A.Streun, PSI, Oct.19, 2010

The simplified Hamiltonian for transverse beam dynamics

1 The curvilinear system

A beam consists of particles travelling closely along a curve in space, therefore it is convenient to introduce a local curvilinear coordinate system traveling along the design orbit of the accelerator, which is a curve in space:

$$\vec{r}(s) = \vec{r}_o(s) + \hat{\vec{x}}(s) \cdot x(s) + \hat{\vec{y}}(s) \cdot y(s) \tag{1}$$

The vectors $\{\hat{\vec{x}}; \hat{\vec{s}}; \hat{\vec{y}}\}$ form a right hand side system of unit vectors [6], with $\hat{\vec{s}}$ tangential to the curve, pointing in forward direction, $\hat{\vec{x}}$ radial to the curvature of the curve, pointing away from the center of curvature, and $\hat{\vec{y}} = \hat{\vec{s}} \times \hat{\vec{x}}$. These unit vectors are defined a priori, i.e. they define the curve of our choice, whereas $x(s)$ and $y(s)$ will be solutions of the equations of motion in the curvilinear system. The independent coordinate s is the length of the curve, measured from some point. We neglect here a possible torsion of the curve.

Derivatives of the vectors with respect to s [5, 2]:

$$\frac{d\vec{r}_o(s)}{ds} = \hat{\vec{s}} \quad \text{because} \quad \vec{r}_o(s + ds) = \vec{r}_o + \hat{\vec{s}} \cdot ds \tag{2}$$

Following the curvature by an angle $d\phi = hds$ with $h = 1/\rho$ the local curvature (inverse radius of curvature) applies a rotation around the y-axis:

$$\vec{a} + d\vec{a} = \begin{pmatrix} \cos d\phi & -\sin d\phi & 0 \\ \sin d\phi & \cos d\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \vec{a} \quad \xrightarrow{d\phi \ll 1} \quad d\vec{a} = \begin{pmatrix} 0 & -h\,ds & 0 \\ h\,ds & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3}$$

The unit vectors are given by

$$\hat{\vec{x}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \hat{\vec{s}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \hat{\vec{y}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

So the derivatives follow as

$$\frac{d\hat{\vec{x}}}{ds} = h\hat{\vec{s}} \qquad \frac{d\hat{\vec{s}}}{ds} = -h\hat{\vec{x}} \qquad \frac{d\hat{\vec{y}}}{ds} = 0 \tag{4}$$

2 The relativistic Hamiltonian

In the most general form, the Hamiltonian with electromagnetic fields is given by

$$H_o = c\sqrt{(\underline{\vec{p}} - q\underline{\vec{A}})^2 + m_o^2 c^2} + qV \tag{5}$$

with m, q particle mass and charge, V the electrostatic potential and $\underline{\vec{A}}$ the magnetic vector potential. The underline indicates that the vector is expressed in a global, cartesian coordinate system. $\underline{\vec{p}}$ is the *canonical momentum*, not the kinetic momentum, which would be given by

$$m\underline{\vec{v}} = \underline{\vec{p}} - q\underline{\vec{A}}. \tag{6}$$

The Hamilton-Jacobi equations of motion,

$$\dot{\underline{x}}_k = \frac{\partial H_o}{\partial \underline{p}_k} \qquad \dot{\underline{p}}_k = -\frac{\partial H_o}{\partial \underline{x}_k}, \tag{7}$$

apply only to the canonic coordinates $\underline{\vec{x}}$ and $\underline{\vec{p}}$, but not to the kinetic momentum.

3 Canonical transformations in general

This is well explained in [5]: Canonical transformations perform a coordinate transformation of the Hamiltonian from an old set of canonical variables $\{p; q\}$ to a new set of variables $\{P; Q\}$ which are canonical too. So, a canonical transformation is more than just a coordinate transformation.

The transformation is performed through a *generating function* F. Its explicit time dependency is added to the Hamiltonian:

$$H_{\text{new}}(P, Q, t) = H_{\text{old}}(p, q, t) + \frac{\partial F}{\partial t} \tag{8}$$

Four types of generating functions are distinguished, depending on which set of old and new coordinates they contain:

$$F_1(q, Q, t), \qquad F_2(q, P, t), \qquad F_3(p, Q, t), \qquad F_4(p, P, t)$$

One uses the one which is most convenient for the problem given.

As an example, let's use F_2. We want the equation of motion in the new coordinates:

$$\dot{Q}_k = \frac{\partial H_{\text{new}}(P, Q, t)}{\partial P_k} = \frac{\partial}{\partial P_k}\left(\partial H_{\text{old}}(p, q, t) + \frac{\partial F_2(q, P, t)}{\partial t}\right) = \frac{\partial}{\partial t}\frac{\partial F_2(q, P, t)}{\partial P_k}$$

$$\longrightarrow \qquad Q_k = \frac{\partial F_2(q, P, t)}{\partial P_k}$$

$$\dot{p}_k = -\frac{\partial H_{\rm old}(p,q,t)}{\partial q_k} = -\frac{\partial}{\partial q_k}\left(\partial H_{\rm new}(P,Q,t) - \frac{\partial F_2(q,P,t)}{\partial t}\right) = \frac{\partial}{\partial t}\frac{\partial F_2(q,P,t)}{\partial q_k}$$

$$\longrightarrow \qquad p_k = \frac{\partial F_2(q,P,t)}{\partial q_k}$$

For other cases, see [5] or calculate analogously.
The signs for the four cases are as $+-$, $++$, $--$, $-+$.

4 The contact transformation

Now we want to apply a canonical transformation to transform the general Hamiltonian from eq.5 into the curvilinear system. We have the *new* coordinates $q = \{x; s; y\}$ and the *old* momenta $P = \{\underline{p_x}; \underline{p_y}; \underline{p_z}\}$, so we need a generating function of third type [5]. A possible solution is given by

$$F_3(\underline{\vec{p}}, \vec{r}, t) = -\underline{\vec{p}} \bullet \vec{r}(s) \tag{9}$$

with $\vec{r}(s)$ from eq.1 ($\bullet$ = scalar product). The old coordinates (we are not interested in) are obtained from

$$\underline{x} = -\frac{\partial F_3}{\partial \underline{p_x}} = \vec{r}_x \ \text{ etc,}$$

i.e. $\underline{\vec{r}}(s) = \vec{r}(s)$, which is trivial and was just the definition of our curve in space. What we need are the new momenta in the curvilinear system:

$$p_s = -\frac{\partial F_3}{\partial s} = \underline{\vec{p}} \bullet \left(\frac{\partial \vec{r}_o(s)}{\partial s} + \frac{\partial \hat{\vec{x}}(s)}{\partial s} \cdot x + \frac{\partial \hat{\vec{y}}(s)}{\partial s} \cdot y\right) \tag{10}$$

From eqs.2 and 4 follows:

$$p_s = \underline{\vec{p}} \bullet \hat{\vec{s}} \cdot (1 + hx) \tag{11}$$

For the other components we get directly with eq.1

$$p_x = \underline{\vec{p}} \bullet \hat{\vec{x}} \qquad p_y = \underline{\vec{p}} \bullet \hat{\vec{y}} \tag{12}$$

Since F_3 does not contain an explicit time dependence, the new Hamiltonian is the old:

$$H_1 = H_o \tag{13}$$

It is important to note, that p_s is *not* the tangential component of the momentum (which would be just given by $\underline{\vec{p}} \bullet \hat{\vec{s}}$), but the new $\vec{p}$ is the *canonical* momentum in the curvilinear system.

The next step is handled differently by the different authors: In refs.[6, 2, 3] the vector potential is considered a *canonical vector potential* and transformed in the same way as the momentum, but refs.[1, 5] transform only the canonical momentum, since transforming both would imply a transformation of the kinetic momentum (see eq.6), which is not canonic. We follow the second line, since it will lead to the correct dipole focusing term in the simplified Hamiltonian using an obvious definition of the vector potential. So, we insert the canonical transformation in the components of the momentum, i.e. its projection along the new coordinate axis:

$$\underline{\vec{p}} \bullet \hat{\vec{s}} = \frac{p_s}{1+hx} \qquad \underline{\vec{p}} \bullet \hat{\vec{x}} = p_x \qquad \underline{\vec{p}} \bullet \hat{\vec{y}} = p_y$$

and insert this into the Hamiltonian:

$$H_1 = c\sqrt{\left(\frac{p_s}{1+hx} - qA_s\right)^2 + (p_x - qA_x)^2 + (p_y - qA_y)^2 + m_o^2c^2} + qV \tag{14}$$

5 Change of variable $t \to s$

For beam transformation we would like to use s, the length along the curve describing the accelerator as the independent variable rather than the time t [3]. The equations of motion are given by (y not shown):

$$\dot{x} = \frac{\partial x}{\partial t} = \frac{\partial H_1}{\partial p_x}, \quad \dot{p_x} = -\frac{\partial H_1}{\partial x} \qquad \dot{s} = \frac{\partial H_1}{\partial p_s}, \quad \dot{p_s} = -\frac{\partial H_1}{\partial s}$$

Just inverting the last two equations we get

$$t' = \frac{\partial t}{\partial s} = \frac{1}{\dot{s}} = \frac{\partial p_s}{\partial H_1} \qquad H_1' = \frac{\partial H_1}{\partial s} = -\frac{\partial p_s}{\partial t}$$

and for x (and for y the same way)

$$x' = \frac{\partial x}{\partial s} = \frac{\partial x/\partial t}{\partial s/\partial t} = \frac{\partial H_1/\partial p_x}{\partial H_1/\partial p_s}$$

H_1 must not change if we vary the canonical momenta, i.e.

$$0 = dH_1 = \left.\frac{\partial H_1}{\partial p_x}\right|_{p_s} dp_x + \left.\frac{\partial H_1}{\partial p_s}\right|_{p_x} dp_s \qquad \longrightarrow \qquad \left.\frac{\partial p_s}{\partial p_x}\right|_{H_1} = -\frac{\partial H_1/\partial p_x}{\partial H_1/\partial p_s}$$

$$\longrightarrow \qquad x' = -\frac{\partial p_s}{\partial p_x} \qquad p_x' = \frac{\partial p_s}{\partial x} \tag{15}$$

Now $-p_s$ plays the role of a new Hamiltonian for a new set of coordinates, given by $\{x, p_x, y, p_y, H_1, t\}$, so we define $H_2 = -p_s$, with the equations of motion

$$x' = \frac{\partial H_2}{\partial p_x}, \quad p_x' = -\frac{\partial H_2}{\partial x}, \qquad H_1' = \frac{\partial H_2}{\partial t}, \quad t' = -\frac{\partial H_2}{\partial H_1}, \tag{16}$$

The new Hamiltonian is obtained by solving eq.14 for $-p_s$:

$$H_2 = -p_s = -(1+hx) \cdot \left(\sqrt{\left(\frac{H_1 - qV}{c}\right)^2 - m_o^2c^2 - (p_x - qA_x)^2 - (p_y - qA_y)^2} + qAs\right) \tag{17}$$

6 Approximations

Several approximations will be introduced, which are justified at least for large high energy machines [3], and make it much simpler for us to proceed:

- High relativistic: $E = pc$
- Static: $\partial \vec{A}/\partial t = 0$
- No electric fields, i.e. no energy gain: $V = 0$. The H_1 from eq.14 is just the total energy of the particle: $H_1 = E$.
- No fringe fields, i.e. piece-wise constant, hard-edge fields: $B_s = 0$. In cartesian and cylindrical coordinates, which are the two cases of a tangential coordinate system to a torsion-free curve in space we consider, the curl operator gives

$$B_s = \frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}$$

$$\text{This allows} \quad A_x = A_y = 0 \qquad \longrightarrow \qquad B_x = \frac{\partial A_s}{\partial y} \qquad B_y = \frac{\partial A_s}{\partial x}$$

- Paraxial approximation: particles stay close to the curve, and their transverse momenta are small compared to the total momentum, i.e. $p_{x,y} \ll p$.

The first four items simplify the Hamiltonian from eq.17 as

$$H_2 = -(1 + hx) \left(\sqrt{\underbrace{E^2/c^2 - m_o^2 c^2}_{p^2} - p_x^2 - p_y^2} + qA_s \right) \tag{18}$$

The fifth item allows to expand the square root, since $p_{x,y}/p \ll 1$:

$$H_2 = -(1 + hx) \left(p \cdot \left(1 - \frac{p_x^2 + p_y^2}{2p^2} \right) + qA_s \right) \tag{19}$$

Finally we refer to a reference momentum p_o the machine is designed for, and consider particles with relative momentum deviations

$$\delta = \frac{p - p_o}{p_o} \tag{20}$$

Introducing p_o and δ in the Hamiltonian gives

$$\tilde{H}_2 = -(1 + hx) \left(1 + \delta - \frac{\tilde{p}_x^2 + \tilde{p}_y^2}{2(1 + \delta)} + \frac{q}{p_o} A_s \right) \tag{21}$$

Here we introduced dimensionless momenta normalized to the reference momentum $\tilde{p}_{x,y} = p_{x,y}/p_o$ and also normalized the Hamiltonian, such that $\tilde{H}_2 = H_2/p_o$ is dimensionless.

7 Modeling the vector potential

The curve in space could be any curve, but we restrict it to a curve which is composed from pieces which are either straight (cartesian geometry) or curved with constant curvature (cylindrical geometry) - like tracks of a toy train. In both geometries the magnetic field, given by $\vec{B} = \nabla \times \vec{A}$, simplifies due to $A_x = A_y = 0$:

In **cartesian** geometry this gives

$$B_x = -\frac{\partial A_s}{\partial y} \qquad B_y = +\frac{\partial A_s}{\partial x} \tag{22}$$

In **cylindrical** geometry we have to use the actual radius, given by $\rho = \rho_o + x = \frac{1+hx}{h}$, with h the local *design* curvature of our curve. Using the curl in cylindrical coordinates from any mathematics book, we have to identify $\hat{\vec{\rho}} = \hat{\vec{x}}$, $\hat{\vec{\phi}} = \hat{\vec{s}}$, $\hat{\vec{z}} = \hat{\vec{y}}$. This yields

$$B_x = -\frac{\partial A_s}{\partial y} \qquad B_y = \frac{h}{1+hx}A_s + \frac{\partial A_s}{\partial x} \tag{23}$$

Since there are no currents, we have $\nabla \times \vec{B} = 0$.

In **cartesian** geometry a possible expression for $\vec{A}$ is given by a *multipole expansion*:

$$\frac{q}{p_o}As = -\Re \sum_{n=1}^{\infty} \frac{ia_n + b_n}{n}(x+iy)^n \tag{24}$$

We prove that $\nabla \times \vec{B} = 0$:

$$\nabla \times \vec{B} = \nabla \times \nabla \times \vec{A} = \frac{\partial^2 A_s}{\partial x^2} + \frac{\partial^2 A_s}{\partial y^2}$$

$$\frac{\partial^2 A_s}{\partial x^2} = -\frac{p_o}{q}\Re \sum (n-1)(ia_n + b_n)(x+iy)^{n-2} = -\frac{\partial^2 A_s}{\partial y^2} \qquad \longrightarrow \qquad \text{ok.}$$

Here b_n refer to regular multipoles ($n = 1$ dipole, $n = 2$ quadrupole etc.) and a_n to the corresponding *skew* multipoles, which we will not consider further for now. For the field components we obtain immediately

$$\begin{aligned} B_y &= -\tfrac{p_o}{q}(b_1 \quad + b_2 x + b_3(x^2 - y^2)\ldots) \\ B_x &= -\tfrac{p_o}{q}(\quad\quad + b_2 y + b_3 xy \ldots) \end{aligned} \tag{25}$$

The **cylindrical** geometry is much more complicated. Calculating $\nabla \times \vec{B}$ in cylindrical coordinates is simplified by the assumption of a piece-wise constant field, i.e. $B_s = 0$ and $\partial B_(x,y)/\partial s = 0$:

$$\nabla \times \vec{B} = \left(\frac{\partial B_x}{\partial y} - \frac{\partial B_y}{\partial x}\right) \cdot \hat{\vec{s}} \overset{!}{=} 0$$

Inserting eq.23 we obtain

$$\nabla \times \vec{B} = -\hat{\vec{s}} \cdot \left(\frac{\partial^2 A_s}{\partial x^2} + \frac{\partial^2 A_s}{\partial y^2} + \frac{h}{1+hx} \frac{\partial A_s}{\partial x} - \left(\frac{h}{1+hx} \right)^2 A_s \right) \stackrel{!}{=} 0 \tag{26}$$

A solution of this differential equation for A_s is given by [1]

$$\frac{q}{p_o} A_s = -\frac{b_1}{2} \frac{1+hx}{h} \tag{27}$$

Calculating the magnet field from eq.23 gives

$$B_y = -\frac{p_o}{q} b_1 \tag{28}$$

which is just dipole field as it exists in a sector magnet.

Considering combined function magnets (i.e. bending magnets with gradients as commonly used in accelerators), a similar multipole representation of the vector potential can be calculated, but is rather complicated (see p.362 in [4]).

Assuming a separate function lattice, we restrict ourselves to pure dipole fields so far. The simplified Hamiltonian we will find thus is not valid for combined function magnets. However in large rings, with strong gradients and low curvatures, the common but inconsistent procedure to treat the dipole component in cylindrical geometry and the higher multipoles in cartesian geometry, is good enough.

8 The simplified Hamiltonian

Assuming a separate function machine with pure sector magnets (cylindrical geometry) and cartesian quadrupoles and higher multipoles, we express the vector potential A_s by the sum of eqs.27 and 24 and introduce it into eq.21. Further, we consider only regular multipoles and show only the first terms of the multipole sum:

$$\tilde{H}_2 = -(1+hx) \left(1 + \delta - \frac{\tilde{p}_x^2 + \tilde{p}_y^2}{2\,(1+\delta)} \right) + \frac{b_1}{2} \frac{(1+hx)^2}{h} + \frac{b_2}{2}(x^2 - y^2) + \ldots \tag{29}$$

Here the multiplication of all terms with $(1+hx)$ was not executed for the multipole sum since we assumed all these multipoles as cartesian, i.e. $h = 0$.

Now the dipole field is *adjusted* to provide the curvature h, i.e. we set $h = b_1$, but note that these two quantities are not necessarily identical. Subtracting constant terms, which are irrelevant for the equations of motion from the Hamiltonian, we thus arrive at

$$\mathcal{H} = \tilde{H}_2 + \frac{1}{2} + \delta = (1+hx) \frac{\tilde{p}_x^2 + \tilde{p}_y^2}{2\,(1+\delta)} - b_1 x \delta + \frac{b_1^2}{2} x^2 + \frac{b_2}{2}(x^2 - y^2) + \ldots \tag{30}$$

In a last step, the curvature at the kinematic term is dropped, which is justified for large rings. Also including again the sextupole term, we finally arrive at the simplified Hamiltonian useful for transverse dynamics in large rings:

$$\mathcal{H} = \frac{\tilde{p}_x^2 + \tilde{p}_y^2}{2\,(1+\delta)} - b_1 x \delta + \frac{b_1^2}{2} x^2 + \frac{b_2}{2}(x^2 - y^2) + \frac{b_3}{3}(x^3 - xy^2) + \ldots \tag{31}$$

From this follow the equations of motion

$$x' = \frac{\partial \mathcal{H}}{\partial \tilde{p}_x} = \frac{\tilde{p}_x}{1+\delta} = \frac{p_x}{p}$$

which is just the divergence angle (for $x' \ll 1$) (same for y'), and

$$\tilde{p}_x' = -\frac{\partial \mathcal{H}}{\partial x} = b_1 \delta - (b_1^2 + b_2) x - b_3 (x^2 - y^2)$$

$$\tilde{p}_y' = -\frac{\partial \mathcal{H}}{\partial y} = b_2 y + 2 b_3 x y$$

References

[1] J.Bengtsson, The sextupole scheme for the Swiss Light Source, SLS-Note 9/97. http://slsbd.psi.ch/pub/slsnotes/sls0997.pdf

[2] E.D.Courant & H.S.Snyder, Theory of the alternating gradient synchrotron, Ann. Phys. 3,1(1958). http://ab-abp-rlc.web.cern.ch/ab-abp-rlc/AP-literature/Courant-Snyder-1958.pdf

[3] D.A.Edwards & M.J.Syphers, An Introduction to the physics of high energy accelerators, Wiley-VCH, 1992.

[4] E.Forrest, Beam Dynamics, Harwood Academic Publishers, Amsterdam 1998.

[5] J.R.Rees, Symplecticity in beam dynamics, SLAC-PUB 9939, 2003. http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-pub-9939.pdf

[6] R.D.Ruth, Single particle dynamics in circular accelerators, SLAC-PUB 4103, 1986. http://www.slac.stanford.edu/cgi-wrap/getdoc/slac-pub-4103.pdf