End Effects of Beam Transport Elements

R. Baartman

TRIUMF

July, 2001

Introduction

Since we now have excellent higher order codes like `MARYLIE` and `COSY-`∞, why do we need to know more about fringe fields?

- Keeps us from wasting time trying to do impossible things.
- Deeper understanding results in more efficiency in designing transport systems.
- Finding the real sources of end effects allows development of simple formulas for inclusion in first order codes to determine whether or not aberrations will be a problem.

We can understand beam optics by using $F = ma$, or more deeply by using Hamiltonians. Knowing the details of the forces gives us a secure feeling, but in the end gives us too much detail. Why too much? Because beam particles are nothing more than integrators of a certain type. Concentrating on the bare forces tricks us into thinking the motion depends upon all the details of those forces. It does not.

I want to show some simple results from using a Hamiltonian formalism to understand beam optics. In the process, I hope to kill a few long-standing myths such as

- Occupancy: “don’t allow beam to fill more than x% of aperture”
- Aspect Ratio: “don’t allow elements to be shorter than x times the aperture”.

Illustrative Example: The Solenoid

$\vec{F} = q\vec{v} \times \vec{B}$, so with $\vec{B}$ in the z-direction, $F_x = qv_y B$, $F_y = -qv_x B$. We all know the equations of motion are easily integrated to give the following (incorrect, non-symplectic) transfer matrix for the solenoid:

$$\begin{pmatrix} x_\mathsf{f} \\ x'_\mathsf{f} \\ y_\mathsf{f} \\ y'_\mathsf{f} \end{pmatrix} = \begin{pmatrix} 1 & \rho\sin\theta & 0 & \rho(1-\cos\theta) \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & -\rho(1-\cos\theta) & 1 & \rho\sin\theta \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} x_\mathsf{i} \\ x'_\mathsf{i} \\ y_\mathsf{i} \\ y'_\mathsf{i} \end{pmatrix}$$

($\rho \equiv (B\rho)/B$, and $\theta = L/\rho$ where L is the length.)

This is incorrect because we haven't included the end effects.

At the ends, the magnetic field has a radial component which can be considered as arising from $\nabla \cdot \vec{B} = 0$.

$$B_r = -\frac{r}{2}\frac{dB}{dz} \equiv -\frac{r}{2}B'$$

The Solenoid –cont'd

This yields an additional transverse force $F_x = qv_0 B' y/2$, $F_y = -qv_0 B' x/2$. This seems problematic to integrate because x, y and B' change simultaneously in the fringe field region. But in the hard-edge case, it is easily done. For entry,

$$\Delta x' = \frac{y}{2\rho}, \, \Delta y' = -\frac{x}{2\rho}$$

with opposite signs for exit. When applied to the matrix above, the result is a matrix (recall $\theta = L/\rho$)

$$\begin{pmatrix} \cos^2\frac{\theta}{2} & \rho\sin\theta & \frac{1}{2}\sin\theta & 2\rho\sin^2\frac{\theta}{2} \\ \frac{-1}{4\rho}\sin\theta & \cos^2\frac{\theta}{2} & \frac{-1}{2\rho}\sin^2\frac{\theta}{2} & \frac{1}{2}\sin\theta \\ \frac{-1}{2}\sin\theta & -2\rho\sin^2\frac{\theta}{2} & \cos^2\frac{\theta}{2} & \rho\sin\theta \\ \frac{1}{2\rho}\sin^2\frac{\theta}{2} & -\frac{1}{2}\sin\theta & \frac{-1}{4\rho}\sin\theta & \cos^2\frac{\theta}{2} \end{pmatrix}$$

which can be decomposed into a rotation by angle $\frac{\theta}{2}$ and a decoupled, stigmatic focusing element described by

$$\begin{pmatrix} \cos\frac{\theta}{2} & 2\rho\sin\frac{\theta}{2} & 0 & 0 \\ -\frac{1}{2\rho}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} & 0 & 0 \\ 0 & 0 & \cos\frac{\theta}{2} & 2\rho\sin\frac{\theta}{2} \\ 0 & 0 & -\frac{1}{2\rho}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$$

With this derivation, the student would be forgiven for the incorrect conclusion that the transfer matrix adopts such a simple form only in the hard-edged case.

Solenoid: Hamiltonian Derivation

The canonical momentum $\vec{P}$ is not the 'kinetic' momentum $\vec{p}$, but is instead $\vec{p} + q\vec{A}$, where $\vec{A}$ is the vector potential. So the Hamiltonian is $H = c\sqrt{m^2c^2 + |\vec{P} - q\vec{A}|^2}$. In the usual way, we note H is independent of t, and so is constant $H = E$, switch to z as independent variable by solving for P_z, and write the new Hamiltonian as $H = -P_z$.

For the solenoid, by symmetry, $\vec{A}$ has only a component in the azimuthal direction. To first order, $A_\theta = rB(z)/2$. So to first order (second degree in H),

$$H = \frac{1}{2}\left[\left(P_x + \frac{y}{2\rho}\right)^2 + \left(P_y - \frac{x}{2\rho}\right)^2\right]$$

(We have normalized by dividing by $p_0 = q(B\rho) = \sqrt{E^2/c^2 - m^2c^2}$, and B appears in $1/\rho = qB(z)/p_0 = B(z)/(B\rho)$.)

We note that only B appears. Though we tend to think of B' as the source of all our end effects, it doesn't even appear.

Solenoid: Hamiltonian Derivation – cont'd

The equations of motion are

$$
\begin{aligned}
x' = && \tfrac{\partial H}{\partial P_x} = && P_x + \frac{y}{2\rho} \\
P_x' = && -\tfrac{\partial H}{\partial x} = && \frac{1}{2\rho}\left(P_y - \frac{x}{2\rho}\right) \\
y' = && \tfrac{\partial H}{\partial P_y} = && P_y - \frac{x}{2\rho} \\
P_y' = && -\tfrac{\partial H}{\partial y} = && -\frac{1}{2\rho}\left(P_x + \frac{y}{2\rho}\right)
\end{aligned}
$$

These are easily solved for the hard-edge case, and without any additional physics yield the matrix we found previously. We had thought that matrix depended upon a composition of entrance field, body field, exit fringe field.

For tidiness, we can also write the equations of motion as a matrix: $\mathbf{x}' = \mathbf{F}\mathbf{x}$ where $\mathbf{x} = (x, P_x, y, P_y)^T$, and

$$
\mathbf{F} = \begin{pmatrix}
0 & 1 & \frac{1}{2\rho} & 0 \\
-\left(\frac{1}{2\rho}\right)^2 & 0 & 0 & \frac{1}{2\rho} \\
-\frac{1}{2\rho} & 0 & 0 & 1 \\
0 & -\frac{1}{2\rho} & -\left(\frac{1}{2\rho}\right)^2 & 0
\end{pmatrix}
$$

$\mathbf{F}$ is simply the transfer matrix of an infinitesimal distance dz, with the identity matrix subtracted, and divided by dz.

Solenoid: Hamiltonian Derivation – cont'd

Now do the following exercise. Take the hard-edged solenoid transfer matrix derived previously, in the limit of infinitesimal length dz,

$$\begin{pmatrix} 1 & dz & \frac{dz}{2\rho} & 0 \\ \frac{-dz}{4\rho^2} & 1 & 0 & \frac{dz}{2\rho} \\ \frac{-dz}{2\rho} & 0 & 1 & dz \\ 0 & -\frac{dz}{2\rho} & \frac{-dz}{4\rho^2} & 1 \end{pmatrix}$$

subtract I, divide by dz. **Remarkably, this is identical to the infinitesimal transfer matrix above.** There simply are no 'end effects'.

Closer investigation shows that what we thought of as the end effects are replaced in this picture by transformations out of and back into canonical coordinates $p_x = P_x + y/(2\rho)$, $p_y = P_y - x/(2\rho)$.

Can this trick apply to end effects of other elements as well? Surprisingly, the answer is **YES**.

Quadrupole End Effects: Magnetic

In quadrupoles, there are no end effects in first order: they only appear in third order. The Hamiltonian is as before, but the vector potential is

$$A_x = -\frac{k'}{4}xy^2, \quad A_y = \frac{k'}{4}x^2y$$
$$A_z = -\frac{k}{2}(x^2 - y^2) + \frac{k''}{48}(x^4 - y^4),$$

Note that this is for a gauge which gives the simplest $\vec{A}$ while still satisfying $\nabla \times \nabla \times \vec{A} = 0$; it is not the Coulomb gauge.

The Hamiltonian can be written:

$$\begin{aligned} H &= \frac{1}{2}\left[k\,(x^2 - y^2) - \frac{k''}{24}(x^4 - y^4) + P_x^2 + P_y^2\right] + \\ &+ \frac{k'xy}{4}(yP_x - xP_y) + \frac{1}{8}(P_x^2 + P_y^2)^2. \end{aligned}$$

Now it looks like we really do have end effects which depend on the details of the fringe field, since both first and second derivatives of $k(z)$ appear, even though we are already using canonical momenta.

Magnetic Quad – cont'd

Surprisingly, it turns out to be possible to find a canonical transformation which eliminates these derivatives of $k(z)$. In our case, we wish to retain terms to 4$^{\text{th}}$ degree in the Hamiltonian (3$^{\text{rd}}$ order on force). The transformation $(x, P_x, y, P_y) \to (X, P_X, Y, P_Y)$ with generating function

$$G(x, P_X, y, P_Y) = xP_X + yP_Y + \frac{k'}{48}(x^4 - y^4) + \\ -\frac{k}{12}\left[(x^3 + 3xy^2)P_X - (3x^2y + y^3)P_Y\right]$$

accomplishes exactly this. To 3$^{\text{rd}}$ order it yields transformation

$$\begin{aligned} x &= X + \frac{k}{12}(X^3 + 3XY^2) \\ P_x &= P_X - \frac{k}{4}\left[(X^2 + Y^2)P_X - 2XYP_Y\right] + \frac{k'}{12}X^3, \end{aligned}$$

The y-transformation is obtained by replacing x, P_x, X, P_X with y, P_y, Y, P_Y and k with $-k$. Note that outside the quadrupole, the transformed coordinates are the same as the original ones.

The transformed Hamiltonian is

$$\begin{aligned} H^* &= \frac{k}{2}(X^2 - Y^2) + \frac{1}{2}(P_X^2 + P_Y^2) + \\ &+ \frac{1}{8}(P_X^2 + P_Y^2)^2 - \frac{k}{4}(X^2 + Y^2)(P_X^2 - P_Y^2) \\ &+ \frac{k^2}{12}(X^4 + Y^4) + \frac{k^2}{2}X^2Y^2. \end{aligned}$$

Magnetic Quad – cont'd

We can identify the terms: the first two are the usual linear ones; the third term is not related to the quadrupole field; the 4^{th} term is also small and arises because a particle going through the quadrupole at an angle is inside the quad for slightly longer than one which remains on axis.

The dominating higher order terms are the last two terms in the Hamiltonian. Since there are no derivatives of k, we can directly write down the aberrations in the thin-lens limit:

$$\Delta P_x \;\; = -\int \tfrac{\partial H^*}{\partial y}\, ds \approx \;\; \frac{-1}{f^2 L}\left(\frac{1}{3}x^3 + xy^2\right),$$

with a similar expression for ΔP_y. L and f are the quadrupole's effective length and focal length. The fractional focal error is found by dividing by the linear part $\Delta_0 P_x = -x/f$:

$$\frac{\Delta f_x}{f} = \frac{1}{fL}\left(\frac{1}{3}x^2 + y^2\right)$$

for x, and similarly for y.

It is important to realize that since there are no derivatives of k in the Hamiltonian, there is no essential difference between the thin-lens limit and a real quadrupole. A real quadrupole can be thought of as simply a string of many thin quadrupoles with strengths varying as $k(z)$. This is just as we found for first-order end effects in the solenoid.

Does the transforming away of end effects depend on some fluky property of magnetic quads? Well, let's try electric quads.

Electrostatic Quads

Here the end effects show up quite differently in the Hamiltonian:

$$
\begin{aligned}
H &= \frac{1}{2}\left[V\left(x^2-y^2\right)-\frac{V''}{12}\left(x^4-y^4\right)+p_x^2+p_y^2\right] \\
&\quad +\frac{1}{8}\left[V\left(x^2-y^2\right)+p_x^2+p_y^2\right]^2 .
\end{aligned}
$$

This follows from using Laplace's equation to expand the quad potential $V(z)(x^2-y^2)$ to 4^{th} order, thereby bringing in the dependence on V''.

We try the generating function:

$$
G(x,P_X,y,P_Y)=xP_X+yP_Y+\frac{V'}{24}(x^4-y^4)+ \\
-\frac{V}{6}(x^3P_X-y^3P_Y).
$$

which gives transformation

$$
\begin{aligned}
x &= X+\frac{V}{6}X^3 \\
p_x &= P_X-\frac{V}{2}X^2P_X+\frac{V'}{6}X^3.
\end{aligned}
$$

The new Hamiltonian is

$$
\begin{aligned}
H^* &= \frac{V}{2}(X^2-Y^2)+\frac{1}{2}(P_X^2+P_Y^2)+ \\
&+ \frac{1}{8}(P_X^2+P_Y^2)^2-\frac{V}{4}(X^2+Y^2)(P_X^2-P_Y^2) \\
&+ \frac{7V^2}{24}(X^4+Y^4)-\frac{V^2}{4}X^2Y^2.
\end{aligned}
$$

This is the same as for the magnetic case except for the coefficients of the last 2 terms. The main aberration is given by

$$
\frac{\Delta f_x}{f}=\frac{1}{fL}\left(\frac{7}{6}x^2-\frac{1}{2}y^2\right)
$$

for x, and similarly for y.

Discussion – Quads

The new equations of motion really are correct to third order (I've checked against COSY-∞). But they are also way more convenient computationally. Only $k(z)$ needs to be known. One can use survey data for a real quad, without the complication of calculating numerical derivatives.

Moreover, for most purposes, it is sufficient to use the simple first order transfer matrix, augmented by third order kicks applied at the entrance and exit. As with the solenoid, these kicks are simply those given by the canonical transformation. This procedure is much simpler than calculating 'fringe field integrals'

Lastly, we can see that the third order aberrations do not depend upon aperture, aperture occupancy, length-to-aperture aspect ratio, or any of these kinds of considerations. They only depend upon strength and length. In any given application with f imposed, these aberrations are reduced by increasing the quad's effective length. There is no other choice.

Dipoles: Sector Magnet

Now let's try a magnetic dipole. A complicating feature here is that we have to use a curvilinear coordinate system. For emphasis of this fact, for the independent variable, we replace the Cartesian z with s, the distance along the reference trajectory. We are interested in the lowest-order effects of the end fields. Unlike quads, where this is third order, in dipoles the lowest order effect is second order (third degree in H). The Hamiltonian is

$$H = -(1+hx)\left[1 + \frac{qA_s}{p_0} - \left(P_x - \frac{qA_x}{p_0}\right)^2 - \left(P_y - \frac{qA_y}{p_0}\right)^2\right]$$

where $h = h(s)$ is the local curvature on the reference trajectory.

The vector potential is

$$\begin{aligned}(1+hx)A_x &= -\frac{1}{2}B'y^2 \\ A_y &= 0 \\ (1+hx)A_s &= -\left(1+\frac{hx}{2}\right)Bx\end{aligned}$$

As anticipated, insertion into H yields a condition $qB/p_0 = h$ to make first degree terms disappear, so that $x = y = 0$ is indeed the reference trajectory. The final H is

$$H = \frac{1}{2}\left[h^2 x^2 + P_x^2 + P_y^2 + hx(P_x^2 + P_y^2) + h'\, y^2 P_x\right]$$

Note in particular the appearance of the problematic derivative of h.

Magnetic Dipole – cont'd

This is in fact the easiest example. The generating function to get rid of h' almost suggests itself:

$$G(x, P_X, y, P_Y) = xP_X + yP_Y - \frac{h}{2}y^2\, P_X$$

The transformation is

$$\begin{aligned} x &= X + h\,Y^2/2 \\ y &= Y \\ P_x &= P_X \\ P_y &= P_Y - h\,yP_X, \end{aligned}$$

and the transformed H is

$$\begin{aligned} H^* &= \frac{h^2}{2}X^2 + \frac{1}{2}(P_X^2 + P_Y^2) + \frac{h}{2}X(P_X^2 + P_Y^2) \\ & -h\,Y P_X P_Y + \frac{h^3}{2}XY^2 \end{aligned}$$

Again, the end-fields have been magically transformed away: the last 2 terms have taken the place of the term containing h'.

Magnetic Dipole – cont'd

Let us try to get a deeper understanding of this transformation.

The important last term in the new Hamiltonian gives for the thin-lens approximation,

$$\Delta p_y = \quad -\int \frac{\partial H^*}{\partial y}\, ds \quad \approx -h^3 Lxy.$$

Now look at the old Hamiltonian:

$$\Delta p_y = \quad -\int \frac{\partial H}{\partial y}\, ds \quad \approx -y \int h' P_x \, ds = y \int h P_x' \, ds,$$

where we have integrated by parts. We use the first-order expression $P_x' = -h^2 x$, to get again $\Delta p_y \approx -h^3 Lxy$.

The symplectic counterpart

$$\Delta p_x = \quad -\int \frac{\partial H^*}{\partial x}\, ds \quad \approx -h^3 L y^2/2.$$

is not as easily found from the original H. We note that

$$\Delta p_x = \quad -\int \frac{\partial H}{\partial x}\, ds \quad \approx -\int h^2 x \, ds,$$

which is the first order effect, but x is shifted because $x' = \frac{\partial H}{\partial P_x}$ and this has an additional term $h' y^2/2$, giving $\Delta x = h y^2/2$ (seem familiar?). Combined, we recover $\Delta p_x \approx -h^3 L y^2/2$.

This 'traditional' kind of approach works for quads as well.

Electrostatic Bend with Field Index

(Use Mathematica...)

Electrostatic Dipole – cont'd

Reminder: $h = h(z)$

$$
\begin{aligned}
H(X, P_X, Y, P_Y; z) &= \frac{P_X^2}{2} + \frac{P_Y^2}{2} \\
&+ \left(-\frac{kh}{2} + h^2\right) X^2 + \frac{kh}{2} Y^2 \\
&+ \left(\frac{4}{3}h^3 - \frac{7}{6}kh^2 + \frac{1}{3}k^2 h\right) X^3 \\
&+ \left(\frac{1}{2}kh^2 - k^2 h\right) XY^2 \\
&+ h\, XP_Y^2
\end{aligned}
$$

$h = 1/\rho$, let $c \equiv k/h$. Nonlinear part of $P_X' = -\partial H/\partial X$ is integrated through the bend length $L = \rho\theta$ to give

$$
\begin{aligned}
\Delta p_x &= \frac{\theta}{\rho^2}\left[\left(-4 + \frac{7}{2}c - c^2\right) x^2 + \left(-\frac{1}{2}c + c^2\right) y^2\right] \\
\Delta p_y &= \frac{\theta}{\rho^2}\left(-c + 2c^2\right) xy
\end{aligned}
$$

Electrostatic Dipole – cont'd

We test in COSY-∞, using parameters such that

$$\theta/\rho^2 = 10^{-4}.$$

Cylindrical bend $c = 0$:

$$\Delta p_x = -\frac{4\theta}{\rho^2}\, x^2, \text{ and } \Delta p_y = 0$$

x	p_x	y	p_y	
-0.199665E-04	-0.400046E-03	0.00000	0.00000	2000
0.00000	0.00000	-0.324334E-07	0.182194E-06	1010
0.288633E-08	0.911004E-07	0.00000	0.00000	0020

Spherical bend $c = 1$:

$$\Delta p_x = \frac{\theta}{\rho^2}\left(-\frac{3}{2}x^2 + \frac{1}{2}y^2\right), \text{ and } \Delta p_y = \frac{\theta}{\rho^2}\, xy$$

x	p_x	y	p_y	
-0.748504E-05	-0.150067E-03	0.00000	0.00000	2000
0.00000	0.00000	-0.499447E-05	0.101086E-03	1010
0.250319E-05	0.505384E-04	0.00000	0.00000	0020

Higher Order End Effects

Though the lowest-order end-fields can be transformed away (i.e., do not depend upon aperture), this is not true of the next-to-lowest order end fields. They do depend upon aperture [A. Dragt]. This can be checked with COSY-∞. In terms of canonical transformations, it means that no transformation exists which eliminates the derivatives of the strength function. Our results can be summarized in the table

Element \ Order	1	2	3	4	5
Solenoid	1	0	a^{-2}	0	a^{-4}
Dipole	-	1	a^{-1}	a^{-2}	a^{-3}
Quadrupole	-	0	1	0	a^{-2}

There seems to be a general rule that lowest order end effects are integrable, and higher orders are aperture-dependent as $\sim$ $(x/\text{aperture})^n$, where n=the appropriate power, and so diverge in the hard-edge limit.

Summary

- Lowest-order end effects do not depend upon aperture: they can be transformed away for all elements checked so far. Rigorous proof anyone?

- These effects are often adequately approximated by the extremely simple formulas obtained in the hard-edge approximation.

- Next-to-lowest-order end effects do depend on aperture and therefore diverge in the hard-edge limit.