi_3^2}{2} + K_1 \frac{y^2}{2} \right) + \frac{\pi_1^2}{2p_o} + \left[K^2 - K_1 \right] \frac{x^2}{2} + \frac{\pi_\delta^2}{2} \cdot \frac{m^2 c^2}{p_o^2} - K \frac{c}{v_o} x \pi_\delta; \quad K_1 = \frac{e}{p_o c} \frac{\partial B_y}{\partial x};\tag{L3-42}$$

with a clearly separated vertical (y) part of motion. In the presence of the curvature \mathbf{K} , i.e., a non-zero dipole field in the reference orbit, both the longitudinal and horizontal (x) degrees of freedom remain coupled. In a quadrupole $\mathbf{K}=0$, the Hamiltonian is completely decoupled into three degrees of freedom:

$$\tilde{h} = \left(\frac{\pi_3^2}{2} + K_1 \frac{y^2}{2} \right) + \left(\frac{\pi_1^2}{2} - K_1 \frac{x^2}{2} \right) + \frac{\pi_\delta^2}{2} \cdot \frac{m^2 c^2}{p_o^2}; \quad K_1 = \frac{e}{p_o c} \frac{\partial B_y}{\partial x};\tag{L3-43}$$

The matrix in the longitudinal direction is the same as that for a drift (L3-29), while the x and y matrices are given by (L3-39). Depending on the sign of the gradient $\partial B_y / \partial x$, the quadrupole focuses in x and defocuses in y , or vice versa:

$$D_x = \begin{bmatrix} 0 & 1 \\ K_1 & 0 \end{bmatrix}; \quad D_y = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix}; \quad \phi = s\sqrt{K_1}$$

$$M_F = \begin{bmatrix} \cos \phi & \sin \phi / \sqrt{K_1} \\ -\sqrt{K_1} \sin \phi & \cos \phi \end{bmatrix}; \quad M_D = \begin{bmatrix} \cosh \phi & \sinh \phi / \sqrt{K_1} \\ \sqrt{K_1} \sinh \phi & \cosh \phi \end{bmatrix}$$
(L3-44)

It is worth noting that there is no difference if we use momentum and coordinates, not x, x' .

$$\tilde{h} = \left(\frac{P_3^2}{2p_o} + p_o K_1 \frac{y^2}{2} \right) + \left(\frac{P_1^2}{2p_o} - p_o K_1 \frac{x^2}{2} \right) + \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2}; \quad K_1 = \frac{e}{p_o c} \frac{\partial B_y}{\partial x};$$
(L3-45)

$$D_x = \begin{bmatrix} 0 & 1/p_o \\ p_o K_1 & 0 \end{bmatrix}; \quad D_y = \begin{bmatrix} 0 & 1/p_o \\ -p_o K_1 & 0 \end{bmatrix}; \quad \phi = s\sqrt{K_1}$$

$$M_F = \begin{bmatrix} \cos \phi & \sin \phi / p_o \sqrt{K_1} \\ -p_o \sqrt{K_1} \sin \phi & \cos \phi \end{bmatrix}; \quad M_D = \begin{bmatrix} \cosh \phi & \sinh \phi / p_o \sqrt{K_1} \\ p_o \sqrt{K_1} \sinh \phi & \cosh \phi \end{bmatrix}$$
(L3-46)

As we can see, this is not a more complicated that using x, x' , but definitely correct for any accelerator.

Matrix of general DC accelerator element (including twisted quads or helical wiggler) can be found using our recipe. With all diversity of possible elements on accelerators, DC (or almost DC) magnets play the most prominent role. In this case energy of the particle stays constant and we can use reduced variables. Furthermore, large number of terms in the Hamiltonian simply disappear and from the previous lecture we have:

$$\tilde{h}_n = \frac{\pi_1^2 + \pi_3^2}{2} + f \frac{x^2}{2} + n \cdot xy + g \frac{y^2}{2} + L(x\pi_3 - y\pi_1) + \frac{\pi_o^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + g_x x \pi_o + g_y y \pi_o; \quad (\text{L2-46-n})$$

Even though it is tempting to remove electric field, it does not either help or hurt in general case of an element. Hence, we will keep DC transverse electric fields. We also assume that these fields are in vacuum and $\frac{\partial B_y}{\partial x} = \frac{\partial B_x}{\partial y}$, $\frac{\partial E_x}{\partial x} + K E_x + \frac{\partial E_y}{\partial y} = 0$:

$$f = K^2 - \frac{e}{p_o c} \frac{\partial B_y}{\partial x} - \frac{e}{p_o v_o} \frac{\partial E_y}{\partial y} + \left(\frac{e B_s}{2 p_o c} \right)^2 + \left(\frac{m e E_x}{p_o^2} \right)^2;$$

$$g = \frac{e}{p_o c} \frac{\partial B_y}{\partial x} + \frac{e}{p_o v_o} \frac{\partial E_y}{\partial y} + \left(\frac{e B_s}{2 p_o c} \right)^2 + \left(\frac{m e E_z}{p_o^2} \right)^2; \quad ; (\text{L3-47})$$

$$2n = \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial x} - \frac{e}{p_o c} \frac{\partial B_y}{\partial y} \right] - K \cdot \frac{e}{p_o c} B_x - \frac{e}{p_o v_o} \left(\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} \right) - 2K \frac{e E_y}{p_o v_o} + \left(\frac{m e E_z}{p_o^2} \right) \left(\frac{m e E_x}{p_o^2} \right)$$

$$L = \kappa + \frac{e}{2 p_o c} B_s; \quad g_x = \frac{(m c)^2 \cdot e E_x}{p_o^3} - K \frac{c}{v_o}; \quad g_y = \frac{(m c)^2 \cdot e E_y}{p_o^3};$$

In the absence of longitudinal electric field, the momentum P_2 is constant as well $\pi_o = const$, $\delta = const$. The fact that particle's energy does not change in such element is rather obvious¹⁰:

¹⁰ It is completely correct for magnetic elements. Presence of electric field makes it less obvious, but it comes from the fact that Hamiltonian does not depend on time!

$$\pi_o' = -\frac{\partial h}{\partial \tau} = 0.$$

Equations of motion become specific:

$$\mathbf{X}^T = [x, \pi_1, y, \pi_3, \tau, \pi_o] = [X^T, \tau, \pi_o]; \quad X^T = [x, \pi_1, y, \pi_3], \quad (\text{L3-48})$$

$$\frac{d\mathbf{X}}{ds} = \mathbf{D}(s) \cdot \mathbf{X}; \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{H}(s) = \begin{bmatrix} 0 & 1 & -L & 0 & 0 & 0 \\ -f & 0 & -n & -L & 0 & g_x \\ L & 0 & 0 & 1 & 0 & 0 \\ -n & L & -g & 0 & 0 & g_y \\ g_x & 0 & g_y & 0 & 0 & m^2 c^2 / p_o^2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad (\text{L3-49})$$

and can be rewritten in a slightly different (just deceptively looking better) way:

$$\frac{dX}{ds} = D \cdot X + \pi_o \cdot C; \quad (\text{L3-50})$$

$$\frac{d\tau}{ds} = g_x x + g_y y + \pi_o \cdot m^2 c^2 / p_o^2; \quad D = \begin{bmatrix} 0 & 1 & -L & 0 \\ -f & 0 & -n & -L \\ L & 0 & 0 & 1 \\ -n & L & -g & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 0 \\ g_x \\ 0 \\ g_y \end{bmatrix}.$$

Hence, solution for transverse motion (4-vector) in such an element can be written as combination general solution of homogeneous equation plus specific solution of inhomogeneous one:

$$X = M(s) \cdot X_o + \pi_o \cdot R(s); \quad M = e^{D(s-s_o)}; \quad R' = D \cdot R + C; \quad R(s_o) = 0. \quad (\text{L3-51})$$

It worth noting that C=0 as soon as there is no field on the orbit – E=0, B=0. In this case R=0.

Before finding 4x4 matrixes M and vector R, let's see what we will know about the 6x6 matrix after that. First, the obvious:

$$\mathbf{M}_{6 \times 6} = \begin{bmatrix} \mathbf{M}_{4 \times 4} & 0 & R \\ R_{51} & R_{52} & R_{53} & R_{54} & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{L3-52})$$

with a natural question of what are non-trivial R_{5k} elements? Usually these elements, with exception of R_{56} are not even mentioned in most of textbooks. Fortunately for us, Mr. Hamiltonian gives us a hand in the form of symplecticity of transport matrixes. Using (L3-18) and (L3-18-1) we can find that:

$$\mathbf{M}_{6 \times 6}^T \mathbf{S} \mathbf{M}_{6 \times 6} = \begin{bmatrix} \mathbf{M}_{4 \times 4}^T & L^T & 0 \\ 0 & 1 & 0 \\ R^T & R_{56} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{S}_{4 \times 4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{4 \times 4} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{4 \times 4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{M}_{4 \times 4}^T \mathbf{S}_{4 \times 4} & 0 & L^T \\ 0 & 0 & 1 \\ R^T \mathbf{S}_{4 \times 4} & -1 & R_{56} \end{bmatrix} \begin{bmatrix} \mathbf{M}_{4 \times 4} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{4 \times 4}^T \mathbf{S}_{4 \times 4} \mathbf{M}_{4 \times 4} & 0 & 0 \\ 0 & 0 & 1 \\ R^T \mathbf{S}_{4 \times 4} \mathbf{M}_{4 \times 4} - L & -1 & R^T \mathbf{S}_{4 \times 4} R \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{4 \times 4} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

where we used $L = [R_{51}, R_{52}, R_{53}, R_{54}]$. We should note what $\mathbf{X}^T \mathbf{S} \mathbf{X} = \mathbf{0}$ for any vector, $\mathbf{M}^T_{4 \times 4} \mathbf{S}_{4 \times 4} \mathbf{M}_{4 \times 4} = \mathbf{S}_{4 \times 4}$ and only non-trivial condition from the equation above is:

$$R^T \mathbf{S}_{4 \times 4} \mathbf{M}_{4 \times 4} - L = 0$$

which gives us very valuable dependence of the arrival time on the transverse motions:

$$L = R^T \mathbf{S}_{4 \times 4} \mathbf{M}_{4 \times 4}; \text{ or } L^T = -\mathbf{M}^T_{4 \times 4} \mathbf{S}_{4 \times 4} R. \quad (\text{L3-53})$$

Element R_{56} is decoupled from the symplectic condition in this case and should be determined by direct integration - no magic here:

$$\tau(s) = \tau(s_o) + \pi_o \cdot \left\{ m^2 c^2 / p_o (s - s_o) + 2 \int_{s_o}^s (g_x R(s)_{16} + g_y R_{36}(s)) ds \right\} \quad (\text{L3-54})$$

$$R_{56} = m^2 c^2 / p_o (s - s_o) + 2 \int_{s_o}^s (g_x R(s)_{16} + g_y R_{36}(s)) ds$$

Let's find the solutions for 4x4 matrixes of arbitrary element. First, let solve characteristic equation for D:

$$\det[D - \lambda I] = \lambda^4 + \lambda^2 (f + g + 2L^2) + fg + L^4 - L^2 (f + g) - n^2 = 0 \quad (\text{L3-55})$$

with easy roots:

$$\lambda^2 = a \pm b; \quad a = -\frac{f + g + 2L^2}{2}; \quad b^2 = \frac{(f - g)^2}{4} + 2L^2 (f + g) + n^2 \quad (\text{L3-56})$$

Before starting classification of the cases, let's note that

$$f + g = K^2 + 2 \left(\frac{eB_s}{2p_o c} \right)^2 + \left(\frac{meE_x}{p_o^2} \right)^2 + \left(\frac{meE_z}{p_o^2} \right)^2 \geq 0$$

i.e. $a \leq 0$; $b^2 \geq 0$; $\text{Im}(b) = 0$. The solutions can be classified as following: remember that the full set of eigen values is $\lambda_1, -\lambda_1, \lambda_2, -\lambda_2$:

- I. $\lambda_1 = \lambda_2 = 0$; $a = 0$; $b = 0$;
- II. $\lambda_1 = \lambda_2 = i\omega$; $a = -\omega^2$; $b = 0$;
- III. $\lambda_1 = 0$; $\lambda_2 = i\omega$; $a + b = 0$; $2b = \omega^2$
- IV. $\lambda_1 = i\omega_1$; $\lambda_2 = i\omega_2$; $\omega_1^2 = -a - b$; $\omega_2^2 = -a + b$; $|a| > b$
- V. $\lambda_1 = i\omega_1$; $\lambda_2 = \omega_2$; $\omega_1^2 = -a - b$; $\omega_2^2 = b - a$; $b > |a|$

Before going to case-by-case calculations, lets use Sylvester's formulae and try to find solution of inhomogeneous equation:

$$\frac{d\mathbf{R}}{ds} = \mathbf{D} \cdot \mathbf{R} + \mathbf{C}; \quad \mathbf{R}(0) = \mathbf{0}. \quad (\text{L3-57})$$

When matrix $\det \mathbf{D} \neq 0$, (L3-57) can be inversed using a $\mathbf{R} = A + e^{Ds} \cdot B$ as a guess and the boundary condition $\mathbf{R}(0) = \mathbf{0}$:

$$\mathbf{R} = (\mathbf{M}_{4 \times 4}(s) - \mathbf{I}) \cdot \mathbf{D}^{-1} \cdot \mathbf{C} \quad (\text{L3-58})$$

is the easiest solution. Prove is just straight forward:

$$\mathbf{R}' = \mathbf{D} \cdot \mathbf{M}_{4 \times 4}^{-1} \cdot \mathbf{C};$$

$$\mathbf{D} \cdot (\mathbf{M} - \mathbf{I}) \cdot \mathbf{D}^{-1} \cdot \mathbf{C} + \mathbf{C} = \mathbf{D} \cdot \mathbf{M}_{4 \times 4}^{-1} \cdot \mathbf{C} \quad \#$$

In all cases we can use method of variable constants to find it:

$$\frac{dR}{ds} = R' = \mathbf{D} \cdot R + \mathbf{C}; \quad \mathbf{M}' = \mathbf{D}\mathbf{M};$$

$$R = \mathbf{M}(s)A(s) \Rightarrow \mathbf{M}'A + \mathbf{M}A' = \mathbf{D}\mathbf{M}A + \mathbf{C}; \quad R(0) = 0 \Rightarrow A_0 = 0 \quad (\text{L-53})$$

$$A' = \mathbf{M}^{-1}(s)\mathbf{C} \Rightarrow A = \int_0^s \mathbf{M}^{-1}(z)\mathbf{C}dz = \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot \mathbf{C}; \quad R = e^{\mathbf{D}s} \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot \mathbf{C}$$

It is important to remember that $\mathbf{M}^{-1}(\mathbf{s})$ is just the $\mathbf{M}(-\mathbf{s}) = e^{-\mathbf{D}s}$. Hence in all our formulae for matrixes from previous lectures we need to replace s by $-\mathbf{s}$ to get $\mathbf{M}^{-1}(\mathbf{s})$. Other vice, we have to use general formula (L3-33) for the homogeneous solution and use method of variable constants (see Appendix F) to find it:

$$R(s) = \sum_{k=1}^m \left\{ \prod_{i \neq k} \left[\frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right] \sum_{j=0}^{n_k-1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\} \sum_{n=0}^{n_k-1} (\mathbf{D} - \lambda_k \mathbf{I})^n \frac{s^n}{n!} \cdot \sum_{p=0}^{n_k-1} (-1)^{p+1} (\mathbf{D} - \lambda_k \mathbf{I})^p \cdot \mathbf{C} \cdot \left[\sum_{q=0}^{p1} \frac{s^{p-q}}{(p-q)! \lambda_k^{q+1}} - \frac{e^{\lambda_k}}{\lambda_k^{p+1}} \right] \quad (\text{L3-60})$$

In all specific cases I, II, III, IV and V, integrating (L-53) directly is usually easier that using general form of (L3-60).

$$\frac{f + g + 2L^2}{2} = 0; \quad \frac{(f - g)^2}{4} + 2L^2(f + g) + n^2 = 0;$$

Case I.

$$f + g = pos^2 \geq 0 \Rightarrow (f - g)^2 = 0; \quad L^2(f + g) = 0; \quad n^2 = 0$$

$$f + g + 2L^2 = pos^2 + 2L^2 = 0 \Rightarrow L = 0; \quad f + g = 0 \Rightarrow$$

$$f - g = 0 \Rightarrow f = g = L = n = 0!!!$$

means that there is nothing in the Hamiltonian but p^2 – is this the drift section matrix of which we already know. Hence, there is not curvature as well and $R=0$.

$$\mathbf{M}_{4 \times 4} = \begin{bmatrix} 1 & s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & s \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad \mathbf{R} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (\text{L3-I-1})$$

The only not trivial (ha-ha – it is also as trivial as it can be) is R_{56} :

$$R_{56} = \frac{m^2 c^2}{p_o^2} s \quad (\text{L3-I-2})$$

we already had seen it when studied nilpotent case...

Case II: $b = \frac{(f-g)^2}{4} + 2L^2(f+g) + n^2 = 0;$

$$f = g; n = 0 \text{ and } L^2(f+g) = L^2(K^2 + \Omega^2 + El^2) = 0; \Omega = eB_s / p_o c; E_{\perp} = 0.$$

i.e. there are two cases: $L=0$ or $f+g=0$.

If both are equal zero, i.e. $f+g=0; L=0$, this is equivalent to the case I above.

Case II a: $f+g=0, K \neq 0, B_s=0 \rightarrow L=K$. Thus, this is just a drift (straight section) with rotation, whose matrix is trivial: Drift + rotation. There is not transverse force – hence $R=0$.

$$\mathbf{M}_{4 \times 4} = \begin{bmatrix} M_d \cdot \cos \kappa s & -M_d \cdot \sin \kappa s \\ M_d \cdot \sin \kappa s & M_d \cdot \cos \kappa s \end{bmatrix}; M_d = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}; \mathbf{R} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (\text{L3-IIa-1})$$

R_{56} is as for a drift:

$$R_{56} = \frac{m^2 c^2}{2} s \quad (\text{L3-IIa-2})$$

Case II b: $L=0; f = g = (K^2 + \Omega^2)/2; \kappa = -\Omega$; i.e. the motion is uncoupled:

$$D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -f & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -f & 0 \end{bmatrix}; C = \begin{bmatrix} 0 \\ g_x \\ 0 \\ g_y \end{bmatrix}.$$

$$\mathbf{M}_{4 \times 4} = \begin{bmatrix} M & 0 \\ 0 & M \end{bmatrix}; M = \begin{bmatrix} \cos \omega s & \sin \omega s / \omega \\ -\omega \sin \omega s & \cos \omega s \end{bmatrix} \quad (\text{L3-IIb-1})$$

Here we may have non-zero R: yes, it may be! It is simple integrals to be taken care of:

$$C_{x,y} = -g_{x,y} \begin{bmatrix} 0 \\ 1 \end{bmatrix}; M^{-1}(z) = \begin{bmatrix} \cos \omega z & -\sin \omega z / \omega \\ \omega \sin \omega z & \cos \omega z \end{bmatrix} C_{x,y} = g_{x,y} \begin{bmatrix} \sin \omega z / \omega \\ -\cos \omega z \end{bmatrix};$$

$$\int_0^s \mathbf{M}^{-1}(z) C_{x,y} dz = g_{x,y} \begin{bmatrix} \int_0^s \sin(\omega z) dz / \omega \\ -\int_0^s \cos(\omega z) dz \end{bmatrix} = \frac{g_{x,y}}{\omega} \begin{bmatrix} (1 - \cos \omega s) / \omega \\ -\sin(\omega s) \end{bmatrix}$$

$$C_{x,y} = -g_{x,y} \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \int_0^s \mathbf{M}^{-1}(z) dz \cdot C_{x,y} = -g_{x,y} \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

$$M(s) \int_0^s \mathbf{M}^{-1}(z) C_{x,y} dz = \frac{g_{x,y}}{\omega} \begin{bmatrix} \cos \omega s & \sin \omega s / \omega \\ -\omega \sin \omega s & \cos \omega s \end{bmatrix} \cdot \begin{bmatrix} (1 - \cos \omega s) / \omega \\ -\sin(\omega s) \end{bmatrix} = \frac{g_{x,y}}{\omega^2} \begin{bmatrix} \cos \omega s - 1 \\ -\omega \sin \omega s \end{bmatrix}$$

$$R_{56} = s \cdot m^2 c^2 / p_o + \int_0^s (g_x R(z)_{16} + g_y R_{36}(z)) dz =$$

$$\int_0^s (g_x R(z)_{16} + g_y R_{36}(z)) dz = \frac{g_x^2 + g_y^2}{\omega^2} \int_0^s (\cos \omega z - 1) dz = \frac{g_x^2 + g_y^2}{\omega^2} \left(\frac{\sin \omega s}{\omega} - s \right)$$

with the result:

$$R = \begin{bmatrix} \frac{g_x}{\omega^2} (\cos \omega s - 1) \\ -\frac{g_x}{\omega} \sin \omega s \\ \frac{g_y}{\omega^2} (\cos \omega s - 1) \\ -\frac{g_x}{\omega} \sin \omega s \end{bmatrix}; \quad R_{56} = \frac{m^2 c^2}{p_o^2} s + \frac{g_x^2 + g_y^2}{\omega^2} \left(\frac{\sin \omega s}{\omega} - s \right) \quad (\text{L3-IIb-2})$$

Case III: $a + b = 0$; $\det D = 0$; $\omega^2 = 2b$; $\lambda_{1,2} = \pm i\omega$; $\lambda_3 = 0$; $m = 3$.

We have to use degenerated case formula, but the maximum height of the eigen vector is 2 and only for 3-rd eigen value. Since it is not scary at all: $n_1=1; n_2=1; n_3=2$

Because of the Hamilton-Kelly theorem, $\mathbf{D}^2(\mathbf{D}^2 + \omega^2 I) = 0$. Let's do it

$$\exp[\mathbf{D}s] = \sum_{k=1}^3 \left[e^{\lambda_k s} \prod_{i \neq k} \left\{ \frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \sum_{j=0}^{n_i-1} \left(\frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\}^{n_i} \sum_{p=0}^{n_k-1} \frac{s^p}{p!} (\mathbf{D} - \lambda_k \mathbf{I})^p \right] =$$

$$\lambda_1 \lambda_2 = \omega^2; \quad i\omega$$

$$k = 3; \quad \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right)^2 (\mathbf{I} + s\mathbf{D}); \quad \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right)^2 = \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) + \frac{\mathbf{D}^2}{\omega^2} \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) \downarrow_0 = \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right)$$

$$k = 3; \quad \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) (\mathbf{I} + s\mathbf{D})$$

$$k = 1 + 2; \quad e^{i\omega s} \frac{\mathbf{D} + i\omega \mathbf{I}}{-2i\omega} \frac{\mathbf{D}^2}{\omega^2} + c.c. = -\frac{\mathbf{D}^2}{\omega^2} \left(\mathbf{I} \cos \omega s + \frac{\mathbf{D}}{\omega} \sin \omega s \right)$$

$$M = \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) (\mathbf{I} + s\mathbf{D}) - \frac{\mathbf{D}^2}{\omega^2} \left(\mathbf{I} \cos \omega s + \frac{\mathbf{D}}{\omega} \sin \omega s \right)$$

$$\mathbf{M}_{4 \times 4} = \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) (\mathbf{I} + s\mathbf{D}) - \frac{\mathbf{D}^2}{\omega^2} \left(\mathbf{I} \cos \omega s + \frac{\mathbf{D}}{\omega} \sin \omega s \right) \quad (\text{L3-III-1})$$

Similarly

$$R = \left\{ \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) \left(\mathbf{I} s + \mathbf{D} \frac{s^2}{2} \right) + \frac{\mathbf{D}^2}{\omega^4} (\mathbf{D} (\cos \omega s - 1) - \mathbf{I} \omega \sin \omega s) \right\} \mathbf{C} \quad (\text{L3-III-2})$$

Next is just

$$\int_0^s C^T \left\{ \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) \left(\mathbf{I}z + \mathbf{D} \frac{z^2}{2} \right) + \frac{\mathbf{D}^2}{\omega^4} (\mathbf{D}(\cos \omega z - 1) - \mathbf{I} \omega \sin \omega z) \right\} C dz =$$

$$C^T \left\{ \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) \left(\mathbf{I} \frac{s^2}{2} + \mathbf{D} \frac{s^3}{6} \right) + \frac{\mathbf{D}^2}{\omega^4} \left(\mathbf{D} \left(\frac{\sin \omega s}{\omega} - s \right) \mathbf{I} (\cos \omega z - 1) \right) \right\} C$$

with result of:

$$R_{56} = m^2 c^2 / p_o s + C^T \left\{ \left(\mathbf{I} + \frac{\mathbf{D}^2}{\omega^2} \right) \left(\mathbf{I} \frac{s^2}{2} + \mathbf{D} \frac{s^3}{6} \right) + \frac{\mathbf{D}^2}{\omega^4} \left(\mathbf{D} \left(\frac{\sin \omega s}{\omega} - s \right) \mathbf{I} (\cos \omega z - 1) \right) \right\} C \quad (\text{L3-III-3})$$

Case IV: all roots are different, no degeneration. Use formula (L3-36)

$$\exp[\mathbf{D}s] = \sum_{k=1}^2 \left(\frac{e^{\lambda_k s} + e^{-\lambda_k s}}{2} \mathbf{I} + \frac{e^{\lambda_k s} - e^{-\lambda_k s}}{2\lambda_k} \mathbf{D} \right) \prod \left(\frac{\mathbf{D}^2 - \lambda_j^2 \mathbf{I}}{\lambda_k^2 - \lambda_j^2} \right)$$

with only one term in the product:

$$\mathbf{M}_{4 \times 4} = \frac{1}{\omega_1^2 - \omega_2^2} \left\{ \left(\mathbf{I} \cos \omega_1 s + \mathbf{D} \frac{\sin \omega_1 s}{\omega_1} \right) (\mathbf{D}^2 + \omega_2^2 \mathbf{I}) - \left(\mathbf{I} \cos \omega_2 s + \mathbf{D} \frac{\sin \omega_2 s}{\omega_2} \right) (\mathbf{D}^2 + \omega_1^2 \mathbf{I}) \right\} \quad (\text{L3-IV-1})$$

For R we invoke a simplest formula:

$$\mathbf{R} = (\mathbf{M}_{4 \times 4}(s) - \mathbf{I}) \mathbf{D}^{-1} \cdot \mathbf{C} \quad (\text{L3-IV-2})$$

For R56 it is tedious but easy:

$$R_{56} = m^2 c^2 / p_o s + C^T \mathbf{M} \mathbf{D}^{-1} \mathbf{C};$$

$$\mathbf{M} = \frac{1}{\omega_1^2 - \omega_2^2} \left\{ \begin{aligned} & \left(\mathbf{I} \frac{\sin \omega_1 s}{\omega_1} + \mathbf{D} \frac{1 - \cos \omega_1 s}{\omega_1^2} \right) (\mathbf{D}^2 + \omega_2^2 \mathbf{I}) - \\ & \left(\mathbf{I} \frac{\sin \omega_2 s}{\omega_2} + \mathbf{D} \frac{1 - \cos \omega_2 s}{\omega_2^2} \right) (\mathbf{D}^2 + \omega_1^2 \mathbf{I}) - \mathbf{I} \cdot s \end{aligned} \right\} \quad (\text{L3-IV-3})$$

Case V: all roots are different, no degeneration. Use formula (L3-36) again

$$\mathbf{M}_{4 \times 4} = \frac{1}{\omega_1^2 + \omega_2^2} \left\{ \left(\mathbf{I} \cos \omega_1 s + \mathbf{D} \frac{\sin \omega_1 s}{\omega_1} \right) (\mathbf{D}^2 - \omega_2^2 \mathbf{I}) - \left(\mathbf{I} \cosh \omega_2 s + \mathbf{D} \frac{\sinh \omega_2 s}{\omega_2} \right) (\mathbf{D}^2 + \omega_1^2 \mathbf{I}) \right\} \quad (\text{L3-V-1})$$

$$\mathbf{R} = (\mathbf{M}_{4 \times 4}(s) - \mathbf{I}) \mathbf{D}^{-1} \cdot \mathbf{C} \quad (\text{L3-V-2})$$

$$R_{56} = m^2 c^2 / p_o s + C^T \mathbf{M} \mathbf{D}^{-1} \mathbf{C};$$

$$\mathbf{M} = \frac{1}{\omega_1^2 + \omega_2^2} \left\{ \begin{aligned} & \left(\mathbf{I} \frac{\sin \omega_1 s}{\omega_1} + \mathbf{D} \frac{1 - \cos \omega_1 s}{\omega_1^2} \right) (\mathbf{D}^2 - \omega_2^2 \mathbf{I}) - \\ & \left(\mathbf{I} \frac{\sinh \omega_2 s}{\omega_2} + \mathbf{D} \frac{\cosh \omega_2 s - 1}{\omega_2^2} \right) (\mathbf{D}^2 + \omega_1^2 \mathbf{I}) - \mathbf{I} \cdot s \end{aligned} \right\} \quad (\text{L3-V-3})$$

Before going into the discussion of the parameterization of the motion, we need to finish discussion of few remaining topics for 6x6 matrix of an accelerator. First is multiplication of the 6x6 matrixes for purely magnetic elements:

$$\mathbf{M}_k^{(6 \times 6)} = \begin{bmatrix} \mathbf{M}_k^{(4 \times 4)} & 0 & R_k \\ L_k & 1 & R_{56_k} \\ 0 & 0 & 1 \end{bmatrix}; \quad (L3-61)$$

$$\mathbf{M}_2^{(6 \times 6)} \mathbf{M}_1^{(6 \times 6)} = \begin{bmatrix} \mathbf{M}_2^{(4 \times 4)} & 0 & R \\ L & 1 & R_{56} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_2 \mathbf{M}_1 & 0 & R_2 + \mathbf{M}_2 R_1 \\ L_2 + L_1 \mathbf{M}_2 & 1 & R_{56_1} + R_{56_2} + L_2 R_1 \\ 0 & 0 & 1 \end{bmatrix}$$

i.e. having transformation rules for mixed size objects: a 4x4 matrix M, 4-element column R, 4 element line L, and a number R_{56} . As you remember, L is dependent (L4-7) and expressed as $L = \mathbf{R}^T \mathbf{S} \mathbf{M}$. Thus:

$$\mathbf{M}_{(4 \times 4)} = \mathbf{M}_2 \mathbf{M}_1; \quad R = \mathbf{M}_2 R_1 + R_2; \quad L = L_2 \mathbf{M}_1 + L_1; \quad R_{56} = R_{56_1} + R_{56_2} + L_2 R_1 \quad (L3-62)$$

One thing is left without discussion so far – the energy change. Thus, we should look into a particle passing through an RF cavity, which has alternating longitudinal field. Again, for simplicity we will assume that equilibrium particle does not gain energy, i.e. p_0 stays constant and we can continue using reduced variables. We will also assume that there is no transverse field, neither AC or DC. In this case the Hamiltonian reduces to a simple, fully decoupled:

$$\tilde{h} = \frac{\pi_1^2 + \pi_3^2}{2} + \frac{\pi_o^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + u \frac{\tau^2}{2}; \quad (L2-46)$$

$$\frac{dX}{ds} = \mathbf{D} \cdot X; \quad \mathbf{D} = \begin{bmatrix} D_x & 0 & 0 \\ 0 & D_y & 0 \\ 0 & 0 & D_l \end{bmatrix}; \quad D_x = D_y = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad D_l = \begin{bmatrix} 0 & \frac{m^2 c^2}{p_o^2} \\ -u & 0 \end{bmatrix};$$

$$\mathbf{M} = \begin{bmatrix} M_x & 0 & 0 \\ 0 & M_y & 0 \\ 0 & 0 & M_l \end{bmatrix}; \quad M_x = M_y = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}; \quad \omega = \sqrt{|\det D_l|} = \frac{mc}{p_o} \sqrt{|u|} \quad (L3-63)$$

$$M_l = \begin{bmatrix} \cos \omega s & \frac{m^2 c^2}{p_o^2} \sin \omega s / \omega \\ -u \sin \omega s / \omega & \cos \omega s \end{bmatrix}; \quad u > 0; \quad M_l = \begin{bmatrix} \cosh \omega s & \frac{m^2 c^2}{p_o^2} \sinh \omega s / \omega \\ -u \sinh \omega s / \omega & \cosh \omega s \end{bmatrix}; \quad u < 0;$$

In majority of the cases $\omega s \ll 1$ ($mc/p_0 \sim 1/\gamma$) and RF cavity can be represented as a thin lens located in its center:

$$\mathbf{M} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_l \end{bmatrix}; \quad M_l = \begin{bmatrix} 1 & 0 \\ -q & 1 \end{bmatrix}; \quad q = u \cdot l_{RF} = -\frac{e}{p_o c} \frac{\partial V_{rf}}{\partial t} \quad (L3-64)$$

Lecture 5. Parameterization of motion

5.1 Parameterization of motion

A periodic system can be a circular pass, or the cells of an accelerator (FODO is an example) repeating themselves again and again with period C , i.e.,

$$H = \frac{1}{2} \sum_{i=1}^{2n} \sum_{j=1}^{2n} h_{ij}(s) x_i x_j \equiv \frac{1}{2} X^T \cdot \mathbf{H}(s) \cdot X, \quad \mathbf{H}(s+C) = \mathbf{H}(s); \quad (\text{L5-1})$$

In this case, a one-turn (or one period) transport matrix

$$\mathbf{T}(s) = \mathbf{M}(s|s+C) \quad (\text{L5-2})$$

plays a very important role. Its eigen values, λ_i ,

$$\det[\mathbf{T} - \lambda_i \cdot \mathbf{I}] = 0 \quad (\text{L5-3})$$

determine its motion is stable, that all $|\lambda_i| \leq 1$ or are unstable, some $|\lambda_i| > 1$. Before making specific statements about the stability, we look at the properties of the eigen vectors.

First, eigen values are a function of periodic system and do not depend on the azimuth, s . It is easy to show that a one-turn matrix is transformed by the transport matrix as

$$\mathbf{T}(s_1) = \mathbf{M}(s|s_1) \mathbf{T}(s) \mathbf{M}^{-1}(s|s_1) \quad (\text{L5-4})$$

$$\begin{aligned} \mathbf{T}(s_1) &= \mathbf{M}(s_1|s_1+C) = \mathbf{M}(s+C|s_1+C) \mathbf{M}(s_1|s+C) = \mathbf{M}(s+C|s_1+C) \mathbf{M}(s|s+C) \mathbf{M}(s_1|s) \\ \mathbf{M}(s+C|s_1+C) &\equiv \mathbf{M}(s|s_1); \quad \mathbf{M}(s|s+C), \quad \mathbf{M}(s_1|s) \equiv \mathbf{M}^{-1}(s|s_1) \Rightarrow \mathbf{T}(s_1) = \mathbf{M}(s|s_1) \mathbf{T}(s) \mathbf{M}^{-1}(s|s_1) \end{aligned}$$

It means that $\mathbf{T}(s_1)$ has the same eigen values (L5-3); thus, the eigen values of $\mathbf{T}(s)$ do not depend upon s because

$$\begin{aligned} \det[\mathbf{M} \mathbf{T} \mathbf{M}^{-1} - \lambda_i \cdot \mathbf{I}] &= \det[\mathbf{M}(\mathbf{T} - \lambda_i \cdot \mathbf{I}) \mathbf{M}^{-1}] = \det[\mathbf{T} - \lambda_i \cdot \mathbf{I}] \\ \Rightarrow [\mathbf{T}(s_1) - \lambda_i \cdot \mathbf{I}] &= [\mathbf{T}(s) - \lambda_i \cdot \mathbf{I}] = 0 \end{aligned} \quad (\text{L5-5})$$

The matrix \mathbf{T} is a real, complex conjugate of eigen value λ_i^* which is also eigen value of \mathbf{T}

$$[\mathbf{T} - \lambda_i \cdot \mathbf{I}]^* = [\mathbf{T} - \lambda_i^* \cdot \mathbf{I}] = 0$$

Furthermore, the symplecticity of \mathbf{T} requires that λ_i^{-1} also is eigen value of \mathbf{T} . Proving that the inverse matrix \mathbf{T}^{-1} has λ_i^{-1} as a eigen value is easy. At the same time

$$0 = \det[\mathbf{T}^{-1} - \lambda_i^{-1} \mathbf{I}] = \det(\mathbf{S}[\mathbf{T}^T - \lambda_i^{-1} \mathbf{I}]\mathbf{S}^{-1}) = \det[\mathbf{T}^T - \lambda_i^{-1} \mathbf{I}] = \det[\mathbf{T} - \lambda_i^{-1} \mathbf{I}] \quad (\text{L5-5'})$$

and here, symplectic conditions help us again. Thus, the real symplectic matrix has n pairs of eigen values as follows: a) inverse $\{\lambda_i, \lambda_i^{-1}\}$, and b) complex conjugate $\{\lambda_i, \lambda_i^*\}$. We assume that matrix \mathbf{T} can be diagonalized (see note on the following page for the general case of Jordan normal form).

Therefore, repeating the matrix \mathbf{T} again and again undoubtedly will cause an exponentially growing solution if $|\lambda_i| > 1$. This statement is readily verified, but in so, we introduce some useful

term and matrices. The set of eigen vectors Y_i of matrix \mathbf{T} ¹¹

$$\mathbf{T} \cdot Y_i = \lambda_i \cdot Y_i; \quad i = 1, 2, \dots, 2n \quad (\text{L5-6})$$

is complete and an arbitrary vector X can be expanded about this basis:

$$X = \sum_{i=1}^{2n} a_i Y_i \equiv \mathbf{U} \cdot \mathbf{A}, \quad \mathbf{U} = [Y_1, \dots, Y_{2n}], \quad \mathbf{A}^T = [a_1, \dots, a_{2n}]. \quad (\text{L5-7})$$

where we introduces matrix \mathbf{U} built from eigen vector of the matrix \mathbf{T} :

$$\mathbf{T} \cdot \mathbf{U} = \mathbf{U} \cdot \mathbf{\Lambda}, \quad \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_{2n} \end{bmatrix} \quad (\text{L5-8})$$

The later equation is equivalent to diagonalization of the matrix \mathbf{T} :

$$\mathbf{U}^{-1} \cdot \mathbf{T} \cdot \mathbf{U} = \mathbf{\Lambda}, \quad \text{or} \quad \mathbf{T} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{U}^{-1} \quad (\text{L5-9})$$

Multiple application of matrix \mathbf{T} (i.e., passes around the ring)

$$\mathbf{T}^n \cdot X = \sum_{i=1}^{2n} \lambda_i^n a_i Y_i \quad (\text{L5-10})$$

exhibit exponentially growing terms if the module of even one eigen value is larger than 1 $\lambda_k = |\lambda| e^{i\mu}$, $|\lambda| > 1$; we easily observe that a solution with the initial condition $X_o = \text{Re } a_k Y_k$ grows exponentially:

$$\mathbf{T}^n X_o = |\lambda|^n \text{Re } a_k Y_k e^{i\mu n}.$$

Immediately this suggests that the only possible stable system is when all eigen values are unimodular

$$|\lambda_i| = 1. \quad (\text{L5-11})$$

otherwise assuming $|\lambda_i| < 1$ means that there is eigen value $\lambda_k = \lambda_i^{-1}$; $|\lambda_k| = 1/|\lambda_i| > 1$. We also consider only cases when all eigen vectors differ. Then, there are n pairs of eigen vectors, which we term modes of oscillations:

$$\lambda_k \equiv 1/\lambda_{k+n} \equiv \lambda_{k+n}^* \equiv e^{i\mu_k}; \quad \mu_k \equiv 2\pi\nu_k, \quad \{k = 1, \dots, n\}. \quad (\text{L5-12})$$

where the complex conjugate pairs are identical to the inverse pairs. Eq. (L5-9) can be rewritten as

$$\mathbf{T}(s) = \mathbf{U}(s) \mathbf{\Lambda} \mathbf{U}^{-1}(s); \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1^* & 0 \\ & & \dots & 0 \\ 0 & 0 & 0 & \lambda_n^* \end{pmatrix}; \quad \mathbf{T}(s) \cdot \mathbf{U}(s) = \mathbf{U}(s) \cdot \mathbf{\Lambda} \quad (\text{L5-13})$$

and matrix \mathbf{U} built from complex conjugate eigen vectors of \mathbf{T} :

¹¹ In general case of multiplicity of eigen vectors, the matrix cannot be diagonalized but can be brought to Jordan normal form http://en.wikipedia.org/wiki/Jordan_normal_form#Generalized_eigenvectors, Glenn James and Robert C. James, Mathematics. In this case, there is a subset of generalized eigen vectors $\{Y_{k,1}, \dots, Y_{k,h}\}$ that belong to a eigen value λ_k with multiplicity h :

$$\mathbf{T} \cdot Y_{k,h} = \lambda_k Y_{k,h}; \quad \mathbf{T} \cdot Y_{k,m} = \lambda_k Y_{k,m} + Y_{k,m+1}; \quad m = 1 \dots h-1.$$

The result is even stronger than in the diagonal case; motion is unstable even when $\lambda_k = 1$:

$$\mathbf{T} \cdot Y_{k,h-1} = \lambda_k Y_{k,h-1} + Y_{k,h} \Rightarrow \mathbf{T}^n \cdot Y_{k,h-1} = Y_{k,h-1} + n \cdot Y_{k,h}$$

$$\mathbf{U}(s) = [Y_1, Y_1^* \dots Y_n, Y_n^*]; \quad \mathbf{T}(s)Y_k(s) = \lambda_k Y_k(s) \Leftrightarrow \mathbf{T}(s)Y_k^*(s) = \lambda_k^* Y_k^*(s) \quad (\text{L5-14})$$

Thus, eigen vectors can be transported from one azimuth to another by the transport matrix:

$$\tilde{Y}_k(s_1) = \mathbf{M}(s|s_1)\tilde{Y}_k(s) \Leftrightarrow \frac{d}{ds}\tilde{Y}_k = \mathbf{D}(s) \cdot \tilde{Y}_k \quad (\text{L5-15})$$

It is eigen vector of $\mathbf{T}(s_1)$. - just add (L5-4) to (L5-14):

$$\mathbf{T}(s_1)\tilde{Y}_k(s_1) = \mathbf{M}(s|s_1)\mathbf{T}(s)\mathbf{M}^{-1}(s|s_1)\mathbf{M}(s|s_1)\tilde{Y}_k(s) = \mathbf{M}(s|s_1)\mathbf{T}(s)\tilde{Y}_k(s) = \lambda_k\mathbf{M}(s|s_1)\tilde{Y}_k(s) = \lambda_k\tilde{Y}_k(s_1)\#$$

Similarly,

$$\tilde{\mathbf{U}}(s_1) = \mathbf{M}(s|s_1)\tilde{\mathbf{U}}(s) \Leftrightarrow \frac{d}{ds}\tilde{\mathbf{U}} = \mathbf{D}(s) \cdot \tilde{\mathbf{U}} \quad (\text{L5-16})$$

with the obvious follow-up by

$$\tilde{\mathbf{U}}(s+C) = \tilde{\mathbf{U}}(s) \cdot \Lambda, \quad \tilde{Y}_k(s+C) = \lambda_k \tilde{Y}_k(s) = e^{i\mu_k} \tilde{Y}_k(s) \quad (\text{L5-17})$$

The k^{th} eigen vectors are multiplied by $e^{i\mu_k}$ after each pass through the period. Hence, we can write

$$\tilde{Y}_k(s) = Y_k(s)e^{\psi_k(s)}; \quad Y_k(s+C) = Y_k(s); \quad \psi_k(s+C) = \psi_k(s) + \mu_k \quad (\text{L5-18})$$

$$\tilde{\mathbf{U}}(s) = \mathbf{U}(s) \cdot \Psi(s), \quad \Psi(s) = \begin{pmatrix} e^{i\psi_1(s)} & 0 & 0 \\ 0 & e^{-i\psi_1(s)} & 0 \\ & & \dots & 0 \\ 0 & 0 & 0 & e^{-i\psi_n(s)} \end{pmatrix} \quad (\text{L5-19})$$

It is remarkable that the symplectic products (L5-12) of the eigen vectors are non-zero-only complex conjugate pairs: in other words, the structure of the Hamiltonian metrics is preserved here. $Y_k^T \cdot \mathbf{S} \cdot Y_k^T = 0$ is obvious. Using only the symplecticity of \mathbf{T} gives us desirable yields

$$Y_k^T \cdot \mathbf{S} \cdot Y_j^T = Y_k^T \cdot \mathbf{T}^T \mathbf{S} \mathbf{T} \cdot Y_j^T = \lambda_k \lambda_j (Y_k^T \cdot \mathbf{S} \cdot Y_j^T) \Rightarrow (1 - \lambda_k \lambda_j) (Y_k^T \cdot \mathbf{S} \cdot Y_j^T) = 0$$

for $\lambda_k \lambda_j \neq 1$

$$Y_k^{T*} \cdot \mathbf{S} \cdot Y_{j \neq k} = 0; \quad Y_k^T \cdot \mathbf{S} \cdot Y_j = 0; \quad . \quad (\text{L5-20})$$

and only the nonzero products for $\lambda_k = 1/\lambda_j = \lambda_j^*$ are clearly pure imaginary¹²:

$$Y_k^{T*} \cdot \mathbf{S} \cdot Y_k = 2i, \quad (\text{L5-21})$$

where we chose the calibration of purely imaginary values as $2i$ for the following expansion to be symplectic. Eqs. (L5-20-21) in compact matrix form is

$$\mathbf{U}^T \cdot \mathbf{S} \cdot \mathbf{U} = \tilde{\mathbf{U}}^T \cdot \mathbf{S} \cdot \tilde{\mathbf{U}} = -2i\mathbf{S}, \quad \mathbf{U}^{-1} = \frac{1}{2i} \mathbf{S} \cdot \mathbf{U}^T \cdot \mathbf{S}. \quad (\text{L5-22})$$

The expressions for the transport matrices through β , α -functions, and phase advances often derived as a ‘‘miraculous’’ result, and hence called matrix gymnastics, is just a trivial consequence of equations (L5-16), (L5-19), and (L5-22):

$$\mathbf{M}(s|s_1) = \tilde{\mathbf{U}}(s_1)\tilde{\mathbf{U}}^{-1}(s) = \frac{1}{2i}\tilde{\mathbf{U}}(s_1) \cdot \mathbf{S} \cdot \tilde{\mathbf{U}}^T(s) \cdot \mathbf{S} = \frac{1}{2i}\mathbf{U}(s_1) \cdot \Psi(s_1) \cdot \mathbf{S} \cdot \Psi^{-1}(s) \cdot \mathbf{U}^T(s_1) \quad (\text{L5-16}')$$

¹² $(A^{*T} \cdot \mathbf{S} \cdot A)^* = (A^T \cdot \mathbf{S} \cdot A^*) = -(A^{*T} \cdot \mathbf{S} \cdot A)^T = -(A^{*T} \cdot \mathbf{S} \cdot A)$

with a specific case of a one-turn matrix:

$$\mathbf{T} = \mathbf{U}\Lambda\mathbf{U}^{-1} = \frac{1}{2i}\mathbf{U}\Lambda\mathbf{S}\mathbf{U}^T\mathbf{S} \quad (\text{L5-13}')$$

S-orthogonality (L5-20) provides an excellent tool of finding complex coefficients in the expansion eq. (L5-7) of an arbitrary solution $\mathbf{X}(s)$

$$X_o = \sum_{i=1}^{2n} a_i Y_i \Rightarrow X(s) = \frac{1}{2} \sum_{k=1}^n (a_k \tilde{Y}_k + a_k^* \tilde{Y}_k^*) \equiv \text{Re} \sum_{k=1}^n a_k Y_k e^{i\psi_k} \equiv \frac{1}{2} \tilde{\mathbf{U}} \cdot \mathbf{A} = \frac{1}{2} \mathbf{U} \cdot \boldsymbol{\Psi} \cdot \mathbf{A} = \frac{1}{2} \mathbf{U} \cdot \tilde{\mathbf{A}} \quad (\text{L5-23})$$

where $2n$ complex coefficients, which are constants of motion¹³ ! for linear Hamiltonian system, can be found by a simple multiplications (instead of solving a system of $2n$ linear equations (L5-7))

$$a_i = \frac{1}{2i} Y_i^{*T} \mathbf{S} \mathbf{X}; \quad \tilde{a}_i \equiv a_i e^{i\psi_i} = \frac{1}{2i} Y_i^{*T} \mathbf{S} \mathbf{X}; \quad (\text{L5-24})$$

$$\mathbf{A} = 2\tilde{\mathbf{U}}^{-1} \cdot \mathbf{X} = -i\boldsymbol{\Psi}^{-1} \cdot \mathbf{S} \cdot \mathbf{U}^{T*} \cdot \mathbf{S} \cdot \mathbf{X}; \quad \tilde{\mathbf{A}} = \boldsymbol{\Psi} \mathbf{A} = -i \cdot \mathbf{S} \cdot \mathbf{U}^{T*} \cdot \mathbf{S} \cdot \mathbf{X}.$$

Equation (L5-23) is nothing else but a general parameterization of motion in the linear Hamiltonian system. It is very powerful tool and we will use this many times in this course.

We consider next a specific case of a 1D system with a linear periodical Hamiltonian:

$$\tilde{h} = \frac{p^2}{2} + K_1(s) \frac{y^2}{2}; \quad \mathbf{H} = \begin{bmatrix} K_1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{D} = \mathbf{S}\mathbf{H} = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix}. \quad (\text{L5-25})$$

The equations of motion are simple

$$\frac{d}{ds} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} p \\ -K_1 x \end{bmatrix} \quad (\text{i.e. } x' \equiv p). \quad (\text{L5-26})$$

A one-turn matrix within its determinant ($ad-bc=1$)

$$\mathbf{T}(s) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathbf{U}(s)\Lambda\mathbf{U}^{-1}(s); \quad \Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} = \begin{pmatrix} e^{i\mu} & 0 \\ 0 & e^{-i\mu} \end{pmatrix}; \quad (\text{L5-27})$$

$$\mathbf{Y} = \begin{bmatrix} w \\ u + i/w \end{bmatrix}; \quad \tilde{\mathbf{Y}} = \begin{bmatrix} w \\ u + i/w \end{bmatrix} e^{i\psi}; \quad \mathbf{U} = \begin{bmatrix} w & w \\ u + i/w & u - i/w \end{bmatrix}; \quad \tilde{\mathbf{U}} = \mathbf{U} \cdot \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix} \quad (\text{L5-28})$$

where $w(s)$ ¹⁴ and $u(s)$ are real functions and calibration was used for (L5-21). \mathbf{T} has a trace

$$\text{Trace}(\mathbf{T}) = \text{Trace}(\Lambda) = 2 \cos \mu \quad (\text{L5-29})$$

(because $\text{Trace}(ABA^{-1}) = \text{Trace}(B)$). Thus, the stability of motion (when μ is real!) is easy to check:

$$-2 < \text{Trace}(\mathbf{T}) < 2 \quad (\text{L5-30})$$

where some well-know resonances are excluded: The integer $\mu = 2\pi m$, and the half-integer $\mu = 2(m+1)\pi$ as being unstable (troublesome!).

Combining (L5-28) into the equations of motion (L5-25)

¹³ in matrix form using (L5-16) we have $X = \frac{1}{2} \tilde{\mathbf{U}} \mathbf{A}$, $X' = \frac{1}{2} (\tilde{\mathbf{U}}' \mathbf{A} + \tilde{\mathbf{U}} \mathbf{A}')$, $= \mathbf{D} \mathbf{X} = \frac{1}{2} \mathbf{D} \tilde{\mathbf{U}} \cdot \mathbf{A} = \frac{1}{2} \tilde{\mathbf{U}}' \cdot \mathbf{A} \Rightarrow \mathbf{A}' = 0$

¹⁴ We are free to multiply the eigen vector \mathbf{Y} by $e^{i\phi}$ to make \mathbf{a} real number. In other words we define the choice of our phase as $\tilde{\mathbf{Y}}(s) = \begin{pmatrix} \tilde{y}_1(s) \\ \tilde{y}_2(s) \end{pmatrix}$; $w(s) = |\tilde{y}_1(s)|$; $\psi(s) = \arg(\tilde{y}(s))$.

$$\frac{d}{ds} \begin{bmatrix} w \\ u + i/w \end{bmatrix} e^{i\psi} = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix} \cdot \tilde{Y} = \begin{bmatrix} w \\ u + i/w \end{bmatrix} e^{i\psi} \Rightarrow \begin{matrix} w' + iw\psi' = u + i/w \\ u' - iw'/w^2 + i\psi'(u + i/w) = -K_1 w \end{matrix} \quad (\text{L5-31})$$

Then, separating the real and imaginary parts, we have from the first equation:

$$u = w'; \quad \psi' = 1/w^2. \quad (\text{L5-32})$$

Plugging these into the second equation yields one nontrivial equation on the envelope function, $w(s)$:

$$w'' + K_1(s)w = \frac{1}{w^3}. \quad (\text{L5-33})$$

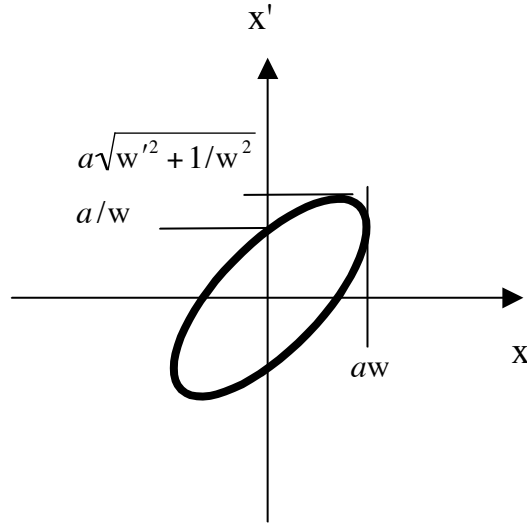
Thus, the final form of the eigen vector can be rewritten as

$$Y = \begin{bmatrix} w \\ w' + i/w \end{bmatrix}; \quad \psi' = \frac{1}{w^2}; \quad \tilde{Y} = Y e^{i\psi} \quad (\text{L5-34})$$

The parameterization of the linear 1D motion is

$$\begin{aligned} \begin{bmatrix} x \\ x' \end{bmatrix} &= \text{Re} \left(a e^{i\varphi} \begin{bmatrix} w \\ w' + i/w \end{bmatrix} e^{i\psi} \right); \\ x &= a \cdot w(s) \cdot \cos(\psi(s) + \varphi) \\ x' &= a \cdot (w'(s) \cdot \cos(\psi(s) + \varphi) - \sin(\psi(s) + \varphi)/w(s)) \end{aligned} \quad (\text{L5-35})$$

where a and φ are the constants of motion.



Tradition in accelerator physics calls for using the so-called β -function, which simply a square of the envelope function:

$$\beta \equiv w^2 \Rightarrow \psi' = 1/\beta. \quad (\text{L5-36})$$

and a wavelength of oscillations divided by 2π . Subserving functions are defined as

$$\alpha \equiv -\beta' \equiv -w w', \quad \gamma \equiv \frac{1 + \alpha^2}{\beta}. \quad (\text{L5-37})$$

Manipulations with them is much less transparent, and oscillation (L5-35) looks like

$$\begin{aligned}
x &= a \cdot \sqrt{\beta(s)} \cdot \cos(\psi(s) + \varphi) \\
x' &= -\frac{a}{\sqrt{\beta(s)}} \cdot (\alpha(s) \cdot \cos(\psi(s) + \varphi) + \sin(\psi(s) + \varphi))
\end{aligned} \tag{L5-35'}$$

Finally, (L5-13') gives us a well-known feature in AP parameterization of a one-turn matrix:

$$\mathbf{T} = \mathbf{U}\Lambda\mathbf{U}^{-1} = \mathbf{I}\cos\mu + \mathbf{J}\sin\mu; \quad \mathbf{J} = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix} \tag{L5-13''}$$

Another important parameter in accelerator physics is the transformation to action-angle variables $\left\{ \varphi_k, I_k = \frac{a_k^2}{2} \right\}$. Usually this requires two steps: The first is

$$\left\{ \tilde{q}_k = a_k e^{i\varphi_k} / \sqrt{2}, \tilde{p}_k = i a_k e^{i\varphi_k} / \sqrt{2} \right\}.$$

The second was given as a homework. Let us demonstrate that symplectic transformation $X(s) \Rightarrow \tilde{X}(s)$.

$$X(s) = \mathbf{V}(s)\tilde{X}, \quad \mathbf{V}'(s) = \mathbf{S}\mathbf{H}(s)\mathbf{V}(s). \tag{L5-38}$$

is canonical. Beginning from a Hamiltonian composed of two parts, a linear part and an arbitrary one

$$\mathcal{H} = \frac{1}{2} X^T \mathbf{H}(s) X + \mathcal{H}_1(X, s). \tag{L5-39}$$

The equation of motion

$$\frac{dX}{ds} = \mathbf{S} \cdot \frac{\partial \mathcal{H}}{\partial X} = \mathbf{S}\mathbf{H}(s) \cdot X + \mathbf{S} \cdot \frac{\partial \mathcal{H}_1}{\partial X}. \tag{L5-40}$$

becomes with substitution (L5-38)

$$(\mathbf{V}\tilde{X})' = \mathbf{S}\mathbf{H}\mathbf{V} \cdot \tilde{X} + \mathbf{V}\tilde{X}' = \mathbf{S}\mathbf{H}(s) \cdot \mathbf{V}\tilde{X} + \mathbf{S} \cdot \frac{\partial \mathcal{H}_1}{\partial X} \Rightarrow \mathbf{V}\tilde{X}' = \mathbf{S} \cdot \frac{\partial \mathcal{H}_1}{\partial X}. \tag{L5-41}$$

equivalent to the equations of motion with the new Hamiltonian: $\mathcal{H}_1(\mathbf{V}\tilde{X}, s)$

$$\tilde{X}' = \mathbf{V}^{-1}\mathbf{S} \cdot \frac{\partial \mathcal{H}_1}{\partial X}; \frac{\partial}{\partial X} = \mathbf{V}^{-1T} \frac{\partial}{\partial \tilde{X}} \Rightarrow \tilde{X}' = (\mathbf{V}^{-1}\mathbf{S}\mathbf{V}^{-1T}) \cdot \frac{\partial \mathcal{H}_1}{\partial \tilde{X}} \Rightarrow \tilde{X}' = \mathbf{S} \cdot \frac{\partial \mathcal{H}_1}{\partial \tilde{X}}. \tag{L5-42}$$

This result (even though expected) has long-lasting consequences – the trivial (linear) part in the Hamiltonian can be removed from equations of motion, so allowing one to use this in perturbation theory or at least to focus only on non-trivial part of the motion.

$$\mathbf{V} = \frac{1}{\sqrt{2}} [Y_1, iY_1, \dots] \Rightarrow \mathbf{V}^T \mathbf{S} \mathbf{V} = \mathbf{S} \# \tag{L5-43}$$

One should note for completeness that there is another way of parameterization of coupled motion proposed by Edward and Teng parameterization (*D.A.Edwards, L.C.Teng, IEEE Trans. Nucl. Sci. NS-20 (1973) 885*), which differs we do not discuss here.

Because parameterization is very useful for solving many standard accelerator problems, bellow I summarize our main funding with emphasis on most useful cases. I will not repeat the argument but write all appropriate equations:

General parameterization of motion of an arbitrary linear Hamiltonian system:

$$H = \frac{1}{2} \sum_{i=1}^{2n} \sum_{j=1}^{2n} h_{ij}(s) x_i x_j \equiv \frac{1}{2} X^T \cdot \mathbf{H}(s) \cdot X, \quad X^T = [q^1 \quad P_1 \quad \dots \quad q^n \quad P_n] = [x_1 \quad x_2 \quad \dots \quad x_{2n-1} \quad x_{2n}].$$

$$\frac{dX}{ds} = \mathbf{D}(s) \cdot X; \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{H}(s), \quad X(s) = \mathbf{M}(s|s_0) \cdot X_0. \quad (\text{I})$$

$$\mathbf{M}' \equiv \frac{d\mathbf{M}}{ds} = \mathbf{D}(s) \cdot \mathbf{M}; \quad \mathbf{M}(s_0) = \mathbf{I}, \quad \mathbf{M}^T \cdot \mathbf{S} \cdot \mathbf{M} = \mathbf{S}, \quad \mathbf{M}^{-1} = -\mathbf{S} \cdot \mathbf{M}^T \cdot \mathbf{S}. \quad (\text{II})$$

Parameterization of motion: a periodic system with period C $\mathbf{H}(s+C) = \mathbf{H}(s)$

$$\mathbf{T}(s) = \mathbf{M}(s|s+C), \quad \det[\mathbf{T} - \lambda_i \cdot \mathbf{I}] = 0 \quad (\text{III})$$

$$\text{Stable system } |\lambda_i| = 1, \lambda_k \equiv 1/\lambda_{k+n} \equiv \lambda_{k+n}^* \equiv e^{i\mu_k}; \quad \mu_k \equiv 2\pi\nu_k, \quad \{k=1, \dots, n\}. \quad (\text{IV})$$

1D - ACCELERATOR

$$\tilde{h} = \frac{p^2}{2} + K_1(s) \frac{y^2}{2}; \quad \mathbf{H} = \begin{bmatrix} K_1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \mathbf{D} = \mathbf{S}\mathbf{H} = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix}; \quad \frac{d}{ds} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K_1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} p \\ -K_1 x \end{bmatrix} \quad (\text{i.e. } x' \equiv p).$$

$$\mathbf{T}(s) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathbf{U}(s) \Lambda \mathbf{U}^{-1}(s); \quad \Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} = \begin{pmatrix} e^{i\mu} & 0 \\ 0 & e^{-i\mu} \end{pmatrix};$$

$$\mathbf{T} = \mathbf{I} \cos \mu + \mathbf{J} \sin \mu;$$

$$\mathbf{J} = \begin{bmatrix} -ww' & w^2 \\ -w'^2 - \frac{1}{w^2} & ww' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix}; \quad \mathbf{J}^2 = -\mathbf{I}$$

HENCE:

$$\mathbf{T} = \mathbf{I} \cos \mu + \mathbf{J} \sin \mu;$$

$$\mathbf{J} = \begin{bmatrix} -ww' & w^2 \\ -w'^2 - \frac{1}{w^2} & ww' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix}; \quad \mathbf{J}^2 = -\mathbf{I}; \quad \gamma = (1 + \alpha^2)/\beta$$

$$\cos \mu = \text{Trace}(\mathbf{T})/2 = \frac{T_{11} + T_{22}}{2}$$

Stability if : $-1 < \text{Trace}(\mathbf{T})/2 < 1$

$$w^2 \equiv \beta = \frac{T_{12}}{\sin \mu} = \frac{|T_{12}|}{\sqrt{1 - (\text{Trace}(\mathbf{T})/2)^2}}; \quad w = \sqrt{\frac{|T_{12}|}{\sqrt{1 - (\text{Trace}(\mathbf{T})/2)^2}}};$$

$$ww' \equiv -\alpha = \frac{T_{22} - T_{11}}{2 \cos \mu} = -\frac{T_{11} - T_{22}}{T_{11} + T_{22}}$$

$$\mathbf{T}(s) = \mathbf{U}(s)\Lambda\mathbf{U}^{-1}(s); \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1^* & 0 \\ & & \dots & 0 \\ 0 & 0 & 0 & \lambda_n^* \end{pmatrix}; \mathbf{T} \cdot \mathbf{U} = \mathbf{U} \cdot \Lambda; \Lambda = \mathbf{U}^{-1} \cdot \mathbf{T} \cdot \mathbf{U} \dots \quad (\text{V})$$

$$\mathbf{U}(s) = [Y_1, Y_1^* \dots Y_n, Y_n^*]; \quad \mathbf{T}(s)Y_k(s) = \lambda_k Y_k(s) \Leftrightarrow \mathbf{T}(s)Y_k^*(s) = \lambda_k^* Y_k^*(s) \quad (\text{VI})$$

$$\tilde{Y}_k(s_1) = \mathbf{M}(s|s_1)\tilde{Y}_k(s) \Leftrightarrow \frac{d}{ds}\tilde{Y}_k = \mathbf{D}(s) \cdot \tilde{Y}_k; \quad \tilde{\mathbf{U}}(s_1) = \mathbf{M}(s|s_1)\tilde{\mathbf{U}}(s) \Leftrightarrow \frac{d}{ds}\tilde{\mathbf{U}} = \mathbf{D}(s) \cdot \tilde{\mathbf{U}} \quad (\text{VII})$$

$$\tilde{\mathbf{U}}(s+C) = \tilde{\mathbf{U}}(s) \cdot \Lambda, \quad \tilde{Y}_k(s+C) = \lambda_k \tilde{Y}_k(s) = e^{i\mu_k} \tilde{Y}_k(s) \quad (\text{VIII})$$

$$\tilde{Y}_k(s) = Y_k(s)e^{\psi_k(s)}; \quad Y_k(s+C) = Y_k(s); \quad \psi_k(s+C) = \psi_k(s) + \mu_k \quad (\text{IX})$$

$$\tilde{\mathbf{U}}(s) = \mathbf{U}(s) \cdot \Psi(s), \quad \Psi(s) = \begin{pmatrix} e^{i\psi_1(s)} & 0 & 0 \\ 0 & e^{-i\psi_1(s)} & 0 \\ & & \dots & 0 \\ 0 & 0 & 0 & e^{-i\psi_n(s)} \end{pmatrix} \quad (\text{X})$$

$$Y_k^{T*} \cdot \mathbf{S} \cdot Y_{j \neq k} = 0; \quad Y_k^T \cdot \mathbf{S} \cdot Y_j = 0; \quad Y_k^{T*} \cdot \mathbf{S} \cdot Y_k = 2i \quad (\text{XI})$$

$$\mathbf{U}^T \cdot \mathbf{S} \cdot \mathbf{U} \equiv \tilde{\mathbf{U}}^T \cdot \mathbf{S} \cdot \tilde{\mathbf{U}} = -2i\mathbf{S}, \quad \mathbf{U}^{-1} = \frac{1}{2i}\mathbf{S} \cdot \mathbf{U}^T \cdot \mathbf{S} \quad (\text{XII})$$

$$X(s) = \frac{1}{2} \sum_{k=1}^n (a_k \tilde{Y}_k + a_k^* \tilde{Y}_k^*) \equiv \text{Re} \sum_{k=1}^n a_k Y_k e^{i\psi_k} \equiv \frac{1}{2} \tilde{\mathbf{U}} \cdot \mathbf{A} = \frac{1}{2} \mathbf{U} \cdot \Psi \cdot \mathbf{A} = \frac{1}{2} \mathbf{U} \cdot \tilde{\mathbf{A}} \quad (\text{XIII})$$

$$a_i = \frac{1}{2i} Y_i^{*T} S X; \quad \tilde{a}_i \equiv a_i e^{i\psi_i} = \frac{1}{2i} Y_i^{*T} S X; \quad (\text{XIV})$$

$$\mathbf{A} = 2\tilde{\mathbf{U}}^{-1} \cdot X = -i\Psi^{-1} \cdot \mathbf{S} \cdot \mathbf{U}^{T*} \cdot \mathbf{S} \cdot X; \quad \tilde{\mathbf{A}} = \Psi\mathbf{A} = -i \cdot \mathbf{S} \cdot \mathbf{U}^{T*} \cdot \mathbf{S} \cdot X.$$

1D - ACCELERATOR

$$Y = \begin{bmatrix} w \\ w' + i/w \end{bmatrix}; \tilde{Y} = \begin{bmatrix} w \\ w' + i/w \end{bmatrix} e^{i\psi}; \mathbf{U} = \begin{bmatrix} w & w \\ w' + i/w & w' - i/w \end{bmatrix}; \tilde{\mathbf{U}} = \mathbf{U} \cdot \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix}$$

$$w'' + K_1(s)w = \frac{1}{w^3}, \quad \psi' = 1/w^2; \quad \begin{bmatrix} x \\ x' \end{bmatrix} = \text{Re} \left(a e^{i\psi} \begin{bmatrix} w \\ w' + i/w \end{bmatrix} e^{i\psi} \right);$$

$$x = a \cdot w(s) \cdot \cos(\psi(s) + \varphi)$$

$$x' = a \cdot (w'(s) \cdot \cos(\psi(s) + \varphi) - \sin(\psi(s) + \varphi)/w(s))$$

$$\beta \equiv w^2 \Rightarrow \psi' = 1/\beta; \alpha \equiv -\beta' \equiv -w w', \quad \gamma \equiv \frac{1 + \alpha^2}{\beta} \text{ - definitions}$$

$$x = a \cdot \sqrt{\beta(s)} \cdot \cos(\psi(s) + \varphi)$$

$$x' = -\frac{a}{\sqrt{\beta(s)}} \cdot (\alpha(s) \cdot \cos(\psi(s) + \varphi) + \sin(\psi(s) + \varphi))$$

Complex amplitude and real amplitude and phase are easy to calculate. Expression for a^2 is called Currant-Snyder invariant.

$$X = \text{Re } \tilde{a}Y;$$

$$ae^{i\varphi} = -iY^{*T}SX = \begin{bmatrix} w \\ w' - i/w \end{bmatrix}^T \cdot \begin{bmatrix} x' \\ -x \end{bmatrix} = x/w + i(w'x - wx')$$

$$a^2 = \frac{x^2}{w^2} + (w'x - wx')^2 \equiv \frac{x^2 + (\alpha x + \beta x')^2}{\beta} \equiv \gamma x^2 + 2\alpha x x' + \beta x'^2$$

$$\varphi = \arg(x/w + i(w'x - wx')) = \tan^{-1} \frac{ww'x - w^2x'}{x} = -\tan^{-1} \frac{\alpha x + \beta x'}{x};$$

$$\varphi = \sin^{-1} \frac{w'x - wx'}{\sqrt{x^2 + w^2(w'x - wx')^2}} = -\sin^{-1} \frac{\alpha x + \beta x'}{\sqrt{\gamma x^2 + 2\alpha x x' + \beta x'^2}}$$

Inverse ratios – matrices through parameterization: reverse of eq (VII), where U is propagated by M.

$$\mathbf{M}(s_1|s_2) = \tilde{\mathbf{U}}(s_2)\tilde{\mathbf{U}}^{-1}(s_1) = \frac{i}{2}\tilde{\mathbf{U}}(s_2) \cdot \mathbf{S} \cdot \tilde{\mathbf{U}}^T(s_1) \cdot \mathbf{S} = \frac{i}{2}\mathbf{U}(s_2) \cdot \Psi(s_2) \cdot \mathbf{S} \cdot \Psi^{-1}(s_1) \cdot \mathbf{U}^T(s_1) \cdot \mathbf{S} \quad (\text{XV})$$

$$\Psi(s_2) \cdot \mathbf{S} \cdot \Psi^{-1}(s_1) \equiv \Psi(s_2 - s_1) \cdot \mathbf{S}; \quad \mathbf{M}(s_1|s_2) = \frac{i}{2}\mathbf{U}(s_2) \cdot \Psi(s_2 - s_1) \cdot \mathbf{S} \cdot \mathbf{U}^T(s_1) \cdot \mathbf{S}$$

$$\mathbf{T} = \mathbf{U}\mathbf{A}\mathbf{U}^{-1} = \frac{i}{2}\mathbf{U}\mathbf{A}\mathbf{S}\mathbf{U}^T\mathbf{S} \quad \text{Specific case of } s_1=s+C$$

$$\mathbf{T} = \frac{1}{2i}\mathbf{U}\mathbf{A}\mathbf{S}\mathbf{U}^T\mathbf{S} = \frac{1}{2i} \begin{bmatrix} w & w \\ w' + \frac{i}{w} & w' - \frac{i}{w} \end{bmatrix} \cdot \begin{bmatrix} e^{iu} & 0 \\ 0 & e^{-iu} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w & w' + \frac{i}{w} \\ w & w' - \frac{i}{w} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} =$$

$$\frac{1}{2i} \begin{bmatrix} we^{iu} & we^{-iu} \\ \left(w' + \frac{i}{w}\right)e^{iu} & \left(w' - \frac{i}{w}\right)e^{-iu} \end{bmatrix} \cdot \begin{bmatrix} \frac{i}{w} - w' & w \\ \frac{i}{w} + w' & -w \end{bmatrix} = \begin{bmatrix} \frac{e^{iu} + e^{-iu}}{2} - \frac{e^{iu} - e^{-iu}}{2i} ww' & w^2 \frac{e^{iu} - e^{-iu}}{2i} \\ -\left(w'^2 + \frac{1}{w^2}\right) \frac{e^{iu} - e^{-iu}}{2} & \frac{e^{iu} + e^{-iu}}{2} + \frac{e^{iu} - e^{-iu}}{2i} ww' \end{bmatrix}$$

$$\mathbf{T} = \frac{1}{2i}\mathbf{U}\mathbf{A}\mathbf{S}\mathbf{U}^T\mathbf{S} = \begin{bmatrix} \cos \mu - ww' \cdot \sin \mu & w^2 \cdot \sin \mu \\ -\left(w'^2 + \frac{1}{w^2}\right) \cdot \sin \mu & \cos \mu + ww' \cdot \sin \mu \end{bmatrix} = \mathbf{I} \cos \mu + \mathbf{J} \sin \mu;$$

$$\mathbf{J} = \begin{bmatrix} -ww' & w^2 \\ -w'^2 - \frac{1}{w^2} & ww' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix};$$

$$\mathbf{J}^2 = \begin{bmatrix} -ww' & w^2 \\ -\left(w'^2 + \frac{1}{w^2}\right) & ww' \end{bmatrix} \begin{bmatrix} -ww' & w^2 \\ -\left(w'^2 + \frac{1}{w^2}\right) & ww' \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbf{I}$$

1D - ACCELERATOR

$$M(s_1|s_2) = \frac{1}{2i} \begin{bmatrix} w_2 & w_2 \\ w_2' + i/w_2 & w_2' - i/w_2 \end{bmatrix} \cdot \begin{pmatrix} e^{i\Delta\psi} & 0 \\ 0 & e^{-i\Delta\psi} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{bmatrix} w_1 & w_1' + i/w_1 \\ w_1 & w_1' - i/w_1 \end{bmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\Delta\psi = \psi(s_2) - \psi(s_1);$$

$$M(s_1|s_2) = \frac{1}{2i} \begin{bmatrix} w_2 & w_2 \\ w_2' + i/w_2 & w_2' - i/w_2 \end{bmatrix} \cdot \begin{bmatrix} -(w_1' - i/w_1)e^{i\Delta\psi} & w_1 e^{i\Delta\psi} \\ (w_1' + i/w_1)e^{-i\Delta\psi} & -w_1 e^{-i\Delta\psi} \end{bmatrix} =$$

$$\begin{bmatrix} w_2/w_1 \cos\Delta\psi - w_2 w_1' \sin\Delta\psi & w_1 w_2 \sin\Delta\psi \\ (w_2'/w_1 - w_1'/w_2) \cos\Delta\psi & w_1/w_2 \cos\Delta\psi + w_1 w_2' \sin\Delta\psi \\ -(w_1' w_2' + 1/(w_2 w_1)) \sin\Delta\psi & \end{bmatrix}$$

use $w = \sqrt{\beta}$; $w' = -\alpha/\sqrt{\beta}$ to get a standard

$$M(s_1|s_2) = \begin{bmatrix} \frac{\cos\Delta\psi + \alpha_1 \sin\Delta\psi}{\sqrt{\beta_1/\beta_2}} & \sqrt{\beta_1\beta_2} \sin\Delta\psi \\ -\frac{(\alpha_2 - \alpha_1) \cos\Delta\psi + (1 + \alpha_1\alpha_2) \sin\Delta\psi}{\sqrt{\beta_1\beta_2}} & \frac{\cos\Delta\psi - \alpha_2 \sin\Delta\psi}{\sqrt{\beta_2/\beta_1}} \end{bmatrix}$$

with obvious simplification for one-turn matrix:

$$\mathbf{T} = M(s|s+C) = \begin{bmatrix} t11 & t12 \\ t21 & t22 \end{bmatrix} = \begin{bmatrix} \cos\mu + \alpha \sin\mu & \beta \sin\mu \\ -\frac{(1 + \alpha^2) \sin\mu}{\beta} & \cos\mu - \alpha \sin\mu \end{bmatrix} \mathbf{I} \cos\mu + \mathbf{J} \sin\mu; \mathbf{J} = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix}$$

from where one can get easily all functions and constants:

$$\mu = \cos^{-1} \left(\frac{t11 + t22}{2} \right) \text{ with } \text{sign}(\sin\mu) = \text{sign}(t12); \beta = t12/\sin\mu; \alpha = \frac{t11 - t22}{2\sin\mu}$$

A little bit more complex is fully coupled 2D case.

$$T(s) = \begin{bmatrix} N_1 & N_2 \\ N_3 & N_4 \end{bmatrix}; \det[T - \lambda I] = \lambda^4 - \lambda^3 \text{Tr}[T] + \lambda^2 (2 + a) - \lambda \text{Tr}[T] + 1 = 0$$

$$a = \text{Tr}[N_1] \cdot \text{Tr}[N_4] - \text{Tr}[N_2 N_3] - 2 \det N_2$$

$$(\text{note } \det N_2 = \det N_3 = 1 - \det N_1 = 1 - \det N_4)$$

Finding roots:

$$z_k = \lambda_k + \lambda_k^{-1}; z_k = \frac{\text{Tr}[N_1 + N_4]}{2} \pm \left\{ \frac{\text{Tr}^2[N_1 - N_4]}{4} + \text{Tr}[N_2 N_3] + 2 \det N_2 \right\}$$

$$X = \begin{bmatrix} x \\ P_x \\ y \\ P_y \end{bmatrix} = \text{Re } \tilde{a}_1 Y_1 + \text{Re } \tilde{a}_1 Y_2 = \text{Re } a_1 \tilde{Y}_1 + \text{Re } \tilde{a}_1 \tilde{Y}_2$$

$$Y_k = R_k + iQ_k; \quad \tilde{Y}_k = \begin{bmatrix} w_{kx} e^{i\psi_{kx}} \\ (u_{kx} + i v_{kx}) e^{i\psi_{kx}} \\ w_{ky} e^{i\psi_{ky}} \\ (u_{ky} + i v_{ky}) e^{i\psi_{ky}} \end{bmatrix}; \quad \psi_{kx}(s+C) = \psi_{kx}(s) + \mu_k; \quad \psi_{ky}(s+C) = \psi_{ky}(s) + \mu_k;$$

$$w_{kx} v_{kx} + w_{ky} v_{ky} = 1;$$

Conditions: there are

$$Y_k^{*T} S Y_k = 2i; \quad Y_1^{*T} S Y_2 = 0; \quad Y_1^T S Y_2 = 0; \quad \theta_k = \psi_{kx} - \psi_{ky}$$

$$a) \quad w_{1x} v_{1x} = w_{2y} v_{2y} = 1 - q \quad \Rightarrow \quad v_{1x} = \frac{1-q}{w_{1x}}; \quad v_{2y} = \frac{1-q}{w_{2y}}$$

$$b) \quad w_{1y} v_{1y} = w_{2x} v_{2x} = q \quad \Rightarrow \quad v_{2x} = \frac{q}{w_{2x}}; \quad w_{1y} = \frac{q}{w_{1y}}$$

$$c) \quad c = w_{1x} w_{1y} \sin \theta_1 = -w_{2x} w_{2y} \sin \theta_2$$

$$d) \quad d = w_{1x} (u_{1y} \sin \theta_1 - v_{1y} \cos \theta_1) = -w_{2x} (u_{2y} \sin \theta_2 - v_{2y} \cos \theta_2)$$

$$e) \quad e = w_{1y} (u_{1x} \sin \theta_1 + v_{1x} \cos \theta_1) = -w_{2y} (u_{2x} \sin \theta_2 + v_{2x} \cos \theta_2)$$

Conditions are result of symplecticity. Conditions a) and b) are equivalent to Poincaré's invariants conserving sum of projections on (x-px) and (y-py) planes.

$$Y_1 = \begin{bmatrix} w_{1x} e^{i\varphi_{1x}} \\ \left(u_{1x} + i \frac{q}{w_{1x}} \right) e^{i\varphi_{1x}} \\ w_{1y} e^{i\varphi_{1y}} \\ \left(u_{1y} + i \frac{1-q}{w_{1y}} \right) e^{i\varphi_{1y}} \end{bmatrix}; \quad Y_2 = \begin{bmatrix} w_{2x} e^{i\varphi_{2x}} \\ \left(u_{2x} + i \frac{1-q}{w_{2x}} \right) e^{i\varphi_{2x}} \\ w_{2y} e^{i\varphi_{2y}} \\ \left(u_{2y} + i \frac{q}{w_{2y}} \right) e^{i\varphi_{2y}} \end{bmatrix}$$

Lecture 6. Action-phase variables, applications of the parameterization.

6.1 Action-phase variables and “slow” equation of motion

One should remember that

$$\left\{ \varphi_k, I_k = \frac{a_k^2}{2} \right\} \text{ and } \left\{ \tilde{q}_k = \frac{a_k}{\sqrt{2}} e^{i\varphi_k}, \tilde{p}_k = i \frac{a_k}{\sqrt{2}} e^{i\varphi_k} \right\}$$

are Canonical variables

In many practical situations the perturbation is weak

$$\mathcal{H} = \frac{1}{2} X^T \mathbf{H}(s) X + \mathcal{H}_1(X, s). \quad (6-1)$$

and use of the action and angle variables provide the tool needed to solve the equations either exactly or using various perturbation or averaging methods. The transformation to $\left\{ \varphi_k, I_k = \frac{a_k^2}{2} \right\}$ is canonical, and it reduces equation of motion to:

$$\frac{d\varphi_k}{ds} = \frac{\partial H_1(X(I_k, \varphi_k, s))}{\partial I_k}, \quad \frac{dI_k}{ds} = - \frac{\partial H_1(X(I_k, \varphi_k, s))}{\partial \varphi_k} \quad (6-2)$$

where fast regular oscillatory terms are taken care of by our parameterization.

6.2 Applications of parameterization to standard problems

Complete parameterization developed in previous lecture can be used to solve most (if not all) of standard problems in accelerator. Incomplete list is given below:

1. Dispersion
2. Orbit distortions
3. AC dipole (periodic excitation)
4. Tune change with quadrupole (magnets) changes
5. Chromaticity
6. Beta-beat
7. Weak coupling
8. Synchro-betatron coupling
9.

We do not plan to go through all these examples while focusing on general methodology and use selected examples to demonstrate power of the symplectic linear parameterization.

Sample I. Let's start from simplest problems such as dispersion and closed orbit. We found a general form of parameterization of linear motion in Hamiltonian system, which is solution of homogeneous linear equations, where \mathbf{B} is constant vector:

$$\frac{dX}{ds} = \mathbf{D}(s) \cdot X; X = \tilde{\mathbf{U}}(s) \cdot \mathbf{B} \quad (6-I-1)$$

A standards problems is a solution of inhomogeneous equations:

$$\frac{dX}{ds} = \mathbf{D}(s) \cdot X + F(s); \quad (6-I-2)$$

It can be done analytically by varying the constant \mathbf{B} :

$$X = \tilde{\mathbf{U}}(s)B(s) \Rightarrow \tilde{\mathbf{U}} \cdot B' = F(s) \Rightarrow B' = \tilde{\mathbf{U}}^{-1}(s)F(s) \Rightarrow B(s) = B_o + \int_{s_o}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi$$

A general solution is a specific solution of inhomogeneous equation plus arbitrary solution of the homogeneous – result you expect in linear ordinary differential equations (in this case with s-dependent coefficients):

$$X(s) = \tilde{\mathbf{U}}(s)A_o + \tilde{\mathbf{U}}(s) \int_{s_o}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi; \tilde{\mathbf{U}}^{-1} = \frac{i}{2} \mathbf{S} \cdot \tilde{\mathbf{U}}^T \cdot \mathbf{S} \quad (6-I-3)$$

For a periodic force (orbit distortions, dispersion function) $F(s+C) = F(s)$ one can find periodic solution $X(s+C) = X(s)$:

$$\tilde{\mathbf{U}}^{-1}(s) \times \left\{ \tilde{\mathbf{U}}(s)A_o + \tilde{\mathbf{U}}(s) \int_{s_o}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi = \tilde{\mathbf{U}}(s+C)A_o + \tilde{\mathbf{U}}(s+C) \int_{s_o}^{s+C} \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi \right\}$$

$$A_o(\mathbf{I} - \Lambda) = \Lambda \int_s^{s+C} \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi \equiv \int_{s-C}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi \Rightarrow A_o = (\mathbf{I} - \Lambda)^{-1} \int_{s-C}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi \quad (6-I-4)$$

$$X(s) = \tilde{\mathbf{U}}(s)(\mathbf{I} - \Lambda)^{-1} \int_{s-C}^s \tilde{\mathbf{U}}^{-1}(\xi)F(\xi)d\xi$$

It is easy to see that $X(s+C) = X(s)$ exists if none of the eigen values is not equal 1 – otherwise matrix $(\mathbf{I} - \Lambda)$ would have zero determinant and can not be inverted!

Specific examples: Orbit distortions caused by the field errors, transverse dispersion.

When the conditions for the equilibrium particle and the reference trajectory are slightly violated:

$$X^T = \{x, P_1, y, P_3, \tau, \delta\}; F^T = \left\{ 0, \frac{e}{c} \left(\delta B_y + \frac{E_o}{p_o c} \delta E_x \right), 0, \frac{e}{c} \left(\delta B_x - \frac{E_o}{p_o c} \delta E_y \right), 0, 0 \right\}$$

$$K(s) \equiv \frac{1}{\rho(s)} - \frac{e}{p_o c} \left(B_y|_{ref} + \frac{E_o}{p_o c} E_x|_{ref} \right) - f_x; \quad f_x = \frac{e}{p_o c} \left(\delta B_y + \frac{E_o}{p_o c} \delta E_x \right); \quad (6-I-5)$$

$$\frac{e}{p_o c} B_x \left(\left|_{ref} - \frac{E_o}{p_o c} E_y|_{ref} \right. \right) = -f_y = \frac{e}{p_o c} \left(\delta B_x - \frac{E_o}{p_o c} \delta E_y \right)$$

Plugging (6-I-5) into (6-I-4) will give one the periodic closed orbit for such a case. For finding

reduces to

$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_3 - yP_1) + \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + g_x x \delta + g_y y \delta$$

with

$$F = S \frac{\partial H}{\partial X} = \left\{ 0, -g_x, 0, -g_y, 0, -\frac{m^2 c^2}{p_o^3} \right\}^T.$$

1D ACCELERATOR

$$X(s) = \int_{s-C}^s \begin{bmatrix} w(s) & w(s) \\ w'(s) + i/w(s) & w'(s) - i/w(s) \end{bmatrix} \cdot \begin{bmatrix} (w'(\xi) - i/w(\xi)) e^{i\psi(s) - i\psi(\xi)} (1 - e^{i\mu})^{-1} & -w(\xi) e^{i\psi(s) - i\psi(\xi)} (1 - e^{i\mu})^{-1} \\ (-w'(\xi) + i/w(\xi)) e^{i\psi(\xi) - i\psi(s)} (1 - e^{i\mu})^{-1} & w(\xi) e^{i\psi(\xi) - i\psi(s)} (1 - e^{i\mu})^{-1} \end{bmatrix} \frac{iF(\xi)}{2} d\xi$$

$$X^T = \{x, x'\}; F^T = \frac{e}{p_o c} \delta B_y \{0,1\} \text{ - orbit; } F^T = K(s) \{0,1\} \text{ for dispersion, i.e. } F^T = f(s) \{0,1\}$$

$$X(s) = \int_{s-C}^s \left[\frac{\operatorname{Re} \left(w(s) w(\xi) e^{i(\psi(s) - \psi(\xi) - \mu/2)} \left(\frac{e^{-i\mu/2} - e^{i\mu/2}}{-i} \right)^{-1} \right)}{\dots} \right] f(\xi) d\xi$$

$$\text{i.e. } x(s) = \frac{w(s)}{2 \sin \mu/2} \oint_C f(\xi) w(\xi) \cos(\psi(s) - \psi(\xi) - \mu/2) d\xi$$

First example: orbit distortion

$$f_x(s) = -\frac{e \delta B_y(s)}{p_o c}; \quad f_y(s) = \frac{e \delta B_x(s)}{p_o c}$$

$$\delta x(s) = -\frac{w(s)}{2 \sin \mu/2} \oint_C \frac{e \delta B_y(\xi)}{p_o c} w(\xi) \cos(\psi(s) - \psi(\xi) - \mu/2) d\xi$$

$$\delta y(s) = \frac{w(s)}{2 \sin \mu/2} \oint_C \frac{e \delta B_x(\xi)}{p_o c} w(\xi) \cos(\psi(s) - \psi(\xi) - \mu/2) d\xi$$

but this is not the end of the story for horizontal motion! (what about change of the orbiting time?)

Second example: Dispersion

$$f_x(s) = K_o(s) \pi_l = K_o(s) \pi_\tau / \beta_o;$$

$$x(s) = D(s) \cdot \pi_l = D(s) \cdot \pi_\tau / \beta_o;$$

$$D(s) = -\frac{w(s)}{2 \sin \mu/2} \oint_C K_o(\xi) w(\xi) \cos(\psi(s) - \psi(\xi) - \mu/2) d\xi$$

Sample II: Beta-beat – 1D case

It is simple fact that any solution can be expanded upon the eigen vectors of periodic system (FOD cell repeated again and again is an example). Let 's consider that at azimuth $s=s_0$ initial value of “injected” eigen vector \mathbf{V} being different from the periodic solution \mathbf{Y} . We expand it as

$$\begin{aligned}
 V(s_0) &= aY_k(s_0) + bY_k^*(s_0) = \begin{bmatrix} v_o \\ v'_o + \frac{i}{v_o} \end{bmatrix}; Y_k = \begin{bmatrix} w_o \\ w'_o + \frac{i}{w_o} \end{bmatrix} \\
 a &= \frac{1}{2i} Y_k^{*T}(s_0)SV(s_0); \quad b = \frac{1}{-2i} Y_k^T(s_0)SV(s_0) \\
 a &= \frac{1}{2i} \left\{ v_o w'_o - w_o v'_o + i \left(\frac{v_o}{w_o} + \frac{w_o}{v_o} \right) \right\}; \quad b = -\frac{1}{2i} \left\{ v_o w'_o - w_o v'_o + i \left(\frac{v_o}{w_o} - \frac{w_o}{v_o} \right) \right\}; \\
 \frac{d}{ds} \tilde{Y}(s) &= \mathbf{D}(s) \cdot \tilde{Y}(s); \quad \tilde{Y}(s) = Y(s)e^{i\psi(s)}; Y(s+C) = Y(s)
 \end{aligned} \tag{6-II-1}$$

It is self-evident that

$$\begin{aligned}
 \tilde{V}' &= D\tilde{V}; \quad \tilde{V}(s) = a\tilde{Y}_k(s) + b\tilde{Y}_k^*(s) = \begin{bmatrix} v \\ v' + \frac{i}{v} \end{bmatrix} e^{i\psi} = Y_k = a \begin{bmatrix} w \\ w' + \frac{i}{w} \end{bmatrix} e^{i\psi} + b \begin{bmatrix} w \\ w' - \frac{i}{w} \end{bmatrix} e^{-i\psi} \\
 |v|^2 &= \frac{|w|^2}{4} |ae^{i\psi} + be^{-i\psi}|^2 = \frac{|w|^2}{4} (|a|^2 + |b|^2 - 2\text{Re}(ab^* e^{2i\psi}))
 \end{aligned} \tag{6-II-2}$$

i.e. beta-function will beat with double of the betatron phase.

Sample III: Perturbation theory (ala quantum mechanics)

Small variation of the linear Hamiltonian terms (including coupling)

$$\frac{dX}{ds} = (\mathbf{D}(s) + \varepsilon \mathbf{D}_1(s)) \cdot X; \tag{6-III-1}$$

Assuming that changes are very small we can express the changes in the eigen vectors using basis of (VII):

$$\begin{aligned}
 \tilde{Y}_{1k} &= \tilde{Y}_k e^{i\delta\phi_k} + \varepsilon_j \tilde{Y}_j; \quad \frac{d\tilde{Y}_{1k}}{ds} = (\mathbf{D}(s) + \varepsilon \mathbf{D}_1(s)) \cdot \tilde{Y}_{1k} + o(\varepsilon^2); \\
 \tilde{Y}_{1k} &= \tilde{Y}_k e^{i\delta\phi_k} + \varepsilon_j \tilde{Y}_j; \quad \frac{d\tilde{Y}_{1k}}{ds} = (\mathbf{D}(s) + \varepsilon \mathbf{D}_1(s)) \cdot \tilde{Y}_{1k} + o(\varepsilon^2); \\
 \frac{d\tilde{Y}_k}{ds} &= \mathbf{D}(s)\tilde{Y}_k; \quad \frac{d\varepsilon_j}{ds} \tilde{Y}_j = \varepsilon \mathbf{D}_1(s)\tilde{Y}_k; \Rightarrow \varepsilon_j = \varepsilon_{j0} + \frac{\varepsilon}{2i} \int \tilde{Y}_j^{*T}(\xi) \mathbf{SD}_1(\xi) \tilde{Y}_k(\xi) d\xi; \\
 i \frac{d\delta\phi_k}{ds} \tilde{Y}_k &= \varepsilon \mathbf{D}_1(s)\tilde{Y}_k \Rightarrow \delta\phi_k = \frac{\varepsilon}{2} \int \tilde{Y}_k^{*T}(\xi) \mathbf{SD}_1(\xi) \tilde{Y}_k(\xi) d\xi;
 \end{aligned} \tag{6-III-2}$$

$$\begin{aligned}\varepsilon_j(s+C) &= \varepsilon_j(s) \\ \varepsilon_j &= \frac{\varepsilon}{2i(e^{i(\mu_k-\mu_i)}-1)} \int_0^C e^{i(\psi_k(s)-\psi_i(\xi))} Y_j^{*T}(\xi) \mathbf{SD}_1(\xi) Y_k(\xi) d\xi;\end{aligned}\quad (6-III-2)$$

One should be aware of the resonant case $e^{i(\mu_k-\mu_i)} = 1$, when one should solve self-consistently the set of (6-II-2). It is well known case well described in weak coupling resonance case or in the case of parametric resonance.

Sample IV: small variation of the gradient. It can come from errors in quadrupoles or from a deviation of the energy from the reference value. In 1D case (reduced) it is simple addition to the Hamiltonian: (including sextupole term!)

$$\begin{aligned}H_1 &= \delta K_1 \frac{z^2}{2}; \quad z = \{x, y\}; \\ \pi_i &= p/p_o - 1 \\ \delta K_{1\ x,y} &= \mp \delta \left(\frac{e}{pc} \frac{\partial B_y}{\partial x} \right) = \mp \left(\frac{e}{pc} \delta \frac{\partial B_y}{\partial x} \right) - K_1 \pi_i \mp \left(\frac{e}{pc} \frac{\partial^2 B_y}{\partial x^2} D_x \right) \pi_i + o(\pi_i^2)\end{aligned}\quad (6-IV-1)$$

Plugging our parameterization into the residual Hamiltonian we get:

$$\begin{aligned}z &= w(s) \sqrt{2I} \cos(\psi(s) + \varphi) \\ H_1 &= \delta K_1(s) \cdot w^2(s) \cdot I \cdot \cos^2(\psi(s) + \varphi)\end{aligned}\quad (6-IV-2)$$

The easiest way is to average the Hamiltonian (on the phase of fast betatron oscillation – our change is small! And does not effect them strongly) to have a well-know fact that the beta-function is also a Green function (modulo 4π) of the tune response on the variation of the focusing strength.

$$\begin{aligned}\langle H_1 \rangle &= \frac{\langle \delta K_1(s) \cdot w^2(s) \rangle}{2} \cdot I \equiv \frac{\langle \delta K_1(s) \cdot \beta(s) \rangle}{2} \cdot I \\ \langle \varphi' \rangle &= \frac{\partial \langle H_1 \rangle}{\partial I} = \frac{\langle \delta K_1(s) \cdot \beta(s) \rangle}{2}; \\ \Delta\varphi &= \frac{1}{2} \oint \delta K_1(s) \cdot \beta(s) ds; \quad \Delta Q = \frac{\Delta\varphi}{2\pi} = \frac{1}{4\pi} \oint \delta K_1(s) \cdot \beta(s) ds;\end{aligned}\quad (6-IV-3)$$

Direct way will be to put it into the equations (43) and to find just the same, that $\langle I' \rangle = 0$ and the above result.

Finally, putting a weak thin lens as a perturbation gives a classical relation:

$$\begin{aligned}\delta K_1(s) &= \frac{1}{f} \delta(s - s_o) \\ \Delta Q &= \frac{\Delta \varphi}{2\pi} = \frac{1}{4\pi} \frac{\beta_o(s)}{f}\end{aligned}\tag{6-IV-1}$$

In order to find variation of the eigen vector (i.e. bet-function) one will need to use the *Perturbation theory (ala quantum mechanics) – see above.*

Lecture 7. Synchrotron oscillations and effects of synchrotron radiation

First, let's try to finish our 3D oscillator picture, before jumping into a simplified picture usually used in accelerator literature. Passing through a magnetic system will result in as 6x6 matrix

$$\mathbf{T}_{6 \times 6} = \begin{bmatrix} T_{4 \times 4} & 0 & R \\ R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{L3-61})$$

In the absence of RF cavity, the solutions are already known:

$$\begin{aligned} X &= \text{Re } a_1 \mathbf{Y}_1(s) e^{i\psi_1(s)} + \text{Re } a_2 \mathbf{Y}_2(s) e^{i\psi_2(s)} + \mathbf{D}(s) \pi_\tau + \Theta \cdot s \cdot \pi_\tau \\ \mathbf{Y}_1(s+C) &= \mathbf{Y}_1(s); \quad \mathbf{Y}_2(s+C) = \mathbf{Y}_2(s); \\ \psi_1(s+C) &= \psi_1(s) + \mu_1; \quad \psi_2(s+C) = \psi_2(s) + \mu_2; \\ \mathbf{D}(s+C) &= \mathbf{D}(s); \end{aligned} \quad (7-1)$$

which has one growing term – the arrival time continuously grows if energy of particle deviates from the reference value. We already discussed the x and y components of the eigen vectors Y. Let's define what is the form is:

$$\begin{aligned} \mathbf{D}(s) &= \begin{bmatrix} D(s) \\ d_{56}(s) \\ 1 \end{bmatrix}; \quad D(s) = (T - I)^{-1} R \quad ! \\ \Theta &= \begin{bmatrix} 0_4 \\ -\eta \\ 1 \end{bmatrix}; \quad \eta = R_{56}(C)/C = \frac{1}{\beta_o^2 \gamma_o^2} - \alpha_C; \quad \alpha_C = - \int_0^C (g_x(s) D_x(s) + g_y(s) D_y(s)) ds \end{aligned} \quad (7-2)$$

In normal uncoupled case the orbit compaction factor $\alpha_C = \int_0^C K(s) D_x(s) ds / \beta_o$ is determined by the product horizontal dispersion and the orbit curvature. It is usually positive number, because particles in typical accelerator go to outside of the ring ($D_x > 0$). For such ring, there is an energy when time of flight around the ring does not depend (in the first order!) on the particle's energy:

$$\eta = 0 \quad \Rightarrow \quad \beta_t \gamma_t = \frac{1}{\sqrt{\alpha_C}} \quad (7-3)$$

This energy is called transition energy and it can be very important for many hadron rings. An electron storage ring with very low energy electron rings (which is next to impossible to find) would have similar problem, which is called transition.

One can design (and a number of people did build them) a storage ring with negative and zero compaction factor. Zero value of compaction value usually spell trouble... But negative one will work for any ring and will not have transition energy (so called lattice with imaginary transition energy – see (7-3)).

Symplectic condition implies that the term in the 5-th position is defined:

$$\mathbf{D}^T \mathbf{S} \mathbf{Y}_k = 0 \Rightarrow \mathbf{Y}_k(s) = \begin{bmatrix} Y_k(s) \\ D^T \mathbf{S} Y_k \\ 0 \end{bmatrix}; \quad (7-4)$$

Thus, all terms in equation (7-1) are defined. Now we will include a short RF cavity, where we linearized the voltage

$$\mathbf{M}_{RF} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_l \end{bmatrix}; M_l = \begin{bmatrix} 1 & 0 \\ -q & 1 \end{bmatrix}; \quad q = u \cdot l_{RF} = -\frac{e}{p_0 c^2} \frac{\partial V_{rf}}{\partial t} \quad (L3-64)$$

$$\mathbf{T}_{total} = \begin{bmatrix} I_{4 \times 4} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -q & 1 \end{bmatrix} \cdot \begin{bmatrix} T_{4 \times 4} & 0 & R \\ R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & 1 & R_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} T_{4 \times 4} & 0 & R \\ R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & 1 & R_{56} \\ -q R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & -q & 1 - q R_{56} \end{bmatrix} \quad (7-5)$$

It may be surprising but there is no easy way of solving this system and finding detailed solution. We can use our perturbation method to find change in the betatron frequencies caused by the RF cavity using:

$$\begin{aligned} \delta\phi_k &= \frac{1}{2} \int \tilde{Y}_k^{*T}(\xi) \Delta \mathbf{H}(\xi) \tilde{Y}_k(\xi) d\xi; \\ \Delta \mathbf{H}_{55} &= q \delta(s - s_{rf}); \quad \delta Q_k = \frac{q}{4\pi} |Y_{k5}|^2 = \frac{q}{4\pi} |D^T \mathbf{S} Y_k|^2 \\ \delta Q_k &= -\frac{1}{4\pi} \frac{e}{p_0 c^2} \frac{\partial V_{rf}}{\partial t} |D^T \mathbf{S} Y_k|^2 \end{aligned} \quad (7-6)$$

This is probably one of the results you will have hard time finding in the accelerator literature. The term $|D^T \mathbf{S} Y_k|^2$ is actually well know – it is responsible for quantum excitation of the betatron oscillation during the process of synchrotron radiation (see next sections). For uncoupled transverse motion, vertical dispersion is equal zero and only Qx is affected by an RF placed in the straight section with non-zero dispersion:

$$|D^T \mathbf{S} Y_k|^2 = (w'_x D_x - w_x D'_x)^2 + \frac{D_x^2}{w_x^2} \equiv \frac{\gamma_x D_x^2 + 2\alpha_x D_x D'_x + \beta_x D_x'^2}{\beta_x} \quad (7-7)$$

Fortunately, this tune shifts are extremely small. In contrast with transverse motion, longitudinal motion without RF is degenerated (zero tune!). We have two root eigen vectors \mathbf{D} , Θ . Knowing that synchrotron oscillation are slow, we may try to construct a vector, which can serve as a first order approximation:

$$\mathbf{Y}_3 = \begin{bmatrix} iD/w_l \\ w_l \\ i/w_l \end{bmatrix} \quad (7-8)$$

which is symplectically orthogonal to the transverse vectors. Trying it with the matrix (7-5)

$$\begin{bmatrix} T_{4 \times 4} & 0 & R \\ R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & 1 & R_{56} \\ -q R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} & -q & 1 - q R_{56} \end{bmatrix} \cdot \begin{bmatrix} iD/w_l \\ w_l \\ i/w_l \end{bmatrix} = \begin{bmatrix} iD/w_l \\ w_l + i(R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} D + R_{56})/w_l \\ -q w_l - i(1 - q R_{56} - q R^T \mathbf{S}_{4 \times 4} T_{4 \times 4} D)/w_l \end{bmatrix}$$

does not deliver a lot of promise. One case, which is easy to handle is when the cavity is installed in so-called dispersion free section, i.e. $D=0$, $R=0$. In this case, longitudinal matrix is completely separated from the transverse and

$$\mathbf{T}_{total} = \begin{bmatrix} T_{4 \times 4} & 0 \\ 0 & T_l \end{bmatrix}; \quad T_l = \begin{bmatrix} 1 & R_{56} \\ -q & 1 - qR_{56} \end{bmatrix}. \quad (7-9)$$

The motions is stable only when $0 < qR_{56} < 4$ and it will give us (using standard Courant-Snyder parameterization)

$$\cos \mu_l = 1 - qR_{56}/2; \quad w_l^2 = \beta_l = R_{56} / \sqrt{1 - (1 - qR_{56}/2)^2}; \quad \alpha_l = \frac{qR_{56}}{2\sqrt{1 - (1 - qR_{56}/2)^2}} \quad (7-9)$$

What is typical for the majority of the accelerators, that $qR_{56} \ll 1$. In this case we can expand the equations in (7-9) to find:

$$\mu_l \cong \sqrt{qR_{56}}; \quad Q_s \cong \frac{\sqrt{qR_{56}}}{2\pi}; \quad w_l^2 = \beta_l \cong \sqrt{\frac{R_{56}}{q}}; \quad \alpha_l = \frac{\sqrt{qR_{56}}}{2} \sim \mu_l/2 \quad (7-9')$$

The synchrotron tune is usually very low with typical values of $Q_s \sim 0.0001$ to 0.01 . It means also that $\alpha_l \sim \pi Q_s \ll 1$ and the form (7-8) for the eigen vector is actually rather good approximation. We also can conclude that betatron tunes in (7-6) are proportional to square of Q_s :

$$\delta Q_k = \frac{q}{4\pi} |D^T S Y_k|^2 \cong \frac{\pi Q_s^2}{R_{56}} |D^T S Y_k|^2 \quad (7-10)$$

In order to estimate it, we scale it as

$$|D^T S Y_k|^2 \sim D_x^2 / \beta_x; \quad R_{56} \sim 2\pi D_x; \quad \delta Q_x \sim \frac{Q_s^2}{2} \frac{D_x}{\beta_x}, \quad (7-11)$$

i.e. typical change of the betatron tune is below 10^{-4} , i.e. is very-very small. Only when synchrotron tune are approaching values ~ 1 , the synchrotron oscillations become heavily cross-talking with betatron oscillations. In other cases, the synchrotron oscillations can be considered as adiabatically slow as the background for fast betatron oscillations. Hence, the standard treatment from the book.

Lecture 8. Effects of synchrotron radiation

I. Oscillator

Before we embark on detail studies of radiation effects on the beams in accelerators, let's look on a very simple model of harmonic oscillator:

$$H = \frac{P^2}{2m} + k \frac{x^2}{2} \quad \text{or} \quad h = \frac{p^2}{2} + \omega^2 \frac{x^2}{2}; \quad \omega = \sqrt{\frac{k}{m}} \quad (8-1)$$

described by differential equations

$$\begin{aligned} x' &= \frac{\partial h}{\partial p} = p; \quad p' \equiv x'' = -\frac{\partial h}{\partial x} = -\omega^2 x; \quad x = A \cdot \cos(\omega t + \varphi); \quad p = -A\omega \cdot \sin(\omega t + \varphi) \\ X' &= \begin{bmatrix} x' \\ p' \end{bmatrix} = S \frac{\partial h}{\partial X} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \cdot X; \quad X = \text{Re } a Y e^{i(\omega t + \varphi)}; \quad Y = \begin{bmatrix} 1/\sqrt{\omega} \\ i\sqrt{\omega} \end{bmatrix}; \quad \left\{ I = \frac{a^2}{2}, \varphi \right\} \end{aligned} \quad (8-2)$$

Let's add a weak friction $\varepsilon \ll \omega$:

$$\begin{aligned} p &= x'; \quad p' = -\omega^2 x - 2\alpha p; \quad X' = D \cdot X = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\alpha \end{bmatrix} \cdot X; \\ \det \begin{bmatrix} \lambda & -1 \\ \omega^2 & \lambda + 2\alpha \end{bmatrix} &= \lambda(\lambda + 2\alpha) + \omega^2 = 0; \quad \lambda = -\alpha \pm i\omega_1; \quad \omega_1 = \sqrt{\omega^2 - \alpha^2}; \\ x &= A \cdot e^{-\alpha t} \cdot \cos(\omega_1 t); \quad p = -A \cdot e^{-\alpha t} (\alpha \cdot \cos(\omega_1 t) + \sin(\omega_1 t)); \\ X &= \text{Re } a Y e^{\lambda t + i\varphi} = a \cdot e^{-\alpha t} \text{Re } Y e^{i\omega_1 t}; \quad Y = \begin{bmatrix} 1/\sqrt{\lambda} \\ i\sqrt{\lambda} \end{bmatrix}; \end{aligned} \quad (8-3)$$

which make a very small change to the frequency of the oscillations, but make free oscillations slowly decaying. NOTE that damping decrement α is only a half of that of simple decay:

$$p' = -2\alpha p \rightarrow p = p_0 e^{-2\alpha t}.$$

This is the result of oscillations, where, time-averaged, only half energy is in the kinetic energy $p^2/2$, which decays. The potential energy decays only through its coupling to the kinetic energy via oscillations. The action of the oscillator, I , which represent the area of the phase space, decays with the simple decay rate towards zero:

$$I' = \left(\frac{a^2}{2} \right)' = -2\alpha I; \quad I = I_0 e^{-2\alpha t}, \quad (8-4)$$

while the oscillator phase does not stationary point or any decay.

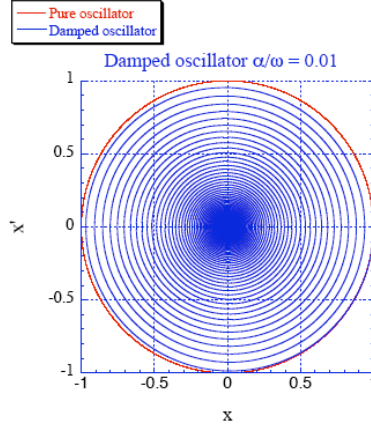


Fig. 1 Poincaré plot of trajectories of normal and damped oscillator in dimensionless coordinates $x/a; x'/\omega a$.

NOTE the second fact, that trace of matrix $D = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\alpha \end{bmatrix}$ gives the damping rate of the oscillator phase space volume. Let's add a random noise to the equations:

$$x' = p + \delta x(t); \quad p' = -\omega^2 x - 2\alpha p + \delta x'(t); \quad X' = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\alpha \end{bmatrix} \cdot X + \begin{bmatrix} \delta x(t) \\ \delta x'(t) \end{bmatrix}; \quad (8-5)$$

$$\langle \delta x(t) \rangle = 0; \quad \langle \delta x'(t) \rangle = 0.$$

where $\delta x(t), \delta x'(t)$ are “sudden” and randomly distributed in time and amplitude jumps.

One can easily calculate change in the amplitude and phase of the oscillator caused by a random kick:

$$\delta(ae^{i\varphi}) = -ie^{-i\omega_1 t} Y^{*T} \cdot S \cdot \begin{bmatrix} \delta x \\ \delta x' \end{bmatrix};$$

$$\delta a + ia\delta\varphi \cong -ie^{-i(\omega_1 t + \varphi)} Y^{*T} \cdot S \cdot (\delta x' / \sqrt{\omega} - i\delta x \sqrt{\omega}) \quad (8-6)$$

$$\langle \delta a \rangle = 0; \langle \delta\delta\varphi \rangle = 0; \quad \delta I = a\delta a + \delta a^2 / 2; \quad \langle \delta I \rangle = \left\langle \frac{\delta a^2}{2} \right\rangle = \frac{\langle \delta x'^2 \rangle / \omega + \omega \langle \delta x^2 \rangle}{2}$$

Thus, the only one thing is well determined – the average change of the action, I . Adding damping term (8-4) to (8-6) we have:

$$\langle I \rangle' = -2\alpha \langle I \rangle + D/2; \quad D = \langle \delta x'^2 \rangle / \omega + \omega \langle \delta x^2 \rangle; \quad (8-7)$$

with stationary solution for average action (emittance) and RMS amplitude of the ensemble of oscillators:

$$\langle I \rangle = \frac{D}{4\alpha}, \quad i.e. \quad \varepsilon = \langle a^2 \rangle = \frac{D}{2\alpha} = \frac{\langle \delta x'^2 \rangle / \omega + \omega \langle \delta x^2 \rangle}{2\alpha}, \quad (8-8)$$

where ε is called emittance – phase space area occupied divided by π - of the ensemble of oscillators ($\varepsilon^2 = \langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2$).

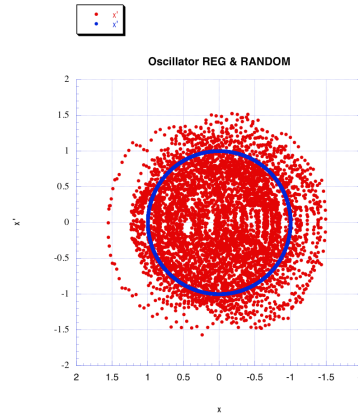


Fig. 2 Poincaré plot of trajectories of normal and damped oscillator with random kicks (in dimensionless coordinates x/a ; $x'/\omega a$).

Figure 2 shows Poincaré plot of few hundreds of such an oscillators starting from the same initial conditions (1,0) and going around for few damping times. Overall, a large ensemble of oscillators (or equivalently distribution of (x,p) for one oscillator in very long time – via Ergodic theorem, see http://en.wikipedia.org/wiki/Ergodic_theory) is described by distribution function. Because (I, φ) is Canonical pair, it is natural to use them as independent variables for the distribution function, $f(I, \varphi, t)$. Few facts are apparent: the phases of oscillators walk randomly and because phase is cyclic function it is distributed evenly in the interval $\{-\pi, \pi\}$. Thus, there is no dependence on φ : $\frac{\partial f}{\partial \varphi} = 0$. Finding distribution function of the action, $f(I, t)$, requires solution of

Fokker-Plank equation (read your favorite stat-mech book or see Appendix H):

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial I} \left(f \frac{dI}{dt} \right) - \frac{1}{2} \frac{\partial}{\partial I} \left(\langle \delta I^2 \rangle \frac{\partial f}{\partial I} \right) = 0. \quad (8-9)$$

We are interested in stationary solution, $\frac{\partial f}{\partial t} = 0$, $\frac{d}{dI} \left(f \frac{dI}{dt} - \frac{1}{2} \langle \delta I^2 \rangle \frac{df}{dI} \right) = 0$

$$\langle \delta I^2 \rangle = 2I \langle \delta I \rangle \Rightarrow (2\alpha f + \langle \delta I \rangle f') I = \text{const} \rightarrow 0 \quad (8-10)$$

$$\ln f' = -\frac{2\alpha}{\langle \delta I \rangle}; \Rightarrow f = a \cdot e^{-I/I_0}; I_0 = \langle \delta I \rangle / 2\alpha$$

Remembering that $I = \frac{x'^2 \omega + \omega \delta x^2}{2}$, it gives us just a trivial Gaussian distribution for the oscillators

$$f(x, x') = \frac{1}{2\pi\varepsilon} e^{-\frac{a^2}{2\varepsilon}} = \frac{1}{2\pi\varepsilon} \exp\left(-\frac{x^2 \omega + x'^2 / \omega}{2\varepsilon}\right) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} \frac{1}{\sqrt{2\pi}\sigma_{x'}} e^{-\frac{x'^2}{2\sigma_{x'}^2}} \quad (8-11)$$

$$\sigma_x = \sqrt{\varepsilon/\omega}; \sigma_{x'} = \sqrt{\varepsilon \cdot \omega}$$

where we normalize it as $\iint f(x,x') dx dx' = 1$.

Conclusions are easy to remember: Position independent diffusion in the presence of linear damping results in stationary Gaussian distribution of the oscillator's amplitudes, positions and velocities. Phases of individual oscillators become random. Naturally, this process takes few damping times $T_d=1/\alpha$, if initial distribution deviates from the stationary.

II. Synchrotron radiation

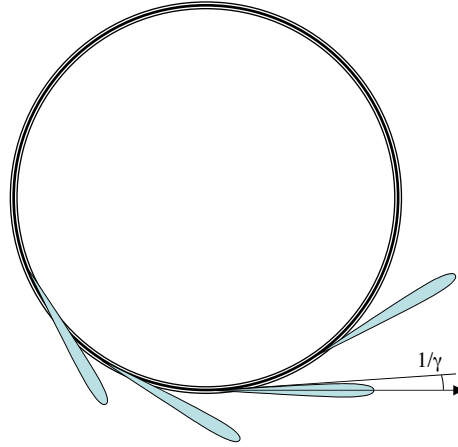


Fig. 3. Synchrotron radiation is a fan of well-directed radiation with vertical opening $\sim 1/\gamma$. Radiation at a certain point of curved trajectory directed along momentum of the particle and is confined within $\sim 1/\gamma$ opening angle in both horizontal and vertical directions.

Detailed description of synchrotron radiation can be found in your favorite E&M book (Jackson?). Some additional material about the synchrotron radiation is in the handout. Here we will need only few specific features of synchrotron radiation, which we will use without derivations. We will also assume that a) our particles are ultra-relativistic, $\gamma \gg 1$; b) losses for synchrotron radiation per turn are a small portion of the particle's energy, i.e. we can treat it as a perturbation which introduces some damping, while not affecting tunes of the particle (remember for the oscillator it is $\sim (\alpha/\omega)^2$). For ultra-relativistic particles E and pc become essentially indistinguishable:

$$E = pc / \sqrt{1 - \gamma^{-2}} \cong pc(1 + 1/2\gamma^2) = pc + o(\gamma^{-2}),$$

hence we will use them this way.

One of the most critical feature for the damping of the transverse of the synchrotron radiation that it is local and is confined within a solid angle $\sim 1/\gamma^2$ around the direction of the particle MECHANICAL momentum. In other words, the loss of the particle momentum is proportional to the total energy loss of the particle with recoil directed against the direction of the momentum:

$$\frac{\vec{p}'}{p} = -\frac{1}{p} \frac{dp_{rad}}{ds} \cdot \frac{\vec{p}}{p} \cong -\frac{1}{E} \frac{dE_{rad}}{ds} \cdot (\vec{\tau} + \vec{n} \cdot x' + \vec{b} \cdot y') + O(x'^2, y'^2, \gamma^{-2}) \quad (8-12)$$

Synchrotron radiation has a white spectrum, which growth as $\omega^{1/3}$ at low frequencies as $\omega < \omega_c$ and falls exponentially at high frequencies $\omega > \omega_c$. The critical frequency, $\omega_c = 2\pi c/\lambda_c$ divides radiated energy by halves: the half $\omega > \omega_c$ and the other half at $\omega < \omega_c$. Critical wavelength of λ_c is $\sim R/\gamma^3$ and

$$\omega_c = \frac{2}{3} \gamma^3 \frac{c}{\rho},$$

where ρ is the radius of curvature. Hence, the most of the radiation happens in the bending magnets. Total radiated power of synchrotron radiation is (Z is the number of unit charges in the particle – $Z=1$ for all elementary particles, but ions):

$$\frac{dE_{rad}}{dt} = \frac{2}{3} \frac{e^2 Z^2 \gamma^2}{m^2 c^3} \left\{ \left(\vec{E} - [\vec{\beta} \times \vec{B}] \right)^2 - \left(\vec{\beta} \cdot \vec{E} \right)^2 \right\}; \vec{\beta} = \vec{v}/c. \quad (8-13)$$

One should note that synchrotron radiation caused by acceleration along the direction of the momentum is energy independent and very weak (to the level that it can be ignored in practical cases). At the same time, acceleration normal to the direction of the motion (bending) causes radiation proportional to the square of the particles energy. It is also worth mentioning that radiation power is inverse proportional to the particles mass squared, thus a proton will radiate $\sim 4 \cdot 10^6$ less power from the same bending magnet compared with an electron or positron with the same γ . For the particles with the same energy there is additional factor $4 \cdot 10^6$. As the result of such dramatic dependence is that synchrotron radiation does not play any significant role in the hadron accelerators (even in 7 TeV LHC damping time is about 13 hours for protons with $\gamma \sim 7,000$!), but extremely important for lepton storage rings.

Returning to s as independent coordinate we can write:

$$\begin{aligned} \frac{\vec{p}'}{p} &\equiv -\frac{1}{E} \frac{dE_{rad}}{ds} \cdot (\vec{\tau} + \vec{n} \cdot x' + \vec{b} \cdot y') + O(x'^2, y'^2, \gamma^{-2}) \\ I &= \frac{1}{E_o} \frac{dE_{rad}}{ds} = \frac{2}{3} \frac{e^2 Z^2 E^2}{m^4 c^8} \left\{ \left(\vec{E} - [\vec{\beta} \times \vec{B}] \right)^2 - \left(\vec{\beta} \cdot \vec{E} \right)^2 \right\} (1 + Kx) \end{aligned} \quad (8-14)$$

The average energy loss for synchrotron radiation is restored by an RF cavity. By design, the RF cavity boosts only longitudinal momentum of the particles (along s), while leaving transverse momenta unchanged. This feature complete the circuit, which is important for understanding of the radiation damping.

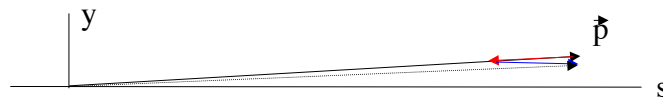


Fig. 4. Particle loses parts of its vertical momentum during radiation process. RF cavity restores (in average) only longitudinal component of the loss momentum. As the result, transverse momentum is damped.

Before we go into full fledged calculations, lets look at a simple picture of what's happening in the vertical plane in a ring without transverse coupling (i.e. majority of the ring's designs). As shown in Fig. 2, the radiation reduces transverse component of the particle's momentum and

$$\frac{p'_y}{p_o} = -\frac{1}{E} \frac{dE_{rad}}{ds} \cdot y'; \quad \alpha = \left\langle \frac{1}{2E} \frac{dE_{rad}}{ds} \right\rangle_c \quad (8-15)$$

$$y \cong \sqrt{a\beta_y} \cos(\psi_y + \varphi) \cdot e^{-\alpha s}$$

where we neglected effect of the distribution of the radiation along the circumference of the machine replacing it by an average damping. Thus, we can conclude that the vertical betatron oscillations in the storage ring will damp e-fold when particle radiated twice its energy.

In a storage ring with separated functions (no gradients in the bending magnets) and rectangular magnets (too many details...- i.e. (1+Kx) term in (1-14) is canceled by the straight edges of the bending magnet) – total radiated power does not depend on the horizontal position, x. In this case, picture shown in Fig. 4 is applicable directly to the horizontal betatron oscillations as well: i.e. the horizontal betatron oscillations in the storage ring will damp e-fold when particle radiated twice its energy. Surprisingly it is almost true for most of modern synchrotron radiation sources.

What about synchrotron (energy) oscillations – they will damp too in the above scenario. Equations of synchrotron motion for radiating ultra-relativistic particle are:

$$\frac{d\delta}{dn} = -\Delta E_{rad}(\delta) + eZ \cdot V_{rf} \sin(\omega_{rf} t_o - \tau/c) \quad (8-16)$$

$$\frac{d\tau}{dn} = -\alpha_c C \frac{\delta}{E_o};$$

with condition of an equilibrium particle of $eZ \cdot V_{rf} \sin \varphi_0 = \Delta E_{rad}(0)$, where $\varphi_0 = \omega_{rf} t_o$ is the equilibrium phase necessary for RF to compensate for average radiation energy loss. Expanding radiation about the design energy and sin about the equilibrium phase, we get for linear motion:

$$\frac{d\delta}{dn} = -\frac{2\Delta E_{rad}}{E_o} \delta - \frac{eZ \cdot V_{rf} \cos(\varphi_o)}{c} \tau; \quad \frac{d\tau}{dn} = -\alpha_c C \frac{\delta}{E_o}; \quad \alpha_{||} = \frac{\Delta E_{rad}}{E_o}; \quad (1-17)$$

This, longitudinal oscillations have twice the damping rate (it is because of the γ^2 dependence of the intensity of the radiation).

How to handle situation in a general case? We fist should expand our linear equations by adding radiation term to our Hamiltonian equations (simplified for $\gamma \gg 1$):

$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_3 - yP_1) + \quad ; \quad (5-46')$$

$$\frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x \delta + g_y y \delta + f_x x \tau + f_y y \tau$$

$$\begin{aligned}
\frac{F}{p_o} &= \left[-K \cdot \frac{e}{E_o} B_y - \frac{e}{E_o} \frac{\partial B_y}{\partial x} + \frac{1}{2} \left(\frac{eB_s}{E_o} \right)^2 \right] - \frac{e}{E_o} \frac{\partial E_x}{\partial x} - 2K \frac{eE_x}{E_o}; \\
\frac{G}{p_o} &= \left[\frac{e}{E_o} \frac{\partial B_x}{\partial y} + \frac{1}{2} \left(\frac{eB_s}{E_o} \right)^2 \right] - \frac{e}{E_o} \frac{\partial E_y}{\partial y} \quad ; \quad (5-47') \\
\frac{2N}{p_o} &= \left[\frac{e}{E_o} \frac{\partial B_x}{\partial x} - \frac{e}{E_o} \frac{\partial B_y}{\partial y} \right] + K \cdot \frac{e}{E_o} B_x - \frac{e}{E_o} \left(\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} \right) - 2K \frac{eE_y}{E_o} \\
\frac{L}{p_o} &= \kappa + \frac{e}{2p_o c} B_s; \quad \frac{U}{p_o} = \frac{e}{p_o c^2} \frac{\partial E_s}{\partial t}; g_x = -K \frac{c}{v_o}; g_y = 0; \\
f_x &= \frac{e}{c} \frac{\partial B_y}{\partial ct} + \frac{e}{v_o} \frac{\partial E_x}{\partial ct}; f_y = -\frac{e}{c} \frac{\partial B_x}{\partial ct} + \frac{e}{v_o} \frac{\partial E_y}{\partial ct}.
\end{aligned}$$

We should focus on additional linear terms. Let's expand the radiation power in (8-14):

$$\begin{aligned}
I &= I_o(1 + c_x \cdot x + c_y \cdot y + c_1 P_1 + c_3 P_3 + c_\tau \cdot \tau + 2\delta/E_o); \\
I_o &= \frac{2}{3} \frac{e^2 Z^2 E^2}{m^4 c^8} \left\{ \left(\vec{E} - [\vec{\beta} \times \vec{B}] \right)^2 - \left(\vec{\beta} \cdot \vec{E} \right)^2 \right\}_{eo} \quad (8-18)
\end{aligned}$$

with

$$\begin{aligned}
c_x &= K - \frac{2e}{KE_o} \frac{\partial}{\partial x} (E_x + B_y) + \frac{e^2 B_s^2}{KE_o^2}; c_y = -\frac{2e}{KE_o} \frac{\partial}{\partial y} (E_x + B_y) - \frac{e^2 B_s E_s}{KE_o^2}; \\
c_1 &= \frac{2eE_s}{KE_o}; c_3 = \frac{2eB_s}{KE_o}; c_\tau = \frac{2e}{KE_o} \frac{\partial}{\partial t} (E_x + B_y)
\end{aligned}$$

Second equation of (8-14) shall be modified using generalized momenta

$$\begin{aligned}
p_x &= p_o x' = P_1 - L p_o y; \quad p_y = p_o y' = P_3 + L p_o x; \\
\frac{dP_1}{ds} &= -\frac{\partial h}{\partial x} - I_o \cdot (P_1 - L p_o y); \quad \frac{dP_3}{ds} = -\frac{\partial h}{\partial y} - I_o \cdot (P_3 + L p_o x);
\end{aligned}$$

Overall, the equations of the motion become:

$$\frac{d}{ds} X = \mathbf{D} \cdot X; \quad \mathbf{D} = (\mathbf{SH} - I_o(s) \cdot \mathbf{G}); \quad \mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -L p_o & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ L p_o & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ c_x & c_1 & c_y & c_3 & c_\tau & 2 \end{bmatrix}. \quad (8-19)$$

First, the trace of the new D matrix is no longer zero, which means that determinants of the transport matrices are no longer unit:

$$\det \mathbf{M}((s_1|s_2)) = \exp \left\{ \int_{s_1}^{s_2} \text{Trace}(\mathbf{D}(s)) ds \right\} = \exp \left\{ 4 \int_{s_1}^{s_2} I_o(s) ds \right\} = \prod_1^6 \lambda_i = \prod_1^3 (\lambda_i^* \lambda_i) \quad (8-20)$$

The above formula is known as a theorem of sums of the decrements in storage rings: i.e. the sum of the decrements of all three eigen modes (oscillations: 2 betatron and one synchrotron) is equal to four times relative loss of energy into synchrotron radiation. This is the rate with which 6D phase space volume shrinks. Surprisingly, it is not very hard to make one of the mode (usually horizontal or synchrotron, which are strongly coupled) to experience anti-damping caused by synchrotron radiation, i.e. to have exponential growth.

Finally, how we can calculate decrements (increments?) of the eigen modes: we will use our parameterization

$$\begin{aligned} X &= \tilde{\mathbf{U}} \cdot A; \quad \tilde{\mathbf{U}}' = \mathbf{S}\mathbf{H}\tilde{\mathbf{U}}; \quad X' = (\mathbf{S}\mathbf{H} - I_o(s) \cdot \mathbf{G})\tilde{\mathbf{U}} \cdot A; \\ A' &= -I_o(s)\{\tilde{\mathbf{U}}^{-1}\mathbf{G}\tilde{\mathbf{U}}\} \cdot A \end{aligned} \quad (8-21)$$

Averaging figure bracket will give us diagonal matrix (not-diagonal terms will vanish because they do oscillate) with real and imaginary part (damping or anti-damping) of the correction to out tunes. Expressed in the terms of the eigen vectors, diagonal terms are:

$$g_i = -I_o(s) \cdot Y_i^{*T}(s) \mathbf{S}\mathbf{G}(s) Y_i(s) \quad (8-22)$$

For a typical plane uncoupled storage ring (L=0), vertical damping is completely decouples and is as we discussed above. Sum of the dimensionless decrements (i.e. divided by I_o) for horizontal betatron and synchrotron oscillations is equal 3.

We should not that for a fixed storage ring where B ~ E, the total loss on synchrotron radiation is proportional γ^4

One should now just use $Y_s = \begin{bmatrix} D \\ D' \\ i\alpha_r \\ 1 \end{bmatrix}$ to calculate what will be the decrement of the synchrotron

oscillations for a case with DC magnetic fields (i.e. no coupling with y, no time dependence, B_s=0, zero electric fields) and: **Do it as a homework problem. Average it over the circumference of the machine. Keep expression analytical. Also give the expression for decrement of the horizontal betatron oscillations using theorem about the sum of the decrements.**

Finally, let's discuss what stops beam oscillations to decay completely? It is quantum fluctuations of the radiation process, i.e. the fact that charged particles radiate energy by limps. It is definitely and directly excites synchrotron oscillations. Let's calculate first the noise in the energy for a simple case of the magnetic field:

$$\langle \Delta \delta^2 \rangle' = D_\delta = N'_{ph} \langle \Delta E_{ph} \rangle = \frac{55}{24\sqrt{3}} e^2 \hbar c K^3 \gamma^7 \quad (8-23)$$

From our oscillator studies, we now know that RMS energy spread will be

$$\sigma_\delta^2 = D_\delta / \alpha_s \quad (8-24)$$

What's happens with betatron oscillations:

One should now just use $\Delta X = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \Delta\delta \end{bmatrix}$; $\Delta(a_i e^{i\varphi_i}) = Y_i^{*T} S \Delta X = \Delta\delta \cdot Y_i^*(5)$; $\langle \Delta I \rangle = \frac{\langle \Delta\delta^2 \cdot |Y_i^*(5)| \rangle}{2}$ to calculate

what will be the decrement of the synchrotron oscillations. Thus, emittances of the beam will be

$$\varepsilon_i = \langle D_\delta |Y_i^*(5)|^2 \rangle / \alpha_i \quad (8-25)$$



Buy-buy linear stuff...

This lecture concluded the theme of using Hamiltonian and eigen vectors approach, similar to that in Kolm's book, for solving a number of standard accelerator problems.

Finally, Appendix J shows one of many possible ways to expand accelerator Hamiltonian analytically using Mathematica.

The other lectures of our course had been more ad-hoc and a subset of them and selected homeworks is collected in the separate PDF file.

Appendix A: 4-D metric of special relativity

“Tensors are mathematical objects - you'll appreciate their beauty by using them”

4-scalars, 4 vectors, 4- tensors. (closely follows [Landau])

An event is fully described by coordinates in 4D-space (time and 3D-space), i.e., by a 4 vector:

$$X^i = (x^0, x^1, x^2, x^3) \equiv (x^0, \vec{r}); x^0 = ct; x^1 = x; x^2 = y; x^3 = z. \quad (\text{A-1})$$

Consider a non-degenerated transformation in 4D space

$$X' = X'(X); \quad (\text{A-2})$$

$$x'^i = x'^i(x^0, x^1, x^2, x^3); i = 0, 1, 2, 3; \quad (\text{A-3})$$

and allowing the inverse transformation

$$X = X(X') \quad (\text{A-4})$$

$$x^i = x^i(x'^0, x'^1, x'^2, x'^3); i = 0, 1, 2, 3$$

Jacobian matrices describe the local deformations of the 4D space:

$$\frac{\partial x'^i}{\partial x^j}, \frac{\partial x^j}{\partial x'^i}; \quad (\text{A-5})$$

and are orthogonal to each other

$$\sum_{j=0}^{j=3} \frac{\partial x'^i}{\partial x^j} \cdot \frac{\partial x^j}{\partial x'^k} = \frac{\partial x'^i}{\partial x^j} \cdot \frac{\partial x^j}{\partial x'^k} = \frac{\partial x'^i}{\partial x'^k} = \delta_k^i; \quad (\text{A-6})$$

Here, we start with the convention to "silently" summate the repeated indexes:

$$a^i b_i \equiv \sum_{i=0}^{i=3} a^i b_i. \quad (\text{A-7})$$

A 4-scalar is defined as any scalar function that preserves its value while undergoing Lorentz transformation (including rotations in 3D space):

$$f(X') = f(X); \forall X' = L \otimes X \quad (\text{A-8})$$

Contravariant 4-vector $A^i = (A^0, A^1, A^2, A^3)$ is defined as an object for which the transformation rule is the same as for the 4D-space vector:

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j \quad (\text{A-9})$$

i.e.,

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j; \quad (\text{A-10})$$

or explicitly

$$A'^i = \frac{\partial x'^i}{\partial x^0} A^0 + \frac{\partial x'^i}{\partial x^1} A^1 + \frac{\partial x'^i}{\partial x^2} A^2 + \frac{\partial x'^i}{\partial x^3} A^3; \quad (\text{A-11})$$

Covariant 4-vector $A_i = (A_0, A_1, A_2, A_3)$ is defined as an object for which the transformation rule is

$$A'_i = \frac{\partial x^j}{\partial x'^i} A_j; \quad (\text{A-12})$$

i.e., the inverse transformation is used for covariant components.

Contravariant F^{jl} and Covariant G_{jl} 4-tensors of rank 2 are similarly defined :

$$F'^{ik} = \frac{\partial x'^i}{\partial x^j} \frac{\partial x'^k}{\partial x^l} F^{jl}; G'_{ik} = \frac{\partial x^j}{\partial x'^i} \frac{\partial x^l}{\partial x'^k} G_{jl}; \quad (\text{A-13})$$

Mixed tensors with co- and contra-variant indexes are transformed by mixed rules:

$$F'^i{}_k = \frac{\partial x'^i}{\partial x^j} \frac{\partial x^l}{\partial x'^k} F^j{}_l; G_i{}^k = \frac{\partial x^j}{\partial x'^i} \frac{\partial x'^k}{\partial x^l} G_j{}^l. \quad (\text{A-14})$$

Tensors of higher rank also are defined in this way. Thus, a tensor of rank n has 4^n components: 4-scalar - $n=0$, $4^0=1$ component; 4-vector - $n=1$, $4^1=4$ components; a tensor of rank 2 - $n=2$, $4^2=16$ components; and so on. Some components may be dependent ones. For example, symmetric- and asymmetric-tensors of rank 2 are defined as $S^{ik} = S^{ki}$; $A^{ik} = -A^{ki}$. A symmetric tensor has 10 independent components: four diagonal terms S^{ii} , and six $S^{i,k \neq i} = S^{k \neq i,i}$ non-diagonal terms. An asymmetric tensor has six independent components: $A^{i,k \neq i} = -A^{k \neq i,i}$, while all diagonal terms are zero $A^{ii} = -A^{ii} \equiv 0$. Any tensor of second rank can be expanded in symmetric- and asymmetric-parts:

$$F^{ik} = \frac{1}{2}(F^{ik} + F^{ki}) + \frac{1}{2}(F^{ik} - F^{ki}). \quad (\text{A-15})$$

The scalar product of two vectors is defined as the product of the co- and contra-variant vectors:

$$A \cdot B = A_i B^i; \quad (\text{A-16})$$

It is the invariant of transformations:

$$A'_i B'^i = \frac{\partial x^j}{\partial x'^i} \frac{\partial x'^i}{\partial x^k} A_j B^k = \frac{\partial x^j}{\partial x^k} A_j B^k = \delta^j_k A_j B^k = A_k B^k; \quad (\text{A-17})$$

where

$$\delta^j_k = \begin{cases} 1; j = k \\ 0; j \neq k \end{cases} \quad (\text{A-18})$$

is the unit tensor. Note that the trace of any tensor is a trivial 4-scalar .

$$\text{Trace}(F) = F^i{}_i \equiv F^0{}_0 + F^1{}_1 + F^2{}_2 + F^3{}_3 = F'^i{}_i; \quad (\text{A-19})$$

The metrics (or norm that must be a 4-scalar) defines the geometry of the 4-space. The traditional (geometric) way is to define it as $ds^2 = dx^i dx_i$. We calculated the 4-scalar defining interval between events.

An infinitesimal interval defines the norm of our "flat" space-time in special relativity:

$$ds^2 = dx^{0^2} - dx^{1^2} - dx^{2^2} - dx^{3^2} = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2; \quad (\text{A-20})$$

and the diagonal metric tensor g^{ik} :

$$ds^2 = g_{ik} dx^i dx^k = g^{ik} dx_i dx_k ;$$

$$g_{ik} = g^{ik} ; g^{00} = 1 ; g^{11} = -1 ; g^{22} = -1 ; g^{33} = -1 ;$$
(A-21)

in which all non-diagonal term are zero ; $g^{i \neq k} = 0$. The metric (A-21) is a consequence of the Euclidean space- frame. In general, it suffices that g^{ik} must be symmetric $g^{ik} = g^{ki}$. Note that the contraction of the metric tensor yield the unit tensor $g_{ij} g^{jk} = \delta_i^k$. Comparing (A-21) and (A-20) we conclude that

$$x^i = g^{ik} x_k ; x_i = g_{ik} x^k ;$$
(A-22)

i.e., the metric tensor g^{ik} raises indexes and g_{ik} lowers them, transforming the co- and contra-variant components

$$F^{...k...} = g^{kj} F_{...j...} = g^{kj} g_{il} F^{...l...} ; etc.$$
(A-23)

For 4-vectors, the lowering or rising indexes change the sign of spatial components. There is no distinction between co- and contra- variants; they can be switched without any consequences. Convention defines them as follows :

$$A^i = (A^0, \vec{A}) = (A^0, A^1, A^2, A^3)$$

$$A_i = (A_0, -\vec{A}) = (A_0, -A_1, -A_2, -A_3) ;$$

$$A \cdot B = A^i \cdot B_i = A^0 B^0 - \vec{A} \cdot \vec{B}$$
(A-24)

The $g^{kj}, g_{il}, g_i^k \equiv \delta_i^k$ tensors are special As they are identical in all inertial frames (coordinate systems). This is apparent for δ_i^k :

$$\delta_j^i = \frac{\partial x^k}{\partial x'^j} \frac{\partial x'^i}{\partial x^l} \delta_l^k = \frac{\partial x^k}{\partial x'^j} \frac{\partial x'^i}{\partial x^k} = \frac{\partial x'^i}{\partial x'^j} = \delta_j^i ;$$
(A-25)

while g^{ik} invariance is obvious from the invariance of the interval (A-20). Hence, it is better to say that the preservation of g^{ik} determines an allowable group of transformations in the 4D-space - the Lorentz group (see Appendix B). There is one more special tensor: the totally asymmetric 4-tensor of rank 4: e^{iklm} . Its components change sign when any if indexes are interchanged:

$$e^{iklm} = -e^{kilm} = -e^{ilk m} = -e^{ikml} .$$
(A-26)

meaning that the components with repeated indexes are zero: $e^{i \dots k \dots} = 0$; $i = k$; and only non-zero components are permutations of $\{0,1,2,3\}$.

By convention

$$e^{0123} = 1 ;$$
(A-27)

So that $e^{1023} = -1$. The tensor e^{iklm} also is invariant of Lorentz transformation that is directly related to the determinant of the Jacobian matrix of Lorentz transformations $J = \det \left[\frac{\partial x'}{\partial x} \right]$. As will be discussed in next lecture $J = 1$.

$$e^{'iklm} = \frac{\partial x'^i}{\partial x^j} \frac{\partial x'^k}{\partial x^n} \frac{\partial x'^l}{\partial x^p} \frac{\partial x'^m}{\partial x^q} e^{jnpq} = \det \left[\frac{\partial x'}{\partial x} \right] e^{jnpq} \delta_j^i \delta_n^k \delta_p^l \delta_q^m = e^{iklm}; \quad (\text{A-28})$$

In the best courses on linear algebra, the above equation is used as the definition of the matrix determinant. For details, see Section 3.4 (pp. 132-134) and section 4.1 in G. Arfken's "Mathematical Methods for Physicists" (where Eq. 4.2 is equivalent to $a_j^i a_n^k a_p^l a_q^m e^{jnpq} = \det[a] e^{jnpq} \delta_j^i \delta_n^k \delta_p^l \delta_q^m$). As mentioned in Landau (footnote in §6), the invariance of a totally asymmetric tensor of rank equal to the dimension of the space with respect to rotations is the general rule. This is very easy to prove for 2D space. The 2D totally asymmetric tensor of rank 2 is $e^{ik} = \begin{Bmatrix} 0 & 1 \\ -1 & 0 \end{Bmatrix}$ has transformations of

$$e^{'ik} = \frac{\partial x'^i}{\partial x^j} \frac{\partial x'^k}{\partial x^n} e^{jn} = \frac{\partial x'^i}{\partial x^1} \frac{\partial x'^k}{\partial x^2} e^{12} + \frac{\partial x'^i}{\partial x^2} \frac{\partial x'^k}{\partial x^1} e^{21} = \frac{\partial x'^i}{\partial x^1} \frac{\partial x'^k}{\partial x^2} - \frac{\partial x'^i}{\partial x^2} \frac{\partial x'^k}{\partial x^1} = \det \left\{ \begin{array}{cc} \frac{\partial x'^i}{\partial x^1} & \frac{\partial x'^i}{\partial x^2} \\ \frac{\partial x'^k}{\partial x^1} & \frac{\partial x'^k}{\partial x^2} \end{array} \right\}; \quad (\text{A-29})$$

Therefore:

$$e^{'ii} = \det \left\{ \begin{array}{cc} \frac{\partial x'^i}{\partial x^1} & \frac{\partial x'^i}{\partial x^2} \\ \frac{\partial x^1}{\partial x'^i} & \frac{\partial x^2}{\partial x'^i} \end{array} \right\} = 0 = e^{ii}; e^{'12} = \det \left\{ \begin{array}{cc} \frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} \\ \frac{\partial x^1}{\partial x'^2} & \frac{\partial x^2}{\partial x'^2} \end{array} \right\} = 1 = e^{12}; e^{'21} = \det \left\{ \begin{array}{cc} \frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} \\ \frac{\partial x^1}{\partial x'^1} & \frac{\partial x^2}{\partial x'^1} \end{array} \right\} = -1 = e^{21}; \quad (\text{A-30})$$

for rotations when $\det \left\{ \frac{\partial x'}{\partial x} \right\} = 1$. Finally, convolution of absolutely asymmetric tensor of rank n is equal n! - a number of permutations. In particular, $e^{iklm} e_{iklm} = 4! = 24$.

Tensors of any rank can be real tensors or pseudo-tensors, i.e., scalars and pseudo-scalars, vectors and pseudo-vectors, and so forth. They follow the same rules for rotations, but have different properties with respect to the sign inversions of coordinates: special transformations that cannot be reduced to rotations. An example of these transformations is the inversion of 3D coordinates signs.

The totally asymmetric tensor e^{iklm} is pseudo-tensor - it does not change sign when the space or time coordinates are inverted: $e^{0123} = 1$; (it is the same as for 3D version of it, $e^{\alpha\beta\gamma}$; $\vec{C} = \vec{A} \times \vec{B} \Rightarrow C^\alpha = e^{\alpha\beta\gamma} A^\beta B^\gamma$, $e^{123} = 1$);. Recall that the vector product in 3D space is a pseudo-vector. Under reflection $\vec{A} \rightarrow -\vec{A}$; $\vec{B} \rightarrow -\vec{B}$; $\vec{C} \Rightarrow \vec{C}$!

We can represent six components of an asymmetric tensor by two 3D-vectors;

$$(A^{ik}) = (\vec{p}, \vec{a}) = \begin{bmatrix} 0 & p_x & p_y & p_z \\ -p_x & 0 & -a_z & a_y \\ -p_y & a_z & 0 & -a_x \\ -p_z & -a_y & a_x & 0 \end{bmatrix}; (A_{ik}) = (-\vec{p}, \vec{a}). \quad (\text{A-31})$$

The time-space components of this tensor change sign under the reflection of coordinates, while purely spatial components do not. Hence, \vec{p} is a real (polar) 3 -D vector , and \vec{a} is 3D pseudo-vector (axial) vector.

$$A^{*ik} = e^{iklm} A_{lm} \quad (\text{A-32})$$

is called the dual tensor to asymmetric tensor A^{ik} , and vice versa. The convolution of dual tensors is pseudo-scalar $ps = A^{*ik} A_{ik}$. Similarly, $e^{iklm} A_m$ is a tensor of rank 3 dual to 4 vector A^i .

Differential operators

Next consider differential operators

$$\frac{\partial}{\partial x^{*i}} = \frac{\partial x^k}{\partial x^{*i}} \frac{\partial}{\partial x^k}; \quad (\text{A-32})$$

that follow the transformation rule for covariant vectors. Therefore, the differentiation with respect to a contravariant component is a covariant vector operator and vice versa! Accordingly, we can now express standard differential operators:

4-gradient:
$$\partial^i \equiv \frac{\partial}{\partial x_i} = \left(\frac{\partial}{\partial x_0}, -\vec{\nabla} \right); \partial_i = \frac{\partial}{\partial x^i} = \left(\frac{\partial}{\partial x_0}, \vec{\nabla} \right);$$
 (A-33)

4-divergence
$$\partial^i A_i = \partial_i A^i = \frac{\partial A^0}{\partial x^0} + \vec{\nabla} \vec{A};$$
 (A-34)

4-Laplacian (De-Lamberdian):
$$\square = \partial^i \partial_i = \frac{\partial^2}{\partial x^{0^2}} - \vec{\nabla}^2.$$
 (A-35)

Using differential operators allows us to construct 4-vectors and 4-tensors from 4-scalars. For example:

$$x^i = \partial^i (s^2). \quad (\text{A-36})$$

Other example is the phase of an oscillator: $\exp[i(\omega t - \vec{k} \vec{r})]$ $\varphi = \omega t - \vec{k} \vec{r}; \omega = kc$. The phase is 4-scalar; it does not depend on the system of observation. It is very important, but not an obvious fact! Imagine a sine wave propagating in space and a detector that registers when the wave intensity is zero. Zero value of wave amplitude is the event and does not depend on the system of observation. Similarly, we can detect any chosen phase. Therefore, the phase is 4-scalar and

$$k^i = \partial^i \varphi = (\omega / c, \vec{k}) \quad (\text{A-37})$$

is a 4-wave-vector undergoing standard transformation. Thus, we readily assessed the transformation of frequency and wave-vector from one system to the other, called the Doppler shift:

$$\omega = \gamma(\omega' + c\vec{\beta}\vec{k}'); \vec{k}_{\parallel} = \gamma(\vec{k}'_{\parallel} + \vec{\beta}\omega' / c); \vec{k}_{\perp} = \vec{k}'_{\perp}. \quad (\text{A-38})$$

then simply applying Lorentz transformations we found as last time:

$$\frac{\partial x^i}{\partial x^j} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \frac{\partial x^i}{\partial x'^j} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{A-39})$$

4-velocity, 4-acceleration

Another way to create new 4-vectors is to differentiate a vector as a function of the scalar function, for example, the interval. Unsurprisingly, 3D velocity transformation rules do not satisfy simple 4-D vector transformation rules; to differentiate over time that is not 4-scalar will be meaningless. 4-velocity is defined as derivative of the coordinate 4-vector x^i over the interval s :

$$u^i = \frac{dx^i}{ds}; \quad (\text{A-40})$$

and ,with simple way to connect it to 3D velocity $dx^i = (c, \vec{v})dt; ds = cdt\sqrt{1 - \frac{v^2}{c^2}} = cdt / \gamma$ we obtain :

$$u^i = \gamma(1, \vec{v} / c); \quad (\text{A-41})$$

that follows all rules of transformation. The first interesting result is that 4-velocity is dimensionless and has unit 4-length:

$$u^i u_i = 1 \quad (\text{A-42})$$

which is evident by taking into account that $ds^2 = dx^i dx_i \equiv u^i u_i ds^2$. Thus, it follows directly that 4-velocity and 4-acceleration

$$w^i = \frac{du^i}{ds} \quad (\text{A-43})$$

are orthogonal to each other:

$$u^i w_i = \frac{d(u^i u_i)}{2ds} = 0 . \quad (\text{A-44})$$

What is more amazing is that simply multiplying 4-velocity by the constant mc yields the 4-momentum:

$$mcu^i = (\gamma mc, \gamma m\vec{v}) = (E / c, \vec{p}) \quad (\text{A-45})$$

, furthermore, gives the simple rules to calculate energy and momentum of particles in arbitrary frame (beware of definition of \mathbf{g} here!):

$$E = \gamma(E' + c\vec{\beta}\vec{p}'); \vec{p}_{\parallel} = \gamma(\vec{p}'_{\parallel} + \vec{\beta}E' / c); \vec{p}_{\perp} = \vec{p}'_{\perp} . \quad (\text{A-46})$$

Integrals and their relations

Transformation rules are needed for elements of hyper-surfaces and for the generalization of Gauss and Stokes theorems. Those who studied have external differential forms in advances math courses will find it trivial, but for those who have not they may not be easy to follow. We

will use all necessary relations during the course when we need them. Here is a simple list:

1. The integral along the 4-D trajectory has an element of integration dx^i i.e., similar to $d\vec{r}$ for the 3D case.

2. An element of the 2D surface in 4D space is defined by two 4-vectors dx_k, dx'_k and an element of the surface is the 2-tensor $df_{ik} = dx_i dx'_k - dx'_i dx_k$. A dual tensor $df^{*ik} = \frac{1}{2} e^{iklm} df_{lm}$; is normal to the surface tensor: $df_{ik} df^{*ik} = 0$. It is similar to 3D case when the surface vector $df_\alpha = \frac{1}{2} e_{\alpha\beta\gamma} f_{\alpha\beta}$; $\alpha, \beta = 1, 2, 3$ is perpendicular to the surface.

3. An element of the 3D surface (hyper-surface or 3D manifold) in 4D space is defined by three 4-vectors dx_k, dx'_k, dx''_k and the three tensor element and dual vector of the 3D surface are

$$dS^{ikl} = \det \begin{bmatrix} dx^i & dx'^i & dx''^i \\ dx^k & dx'^k & dx''^k \\ dx^l & dx'^l & dx''^l \end{bmatrix} = e^{iklm} dS_n; dS^i = \frac{-1}{6} e^{iklm} dS_{klm}. \quad (\text{A-47})$$

Its time component is equal to the elementary 3D-volume $dS^0 = dx dy dz$.

4. The easiest case is that of a 4D-space volume created by four 4-vectors: $dx_i^{(1)}; dx_j^{(2)}; dx_k^{(3)}; dx_l^{(4)}$ which is a scalar

$$d\Omega = e^{iklm} dx_i^1 dx_j^2 dx_k^3 dx_l^4 \Rightarrow d\Omega = dx_0 dx_1 dx_2 dx_3 = c dt dV;$$

5. The rules for generalization of the Gauss and Stokes theorems (actually one general Stokes theorem, expressed in differential forms) are similar to those for 3D theorems, but there more of them:

$$\oint A^i dS_i = \int \frac{\partial A^i}{\partial x^i} d\Omega; \oint A^i dx_i = \int \frac{\partial A^i}{\partial x^k} df_{ik}; \int A^{ik} df_{*ik} = \int \frac{\partial A^{ik}}{\partial x^k} dS_i. \quad (\text{A-48})$$

Appendix B: Lorentz group

http://en.wikipedia.org/wiki/Lorentz_group

Lorentz Group - Matrix representation

Jackson's Classical Electrodynamics, Section 11.7 [Jackson] has an excellent discussion of this topic. Here, we will review it briefly with some attention to the underlying mathematics. Generic Lorentz transformation involves a boost (a transformation from K to K' moving with some velocity \vec{V}) and an arbitrary rotation in 3D space. Matrix representation is well suited to describe 4-vectors transformations. The coordinate vector is defined as

$$X = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}; \quad (\text{B-29})$$

and standard scalar product of 4-vectors is defined by $(a, b) = \tilde{a}b$, where \tilde{a} is the transposed vector. The 4-scalar product involves the metric tensor (matrix):

$$a \cdot b \equiv a^i \cdot b_i = (a, gb) = (ga, b) = \tilde{a}gb; \quad (\text{B-30})$$

$$g = \tilde{g} = \{g^{ik}\} = \{g_{ik}\} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (\text{B-31})$$

Lorentz transformations A (or the group of Lorentz transformations¹) are linear transformations that preserve the interval, or scalar product (B-30):

$$X' = AX; \quad \tilde{X}'gX' = \tilde{X}\tilde{A}gAX = \tilde{X}gX; \Rightarrow \tilde{A}gA = g. \quad (\text{B-33})$$

Using standard ratios for matrices

$$\det(\tilde{A}gA) = \det^2 A \det g = \det g \Rightarrow \det A = \pm 1; \quad (\text{B-34})$$

we find that the matrices of Lorentz transformation have $\det = \pm 1$. We will consider only proper Lorentz transformations with unit determinants $\det A = +1$. Improper Lorentz transformations, like space- and time-inversions, should be considered as special transformations and added to the proper ones.

A 4x4 matrix has 16 elements. Equation (B-33) limited number of independent elements

¹ Group G is defined as a set of elements, with a definition of a product of any two elements of the group; $P = A \cdot B \in G$; $A, B \in G$. The product must satisfy the associative law: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$; there is an unit element in the group $E \in G$; $E \cdot A = A \cdot E = A$; $\forall A \in G$; and inverse elements: $\forall A \in G; \exists B(\text{called } A^{-1}) \in G: A^{-1}A = AA^{-1} = E$.

Matrices NxN with non-zero determinants are examples of the group. Lorentz transformations are other examples: the product of two Lorentz is defined as two consequent Lorentz transformations. Therefore, the product also is a Lorentz transformation whose velocity is defined by rules discussed in previous lectures. The associative law is straightforward: unit Lorentz transformation is a transformation into the same system. Inverse Lorentz transformation is a transformation with reversed velocity. Add standard rotation s, to constitute the Lorentz Group

in matrix A of Lorentz transformations. Matrices on both sides are symmetric. Thus, there are 10 independent conditions on matrix A, leaving six independent elements there. This is unsurprising since rotation in 3D space is represented by 3 angles and a boost is represented by 3 components of velocity. Intuitively, then there are six independent rotations: (xy), (yz), (zx), (t, x), (t, y), and (t, z). No other combinations of 4D coordinates are possible: $C_4^2 = \frac{4!}{2!2!} = 6$.

We next consider the properties of A in standard way, representing A through a generator L:

$$A = e^L; \quad (\text{B-35})$$

where we use matrix exponent defined as the Taylor expansion:

$$e^L \downarrow_{def} \equiv \sum_{n=0}^{\infty} \frac{L^n}{n!}; L^0 = I; I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad (\text{B-36})$$

where I is the unit matrix. Using (B-35) and $g^2 = I$ we find how to compose the inverse matrix for A:

$$\tilde{A}gA = g \Rightarrow A^{-1} = g\tilde{A}g; \quad (\text{B-37})$$

which, in combination with

$$\tilde{A} = \text{transpose}(e^L) = \sum_{n=0}^{\infty} \frac{\tilde{L}^n}{n!} = e^{\tilde{L}}; e^{gUg} = \sum_{n=0}^{\infty} \frac{(gUg)^n}{n!} = \sum_{n=0}^{\infty} g \frac{U^n}{n!} g; \quad (\text{B-38})$$

gives

$$A^{-1} = g\tilde{A}g = e^{g\tilde{L}g}. \quad (\text{B-39})$$

We can that matrix exponent has similar properties as the regular exponent, i.e. $e^U e^{-U} = I$ by explicitly using Taylor expansion to collect the powers of U:

$$e^U e^{-U} = \left(\sum_{n=0}^{\infty} \frac{U^n}{n!} \right) \left(\sum_{k=0}^{\infty} (-1)^k \frac{U^k}{k!} \right) = \sum_{k=0, n=0}^{\infty} (-1)^k \frac{U^{n+k}}{n!k!} = I + \sum_{m=1}^{\infty} c_m U^m; \quad (\text{B-40})$$

and the well-known expansion of $(1-x)^m$. Our goal is to show that all c_m are zero:

$$(1-x)^m = \sum_{n=0}^m \frac{(-1)^n m!}{n!(m-n)!} x^n \Rightarrow m! c_m = \sum_{n=0}^m \frac{(-1)^n m!}{n!(m-n)!} = (1-1)^m = 0. \quad (\text{B-41})$$

Now (B-39) can be rewritten

$$A^{-1} = g\tilde{A}g = e^{g\tilde{L}g} = e^{-L} \Rightarrow g\tilde{L}g = -L; \Rightarrow \tilde{g}\tilde{L} = -gL \quad (\text{B-42})$$

Hence, gL is an asymmetric matrix and has six independent elements as expected:

$$gL = \begin{bmatrix} 0 & L_{01} & L_{02} & L_{03} \\ -L_{01} & 0 & -L_{12} & -L_{13} \\ -L_{02} & L_{12} & 0 & -L_{23} \\ -L_{03} & L_{13} & L_{23} & 0 \end{bmatrix}; L = g(gL) = \begin{bmatrix} 0 & L_{01} & L_{02} & L_{03} \\ L_{01} & 0 & L_{12} & L_{13} \\ L_{02} & -L_{12} & 0 & L_{23} \\ L_{03} & -L_{13} & -L_{23} & 0 \end{bmatrix}. \quad (\text{B-43})$$

Each independent element represents an irreducible (fundamental) element of the Lorentz group or rotations and boosts, as discussed above. The six components of the L can be considered as six components of 3-vectors in the form ("- " is a convention):

$$L = -\vec{\omega}\vec{S} - \vec{\zeta}\vec{K}; A = e^{-\vec{\omega}\vec{S} - \vec{\zeta}\vec{K}}; \quad (\text{B-44})$$

with

$$\vec{S} = \hat{e}_x \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} + \hat{e}_y \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} + \hat{e}_z \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad (\text{B-45})$$

$$\vec{K} = \hat{e}_x \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \hat{e}_y \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \hat{e}_z \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}; \quad (\text{B-46})$$

where $\vec{\omega}\vec{S}$ represents the orthogonal group of rotations in 3D space (O_3^+), and $\vec{\zeta}\vec{K}$ represents the boosts caused by transformation into a moving system. It is easy to check that these matrices satisfy commutation rules of

$$[S_i, S_k] = e_{ikl} S_l; [S_i, K_k] = e_{ikl} K_l; [K_i, K_k] = -e_{ikl} S_l; [A, B] \equiv AB - BA; \quad (\text{B-47})$$

where e_{ikl} is the totally asymmetric 3D-tensor. You should be familiar with 3D rotation $e^{-\vec{\omega}\vec{S}}$ by $\vec{\omega}$: the direction of $\vec{\omega}$ is the axis of rotation and the value of $\vec{\omega}$ is the angle of rotation.

For the arbitrary unit vector \hat{e}

$$(\hat{e}\vec{S})^3 = -\hat{e}\vec{S}; (\hat{e}\vec{K})^3 = \hat{e}\vec{K}. \quad (\text{B-48})$$

Therefore, \vec{S} "behaves" as an imaginary "i" and we should expect *sin* and *cos* to be generated by $\exp(..\vec{S}..)$; $\exp(..\vec{K}..)$ should generate hyperbolic functions *sinh* and *cosh*. It is left for your homework to show, in particular, that boost transformation is:

$$A(\vec{\beta} = \vec{V} / c) = e^{-\vec{\beta}\vec{K} \tanh^{-1} \beta}. \quad (\text{B-49})$$

Finally, all fully relativistic phenomena naturally have six independent parameters. For example, electromagnetic fields are described by two 3D vectors: the vector of the electric field and that of the magnetic field, or in equivalent form of an asymmetric 4-tensor of an electromagnetic field with six components. Not surprisingly, they carry the basic structure of the 4D space into something observable.

Appendix C: The Vlasov equation for an ensemble of particles.

CONSIDERING A LARGE NUMBER OF PARTICLES IN THE PHASE SPACE (Q,P) ALLOWS US TO INTRODUCE THE DISTRIBUTION FUNCTION:

$$f = f(X,t), \quad X = (Q,P): \quad dN_p = f(X,t)dX^{2n} \equiv f(X,t)dV_{2n}. \quad (C-1)$$

The Vlasov equation makes the following assumptions:

The local interaction of the particles in a small volume dV_{2n} is negligible compared with their interaction with the rest of the ensemble;

The system is Hamiltonian, i.e., dissipation is absent; and,

The consequence of (1) is that we can neglect scattering processes between the particles. This feature is important, otherwise we could not state that the number of particles in the phase-space volume remains constant.

A sub-ensemble of particles in the small volume dV_{2n} satisfies the conditions we used earlier for deriving Liouville's theorem. Let' us draw the boundary of the infinitesimal phase-space volume around the particles. **Because the phase space trajectories do not cross**, the particles cannot escape from the volume. The crossing of trajectories contradicts Hamiltonian mechanics, where particles having the same position and momentum at the same time do not have equal position and momenta in the following moment (scattering violates this condition). Therefore, phase density along the trajectory stays constant:

I. The number of the particles dN_p is constant;

II. The volume dV_{2n} is constant.

Thus,
$$f = f(X(t),t) = \frac{dN_p}{dV_{2n}} = const ,$$

(C-2)

when $X(t)$ is the trajectory satisfying the equation of motion. From here, we get the famous! Vlasov equation:

$$\frac{d}{dz} f(X(t),t) = 0 \rightarrow \frac{\partial f}{\partial t} + \frac{\partial f}{\partial X} \frac{dX}{dt} = 0. \quad (C-3)$$

where t is the independent variable, i.e., in our case it can be $t=s!$

Now we employ the Hamiltonian equations to obtain the matrix form:

$$\frac{\partial f}{\partial s} + \frac{\partial f}{\partial X} S \frac{\partial H}{\partial X} = 0 \quad (C-4)$$

or , the more traditional open form

$$\frac{\partial f}{\partial t} + \frac{\partial H}{\partial P_i} \frac{\partial f}{\partial Q_i} - \frac{\partial H}{\partial Q_i} \frac{\partial f}{\partial P_i} = 0. \quad (C-5)$$

The most widely spread form of this equation for 3-D is

$$\frac{\partial f}{\partial t} + \frac{\partial H}{\partial \vec{P}} \frac{\partial f}{\partial \vec{r}} - \frac{\partial H}{\partial \vec{r}} \frac{\partial f}{\partial \vec{P}} = 0. \quad (C-3D)$$

This equation is the most useful tool for studying collective effects in accelerator physics

Appendix D: Recap of Lecture 2

Here we recap what we already discussed in Lecture: the Accelerator Hamiltonian.

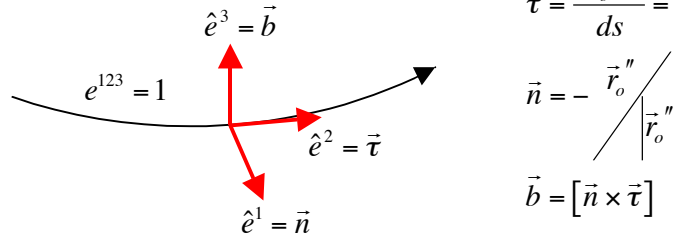
ā

Expressing the Reference particle and its trajectory $\vec{r}_o(t)$ as the function of s :

$$s(t) = \int_{t_i}^t |\dot{\vec{r}}_o(t)| dt = \int_{t_i}^t |\vec{v}_o(t)| dt \quad (\text{L2-1})$$

$$f' = \frac{df}{ds}; \quad f'' = \frac{d^2 f}{ds^2} \dots \dots \quad \dot{f} = \frac{df}{dt}; \quad \ddot{f} = \frac{d^2 f}{dt^2}. \quad (\text{L2-4})$$

The Frenet-Serret coordinate system or natural coordinate system: $\frac{d\vec{r}_o(s)}{ds} = \vec{\tau}$



Three contra-variant coordinates: $q^1 = x; q^2 = s, q^3 = y \quad \vec{r} = \vec{r}_o(s) + x \cdot \vec{n}(s) + y \cdot \vec{b}(s).$ (L2-5)

The Hamiltonian of a charged particle in an EM field

$$H = c \sqrt{\left((1 + Kx)^{-2} \left(\left(P_2 - \frac{e}{c} A_2 \right) + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) \right)^2 + \left(P_1 - \frac{e}{c} A_1 \right)^2 + \left(P_3 - \frac{e}{c} A_3 \right)^2 + m^2 c^2 \right)} + e\varphi \quad (\text{L2-15})$$

Now, we change the independent variable to $s, \{q^1, P_1\}, \{q^3, P_3\}, \{-t, H\}$ + least-action principle:

gives

$$x' = \frac{dx}{ds} = \frac{\partial h^*}{\partial P_1}; \quad \frac{dP_1}{ds} = -\frac{\partial h^*}{\partial x}; \quad y' = \frac{dy}{ds} = \frac{\partial h^*}{\partial P_3}; \quad \frac{dP_3}{ds} = -\frac{\partial h^*}{\partial y} \quad (\text{L2-18})$$

$$t' = \frac{dt}{ds} = \frac{\partial h^*}{\partial P_t} = -\frac{\partial h^*}{\partial H}; \quad \frac{dP_t}{ds} = -\frac{\partial h^*}{\partial t} \rightarrow \frac{dH}{ds} = \frac{\partial h^*}{\partial t}$$

$$h^* = -(1 + Kx) \sqrt{\frac{(H - e\varphi)^2}{c^2} - m^2 c^2 - \left(P_1 - \frac{e}{c} A_1 \right)^2 - \left(P_3 - \frac{e}{c} A_3 \right)^2} + \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) \quad (\text{L2-19})$$

Components of 4-vector potential can be arranged in a form of unique Taylor series:

$$\begin{aligned}
A_1 &= \frac{1}{2} \sum_{n,k=0}^{\infty} \partial_x^k \partial_y^n B_s|_{ro} \frac{x^k}{k!} \frac{y^{n+1}}{(n+1)!}; A_3 = -\frac{1}{2} \sum_{n,k=0}^{\infty} \partial_x^k \partial_y^n B_s|_{ro} \frac{x^{k+1}}{(k+1)!} \frac{y^n}{n!} \\
A_2 &= \sum_{n=1}^{\infty} \left\{ \partial_x^{n-1} \left((1+Kx)B_y + \kappa x B_s \right)_{ro} \frac{x^n}{n!} - \partial_y^{n-1} \left((1+Kx)B_x - \kappa y B_s \right)_{ro} \frac{y^n}{n!} \right\} + \\
&+ \frac{1}{2} \sum_{n,k=1}^{\infty} \left\{ \partial_x^{n-1} \partial_y^k \left((1+Kx)B_y + \kappa x B_s \right)_{ro} \frac{x^n}{n!} \frac{y^k}{k!} - \partial_x^n \partial_y^{k-1} \left((1+Kx)B_x - \kappa y B_s \right)_{ro} \frac{x^n}{n!} \frac{y^k}{k!} \right\}; \\
\varphi &= \varphi_o(s,t) - \sum_{n=1}^{\infty} \partial_x^{n-1} E_x|_{ro} \frac{x^n}{n!} - \sum_{n=1}^{\infty} \partial_y^{n-1} E_y|_{ro} \frac{y^n}{n!} - \frac{1}{2} \sum_{n,k=1}^{\infty} \left(\partial_x^{n-1} \partial_y^k E_x|_{ro} + \partial_x^n \partial_y^{k-1} E_y|_{ro} \right) \frac{x^n}{n!} \frac{y^k}{k!};
\end{aligned} \tag{L2-22}$$

where, $f|_{ro}$; $(f)_{ro}$ - f is taken $x=0$; $y=0$, but in arbitrary t . $f|_{ref}$; $(f)_{ref}$ $\vec{r} = \vec{r}_o(s)$, $t = t_o(s)$.

The equilibrium particle at the reference trajectory

$$\vec{r} = \vec{r}_o(s); \quad t = t_o(s); \quad H = H_o(s) = E_o(s) + \varphi_o(s, t_o(s)), \tag{L2-23}$$

$$K(s) \equiv \frac{1}{\rho(s)} = -\frac{e}{p_o c} \left(B_y|_{ref} + \frac{E_o}{p_o c} E_x|_{ref} \right) \cdot B_x|_{ref} = \frac{E_o}{p_o c} E_y|_{ref}, \tag{L2-26-27}$$

$$\frac{dt_o(s)}{ds} = -\frac{\partial h^*}{\partial H}|_{ref} = \frac{1}{v_o(s)}; \quad \frac{dE_o(s)}{ds} = -e \frac{\partial \varphi}{\partial s}|_{ref} \equiv e E_2(s, t_o(s)). \tag{L2-28-30}$$

Taking the calculation one step further

$$\left\{ \tau = -c(t - t_o(s)), \quad \delta = (H - E_o(s) - e\varphi_o(s, t))/c \right\}, \tag{L2-31}$$

A new Hamiltonian function using up-dated canonical pair (L2-33) is as follows:

$$\tilde{h} = -(1+Kx) \sqrt{p_o^2 + \frac{2E_o}{c} \left(\delta - \frac{e}{c} \varphi_{\perp} \right) + \left(\delta - \frac{e}{c} \varphi_{\perp} \right)^2 - \left(P_1 - \frac{e}{c} A_1 \right)^2 - \left(P_3 - \frac{e}{c} A_3 \right)^2} + \tag{L2-35}$$

$$\begin{aligned}
&+ \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3 \right) - \kappa y \left(P_1 - \frac{e}{c} A_1 \right) + \frac{c}{v_o} \delta - \frac{e}{c} \varphi_{||}(s, \tau) \\
\varphi_{\perp def} &= \varphi(s, x, y, t) - \varphi_o(s, t) \equiv \varphi(s, x, y, t) - \varphi(s, 0, 0, t)
\end{aligned} \tag{L2-36}$$

Expanding the Hamiltonian.:

$$\left| \frac{P_1}{p_o} \right| \ll 1; \quad \left| \frac{P_3}{p_o} \right| \ll 1; \quad \left| \frac{\delta}{p_o} \right| \ll 1;$$

The phase-space vector is

$$\tilde{h} = -p_o + \sum_{v=1}^N \sum_{\sum p_i=v} C_{v\{p_i\}} \prod_{i=1}^6 x_i^{p_i} + O(\alpha^{N+1}); \tag{L2-39}$$

$$X^T = [q^1 \quad P_1 \quad \dots \quad q^n \quad P_n] = [x_1 \quad x_2 \quad \dots \quad x_{2n-1} \quad x_{2n}], \tag{L2-46}$$

where T stands for ‘‘transposed’’. Adopting this notion, the Hamiltonian equations can be written as one:

$$\frac{dX}{ds} = S \cdot \frac{\partial H}{\partial X} \quad \Leftrightarrow \quad \frac{dx_i}{ds} = S_{ij} \cdot \frac{\partial H}{\partial x_j} \equiv \sum_{j=1}^{2n} S_{ij} \cdot \frac{\partial H}{\partial x_j} \tag{L2-42}$$

The first-order terms in the expansion are zeros by design.

Hamiltonian expansion:
$$\tilde{h} = \frac{P_1^2 + P_3^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_3 - yP_1) + \frac{\delta^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x \delta + g_y y \delta + f_x x \tau + f_y y \tau ; \quad (\text{L2-46})$$

The second-order term is of foremost importance in accelerator physics:

$$h = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 h_{ij} x_i x_j + O(\alpha^3) \equiv \frac{1}{2} X^T \cdot \mathbf{H} \cdot X + O(\alpha^3); \quad (\text{L2-45})$$

An important and self-evident feature is

$$\mathbf{H}^T = \mathbf{H}$$

$$\frac{dX}{ds} = \mathbf{D}(s) \cdot X + O(\alpha^2); \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{H}(s) \quad (\text{L2-45-1})$$

i.e. the equations of motion are just a set of s-dependent linear equations

A generator of the symplectic group: $S_{2m-1,2m} = 1 = -S_{2m,2m-1}$, $m = 1, \dots, n$; the other elements are zero. S has n diagonal blocks with a 2x2 matrix σ and the rest is the field of zeros:

$$\mathbf{S} = \begin{bmatrix} \sigma & 0 & \dots & 0 \\ 0 & \sigma & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \sigma \end{bmatrix}_{2n \times 2n} ; \quad \sigma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (\text{L2-43})$$

Important (self-evident?) features are:

$$\mathbf{S}^T = -\mathbf{S}; \quad \mathbf{S}^2 = -\mathbf{I}; \quad \mathbf{S}^T \mathbf{S} = \mathbf{I}; \quad \mathbf{S}^{-1} = -\mathbf{S} \quad (!)$$

Appendix E: Matrix functions and Projection operators

An arbitrary matrix \mathbf{M} can be reduced to an unique matrix, which in general case has a Jordan form: for a matrix with arbitrary height of eigen values the set of eigen values $\{\lambda_1, \dots, \lambda_m\}$ contains only unique eigen values, i.e. $\lambda_k \neq \lambda_j; \forall k \neq j$:

$$size[\mathbf{M}] = M; \{\lambda_1, \dots, \lambda_m\}; m \leq M; \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{U}\mathbf{G}\mathbf{U}^{-1}; \mathbf{G} = \sum_{\oplus \mathbf{k}=1, m} \mathbf{G}_k = \mathbf{G}_1 \oplus \dots \oplus \mathbf{G}_m; \sum size[\mathbf{G}_k] = M \quad (\text{E-1})$$

where \oplus means direct sum of block-diagonal square matrixes \mathbf{G}_k which correspond to the eigen vector sub-space adjacent to the eigen value λ_k . Size of \mathbf{G}_k , which we call l_k , is equal to the multiplicity of the root λ_k of the characteristic equation

$$\det[\lambda \mathbf{I} - \mathbf{M}] = \prod_{k=1, m} (\lambda - \lambda_k)^{l_k}.$$

In general case, \mathbf{G}_k is also a block diagonal matrix comprised of orthogonal sub-spaces belonging to the same eigen value

$$\mathbf{G}_k = \sum_{\oplus j=1, p_k} \mathbf{G}_k^j = \mathbf{G}_k^1 \oplus \dots \oplus \mathbf{G}_k^{p_k}; \sum size[\mathbf{G}_k^j] = l_k \quad (\text{E-2})$$

where we assume that we sorted the matrixes by increasing size: $size[\mathbf{G}_k^{j+1}] \geq size[\mathbf{G}_k^j]$, i.e. the

$$n_k = size[\mathbf{G}_k^{p_k}] \leq l_k \quad (\text{E-3})$$

is the maximum size of the Jordan matrix belonging to the eigen value λ_k . General form of the Jordan matrix is:

$$\mathbf{G}_k^n = \begin{bmatrix} \lambda_k & 1 & 0 & 0 \\ 0 & \lambda_k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix} \quad (\text{E-4})$$

This obviously includes non-degenerate case when matrix \mathbf{M} has M independent eigen values and all is just perfectly simple: matrix is reducible to a diagonal one

$$size[\mathbf{M}] = M; \{\lambda_1, \dots, \lambda_M\}; \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{U}\mathbf{G}\mathbf{U}^{-1}; \mathbf{G} = \begin{bmatrix} \lambda_1 & 0 & & \\ 0 & \dots & & \\ & & \dots & \\ & & & \lambda_M \end{bmatrix}; \mathbf{U} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_M]; \mathbf{M} \cdot \mathbf{Y}_k = \lambda_k \mathbf{Y}_k; k = 1, \dots, M \quad (\text{E-5})$$

An arbitrary analytical matrix function of \mathbf{M} can be expanded into Taylor series and reduced to the function of its Jordan matrix \mathbf{G} :

$$f(\mathbf{M}) = \sum_{i=1}^{\infty} f_i \mathbf{M}^i = \sum_{i=1}^{\infty} f_i (\mathbf{U}\mathbf{G}\mathbf{U}^{-1})^i = \left(\sum_{i=1}^{\infty} f_i \mathbf{U}(\mathbf{G})^i \mathbf{U}^{-1} \right) = \mathbf{U} \left(\sum_{i=1}^{\infty} f_i (\mathbf{G})^i \right) \mathbf{U}^{-1} = \mathbf{U} f(\mathbf{G}) \mathbf{U}^{-1} \quad (\text{E-6})$$

Before embracing complicated things, let's look at trivial case, when Jordan matrix is diagonal:

$$f(\mathbf{G}) = \sum_{i=1}^{\infty} f_i \mathbf{G}^i = \sum_{i=1}^{\infty} f_i \begin{bmatrix} \lambda_1 & 0 & & \\ & \dots & & \\ & & \dots & \\ & & & \lambda_M \end{bmatrix}^i = \begin{bmatrix} \sum_{i=1}^{\infty} f_i \lambda_1^i & 0 & & \\ & \dots & & \\ & & \dots & \\ & & & \sum_{i=1}^{\infty} f_i \lambda_M^i \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & 0 & & \\ & \dots & & \\ & & \dots & \\ & & & f(\lambda_M) \end{bmatrix} \quad (\text{E-7})$$

$$f(\mathbf{M}) = \mathbf{U} \begin{bmatrix} f(\lambda_1) & 0 & & \\ & \dots & & \\ & & \dots & \\ & & & f(\lambda_M) \end{bmatrix} \mathbf{U}^{-1}$$

The last expression can be rewritten as a sum of a product of matrix \mathbf{U} containing only specific eigen vector (other columns are zero!) with matrix \mathbf{U}^{-1} :

$$f(\mathbf{M}) = [\mathbf{Y}_1 \dots \mathbf{Y}_k \dots \mathbf{Y}_M] \cdot \begin{bmatrix} f(\lambda_1) & 0 & & \\ & \dots & & \\ & & \dots & \\ & & & f(\lambda_M) \end{bmatrix} \mathbf{U}^{-1} = \sum_{k=1}^M f(\lambda_k) [0 \dots \mathbf{Y}_k \dots 0] \mathbf{U}^{-1} \quad (\text{E-8})$$

Still both eigen vector and \mathbf{U}^{-1} in is very complicated (and generally unknown) functions of \mathbf{M} Hmmmm! We only need to find a matrix operator, which makes projection onto individual eigen vector. Because all eigen values are different, we have a very clever and simple way of designing projection operators. Operator

$$\mathbf{P}_k^i = \frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \quad (\text{E-9})$$

has two important properties: it is unit operator for \mathbf{Y}_i , it is zero operator for \mathbf{Y}_k and multiply the rest of them by a constant:

$$\begin{aligned} \mathbf{P}_k^i \mathbf{Y}_k &= \frac{\mathbf{M} \cdot \mathbf{Y}_k - \lambda_k \mathbf{I} \cdot \mathbf{Y}_k}{\lambda_i - \lambda_k} = \frac{\lambda_k - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_k \equiv \mathbf{0}; \\ \mathbf{P}_k^i \mathbf{Y}_i &= \frac{\mathbf{M} \cdot \mathbf{Y}_i - \lambda_k \mathbf{I} \cdot \mathbf{Y}_i}{\lambda_i - \lambda_k} = \frac{\lambda_i - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_i \equiv \mathbf{Y}_i; \\ \mathbf{P}_k^i \mathbf{Y}_j &= \frac{\mathbf{M} \cdot \mathbf{Y}_j - \lambda_k \mathbf{I} \cdot \mathbf{Y}_j}{\lambda_i - \lambda_k} = \frac{\lambda_j - \lambda_k}{\lambda_i - \lambda_k} \mathbf{Y}_j \end{aligned} \quad (\text{E-10})$$

I.e. it project \mathbf{U} into a subspace orthogonal to \mathbf{Y}_k . We should note the most important quality of this operator: it comprises of known matrixes: \mathbf{M} and unit one. Also, zero operators for two eigen vectors commute with each other – being combination of \mathbf{M} and \mathbf{I} makes it obvious. Constructing unit projection operator \mathbf{Y}_i which is also zero for remaining eigen vectors is straight forward from here: it is a product of all $M-1$ projection operators

$$\begin{aligned} \mathbf{P}_{unit}^i &= \prod_{k \neq i} \mathbf{P}_k^i = \prod_{k \neq i} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right); \\ \mathbf{P}_{unit}^i \mathbf{Y}_j &= \delta_j^i \mathbf{Y}_j = \begin{cases} \mathbf{Y}_i, & j = i \\ \mathbf{0}, & j \neq i \end{cases} \end{aligned} \quad (\text{E-11})$$

Observation that

$$\mathbf{P}_{unit}^k \mathbf{U} = \mathbf{P}_{unit}^k [\mathbf{Y}_1 \dots \mathbf{Y}_k \dots \mathbf{Y}_M] = [0 \dots \mathbf{Y}_k \dots 0] \quad (\text{E-12})$$

allows us to rewrite eq. (E-8) in the form which is easy to use:

$$f(\mathbf{M}) = \sum_{k=1}^M f(\lambda_k) [0 \dots \mathbf{Y}_k \dots 0] \mathbf{U}^{-1} = \sum_{k=1}^M f(\lambda_k) \mathbf{P}_{unit}^k \mathbf{U} \cdot \mathbf{U}^{-1} = \sum_{k=1}^M f(\lambda_k) \mathbf{P}_{unit}^k; \quad (\text{E-13})$$

which with (E-11) give final form of Sylvester formula (E-for non-degenerated matrixes):

$$f(\mathbf{M}) = \sum_{k=1}^M f(\lambda_k) \prod_{i \neq k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right); \quad (\text{E-14})$$

One can see that this is a polynomial of power M-1 of matrix M, as we expected from the theorem of Jordan and Kelly that matrix is a root of its characteristic equation:

$$g(\lambda) = \det[\mathbf{M} - \lambda \mathbf{I}]; \quad g(\mathbf{M}) \equiv 0; \quad (\text{E-15})$$

which is polynomial of power M. It means that any polynomial of higher order of matrix \mathbf{M} can be reduced to M-1 order. Equation (E-14) gives specific answer how it can be done for the arbitrary series.

If matrix \mathbf{M} is reducible to diagonal form, where some eigen values have multiplicity, we need to sum only by independent eigen values:

$$f(\mathbf{M}) = \sum_{k=1}^m f(\lambda_k) \prod_{\lambda_j \neq \lambda_k} \left(\frac{\mathbf{M} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \right); \quad (\text{E-14-red})$$

and it has maximum power of \mathbf{M} of m-1. Prove it trivial using the above.

Let's return to most general case of Jordan blocks, i.e. a degenerated case when eigen values have non-unit multiplicity. For a general form of the Jordan matrix we can only say that it is direct sum of the function of the Jordan blocks:

$$\begin{aligned} f(\mathbf{G}) &= \sum_{i=0}^{\infty} f_i \mathbf{G}^i = \sum_{i=0}^{\infty} f_i \begin{bmatrix} \mathbf{G}_1^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_m^{p_m} \end{bmatrix}^i = \begin{bmatrix} \sum_{i=0}^{\infty} f_i (\mathbf{G}_1^1)^i & \mathbf{0} \\ \mathbf{0} & \dots \\ & & \sum_{i=0}^{\infty} f_i (\mathbf{G}_m^{p_m})^i \end{bmatrix} \\ &= \begin{bmatrix} f(\mathbf{G}_1^1) & \mathbf{0} \\ \mathbf{0} & \dots \\ & & f(\mathbf{G}_m^{p_m}) \end{bmatrix} = \sum_{\oplus k=1, m, \quad j=1, p_k} f(\mathbf{G}_k^j) = f(\mathbf{G}_1^1) \oplus \dots \oplus f(\mathbf{G}_m^{p_m}); \quad (\text{E-16}) \end{aligned}$$

Function of a Jordan block of size n contains not only the function of corresponding eigen value λ , but also its derivatives to (n-1)th order:

$$\mathbf{G} = \begin{bmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & \lambda \end{bmatrix}; \quad f(\mathbf{G}) = \begin{bmatrix} f(\lambda) & f'(\lambda)/1! & \dots & f^{(k)}(\lambda)/k! & \dots & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & \dots & \dots & \dots & f^{(n-2)}(\lambda)/(n-2)! \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & f'(\lambda)/1! \\ 0 & 0 & \dots & \dots & \dots & f(\lambda) \end{bmatrix} \quad (\text{E-17})$$

The prove is attached in Appendix Eq. 17. We are half-way through.

There is sub-space of eigen vectors \mathcal{Y}_{κ}^n which corresponds to the eigen value λ_{κ} and the block \mathbf{G}_{κ}^n :

$$\mathcal{Y}_{\kappa}^n \in \{\mathbf{Y}_{\kappa}^{n,1}, \dots, \mathbf{Y}_{\kappa}^{n,q}\}; \quad q = \text{size}(\mathbf{G}_{\kappa}^n) \quad (\text{E-18})$$

$$\mathbf{M} \cdot \mathbf{Y}_{\kappa}^{n,1} = \lambda_{\kappa} \mathbf{Y}_{\kappa}^{n,1}; \quad \mathbf{M} \cdot \mathbf{Y}_{\kappa}^{n,l} = \lambda_{\kappa} \mathbf{Y}_{\kappa}^{n,l} + \mathbf{Y}_{\kappa}^{n,l-1}; \quad 1 < l \leq q \quad (\text{E-19})$$

It is obvious from equation (E-17) that projection operator (E-11) will not be zero operator for \mathcal{Y}_{κ}^n , and it also will not be unit operator for \mathcal{Y}_i^n . Now, let's look on how we can project on individual sub-spaces, eigen vectors, including zero-operator for specific sub-spaces. Just step by step (from eq. (E-6) and (E-17)):

$$f(\mathbf{M}) = \mathbf{U}f(\mathbf{G})\mathbf{U}^{-1}$$

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^m \sum_{i=1}^{n_k-1} \frac{f^{(i)}(\lambda_{\kappa})}{i!} \begin{bmatrix} \underset{\lambda_1}{0} & \underset{\lambda_2}{0} & \dots & \underset{\lambda_{k-1}}{0} & \underbrace{A_{\kappa}^i}_{\lambda_{\kappa}} & \dots 0 \dots & \underset{\lambda_m}{0} \end{bmatrix} \quad (\text{E-20})$$

$$A_{\kappa}^i = \begin{bmatrix} B_1^{i \ k} & \dots & B_{p_k}^{i \ k} \end{bmatrix}; \quad B_n^{i \ k} = \begin{bmatrix} \underbrace{0 \dots 0}_{i \text{ collumns}} & Y_{\kappa}^{n,1} & Y_{\kappa}^{n,q_n-1} \end{bmatrix} \quad (\text{E-21})$$

i.e.

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^m \sum_{i=1}^{n_k-1} \frac{f^{(i)}(\lambda_{\kappa})}{i!} \begin{bmatrix} \underset{\lambda_1}{0} & \underset{\lambda_2}{0} & \dots & \underset{\lambda_{k-1}}{0} & \underbrace{\begin{bmatrix} \dots & \dots & \underbrace{0 \dots 0}_{i \text{ collumns}} & Y_{\kappa}^{n,1} & Y_{\kappa}^{n,q_n-1} & \dots \\ \vdots & & & \underbrace{\hspace{2cm}}_{n\text{-th}} & & \vdots \\ & & & & & p_k \end{bmatrix}}_{\lambda_{\kappa}} & \dots 0 \dots & \underset{\lambda_m}{0} \end{bmatrix} \quad (\text{E-22})$$

From (E-19) we get:

$$\begin{aligned} [\mathbf{M} - \lambda_{\kappa} \mathbf{I}] \cdot \mathbf{Y}_{\kappa}^{n,q} &= 0; \quad [\mathbf{M} - \lambda_{\kappa} \mathbf{I}] \cdot \mathbf{Y}_{\kappa}^{n,k} = \mathbf{Y}_{\kappa}^{n,k-1}; \quad 1 < k \leq q \\ U_1^{n \ k} &= [\mathbf{Y}_{\kappa}^{n,1} \ \dots \ \mathbf{Y}_{\kappa}^{n,l} \ \dots \ \mathbf{Y}_{\kappa}^{n,q}]; \\ [\mathbf{M} - \lambda_{\kappa} \mathbf{I}] \cdot U_1^{n \ k} &= U_2^{n \ k} = [0, \mathbf{Y}_{\kappa}^{n,1} \ \dots \ \mathbf{Y}_{\kappa}^{n,l} \ \dots \ \mathbf{Y}_{\kappa}^{n,q-1}] \\ \dots & \\ [\mathbf{M} - \lambda_{\kappa} \mathbf{I}]^j \cdot U_1^{n \ k} &= U_j^{n \ k} = \begin{bmatrix} \underbrace{0 \dots 0}_j, \mathbf{Y}_{\kappa}^{n,1} \ \dots \ \mathbf{Y}_{\kappa}^{n,l} \ \dots \ \mathbf{Y}_{\kappa}^{n,q-j} \end{bmatrix} \\ \dots & \\ [\mathbf{M} - \lambda_{\kappa} \mathbf{I}]^q \cdot U_1^{n \ k} &= 0 \end{aligned} \quad (\text{E-23})$$

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^m \sum_{i=1}^{n_k-1} \frac{f^{(i)}(\lambda_{\kappa})}{i!} [\mathbf{M} - \lambda_{\kappa} \mathbf{I}]^i \begin{bmatrix} \underset{\lambda_1}{0} & \underset{\lambda_2}{0} & \dots & \underset{\lambda_{k-1}}{0} & \underbrace{\begin{bmatrix} U^{k \ 1} & \dots & U^{k \ n} & \dots & U^{k \ p_k} \\ & & \underbrace{\hspace{2cm}}_{n\text{-th}} & & \end{bmatrix}}_{\lambda_{\kappa}} & \dots 0 \dots & \underset{\lambda_m}{0} \end{bmatrix} \quad (\text{E-24})$$

i.e. we collected all eigen vectors belonging to the eigen value λ_k . Now we need a projection non-distorting operator on the sub-space of λ_k . First, let's find zero operator for subspace of λ_i :

$$O_i = [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} \Rightarrow [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} U_1^r = [\mathbf{M} - \lambda_i \mathbf{I}]^{n_i} [\mathbf{Y}_k^{r,1} \dots \mathbf{Y}_k^{r,l} \dots \mathbf{Y}_k^{r,q}] = 0;$$

$$T_k = \prod_{i \neq k} \frac{O_i}{(\lambda_k - \lambda_i)^{n_i}} = \prod_{i \neq k} \left(\frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right)^{n_i} \quad (\text{E-25})$$

T_k is projection operator of sub-space of λ_k , but it is not unit one! To correct that we need an operator which we create as follows:

$$R = \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i}; \quad T = \mathbf{M} - \lambda_k \mathbf{I}; \quad \alpha = \alpha_{k,i} = 1/(\lambda_k - \lambda_i)$$

$$RU_1 = U_1 + \alpha U_2 \quad U_1 = U_1$$

.....

$$RU_{q-1} = U_{q-1} + \alpha U_q \quad U_{q-1} = T^{q-2} U_1$$

$$RU_q = U_q \quad U_q = T^{q-1} U_1$$

$$Q = \alpha T$$

$$U_q = RU_q = RT^{q-1} U_1$$

$$U_{q-1} = R(I + Q)U_{q-1} = RQT^{q-2} U_1$$

$$U_{q-1} = RQU_{q-1} = RQT^{q-2} U_1$$

.....

$$U_1 = R \left(\sum_j^{q-1} Q^j \right) U_1$$

so, we get it:

$$P_k^i = \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k-1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \quad (\text{E-26})$$

The final stroke is:

$$P_k = \prod_{i \neq k} (P_k^i)^{n_i} = \prod_{i \neq k} \left\{ \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k-1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \right\}^{n_i} \quad (\text{E-27})$$

and

$$F(M) = \sum_{k=1}^m \left[\prod_{i \neq k} \left\{ \frac{\mathbf{M} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \left(\mathbf{I} + \sum_{j=1}^{n_k-1} \left(\frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right) \right\} \sum_{i=1}^{n_k-1} \frac{f^{(i)}(\lambda_k)}{i!} [\mathbf{M} - \lambda_k \mathbf{I}]^i \right] \quad (\text{E-28})$$

Proof of eq. (E-17):

$$\mathbf{G}^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & & 1 \end{bmatrix}; \quad \mathbf{G}^1 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^2 & \lambda & 1 \dots & 0 \\ 0 & \lambda^2 & \lambda \dots & 0 \\ 0 & 0 & \dots & \lambda \\ 0 & 0 & 0 & \lambda^2 \end{bmatrix}$$

Induction:

$$\mathbf{G}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-2}/2! & \dots & \dots \\ 0 & \lambda^n & n\lambda^{n-1}/1! & \dots & \dots \\ 0 & 0 & \dots & n\lambda^{n-1}/1! & \dots \\ 0 & 0 & 0 & \lambda^n & \dots \end{bmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-2}/2! & \dots & \dots \\ 0 & \lambda^n & n\lambda^{n-1}/1! & \dots & \dots \\ 0 & 0 & \dots & n\lambda^{n-1}/1! & \dots \\ 0 & 0 & 0 & \lambda^n & \dots \end{bmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (n+1)\lambda^n/1! & (n(n-1)+2n)\lambda^{n-1}/2! & \dots & \dots \\ 0 & \lambda^{n+1} & (n+1)\lambda^n/1! & \dots & \dots \\ 0 & 0 & \dots & (n+1)\lambda^n/1! & \dots \\ 0 & 0 & 0 & \lambda^{n+1} & \dots \end{bmatrix}$$

$$\begin{bmatrix} \lambda^n & C_1^n \lambda^{n-1} & C_2^n \lambda^{n-2} & \dots & C_k^n \lambda^{n-k} & C_{k+1}^n \lambda^{n-k-1} & \dots \\ 0 & \lambda^n & C_1^n \lambda^{n-1} & \dots & C_{k-1}^n \lambda^{n+1-k} & C_k^n \lambda^{n-k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{bmatrix}$$

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda^n & C_1^n \lambda^{n-1} & C_2^n \lambda^{n-2} & \dots & C_k^n \lambda^{n-k} & C_{k+1}^n \lambda^{n-k-1} & \dots \\ 0 & \lambda^n & C_1^n \lambda^{n-1} & \dots & C_{k-1}^n \lambda^{n+1-k} & C_k^n \lambda^{n-k} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{bmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (C_1^n + 1)\lambda^n & (C_2^n + C_1^n)\lambda^{n-2} & \dots & (C_k^n + C_{k-1}^n)\lambda^{n-k+1} & (C_{k+1}^n + C_k^n)\lambda^{n-k} & \dots \\ 0 & \lambda^{n+1} & (C_1^n + 1)\lambda^n & \dots & \dots & (C_k^n + C_{k-1}^n)\lambda^{n-k+1} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^{n+1} \end{bmatrix}$$

polynomial coefficients: $C_k^{n+1} = C_k^n + C_{k-1}^n$; $C_k^n = n!/k!(n-k)!$ proves the point.

Hence, we can now calculate a polynomial functions or any function expandable into a Taylor series:

$$f(\mathbf{G}) = \sum_{n=0}^{\infty} f_n \mathbf{G}^n = \sum_{n=0}^{\infty} f_n \begin{bmatrix} \lambda^n & C_1^n \lambda^{n-1} & \dots & C_k^n \lambda^{n-k} & \dots \\ 0 & & & & \dots \\ 0 & & & & \lambda^n \end{bmatrix}^i = \begin{bmatrix} \sum_{n=0}^{\infty} f_n \lambda^n & \dots & \sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k} & \dots \\ 0 & & & \dots \\ 0 & & 0 & \dots \\ & & & \sum_{n=0}^{\infty} f_n \lambda^n \end{bmatrix}$$

The final stroke is noting that

$$\begin{aligned} \sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k} &= \sum_{n=0}^{\infty} f_n \cdot \frac{n! \lambda^{n-k}}{k! \cdot (n-k)!} = \frac{1}{k!} \sum_{n=0}^{\infty} f_n \cdot \frac{n! \lambda^{n-k}}{(n-k)!} = \frac{1}{k!} \sum_{n=0}^{\infty} f_n \cdot \lambda^{n-k} \prod_{j=0}^{k-1} (n-j) \\ &= \frac{1}{k!} \frac{d^k}{d\lambda^k} \sum_{n=0}^{\infty} f_n \cdot \lambda^n = \frac{1}{k!} \frac{d^k f}{d\lambda^k} \quad \# \end{aligned}$$

Good for HW exercise.

Appendix F: Inhomogeneous solution

Even though calculations are tedious, they are also transparent and straightforward. General expression for the inhomogeneous equation of $2n$ ordinary linear differential equations is found by a standard trick of variable constants (method developed by Lagrange), i.e. assuming that $R = \mathbf{M}(s)A(s)$:

$$\begin{aligned} \frac{dR}{ds} &= R' = \mathbf{D} \cdot R + \mathbf{C}; \quad \mathbf{M}' = \mathbf{D}\mathbf{M}; \\ R &= \mathbf{M}(s)A(s) \Rightarrow \mathbf{M}'A + \mathbf{M}A' = \mathbf{D}\mathbf{M}A + \mathbf{C} \\ R(0) &= 0 \Rightarrow A_0 = 0 \\ A' &= \mathbf{M}^{-1}(s)\mathbf{C} \Rightarrow A = \int_0^s \mathbf{M}^{-1}(z)\mathbf{C}dz = \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot \mathbf{C} \end{aligned} \tag{F-1}$$

with well known result of:

$$R = e^{\mathbf{D}s} \left(\int_0^s e^{-\mathbf{D}z} dz \right) \cdot \mathbf{C}. \tag{F-2}$$

or

$$R = M_{4 \times 4}(s) \left(\int_0^s M_{4 \times 4}^{-1}(z) dz \right) \cdot \mathbf{C}. \tag{F-3}$$

If you use computer, eq. (F-3) is one to use. For analytical folks, you should go through a tedious job is combining all terms together into final form:

$$R(s) = \sum_{k=1}^m \left\{ \prod_{i \neq k} \left[\frac{\mathbf{D} - \lambda_i \mathbf{I}}{\lambda_k - \lambda_i} \right] \sum_{j=0}^{n_k-1} \left(\frac{\mathbf{D} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \right)^j \right\} \sum_{n=0}^{n_k-1} (\mathbf{D} - \lambda_k \mathbf{I})^n \frac{s^n}{n!} \cdot \sum_{p=0}^{n_k-1} (-1)^{p+1} (\mathbf{D} - \lambda_k \mathbf{I})^p \cdot \mathbf{C} \cdot \left[\sum_{q=0}^{p1} \frac{s^{p-q}}{(p-q)! \lambda_k^{q+1}} - \frac{e^{\lambda_k}}{\lambda_k^{p+1}} \right] \tag{F-4}$$

Appendix G. Canonical variables – one more touch.

It was brought to my attention in the class that there is a possibility to use better notations for components of the canonical momenta instead of rather awkward P_1 and P_3 . This is result of my exploration of such possibility. I plan to modify my notes appropriately when time permits...

In lecture 2 (Jan 15) we defined our new coordinates in the natural coordinate system defined by the reference orbit $\vec{r}_o(t)$:

$$\vec{r} = \vec{r}_o(s) + x \cdot \vec{n}(s) + y \cdot \vec{b}(s). \quad (\text{L2-5})$$

$$q^1 = x; \quad q^2 = s, \quad q^3 = y. \quad (\text{L2-5-1})$$

together with the co-variant (R_k) and contra-variant (R^k) components of an arbitrary vector:

$$\begin{aligned} \vec{R} &\equiv R_x \vec{n} + R_s \vec{t} + R_y \vec{b} \equiv \sum_k R^k \vec{a}_k \equiv \sum_k R_k \vec{a}^k \\ R_k &= \vec{R} \cdot \vec{a}_k; \quad R_1 = R_x; R_2 = (1 + Kx)R_s + \kappa(R_x y - R_y x); \quad R_3 = R_y; \end{aligned} \quad (\text{L2-11})$$

$$R^k = \vec{R} \cdot \vec{a}^k; \quad R^1 = R_x - \frac{\kappa y}{1 + Kx} R_s; \quad R^2 = \frac{R_s}{1 + Kx} + \kappa(R_x y - R_y x); \quad R^3 = R_y + \frac{\kappa x}{1 + Kx} R_s;$$

We finished with three canonical pairs

$$\{q^1, P_1\}, \{q^3, P_3\}, \{-t, H\}$$

to emphasize co- and contra-variant nature of the pairs, and finished with the Hamiltonian:

$$\begin{aligned} h^* &= -(1 + Kx) \sqrt{\frac{(H - e\varphi)^2}{c^2} - m^2 c^2 - \left(P_1 - \frac{e}{c} A_1\right)^2 - \left(P_3 - \frac{e}{c} A_3\right)^2} \\ &\quad + \frac{e}{c} A_2 + \kappa x \left(P_3 - \frac{e}{c} A_3\right) - \kappa y \left(P_1 - \frac{e}{c} A_1\right) \end{aligned} \quad (\text{L2-19})$$

We quite naturally switch to use of x and y , but continued the use of “the odd” P_1 and P_3 , which are actually can be (see line 2 in L2-11) called P_x and P_y without loss of any generality:

$$P_1 = P_x; \quad P_3 = P_y; \quad (\text{G-1})$$

as soon and we keep them capital and recognize that the capital (upper case) \mathbf{P} is the notation reserved for the canonical momentum, while low case \mathbf{p} is the notation reserved for mechanical momentum:

$$P_x \equiv p_x + \frac{e}{c} A_x; \quad P_y = p_y + \frac{e}{c} A_y; \quad (\text{G-2})$$

with canonical pairs of:

$$\{x, P_x\}, \{y, P_y\}, \{-t, H\} \quad (\text{G-3})$$

and the Hamiltonian

$$h^* = -(1+Kx)\sqrt{\frac{(H-e\varphi)^2}{c^2} - m^2c^2 - \left(P_x - \frac{e}{c}A_x\right)^2 - \left(P_y - \frac{e}{c}A_y\right)^2} + \frac{e}{c}A_2 + \kappa x\left(P_y - \frac{e}{c}A_y\right) - \kappa y\left(P_x - \frac{e}{c}A_x\right) \quad (\text{G-4})$$

There is the danger in these notations, because they look too familiar from Cartesian system, while our coordinate system is actually curvilinear. Hence, one should use proper differential operators and remember that in general case $A_2 \neq A_s$; Next natural step would be switch to $\{-ct, P_o = H/c\}$ to make the dimensionality of all three pair the same {distance, momentum}. The danger here is to mix \mathbf{P}_o and \mathbf{p}_o – we did not used it. Another option is to cal it \mathbf{P}_{ct} , which is awkward. The random choice of δ in new canonical pair

$$\{\tau = -c(t - t_o(s)), \delta P_o = (H - E_o(s) - e\varphi_o(s,t))/c\}, \quad (\text{L2-31})$$

was unfortunate. Calling it

$$\{\tau = -c(t - t_o(s)), P_\tau = (H - E_o(s) - e\varphi_o(s,t))/c\}, \quad (\text{G-5})$$

would be a better choice with the Hamiltonian looking much better:

$$\tilde{h} = -(1+Kx)\sqrt{p_o^2 + \frac{2E_o}{c}\left(P_\tau - \frac{e}{c}\varphi_\perp\right) + \left(P_\tau - \frac{e}{c}\varphi_\perp\right)^2 - \left(P_x - \frac{e}{c}A_x\right)^2 - \left(P_y - \frac{e}{c}A_y\right)^2} + \frac{e}{c}A_2 + \kappa x\left(P_y - \frac{e}{c}A_y\right) - \kappa y\left(P_x - \frac{e}{c}A_x\right) + \frac{c}{v_o}P_\tau - \frac{e}{c}\varphi_\parallel(s,\tau) \quad (\text{G-6})$$

Furthermore, in storage ring literature use of momentum deviation $p_s - p_o$ is preferred to the use of the energy deviation. In this case (fro constant energy only!) the longitudinal variable is also very similar to the longitudinal displacement of the particle from the reference.

$$l = \frac{v_o(s)}{c} \tau, \quad (\text{G-7})$$

Non-identical part of the canonical transformation is easy ¹⁵:

$$\Omega(P_\tau, l) = -P_\tau \cdot l \cdot \frac{c}{v_o(s)} \quad (\text{G-8})$$

with

¹⁵ $d\Omega = -q_i dP_i - \tilde{P}_i d\tilde{q}_i + (H' - H)dt$; $q_i = -\frac{\partial\Omega}{\partial P_i}$; $\tilde{P}_i = -\frac{\partial\Omega}{\partial \tilde{q}_i}$; $H' = H + \frac{\partial\Omega}{\partial t}$;

$$\begin{aligned}
\tau &\equiv -\frac{\partial}{\partial P_\tau} \Omega(P_\tau, l) = l \cdot \frac{c}{v_o(s)}; \\
P_l &= -\frac{\partial}{\partial l} \Omega(P_\tau, l) = P_\tau \cdot \frac{c}{v_o(s)} = (H - E_o(s) - e\varphi_o(s, t)) / v_o(s) \\
\hat{h} &= \tilde{h} + \frac{\partial}{\partial s} \Omega(P_\tau, l) = P_\tau l \cdot \frac{c}{v_o^2} \frac{dv_o}{ds} = l \cdot P_l \left(\frac{1}{v_o} \frac{dv_o}{ds} \right)
\end{aligned} \tag{G-9}$$

With exception of an additional artificial term, the Hamiltonian becomes may-be a bit more appealing:

$$\begin{aligned}
\hat{h} &= -(1 + Kx) \sqrt{p_o^2 + \frac{2p_o}{c} \left(P_l - \frac{e}{v_o} \varphi_\perp \right) + \left(\frac{v_o}{c} \right)^2 \left(P_l - \frac{e}{v_o} \varphi_\perp \right)^2 - \left(P_x - \frac{e}{c} A_x \right)^2 - \left(P_y - \frac{e}{c} A_y \right)^2} + \\
&+ \frac{e}{c} A_2 + \kappa x \left(P_y - \frac{e}{c} A_y \right) - \kappa y \left(P_x - \frac{e}{c} A_x \right) + P_l - \frac{e}{c} \hat{\varphi}_\parallel(s, l) + l \cdot P_l \left(\frac{1}{v_o} \frac{dv_o}{ds} \right)
\end{aligned} \tag{G-10}$$

$$\hat{\varphi}_\parallel(s, l) = \int_0^{-l/v_o(s)} \left(E_2(s, t_o(s) + \xi) - E_2|_{ref} \right) d\xi \tag{G-11}$$

There is no much difference in appearance between (G-10) and (G-6). The difference appears as soon as we are considering particles with constant energy and can use reduces variables:

With canonical pairs of:

$$\{x, \pi_x\}, \{y, \pi_y\}, \{\tau, \pi_\tau\}; \quad \pi_\alpha \equiv P_\alpha / p_o \tag{G-12}$$

$$\begin{aligned}
\tilde{h} &= -(1 + Kx) \sqrt{1 + \frac{2c}{v_o} \left(\pi_\tau - \frac{e}{p_o c} \varphi_\perp \right) + \left(\pi_\tau - \frac{e}{p_o c} \varphi_\perp \right)^2 - \left(\pi_x - \frac{e}{p_o c} A_x \right)^2 - \left(\pi_y - \frac{e}{p_o c} A_y \right)^2} + \\
&+ \frac{e}{p_o c} A_2 + \kappa x \left(\pi_y - \frac{e}{p_o c} A_y \right) - \kappa y \left(\pi_x - \frac{e}{p_o c} A_x \right) + \frac{c}{v_o} \pi_\tau - \frac{e}{p_o c} \varphi_\parallel(s, \tau)
\end{aligned} \tag{G-13}$$

With canonical pairs of:

$$\{x, \pi_x\}, \{y, \pi_y\}, \{l, \pi_l\}; \quad \pi_\alpha \equiv P_\alpha / p_o \tag{G-14}$$

$$\begin{aligned}
\hat{h} &= -(1 + Kx) \sqrt{1 + 2 \left(\pi_l - \frac{e}{E_o} \varphi_\perp \right) + \left(\frac{v_o}{c} \right)^2 \left(\pi_l - \frac{e}{E_o} \varphi_\perp \right)^2 - \left(\pi_x - \frac{e}{p_o c} A_x \right)^2 - \left(\pi_y - \frac{e}{p_o c} A_y \right)^2} + \\
&+ \frac{e}{p_o c} A_2 + \kappa x \left(\pi_y - \frac{e}{p_o c} A_y \right) - \kappa y \left(\pi_x - \frac{e}{p_o c} A_x \right) + \pi_l - \frac{e}{p_o c} \hat{\varphi}_\parallel(s, l)
\end{aligned} \tag{G-15}$$

Linearized stuff is similar. With canonical pairs of: $\{x, P_x\}, \{y, P_y\}, \{\tau, P_\tau\}; \quad \pi_\alpha \equiv P_\alpha / p_o$

$$\begin{aligned}\tilde{h} &= \frac{P_x^2 + P_y^2}{2p_o} + F \frac{x^2}{2} + Nxy + G \frac{y^2}{2} + L(xP_y - yP_x) + \\ &\quad \frac{P_\tau^2}{2p_o} \cdot \frac{m^2 c^2}{p_o^2} + U \frac{\tau^2}{2} + g_x x P_\tau + g_y y P_\tau + F_x x \tau + F_y y \tau\end{aligned}\quad ; \quad (-6-16)$$

or reduced ($p_o = \text{const!}$)

$$\begin{aligned}\tilde{h} &= \frac{\pi_x^2 + \pi_y^2}{2p_o} + f \frac{x^2}{2} + nxy + g \frac{y^2}{2} + L(x\pi_y - y\pi_x) + \\ &\quad \frac{\pi_\tau^2}{2} \cdot \frac{m^2 c^2}{p_o^2} + u \frac{\tau^2}{2} + g_x x \pi_\tau + g_y y \pi_\tau + f_x x \tau + f_y y \tau\end{aligned}\quad ; \quad (-6-16)$$

with

$$\begin{aligned}f &= \frac{F}{p_o} = \left[-K \cdot \frac{e}{p_o c} B_y - \frac{e}{p_o c} \frac{\partial B_y}{\partial x} + \left(\frac{e B_s}{2p_o c} \right)^2 \right] - \frac{e}{p_o v_o} \frac{\partial E_x}{\partial x} - 2K \frac{e E_x}{p_o v_o} + \left(\frac{m e E_x}{p_o^2} \right)^2; \\ g &= \frac{G}{p_o} = \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial y} + \left(\frac{e B_s}{2p_o c} \right)^2 \right] - \frac{e}{p_o v_o} \frac{\partial E_y}{\partial y} + \left(\frac{m e E_y}{p_o^2} \right)^2; \\ 2n &= \frac{2N}{p_o} = \left[\frac{e}{p_o c} \frac{\partial B_x}{\partial x} - \frac{e}{p_o c} \frac{\partial B_y}{\partial y} \right] - K \cdot \frac{e}{p_o c} B_x - \frac{e}{p_o v_o} \left(\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} \right) - 2K \frac{e E_y}{p_o v_o} + \left(\frac{m e E_x}{p_o^2} \right) \left(\frac{m e E_y}{p_o^2} \right) \\ L &= \kappa + \frac{e}{2p_o c} B_s; \quad u = \frac{U}{p_o} = \frac{e}{p_o c^2} \frac{\partial E_x}{\partial t}; \quad g_x = \frac{(mc)^2 \cdot e E_x}{p_o^3} - K \frac{c}{v_o}; \quad g_y = \frac{(mc)^2 \cdot e E_y}{p_o^3}; \\ f_x &= F_x = \frac{e}{c} \frac{\partial B_y}{\partial ct} + \frac{e}{v_o} \frac{\partial E_x}{\partial ct}; \quad f_y = F_y = -\frac{e}{c} \frac{\partial B_x}{\partial ct} + \frac{e}{v_o} \frac{\partial E_y}{\partial ct}.\end{aligned}\quad (G-17)$$

With canonical pairs of: $\{x, P_x\}$, $\{y, P_y\}$, $\{l, P_l\}$; $\pi_\alpha \equiv P_\alpha / p_o$ Hamiltonian has additional term $l \cdot P_l \left(\frac{1}{v_o} \frac{dv_o}{ds} \right)$ for non-reducible case ($p_o \neq \text{const!}$). The rest is just boring replacement of $l = \beta_o \tau$, $P_l = P_\tau / \beta_o$; $\beta_o = v_o / c \dots$ Please not that corresponding coefficients in (G-17) will be modified.

Appendix H. Fokker-Plank Equation

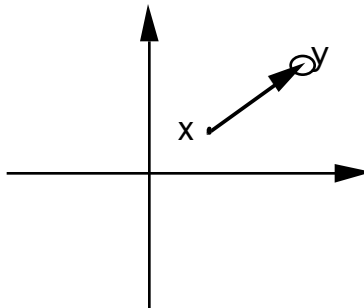
following §83 from *Thermodynamics, Statistical Physics and Kinetics* by Rumer and Ryvkin, Nauka, RAN, 2001 [Rumer]

1. Particles described by distribution function in the Phase Space $x = \{\vec{r}, \vec{P}\}$:

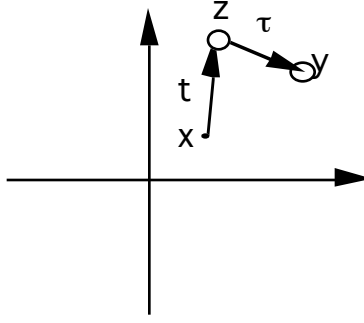
$$f(\vec{r}, \vec{P}, t) \equiv f(x, t); x = \{\vec{r}, \vec{P}\}; \quad \int f(\vec{r}, \vec{P}, t) d\vec{r} d\vec{P} = 1;$$
$$\Rightarrow \rho(\vec{r}, t) = \int f(\vec{r}, \vec{P}, t) d\vec{P}; \quad n(\vec{P}, t) = \int f(\vec{r}, \vec{P}, t) d\vec{r};$$

2. Markov's chain: no dependence on pre-history of the event \Rightarrow **hypothesis**: correlations exist only between two consequent events: the probability to "move" from point x in the phase space to y during time τ depends only on $\{x, y, t, \tau\}$:

$$dw = W(y, x | \tau, t_o) dy; \quad dy = d\vec{r}_y d\vec{P}_y.$$



What probability to move from point x in the phase space to y during time $t + \tau$ through an intermediate point z ?



There t is time for move $x \Rightarrow z$; τ is time for move $z \Rightarrow y$. Two events are independent and total probability is product of two probabilities: $x \Rightarrow z$ and then $z \Rightarrow y$:

$$W(y, z | \tau, t_o + t) dy W(z, x | t, t_o) dz$$

To find probability $W(y, x | t + \tau, t_o)$ it is sufficient to integrate over all z :

$$W(y, x | t + \tau, t_o) = \int dz W(y, z | \tau, t_o + t) dy W(z, x | t, t_o). \quad (\text{H-1})$$

This is Smolukhovsky equation. **Fokker-Plank Equation can be derived** derive from (H-1) in following form ($t_o = 0$):

$$W(y, x | t + \tau, 0) = \int dz W(y, z | \tau, t) dy W(z, x | t, 0). \quad (\text{H-2})$$

Lets consider an analytical (integral-able) function $g(x)$, which is limited in all phase space and goes to zero with all it derivatives at the infinity (i.e. we use a finite system):

$$g(x) \Rightarrow g(x) \rightarrow 0; \frac{\partial^n g(x)}{\prod_{k=1}^n \partial x_{i_k}} \rightarrow 0; |x| \rightarrow \infty.$$

We should keep in mind that $g(x)$ can be an distribution function and these properties are natural for finite system with finite energy: $g(x)$ when $|\vec{r}| > r_{\max}$ or $|\vec{P}| > P_{\max}$ (E_{\max}). Multiplying (H-2) by $g(y)$ and integrating it give us:

$$\int g(y)W(y, x|t + \tau, 0)dy = \iint g(y)W(y, z|\tau, t)dyW(z, x|t, 0)dydz. \quad (H-3)$$

$g(y)$ can be expanded into Taylor series:

$$g(y) = g(z) + \frac{\partial g}{\partial z_i}(y_i - z_i) + \frac{1}{2} \frac{\partial^2 g}{\partial z_i \partial z_k}(y_i - z_i)(y_k - z_k) + \dots$$

(summation is assumed on repeated indexes)

$$\int g(y)W(y, x|t + \tau, 0)dy = \iint \left(g(z) + \frac{\partial g}{\partial z_i}(y_i - z_i) + \frac{1}{2} \frac{\partial^2 g}{\partial z_i \partial z_k}(y_i - z_i)(y_k - z_k) + \dots \right) W(y, z|\tau, t)dyW(z, x|t, 0)dydz. \quad (H-4)$$

Taking into account that:

$$\int g(z)W(y, x|\tau, t)dy = g(z); \quad \int W(y, x|\tau, t)dy = 1;$$

$$\int g(y)W(y, x|t + \tau, 0)dy - \int g(z)W(z, x|t, 0)dz = \int g(y)\{W(y, x|t + \tau, 0) - W(y, x|t, 0)\}dy$$

we can rewrite (H-4) in following from:

$$\int g(y) \frac{W(y, x|t + \tau, 0) - W(y, x|t, 0)}{\tau} dy -$$

$$- \int a_i^{(\tau)}(y, t) \frac{\partial g}{\partial y_i} W(y, x|t, 0) dy$$

$$- \int b_{ik}^{(\tau)}(y, t) \frac{\partial^2 g}{\partial y_i \partial y_k} W(y, x|t, 0) dy - \dots = 0.$$

where we introduce following notations:

$$a_i^{(\tau)}(y, t) = \frac{1}{\tau} \int (z_i - y_i)W(z, y|\tau, t)dz;$$

$$b_{ik}^{(\tau)}(y, t) = \frac{1}{\tau} \int (z_i - y_i)(z_k - y_k)W(z, y|\tau, t)dz;$$

.....

2n-D vector $a = \{a_i^{(\tau)}\}$ is an “average speed” particles’ point on the in Poincaré plot in the phase space. Integration by parts gives us following: $b_{ik}^{(\tau)}$ is 2n-D tensor representing correlations between variations of i 's and k 's components of $x = \{\vec{r}, \vec{p}\}$ with the tensor's trace giving RMS drift of the point $b_{ii}^{(\tau)}(y, t) = \frac{1}{\tau} \int (z_i - y_i)^2 W(z, y | \tau, t) dz$.

Integrating by parts (here we use the boundary condition for finite system!):

$$\begin{aligned} & \int a_i^{(\tau)}(y, t) \frac{\partial g}{\partial y_i} W(y, x | t, 0) dy = \\ & \int \frac{\partial}{\partial y_i} \{a_i^{(\tau)}(y, t) W(y, x | t, 0) g(y)\} dy - \int g(y) \frac{\partial}{\partial y_i} \{a_i^{(\tau)}(y, t) W(y, x | t, 0)\} dy; \\ & \int b_{ik}^{(\tau)}(y, t) \frac{\partial^2 g}{\partial y_i \partial y_k} W(y, x | t, 0) dy = \int \frac{\partial}{\partial y_i} \{b_{ik}^{(\tau)}(y, t) \frac{\partial g}{\partial y_k} W(y, x | t, 0)\} dy - \\ & - \int \frac{\partial}{\partial y_i} \{g(y) \frac{\partial}{\partial y_k} [b_{ik}^{(\tau)}(y, t) W(y, x | t, 0)]\} dy + \int g(y) \frac{\partial^2}{\partial y_i \partial y_k} \{b_{ik}^{(\tau)}(y, t) W(y, x | t, 0)\} dy. \\ & \int \frac{\partial}{\partial y_i} h(y) \prod_{k=1, \dots, 6} dy_k = \int \prod_{k \neq i} dy_k \int \frac{\partial}{\partial y_i} h(y) dy_i = \int \prod_{k \neq i} dy_k \{h(y_{k \neq i}, y_i = +\infty) - h(y_{k \neq i}, y_i = -\infty)\} = 0. \end{aligned}$$

and finally:

$$\int g(y) \left\{ \frac{\partial W(y, x | t, 0)}{\partial t} + \frac{\partial}{\partial y_i} [a_i^{(\tau)}(y, t) W(y, x | t, 0)] - \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_k} [b_{ik}^{(\tau)}(y, t) W(y, x | t, 0)] \right\} dy = 0.$$

$g(y)$ is arbitrary function which requires the expression in the brackets to be zero:

$$\frac{\partial W(y, x | t, 0)}{\partial t} + \frac{\partial}{\partial y_i} [a_i^{(\tau)}(y, t) W(y, x | t, 0)] - \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_k} [b_{ik}^{(\tau)}(y, t) W(y, x | t, 0)] = 0 \quad (\text{H-5})$$

This is called mono-molecular kinetic equation of Fokker and plank. What about the distribution function? The Fokker Plank equation for the distribution function can be derived from this using connection between distribution function $f(x, t) \equiv f(\vec{r}, \vec{p}, t)$ and probability $W(y, x | \tau, t)$: deviation of the particles density in phase space volume dx during time t is equal to the difference between number of particles left this point and arrived into this point:

$$[f(x, t) - f(x, 0)] dx = dx \int [W(x, z | t, 0) f(z, 0) - W(z, x | t, 0) f(x, 0)] dz \quad (\text{H-6})$$

Remembering that $\int W(z, x | t, 0) dz = 1$, we get

$$f(x,t) = \int W(x,z|t,0)f(z,0)dz \quad (\text{H-7})$$

which shows that multiplication on $W(y,x|t,0)$ and integrating over the phase space equivalent to a propagation in the phase space by $(x-z)$ and in time by t .

Thus, multiplying (H-5) by $f(x,0)$ and integrating over x we obtaining Fokker-Plank equation for the distribution function:

$$\frac{\partial f(y,t)}{\partial t} + \frac{\partial}{\partial y_i} [a_i^{(\tau)}(y,t)f(y,t)] - \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_k} [b_{ik}^{(\tau)}(y,t)f(y,t)] = 0 \quad (\text{H-8})$$

This equation also can be written as continuity equations in the phase space:

$$\frac{\partial f(y,t)}{\partial t} + \frac{\partial j_k}{\partial y_k} = 0; \quad j_k = [a_k^{(\tau)}(y,t)f(y,t)] - \frac{1}{2} \frac{\partial^2}{\partial y_i} [b_{ik}^{(\tau)}(y,t)f(y,t)]. \quad (\text{H-8}')$$

This is the final form of the of Fokker-Plank equation, where we just should recognize the terms such as motion of the particle and diffusion coefficients D :

$$\frac{\partial f(y,t)}{\partial t} + \frac{\partial}{\partial y_i} \left[\frac{dy_i(y,t)}{dt} f(y,t) \right] - \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_k} [D_{ik}(y,t)f(y,t)] = 0 \quad (\text{H-8}'')$$

Finally, nobody told us to use time as independent variable, s is as good!

$$\frac{\partial f(y,s)}{\partial s} + \frac{\partial}{\partial y_i} \left[\frac{dy_i(y,s)}{ds} f(y,s) \right] - \frac{1}{2} \frac{\partial^2}{\partial y_i \partial y_k} [D_{ik}(y,s)f(y,s)] = 0 \quad (\text{H-8}''')$$

Appendix I: Standard perturbation method

In 1958 Bogolyubov and Metropolsky developed general asymptotic perturbation methods in the theory of nonlinear oscillators [Bogolyubov]. This methods works till the second order and can be used for many analytical studies.

Mathematics of the method is too involved and interested reader will find it in the English translation of the book: *Bogolyubov and Yu. A. Mitropolsky, Asymptotic Methods in the Theory of Nonlinear Oscillations, Gordon and Breach, New York (1962)*

The system under study is fully described by its n-dimensional vector X. In our case it will be $X^T = [\varphi, I]$ for one-dimension, $X^T = [\varphi_1, I_1, \varphi_2, I_2]$ for two dimensions, etc..

For a slightly perturbed linear oscillator (with $|g|$ to be considered as a small perturbation) described by equation of motion:

$$\frac{dX}{ds} = gF(X, s); \quad (\text{I-1})$$

solution can be written in the first order of perturbation as:

$$(I) \quad X = \xi(s) + g\tilde{F}(\xi, s); \quad \frac{d\xi}{ds} = g\langle F(\xi, s) \rangle; \quad (\text{I-2})$$

where following notations are used:

$$\langle F(X, s) \rangle = \frac{1}{S} \int_s^{s+S} F(X = \text{const}, s) ds; \quad \tilde{F} = \int (F - \langle F \rangle) ds; \quad (\text{I-3})$$

Second order looks a bit more elaborate:

$$(II) \quad X = \xi(s) + g\tilde{F}(\xi, s) + g^2 \left(\overbrace{\left(\tilde{F} \frac{\partial}{\partial \xi} \right) \tilde{F}}^{\sim} \right) - g^2 \frac{\partial \tilde{F}}{\partial \xi} \langle \tilde{F}(\xi, s) \rangle; \quad (\text{I-2})$$

$$\frac{d\xi}{ds} = g\langle F(\xi + g\tilde{F}(\xi, s), s) \rangle \approx g \left\langle \left[1 + g \left(\tilde{F} \frac{\partial}{\partial \xi} \right) \right] F(\xi, s) \right\rangle$$

This method works beautifully and in many cases allows to find finite analytical formulae. There is no systematic approach to extend this method beyond second order perturbation. Lie algebraic theoretical and numerical tools should be used in such cases.

Expansion of 4 – vector in accelerator coordinates

A1 and A3

`n = 4;`

`A1 = Normal[Series[Integrate[ε*Bs[ε*x, ε*y, ε*t], y], {ε, 0, n}]]/2`

$$\begin{aligned} & \frac{1}{2} \left(y \epsilon \text{Bs}[0, 0, 0] + \epsilon^2 \left(t y \text{Bs}^{(0,0,1)}[0, 0, 0] + \frac{1}{2} y^2 \text{Bs}^{(0,1,0)}[0, 0, 0] + x y \text{Bs}^{(1,0,0)}[0, 0, 0] \right) + \right. \\ & \quad \frac{1}{6} \epsilon^3 \left(3 t^2 y \text{Bs}^{(0,0,2)}[0, 0, 0] + 3 t y^2 \text{Bs}^{(0,1,1)}[0, 0, 0] + y^3 \text{Bs}^{(0,2,0)}[0, 0, 0] + \right. \\ & \quad \quad \left. 6 t x y \text{Bs}^{(1,0,1)}[0, 0, 0] + 3 x y^2 \text{Bs}^{(1,1,0)}[0, 0, 0] + 3 x^2 y \text{Bs}^{(2,0,0)}[0, 0, 0] \right) + \\ & \quad \left. \frac{1}{24} \epsilon^4 \left(4 t^3 y \text{Bs}^{(0,0,3)}[0, 0, 0] + 6 t^2 y^2 \text{Bs}^{(0,1,2)}[0, 0, 0] + 4 t y^3 \text{Bs}^{(0,2,1)}[0, 0, 0] + y^4 \text{Bs}^{(0,3,0)}[0, 0, 0] + 12 t^2 x y \text{Bs}^{(1,0,2)}[0, 0, 0] + \right. \right. \\ & \quad \quad \left. \left. 12 t x y^2 \text{Bs}^{(1,1,1)}[0, 0, 0] + 4 x y^3 \text{Bs}^{(1,2,0)}[0, 0, 0] + 12 t x^2 y \text{Bs}^{(2,0,1)}[0, 0, 0] + 6 x^2 y^2 \text{Bs}^{(2,1,0)}[0, 0, 0] + 4 x^3 y \text{Bs}^{(3,0,0)}[0, 0, 0] \right) \right) \end{aligned}$$

`A3 = -Normal[Series[Integrate[ε*Bs[ε*x, ε*y, ε*t], x], {ε, 0, n}]]/2`

$$\begin{aligned} & \frac{1}{2} \left(-x \epsilon \text{Bs}[0, 0, 0] - \epsilon^2 \left(t x \text{Bs}^{(0,0,1)}[0, 0, 0] + x y \text{Bs}^{(0,1,0)}[0, 0, 0] + \frac{1}{2} x^2 \text{Bs}^{(1,0,0)}[0, 0, 0] \right) - \right. \\ & \quad \frac{1}{6} \epsilon^3 \left(3 t^2 x \text{Bs}^{(0,0,2)}[0, 0, 0] + 6 t x y \text{Bs}^{(0,1,1)}[0, 0, 0] + 3 x y^2 \text{Bs}^{(0,2,0)}[0, 0, 0] + \right. \\ & \quad \quad \left. 3 t x^2 \text{Bs}^{(1,0,1)}[0, 0, 0] + 3 x^2 y \text{Bs}^{(1,1,0)}[0, 0, 0] + x^3 \text{Bs}^{(2,0,0)}[0, 0, 0] \right) - \frac{1}{24} \epsilon^4 \\ & \quad \left(4 t^3 x \text{Bs}^{(0,0,3)}[0, 0, 0] + 12 t^2 x y \text{Bs}^{(0,1,2)}[0, 0, 0] + 12 t x y^2 \text{Bs}^{(0,2,1)}[0, 0, 0] + 4 x y^3 \text{Bs}^{(0,3,0)}[0, 0, 0] + 6 t^2 x^2 \text{Bs}^{(1,0,2)}[0, 0, 0] + \right. \\ & \quad \quad \left. 12 t x^2 y \text{Bs}^{(1,1,1)}[0, 0, 0] + 6 x^2 y^2 \text{Bs}^{(1,2,0)}[0, 0, 0] + 4 t x^3 \text{Bs}^{(2,0,1)}[0, 0, 0] + 4 x^3 y \text{Bs}^{(2,1,0)}[0, 0, 0] + x^4 \text{Bs}^{(3,0,0)}[0, 0, 0] \right) \end{aligned}$$

$$-\frac{1}{2} x \epsilon \text{Bs}[0, 0, 0]$$

`Normal[Series[ε*Bs[ε*x, ε*y, ε*t], {ε, 0, n-1}]]`

`-D[A3, x]`

`D[A1, y] - D[A3, x] - Normal[Series[ε*Bs[ε*x, ε*y, ε*t], {ε, 0, n}]]`

`Coefficient[A1, ε]`

$$\frac{1}{2} y \text{Bs}[0, 0, 0]$$

`Coefficient[A3, ε]`

$$-\frac{1}{2} x \text{Bs}[0, 0, 0]$$

A2

```

m = 4;
A21 = Normal[Series[Integrate[ε*(1+K*ε*x)*By[ε*x, 0, ε*t] + K*ε*x*Bs[ε*x, 0, ε*t]], x], {ε, 0, m}]]
A22 = -Normal[Series[Integrate[ε*(Bx[0, ε*y, ε*t] - K*ε*y*Bs[ε*x, 0, ε*t]), y], {ε, 0, m}]]
A23 = Integrate[Integrate[D[Normal[Series[(1+K*ε*x)*By[ε*x, ε*y, ε*t] + K*ε*x*Bs[ε*x, ε*y, ε*t], {ε, 0, m-1}]], y], x], y]/2 -
-Integrate[Integrate[D[Normal[Series[(1+K*ε*x)*Bx[ε*x, ε*y, ε*t] - K*ε*y*Bs[ε*x, ε*y, ε*t], {ε, 0, m-1}]], x], x], y]/2
A2 = A21 + A22 + ε*A23

Simplify[%]
1/24 ε (-y (4 (6 Bx[0, 0, 0] + t ε (6 Bx^{(0,0,1)}[0, 0, 0] + t ε (3 Bx^{(0,0,2)}[0, 0, 0] + t ε Bx^{(0,0,3)}[0, 0, 0])) -
6 y ε (κ (2 Bs[0, 0, 0] + t ε (2 Bs^{(0,0,1)}[0, 0, 0] + t ε Bs^{(0,0,2)}[0, 0, 0])) - 2 Bx^{(0,1,0)}[0, 0, 0] - t ε (2 Bx^{(0,1,1)}[0, 0, 0] + t ε Bx^{(0,1,2)}[0, 0, 0])) +
4 y^2 ε^2 (Bx^{(0,2,0)}[0, 0, 0] + t ε Bx^{(0,2,1)}[0, 0, 0]) + y^3 ε^3 Bx^{(0,3,0)}[0, 0, 0]) + 2 x (12 By[0, 0, 0] +
K y ε (6 Bx[0, 0, 0] + ε (6 t Bx^{(0,0,1)}[0, 0, 0] + 3 t^2 ε Bx^{(0,0,2)}[0, 0, 0] + 3 y Bx^{(0,1,0)}[0, 0, 0] + 3 t y ε Bx^{(0,1,1)}[0, 0, 0] + y^2 ε Bx^{(0,2,0)}[0, 0, 0])) +
ε (2 t^3 ε^2 By^{(0,0,3)}[0, 0, 0] + 3 t^2 ε (2 By^{(0,0,2)}[0, 0, 0] + y ε (By^{(0,1,2)}[0, 0, 0] + Bx^{(1,0,2)}[0, 0, 0])) +
3 t (4 By^{(0,0,1)}[0, 0, 0] + y ε (2 By^{(0,1,1)}[0, 0, 0] + 2 Bx^{(1,0,1)}[0, 0, 0] + y ε (By^{(0,2,1)}[0, 0, 0] + κ Bs^{(1,0,1)}[0, 0, 0] + Bx^{(1,1,1)}[0, 0, 0])) +
y (6 By^{(0,1,0)}[0, 0, 0] + 6 Bx^{(1,0,0)}[0, 0, 0] + 3 y ε (By^{(0,2,0)}[0, 0, 0] + κ Bs^{(1,0,0)}[0, 0, 0] + Bx^{(1,1,0)}[0, 0, 0]) +
y^2 ε^2 (By^{(0,3,0)}[0, 0, 0] - 2 κ Bs^{(1,1,0)}[0, 0, 0] + Bx^{(1,2,0)}[0, 0, 0])) + 3 x^2 ε (4 By^{(1,0,0)}[0, 0, 0] + 4 t ε By^{(1,0,1)}[0, 0, 0] +
2 t^2 ε^2 By^{(1,0,2)}[0, 0, 0] + κ (4 By[0, 0, 0] + ε (2 t^2 ε By^{(0,0,2)}[0, 0, 0] + 2 t (2 By^{(0,0,1)}[0, 0, 0] + y ε (By^{(0,1,1)}[0, 0, 0] + 2 Bx^{(1,0,1)}[0, 0, 0])) +
y (2 By^{(0,1,0)}[0, 0, 0] + 4 Bx^{(1,0,0)}[0, 0, 0] + y ε (By^{(0,2,0)}[0, 0, 0] + 2 Bx^{(1,1,0)}[0, 0, 0])) +
2 y ε By^{(1,1,0)}[0, 0, 0] + 2 t y ε^2 By^{(1,1,1)}[0, 0, 0] + y^2 ε^2 By^{(1,2,0)}[0, 0, 0] + κ (4 Bs[0, 0, 0] + ε (2 t^2 ε Bs^{(0,0,2)}[0, 0, 0] +
2 t (2 Bs^{(0,0,1)}[0, 0, 0] + y ε Bs^{(0,1,1)}[0, 0, 0]) + y (2 Bs^{(0,1,0)}[0, 0, 0] + y ε (Bs^{(0,2,0)}[0, 0, 0] + Bs^{(2,0,0)}[0, 0, 0])) +
2 y ε Bx^{(2,0,0)}[0, 0, 0] + 2 t y ε^2 Bx^{(2,0,1)}[0, 0, 0] + y^2 ε^2 Bx^{(2,1,0)}[0, 0, 0]) +
2 x^2 ε^2 (2 κ (2 Bs^{(1,0,0)}[0, 0, 0] + 2 t ε Bs^{(1,0,1)}[0, 0, 0] + y ε Bs^{(1,1,0)}[0, 0, 0]) +
κ (4 By^{(1,0,0)}[0, 0, 0] + 4 t ε By^{(1,0,1)}[0, 0, 0] + 2 y ε By^{(1,1,0)}[0, 0, 0] + 3 y ε Bx^{(2,0,0)}[0, 0, 0]) +
2 By^{(2,0,0)}[0, 0, 0] + 2 t ε By^{(2,0,1)}[0, 0, 0] + y ε By^{(2,1,0)}[0, 0, 0] + y ε Bx^{(3,0,0)}[0, 0, 0]) +
x^4 ε^3 (3 κ Bs^{(2,0,0)}[0, 0, 0] + 3 κ By^{(2,0,0)}[0, 0, 0] + By^{(3,0,0)}[0, 0, 0]))

Coefficient[A2, ε]
-y Bx[0, 0, 0] + x By[0, 0, 0]

```

Electric field

```

k = 4;
φ1 = -Normal[Series[Integrate[ε*Ex[ε*x, 0, ε*t], x], {ε, 0, k}]] + Normal[Series[Integrate[ε*Ey[0, ε*y, ε*t], y], {ε, 0, k}]]
φ2 = -Integrate[Integrate[D[Normal[Series[Ex[ε*x, ε*y, ε*t], {ε, 0, m-1}]], y], x], y]/2 -
Integrate[Integrate[D[Normal[Series[Ey[ε*x, ε*y, ε*t], {ε, 0, m-1}]], x], x], y]/2
φt = φ1 + ε*φ2

-x ε Ex[0, 0, 0] - y ε Ey[0, 0, 0] - ε^2 (t y Ey^{(0,0,1)}[0, 0, 0] + 1/2 y^2 Ey^{(0,1,0)}[0, 0, 0]) -
1/6 ε^3 (3 t^2 y Ey^{(0,0,2)}[0, 0, 0] + 3 t y^2 Ey^{(0,1,1)}[0, 0, 0] + y^3 Ey^{(0,2,0)}[0, 0, 0]) -
1/24 ε^4 (4 t^3 y Ey^{(0,0,3)}[0, 0, 0] + 6 t^2 y^2 Ey^{(0,1,2)}[0, 0, 0] + 4 t y^3 Ey^{(0,2,1)}[0, 0, 0] + y^4 Ey^{(0,3,0)}[0, 0, 0]) -
ε^2 (t x Ex^{(0,0,1)}[0, 0, 0] + 1/2 x^2 Ex^{(1,0,0)}[0, 0, 0]) - 1/6 ε^3 (3 t^2 x Ex^{(0,0,2)}[0, 0, 0] + 3 t x^2 Ex^{(1,0,1)}[0, 0, 0] + x^3 Ex^{(2,0,0)}[0, 0, 0]) -
1/24 ε^4 (4 t^3 x Ex^{(0,0,3)}[0, 0, 0] + 6 t^2 x^2 Ex^{(1,0,2)}[0, 0, 0] + 4 t x^3 Ex^{(2,0,1)}[0, 0, 0] + x^4 Ex^{(3,0,0)}[0, 0, 0]) +
ε (1/2 (-x y ε Ex^{(0,1,0)}[0, 0, 0] - t x y ε^2 Ex^{(0,1,1)}[0, 0, 0] - 1/2 t^2 x y ε^3 Ex^{(0,1,2)}[0, 0, 0] - 1/2 x y^2 ε^2 Ex^{(0,2,0)}[0, 0, 0] - 1/2 t x y^2 ε^3 Ex^{(0,2,1)}[0, 0, 0] - 1/6 x
y^3 ε^3 Ex^{(0,3,0)}[0, 0, 0] - 1/2 x^2 y ε^2 Ex^{(1,1,0)}[0, 0, 0] - 1/2 t x^2 y ε^3 Ex^{(1,1,1)}[0, 0, 0] - 1/4 x^2 y^2 ε^3 Ex^{(1,2,0)}[0, 0, 0] - 1/6 x^3 y ε^3 Ex^{(2,1,0)}[0, 0, 0]) +
1/2 (-x y ε Ey^{(1,0,0)}[0, 0, 0] - t x y ε^2 Ey^{(1,0,1)}[0, 0, 0] - 1/2 t^2 x y ε^3 Ey^{(1,0,2)}[0, 0, 0] - 1/2 x y^2 ε^2 Ey^{(1,1,0)}[0, 0, 0] - 1/2 t x y^2 ε^3 Ey^{(1,1,1)}[0, 0, 0] - 1/6
x y^3 ε^3 Ey^{(1,2,0)}[0, 0, 0] - 1/2 x^2 y ε^2 Ey^{(2,0,0)}[0, 0, 0] - 1/2 t x^2 y ε^3 Ey^{(2,0,1)}[0, 0, 0] - 1/4 x^2 y^2 ε^3 Ey^{(2,1,0)}[0, 0, 0] - 1/6 x^3 y ε^3 Ey^{(3,0,0)}[0, 0, 0]))

φ1 = Normal[Series[Integrate[ε*(Bs[0, 0, ε*t] - Es[0, 0, 0]), t], {ε, 0, k}]]
1/2 t^2 ε^2 Es^{(0,0,1)}[0, 0, 0] + 1/6 t^3 ε^3 Es^{(0,0,2)}[0, 0, 0] + 1/24 t^4 ε^4 Es^{(0,0,3)}[0, 0, 0]

t = -τ/c
-τ/c

```

Hamiltonian

$$H = -(1 + K \epsilon x) p_0 \sqrt{1 + 2 \frac{(\epsilon \delta - e \phi t / c)}{p_0 / \beta_0} + \frac{(\epsilon \delta - e \phi t / c)^2}{p_0^2} - \frac{(\epsilon P_1 - e A_1 / c)^2}{p_0^2} - \frac{(\epsilon P_3 - e A_3 / c)^2}{p_0^2}} - \frac{e A_2 / c + \kappa \epsilon x (\epsilon P_3 - e A_3 / c) - \kappa \epsilon y (\epsilon P_1 - e A_1 / c) + \epsilon \delta / \beta_0 - e \phi}{p_0}$$

First order terms

```
H1 = Normal[Series[H, {ε, 0, 1}]]
-p0 + 
$$\frac{\epsilon (-c K p_0 x \beta_0 + e y \beta_0 B_x[0, 0, 0] - e x \beta_0 B_y[0, 0, 0] - e x E_x[0, 0, 0] - e y E_y[0, 0, 0])}{c \beta_0}$$

Cx = Coefficient[H1, ε x]

$$\frac{-c K p_0 \beta_0 - e \beta_0 B_y[0, 0, 0] - e E_x[0, 0, 0]}{c \beta_0}$$

Cy = Coefficient[H1, ε y]

$$\frac{e \beta_0 B_x[0, 0, 0] - e E_y[0, 0, 0]}{c \beta_0}$$

```


Second order terms

```

H2 = Normal[Series[H, {ε, 0, 2}]]
Cxx = Coefficient[H2, ε^2 x^2]
Expand[Cxx]

$$\frac{e^2 Bs[0, 0, 0]^2}{8 c^2 p_0} - \frac{e K By[0, 0, 0]}{2 c} - \frac{e K Ex[0, 0, 0]}{c \beta_0} - \frac{e^2 Ex[0, 0, 0]^2}{2 c^2 p_0} + \frac{e^2 Ex[0, 0, 0]^2}{2 c^2 p_0 \beta_0^2} - \frac{e By^{(1,0,0)}[0, 0, 0]}{2 c} - \frac{e Ex^{(1,0,0)}[0, 0, 0]}{2 c \beta_0}$$

Cyy = Coefficient[H2, ε^2 y^2]
Expand[Cyy]

$$\frac{e^2 Bs[0, 0, 0]^2}{8 c^2 p_0} - \frac{e^2 Ey[0, 0, 0]^2}{2 c^2 p_0} + \frac{e^2 Ey[0, 0, 0]^2}{2 c^2 p_0 \beta_0^2} + \frac{e Bx^{(0,1,0)}[0, 0, 0]}{2 c} - \frac{e Ey^{(0,1,0)}[0, 0, 0]}{2 c \beta_0}$$

Cxy = Coefficient[H2, ε^2 x y]
Expand[Cxy]

$$\frac{e K Bx[0, 0, 0]}{2 c} - \frac{e K Ey[0, 0, 0]}{c \beta_0} - \frac{e^2 Ex[0, 0, 0] Ey[0, 0, 0]}{c^2 p_0} + \frac{e^2 Ex[0, 0, 0] Ey[0, 0, 0]}{c^2 p_0 \beta_0^2} - \frac{e By^{(0,1,0)}[0, 0, 0]}{2 c} - \frac{e Ex^{(0,1,0)}[0, 0, 0]}{2 c \beta_0} - \frac{e Bx^{(1,0,0)}[0, 0, 0]}{2 c} - \frac{e Ey^{(1,0,0)}[0, 0, 0]}{2 c \beta_0}$$

CxP1 = Coefficient[H2, ε^2 x P1]
0
CxP3 = Expand[Coefficient[H2, ε^2 x P3]]

$$x + \frac{e Bs[0, 0, 0]}{2 c p_0}$$

CyP1 = Expand[Coefficient[H2, ε^2 y P1]]

$$-x - \frac{e Bs[0, 0, 0]}{2 c p_0}$$

CyP3 = Coefficient[H2, ε^2 y P3]
0
Cxδ = Expand[Coefficient[H2, ε^2 x δ]]

$$-\frac{K}{\beta_0} - \frac{e Ex[0, 0, 0]}{c p_0} + \frac{e Ex[0, 0, 0]}{c p_0 \beta_0^2}$$

Cyδ = Expand[Coefficient[H2, ε^2 y δ]]

$$-\frac{e Ey[0, 0, 0]}{c p_0} + \frac{e Ey[0, 0, 0]}{c p_0 \beta_0^2}$$



---


CP1δ = Expand[Coefficient[H2, ε^2 P1 δ]]
0
CP3δ = Expand[Coefficient[H2, ε^2 P3 δ]]
0
CPδδ = Expand[Coefficient[H2, ε^2 δ^2]]

$$-\frac{1}{2 p_0} + \frac{1}{2 p_0 \beta_0^2}$$

CPtt = Expand[Coefficient[H2, ε^2 τ^2]]

$$-\frac{e Es^{(0,0,1)}[0, 0, 0]}{2 c^2}$$

Cxt = Expand[Coefficient[H2, ε^2 x τ]]

$$\frac{e By^{(0,0,1)}[0, 0, 0]}{c^2} + \frac{e Ex^{(0,0,1)}[0, 0, 0]}{c^2 \beta_0}$$

Cyt = Expand[Coefficient[H2, ε^2 y τ]]

$$-\frac{e Bx^{(0,0,1)}[0, 0, 0]}{c^2} + \frac{e Ey^{(0,0,1)}[0, 0, 0]}{c^2 \beta_0}$$

CP1t = Expand[Coefficient[H2, ε^2 P1 τ]]
0
CP3t = Expand[Coefficient[H2, ε^2 P3 τ]]
0

```

Third order terms

H3 = Normal[Series[H, {ε, 0, 3}]]

Cx3 = Coefficient[H3, ε^3 x^3]; Expand[Cx3]

$$\frac{e^2 K Bs[0, 0, 0]^2}{8 c^2 p_0} - \frac{e^3 Bs[0, 0, 0]^2 Ex[0, 0, 0]}{8 c^3 p_0^2 \beta_0} - \frac{e^2 K Ex[0, 0, 0]^2}{2 c^2 p_0} + \frac{e^2 K Ex^2[0, 0, 0]^2}{2 c^2 p_0 \beta_0^2} - \frac{e^3 Ex[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0^3} +$$

$$\frac{e^3 Ex[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0} - \frac{e K Bs^{(1,0,0)}[0, 0, 0]}{12 c} + \frac{e^2 Bs[0, 0, 0] Bs^{(1,0,0)}[0, 0, 0]}{8 c^2 p_0} - \frac{e K By^{(1,0,0)}[0, 0, 0]}{3 c} - \frac{e K Ex^{(1,0,0)}[0, 0, 0]}{2 c \beta_0}$$

$$\frac{e^2 Ex[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ex[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e By^{(2,0,0)}[0, 0, 0]}{6 c} - \frac{e Ex^{(2,0,0)}[0, 0, 0]}{6 c \beta_0}$$

$$\frac{e^2 K Bs[0, 0, 0]^2}{8 c^2 p_0} - \frac{e^3 Bs[0, 0, 0]^2 Ex[0, 0, 0]}{8 c^3 p_0^2 \beta_0} - \frac{e^2 K Ex[0, 0, 0]^2}{2 c^2 p_0} + \frac{e^2 K Ex^2[0, 0, 0]^2}{2 c^2 p_0 \beta_0^2} - \frac{e^3 Ex[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0^3} +$$

$$\frac{e^3 Ex[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0} - \frac{e K Bs^{(1,0,0)}[0, 0, 0]}{12 c} + \frac{e^2 Bs[0, 0, 0] Bs^{(1,0,0)}[0, 0, 0]}{8 c^2 p_0} - \frac{e K By^{(1,0,0)}[0, 0, 0]}{3 c} - \frac{e K Ex^{(1,0,0)}[0, 0, 0]}{2 c \beta_0}$$

$$\frac{e^2 Ex[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ex[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e By^{(2,0,0)}[0, 0, 0]}{6 c} - \frac{e Ex^{(2,0,0)}[0, 0, 0]}{6 c \beta_0}$$

Cy3 = Coefficient[H3, ε^3 y^3]; Expand[Cy3]

$$\frac{e^3 Bs[0, 0, 0]^2 Ey[0, 0, 0]}{8 c^3 p_0^2 \beta_0} - \frac{e^3 Ey[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0^3} + \frac{e^2 Ey[0, 0, 0]^3}{2 c^3 p_0^2 \beta_0} + \frac{e K Bs^{(0,1,0)}[0, 0, 0]}{4 c} + \frac{e^2 Bs[0, 0, 0] Bs^{(0,1,0)}[0, 0, 0]}{8 c^2 p_0}$$

$$\frac{e^2 Ey[0, 0, 0] Ey^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ey[0, 0, 0] Ey^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} + \frac{e Bx^{(0,2,0)}[0, 0, 0]}{6 c} - \frac{e Ey^{(0,2,0)}[0, 0, 0]}{6 c \beta_0}$$

Cx2y = Coefficient[H3, ε^3 x^2 y]; Expand[Cx2y]

$$\frac{e^3 Bs[0, 0, 0]^2 Ey[0, 0, 0]}{8 c^3 p_0^2 \beta_0} - \frac{e^2 K Ex[0, 0, 0] Ey[0, 0, 0]}{c^2 p_0} + \frac{e^2 K Ex[0, 0, 0] Ey[0, 0, 0]}{c^2 p_0 \beta_0^2} - \frac{3 e^3 Ex[0, 0, 0]^2 Ey[0, 0, 0]}{2 c^3 p_0^2 \beta_0^2} +$$

$$\frac{3 e^3 Ex[0, 0, 0]^2 Ey[0, 0, 0]}{2 c^3 p_0^2 \beta_0} + \frac{e K Bs^{(0,1,0)}[0, 0, 0]}{4 c} + \frac{e^2 Bs[0, 0, 0] Bs^{(0,1,0)}[0, 0, 0]}{4 c^2 p_0} - \frac{e K By^{(0,1,0)}[0, 0, 0]}{4 c}$$

$$\frac{e K Ex^{(0,1,0)}[0, 0, 0]}{2 c \beta_0} - \frac{e^2 Ex[0, 0, 0] Ex^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ex[0, 0, 0] Ex^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e K Bx^{(1,0,0)}[0, 0, 0]}{2 c}$$

$$\frac{e^2 Ey[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ey[0, 0, 0] Ex^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e K Ey^{(1,0,0)}[0, 0, 0]}{2 c \beta_0} - \frac{e^2 Ex[0, 0, 0] Ey^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0} +$$

$$\frac{e^2 Ex[0, 0, 0] Ey^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e By^{(1,1,0)}[0, 0, 0]}{4 c} - \frac{e Ex^{(1,1,0)}[0, 0, 0]}{4 c \beta_0} - \frac{e Bx^{(2,0,0)}[0, 0, 0]}{4 c} - \frac{e Ey^{(2,0,0)}[0, 0, 0]}{4 c \beta_0}$$

Cxy2 = Coefficient[H3, ε^3 x y^2]; Expand[Cxy2]

$$\frac{e^2 K Bs[0, 0, 0]^2}{8 c^2 p_0} - \frac{e^3 Bs[0, 0, 0]^2 Ex[0, 0, 0]}{8 c^3 p_0^2 \beta_0} - \frac{e^2 K Ey[0, 0, 0]^2}{2 c^2 p_0} + \frac{e^2 K Ey[0, 0, 0]^2}{2 c^2 p_0 \beta_0^2} -$$

$$\frac{3 e^3 Ex[0, 0, 0] Ey[0, 0, 0]^2}{2 c^3 p_0^2 \beta_0^3} + \frac{3 e^3 Ex[0, 0, 0] Ey[0, 0, 0]^2}{2 c^3 p_0^2 \beta_0} - \frac{e K Bx^{(0,1,0)}[0, 0, 0]}{4 c} - \frac{e^2 Ey[0, 0, 0] Ex^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0} +$$

$$\frac{e^2 Ey[0, 0, 0] Ex^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e K Ey^{(0,1,0)}[0, 0, 0]}{2 c \beta_0} - \frac{e^2 Ex[0, 0, 0] Ey^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ex[0, 0, 0] Ey^{(0,1,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} -$$

$$\frac{e By^{(0,2,0)}[0, 0, 0]}{4 c} - \frac{e Ex^{(0,2,0)}[0, 0, 0]}{4 c \beta_0} + \frac{e K Bs^{(1,0,0)}[0, 0, 0]}{4 c} + \frac{e^2 Bs[0, 0, 0] Bs^{(1,0,0)}[0, 0, 0]}{4 c^2 p_0} -$$

$$\frac{e^2 Ey[0, 0, 0] Ey^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0} + \frac{e^2 Ey[0, 0, 0] Ey^{(1,0,0)}[0, 0, 0]}{2 c^2 p_0 \beta_0^2} - \frac{e Bx^{(1,1,0)}[0, 0, 0]}{4 c} - \frac{e Ey^{(1,1,0)}[0, 0, 0]}{4 c \beta_0}$$

CP1p3 = Coefficient[H3, ε^3 P1^3]; Expand[CP1p3]

0

CP3p3 = Coefficient[H3, ε^3 P3^3]; Expand[CP3p3]

0

Cδ3 = Coefficient[H3, ε^3 δ^3]; Expand[Cδ3]

$$-\frac{1}{2 p_0^2 \beta_0^3} + \frac{1}{2 p_0^2 \beta_0}$$

Ct3 = Coefficient[H3, ε^3 τ^3]; Expand[Ct3]

$$\frac{e Es^{(0,0,2)}[0, 0, 0]}{6 c^3}$$

Forth order terms

```
H4 = Normal[Series[H, {ε, 0, 4}]]; Expand[H4]
```

A very large output was generated. Here is a sample of it:

$$-p_0 - K p_0 x \epsilon + \frac{p_1^2 \epsilon^2}{2 p_0} + \frac{p_3^2 \epsilon^2}{2 p_0} - \frac{K x \delta \epsilon^2}{\beta_0} - \frac{\delta^2 \epsilon^2}{2 p_0} + \frac{\delta^2 \epsilon^2}{2 p_0 \beta_0^2} + \frac{K p_1^2 x \epsilon^3}{2 p_0} + \llcorner 1156 \gg + \frac{e x^2 y \epsilon^4 \epsilon B y^{(2,0,1)} [0,0,0]}{4 c^2 \beta_0} - \frac{e x^2 y^2 \epsilon^4 B x^{(2,1,0)} [0,0,0]}{8 c} - \frac{e x^3 y \epsilon^4 B y^{(2,1,0)} [0,0,0]}{12 c} - \frac{e x^3 y \epsilon^4 B x^{(2,1,0)} [0,0,0]}{12 c \beta_0} - \frac{e x^2 y^2 \epsilon^4 B y^{(2,1,0)} [0,0,0]}{8 c \beta_0} - \frac{e x^3 y \epsilon^4 B x^{(3,0,0)} [0,0,0]}{12 c} - \frac{e x^4 \epsilon^4 B y^{(3,0,0)} [0,0,0]}{24 c} - \frac{e x^4 \epsilon^4 B x^{(3,0,0)} [0,0,0]}{24 c \beta_0} - \frac{e x^3 y \epsilon^4 B y^{(3,0,0)} [0,0,0]}{12 c \beta_0}$$

```
CP1p4 = Coefficient[H4, ε^4 P1^4]; Expand[CP1p4]
```

$$\frac{1}{8 p_0^3}$$

```
CP3p4 = Coefficient[H4, ε^4 P3^4]; Expand[CP3p4]
```

$$\frac{1}{8 p_0^3}$$

```
CP1p2P3p2 = Coefficient[H4, ε^4 P1^2 P3^2]; Expand[CP1p2P3p2]
```

$$\frac{1}{4 p_0^3}$$

```
CP1p1P3p3 = Coefficient[H4, ε^4 P1 P3^3]; Expand[CP1p1P3p3]
```

$$0$$

Higher orders : just for fun

```
H6 = Normal[Series[H, {ε, 0, 6}]]; Expand[H6]
```

```
CP1p6 = Coefficient[H6, ε^6 P1^6]; Expand[CP1p6]
```

$$\frac{1}{16 p_0^5}$$

```
H14 = Normal[Series[H, {ε, 0, 14}]]; Expand[H14]
```

```
CP1p14 = Coefficient[H14, ε^14 P1^14]; Expand[CP1p14]
```

References:

- [Landau] L.D. Landau, E.M.Lifshitz, The Classical Theory of Field, Pergamon Press, 1975
- [Kolomensky] A.A.Kolomensky, A.N.Lebedev, Theory of Cyclic Particle Accelerators, Wiley, New York, 1962
- [Jackson] J.D. Jackson, Classical Electrodynamics, Second Edition, John Wiley & Sons, New York, 1998
- [Rumer] Thermodynamics, Statistical Physics and Kinetics, Yu.B.Rumer and M.Sm.Ryvkin, 3rd edition, Novosibirsk State University, Nauka, RAN, 2001
- [Bogolyubov] N.Bogolyubov, Yu.A.Mitropolsky, "Asymptotic Methods in the theory of non-linear oscillations", Fizmatgiz, Moscow, 1977
- [Courant] E.E.Courant, H.S.Snyder, "Theory of the alternating-gradient synchrotron", Annals of Physics, 3, 1, 1958