

Low-Energy Beam Transport Design

R. Baartman

TRIUMF, Vancouver, Canada

Abstract

There is a trick to designing beam transport systems. In most cases, we are more interested in assuring that higher-order aberrations are small than we are in finding accurate values for all the transfer map elements. I show a simple way to do this, using an envelope technique and modifying the usual optimization parameter.

Keywords

Solenoid; einzel lens; electrostatic; quadrupole; aberration.

1 Introduction

In most cases, the goal of a design is to transport a well-matched particle beam with no losses and no decrease in quality. ‘Quality’ here, of course, means emittance, and maintaining it means that higher-order effects are negligible. As ‘higher-order’ is difficult to calculate, requiring seemingly much more input, such as Enge coefficients or field maps, the tendency is to design ‘linear’ first and then, if at all, run it through a time-consuming higher-order calculation using a multiparticle code and field maps or Enge coefficients. At this point, the elements have usually already been designed and there is resistance to changing them.

I intend to show that there is a better way. ‘Better’ means not only that it is quicker to calculate but also that, being very quick, it is a simple matter to optimize acceptance, usually bringing the emittance growths down to the level of negligible. The technique uses simple formulae that find the emittance growths at each element and simply add them in quadrature to that arising from mismatch, minimizing this sum while varying gross features, such as quadrupole lengths and solenoid apertures.

A major issue with the usual approach is that field maps are found, and perhaps summarized, by fitting their form in the ‘fringe’ region to a high-order polynomial only after the elements have already been designed. They are not designed as part of the optimization process. Very often, they are designed using folkloric rules-of-thumb regarding aperture occupation or field flatness, in place of a real understanding of the origin of non-linear forces. This has the effect of perpetuating such rules.

The standard beam transport elements for low energy are: dipoles, both electrostatic and magnetic; quadrupoles, also of both types, and solenoids and einzel lenses.

History: the mathematical formalism for describing charged-particle beam transport is only 60 years old. First came the understanding that particle optics of individual particles could be described by transfer matrices [1], then applied to ensembles of particles with the σ -matrix formalism developed by Brown [2]. The first useful codes were TRANSPORT [3], TRACE3D [4],

and GIOS [5], whose first versions were written in the mid-1960s. These used highly idealized ‘hard-edged’ elements and so were often not sufficiently accurate. The result was that beamlines designed in this way were expected to work only approximately and allowed to depart from ‘theoretical’ settings, sometimes by wide margins. There developed a prejudice that the σ -matrix or ‘envelope’ technique was not accurate. Further, it was thought that this technique could never handle space charge very well. Both these turned out to be incorrect but the prejudice remains.

When computing power grew by many orders of magnitude, designers tended towards codes that integrated motions of millions of individual particles. So many particles, that even with today’s computing power, it can take minutes or hours of CPU time on a personal computer to simulate beam transport. In making this step, from simple matrix-based codes that calculated ‘beam envelopes’, to codes that integrate millions of particles through stored magnetic and electric fields, the ‘baby’ (envelope technique) was thrown out with the ‘bathwater’ (hard-edged elements with no fringe-field effects).

The thinking behind this is that if we calculate through well-developed real fields, we automatically take all relevant effects (fringe fields, space charge, time-varying fields) into account. This is true, but it comes at a cost, namely, a loss of any feeling for basic effects, such as changing aperture or shaping fringe fields.

A second, well-regarded technique is to calculate transfer maps to high order, using, for example, differential algebraic tools, as in COSY- ∞ [6]. Again, this is an extremely helpful tool but its use comes at the cost of loss of ‘feel’: for any given system, one can find accurate maps and even minimize the worst higher-order effects but the reverse problem, of knowing how small the aberrations need to be and then finding an optimum transport system, is not tractable.

So both techniques (integrating through stored fields and higher-order transfer maps) give good accurate and trustworthy results ‘after the fact’, i.e., after a beamline design has already been completed, perhaps even built, but are not used in guiding a beamline design from scratch.

The key is to realize that, except in special cases like spectrometers, where higher-order aberrations are inevitable, we don’t actually ‘need’ exact higher-order maps, we only need to be assured that their effects are ‘sufficiently small’. As I will show, the minimization techniques to make this assurance are surprisingly simple.

Let us first determine what is meant by ‘sufficiently small’.

The key parameter is, of course, the emittance. A higher-order distortion is shown in Fig. 1. This may be a cubic (as is common for focusing devices where the fields are odd functions of transverse co-ordinates) or parabolic in the case of the dominant error from dipoles. (Early versions of TRANSPORT extended it to only second-order and so missed this effect entirely. More on this later.)

The order of the dominant effect is not particularly important for our purposes. Due to the higher orders the emittance ellipse starts to deform, Fig. 1 red curve. In case of a misalignment of the beam this effect can lead to an additional outspreading and deformation of the ellipse, which is called filamentation [Edwards, "An introduction to the physics of high energy accelerators", chap.7 "Emittance Preservation"]. The order of the dominant effect is not particularly important for our purposes; only the location of the ‘edge’ of the desired emittance matters. If desired, the

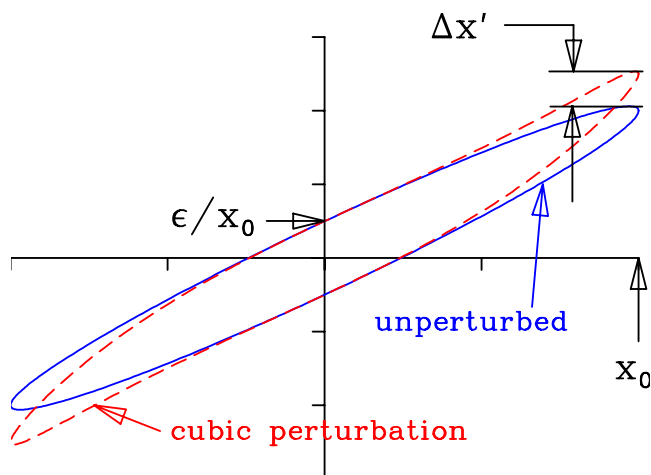


Fig. 1: Phase-space ellipse with cubic distortion

expected misalignment can be included in the ‘size’. A high-quality beam may have negligible emittance growth if it is aligned to the beamline symmetry axis, but not so when misaligned.

Let us call the integrated higher-order error $\Delta x'$. Figure 1 makes clear that this x' is to be compared not with the beam divergence \hat{x}' , but with the local divergence ϵ/\hat{x} , which may be significantly smaller: the two versions agree only at a waist.

The effective fractional growth in emittance is

$$\frac{\Delta\epsilon}{\epsilon} \sim \frac{\Delta x' \hat{x}}{\epsilon}, \quad (1)$$

and we want this to be small compared with one. How small depends on context: typically 1% is sufficient but clearly if there are thousands of elements, we would want this to be much smaller than 1%. Most often, the emittance growth is dominated by one or two focusing elements; the ones that create the final focus.

In any beamline design, there are many constraints besides minimizing emittance growth. The main constraint is to achieve a desired match or ‘Twiss’ parameters at the final focus. This is often done by calculating the mismatch factor, which also constitutes an emittance growth. Clearly, it matters not whether growth comes from mismatch or from aberrations so these two are simply added in quadrature and the sum is minimized. This is the whole ‘trick’.

But it remains to find $\Delta x'$. This turns out to be surprisingly simple: for any of the standard beamline elements, simple formulae can be found.

2 Transport elements

Solenoid third-order aberration depends on aperture, or more specifically, fringe-field ‘hardness’. Quadrupole third-order aberrations do not depend on fringe field, but fifth-order aberrations do. Dipole bender second-order aberrations do not depend on fringe-field hardness, but third-order aberrations do.

2.1 Solenoid

A solenoid's focal strength scales as the square of the ratio of magnetic field to particle momentum, while in quadrupoles the scaling is linear, so solenoids are much weaker focusing elements than quadrupoles. For this reason, they are most commonly used for low-energy electron transport. They have the great convenience that they focus in both transverse directions.

It is useful to know that the magnetic field from a solenoid is completely given by the following expansions:

$$B_z(r, z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!^2} \left(\frac{r}{2}\right)^{2n} \left(\frac{d}{dz}\right)^{2n} B_z(0, z), \quad (2)$$

$$B_r(r, z) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(n+1)!n!} \left(\frac{r}{2}\right)^{2n+1} \left(\frac{d}{dz}\right)^{2n+1} B_z(0, z). \quad (3)$$

These follow directly from Maxwell's equations plus the symmetry and establish that the fields can be completely derived in all of space, knowing only the on-axis field function.

For small r , we have

$$B_z = B_0 - \frac{r^2}{4} B'', \quad (4)$$

$$B_r = -\frac{r}{2} B' + \frac{r^3}{16} B''', \quad (5)$$

where $B_0(z)$ is the on-axis field $B_z(0, z)$ and B' etc. are derivatives of $B_0(z)$.

To the same order,

$$B_z^2 = B_0^2 - \frac{r^2}{2} B_0 B''. \quad (6)$$

The equation of motion through the solenoid has the radial part:

$$r'' + K B_z^2 r = 0, \quad (7)$$

where K is a constant containing the magnetic rigidity: $K = 1/(2B\rho)^2$.

Expanding to cubic force order gives

$$r'' + K \left(B^2 r - \frac{B B''}{2} r^3 \right) = 0. \quad (8)$$

The non-linear term gives an r' error, denoted $\Delta r'$,

$$\begin{aligned} \Delta r' &= \frac{K}{2} \int r^3 B B'' dz = \frac{K}{2} r^3 \int B B'' dz \text{ (thin-lens approximation)} \\ &= -\frac{K}{2} r^3 \int B'^2 dz. \end{aligned} \quad (9)$$

Thus we see that optimizing a solenoid by minimizing the off-axis focal deviation is exactly the same as minimizing the mean-squared value of B' . Solenoids' spherical aberrations

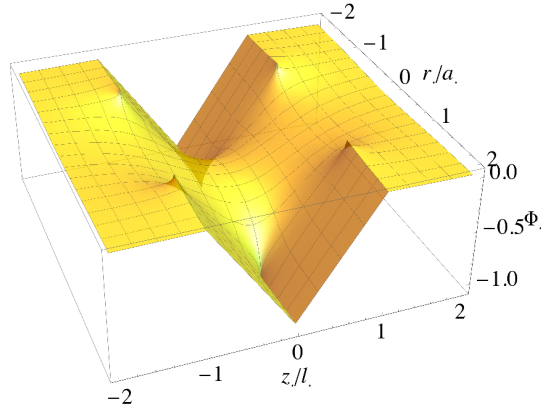


Fig. 2: Potential function plot (r, z) of three-aperture einzel lens, where apertures are at $z = -1, 0, 1$ in units of aperture radius.

diverge in the hard-edge limit. This puts to rest the notion that solenoids are improved by flattening their fields. In fact, the opposite is the case: the softer the edges, the smaller the aberrations. For a long solenoid, it is beneficial to have more turns at the centre than at the entrance and exit. Again, this is contrary to common practice. An example is the Rutherford front end test stand [7]. If they had not flattened their solenoid fields, and had put extra turns near the centre rather than near the ends, the third-order aberration would have been roughly only half as large. See also Biswas [8] on this point.

Clearly, the integral (Eq. (9)) depends on the form of the fringe field. However, if we can characterize it as having an effective width w , then

$$\Delta r' \sim \frac{r^3}{fwL} = \frac{r^3}{4w\rho^2}, \quad (10)$$

as the focal length for solenoids is given by

$$\frac{1}{f} = \frac{\int B^2 dz}{(2B\rho)^2} = \frac{L}{4\rho^2}.$$

Usually, if no special shaping is done and there is a uniform number of turns per unit length, $w \sim 3a$, where a is aperture radius.

Roughly speaking,

$$\frac{\Delta\epsilon}{\epsilon} = \frac{\Delta r'}{\epsilon/\hat{r}} \sim \frac{\hat{r}^4}{\epsilon fwL}. \quad (11)$$

2.2 Einzel lens

As with solenoids, the axial symmetry of einzel or aperture lenses means that the field in all space can be determined if the on-axis field is known. In fact, the equations for the electric field are identical to those for the magnetic field in a solenoid (Eqs. 2 and 3). Integrating the latter

equation as the radial force, in the thin-lens approximation, results in zero focal effect and zero aberrations. It is only when the particle's changing speed is taken into account that there is a net effect.

An einzel lens can be constructed from as few as three apertures; a potential plot in $\Phi(r, z)$ is shown in Fig. 2. Near the axis, the potential is parabolic, but near the aperture edge, it is clearly not. This can be investigated approximately by considering one aperture at a time. The grounded conducting circular aperture separating two regions that have asymptotically constant electric fields on either side is a textbook boundary-value problem, for example by Jackson [9]. From this we find that in the plane of the aperture, the potential is actually elliptical, given by

$$\Phi(r, 0) = \frac{E_1}{\pi} \sqrt{a^2 - r^2}, \quad (12)$$

where E_1 is the electric field on the non-ground side and a is the aperture radius. Thus, the radial electric field diverges at the aperture edge; this, in itself, suggests that the beam should not be allowed to fill the aperture. A quadrupole's focusing force, for example, is not singular in this way.

One can find the focal length of the single aperture by using Eq. 5 as $E_r = -rE'/2$ and integrating in the thin-lens limit:

$$\Delta r' = \int r'' dz = \int \frac{1}{v^2} \frac{d^2 r}{dt^2} dz = \int \frac{q}{mv^2} E_r dz \approx -\frac{r}{4V_0} E_1, \quad (13)$$

where $qV_0 = mv^2/2$ and V_0 is the voltage through which the beam particle of charge q has been accelerated. This a focal length f :

$$\frac{1}{f} = \frac{E_1}{4V_0}. \quad (14)$$

Clearly, this only applies in the non-relativistic regime: einzel lenses are not used at relativistic energies.

Two such apertures, separated by a distance l , the first kept at the source potential V_0 and the second at $V_1 + V_0$, create the field $E_1 = V_1/l$. This constitutes an accelerator column and the focal power is known:

$$\frac{1}{f} \Big|_{\text{acc.column}} = \frac{3\kappa^2(2 + \kappa)}{8l(1 + \kappa)^2}, \quad (15)$$

where

$$\kappa = \sqrt{\frac{V_1 + V_0}{V_0}} - 1$$

is the change in momentum relative to initial momentum. This can be derived from three transfer matrices: entrance aperture, with f given by Eq. 14 ($f = 4V_0l/V_1$); body of column where $E'_z = 0$; and exit aperture, with $f = -4V_1l/V_0$ [10].

We create an einzel lens by placing two acceleration columns back-to-back, bringing beam energy back to source energy. The focal length of this is found to be

$$\frac{1}{f} \Big|_{\text{einzel}} = \frac{3\kappa^2(4 - \kappa^2)}{8l(1 + \kappa)}. \quad (16)$$

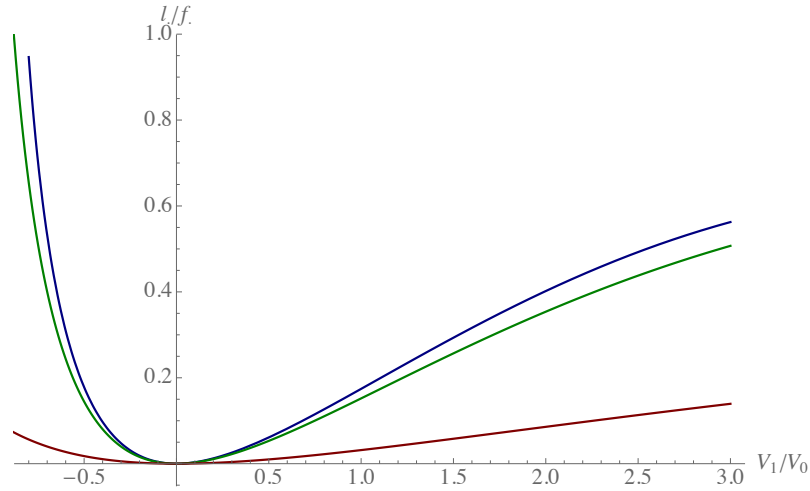


Fig. 3: Focal power in units of $1/l$ for einzel lenses. From top down: blue, infinitesimal aperture; green, $a = l/10$; red, $a = l$, as drawn in Fig. 2.

The plot is the blue curve shown in Fig. 3.

Unfortunately, this expression is only accurate for cases where $l \gg a$, while in common practice, $l/a \sim 4$ or less. The reason is that the superposition of apertures is not valid for small aspect ratios. This is clear from Fig. 2. The three apertures are not actually equipotential surfaces, and the on-axis potential does not reach $V_1 + V_0$ at the lens centre.

The green and red curves of Fig. 3 were calculated using the CEA element of COSY- ∞ [6]. This code is an invaluable tool for calculating higher-order effects in general, and in particular for realistic einzel lenses, where $a \sim l$. COSY- ∞ uses the single-aperture on-axis potential as given by Jackson [9],

$$\Phi(0, z) = \frac{E_1 a}{\pi} \left[1 - \frac{z}{a} \cot^{-1} \left(\frac{z}{a} \right) \right], \quad (17)$$

and adds apertures together to form the lens.

In spite of there being no closed-form expression for focal length f , we can still write the scaling for higher-order effects, in terms of f . For comparison with other lenses, we write in terms of the total length $L = 2l$. From COSY, we find

$$\Delta r' \approx \frac{r^3}{\eta f a L}, \quad (18)$$

where η varies according to aspect ratio: for $L \gg a$, $\eta = 3/2$, but for example for $L = a$, $\eta = 1/2$. Unsurprisingly, the aberrations scale similarly to solenoids as $(faL)^{-1}$, but for einzel lenses the proportionality factor is two to six times worse.

2.3 Quadrupole

Like solenoids, the lowest-order aberration is cubic or third-order. But unlike solenoids, this aberration cannot be reduced by shaping the field strength function, i.e. softening the edges. It turns out that the third-order aberration is independent of fringe-field shape or extent. (This is not, however, true of fifth and higher order.)

The electrostatic quadrupole potential field

$$V(x, y) = \frac{k}{2}(x^2 - y^2) \quad (19)$$

is a solution to Laplace's equation, but only if the quadrupole is infinitely long ($k = \text{constant}$). For finite quadrupoles, we use the expansion

$$V(x, y, z) = \frac{k}{2}(x^2 - y^2) - \frac{k''}{24}(x^4 - y^4) + \frac{k''''}{720}(x^6 - y^6) - \dots \quad (20)$$

The quartic term gives a cubic force term, which leads to the following focusing error,

$$\Delta x' = \frac{-1}{f^2 L_Q} \left(\frac{7}{6} x^3 - \frac{1}{2} x y^2 \right), \quad (21)$$

where L_Q is the quad length and f the focal length. It is important to note that this is independent of aperture size or fringe-field hardness: indeed, the aberration is not affected by changing the fringe-field shape.

For the Hamiltonian technique used to derive these, see Ref. [11]. I further refine electrostatic for the relativistic case [12]:

$$\Delta x' = \frac{-1}{f^2 L_Q} \left(\frac{7 - 3\beta^2}{6} x^3 - \frac{1 - \beta^2}{2} x y^2 \right). \quad (22)$$

The formula for magnetic quadrupoles is similar:

$$\Delta x' = \frac{-1}{f^2 L_Q} \left(\frac{1}{3} x^3 + x y^2 \right), \quad (23)$$

So electrostatic quadrupoles are not really worse than magnetic quadrupoles; they might actually be better, e.g. for ribbon beams.

Roughly speaking,

$$\frac{\Delta\epsilon}{\epsilon} = \frac{\Delta x'}{\epsilon/\hat{x}} \sim \frac{\hat{x}^4}{\epsilon f^2 L_Q}. \quad (24)$$

This formula has an obvious consequence. In any transport system where optics strengths and beam sizes are given, the only way to reduce emittance growth due to aberrations is to lengthen the 'worst-offending' quadrupoles.

In a periodic transport channel with β -function β_T , $x^2 = \epsilon\beta_T \sim \epsilon L_c$ and $f \sim L_c$, so somewhat surprisingly the optics length scale cancels and

$$\frac{\Delta\epsilon}{\epsilon} \sim \frac{\epsilon}{L_Q}. \quad (25)$$

Even if the emittance is as large as $100 \mu\text{m}$, this gives the easily-satisfied constraint $L_Q \gg 0.1 \text{ mm}$. (Commonly, $100 \mu\text{m}$ is quoted as $100\pi \text{ mm-mrad}$. I use the convention that emittance is phase-space area divided by π [13].)

However, this is for a regular FODO transport case. Matching sections often have much tighter constraints. An example is when matching to an RF quadrupole. Since the RF quadrupole requires a very small beam, the matching system must transform a relatively large beam with small divergence to a small beam with large divergence. The ‘worst-offending’ quadrupole in these situations is always the next-to-last quadrupole where the beam is much larger in the defocusing plane of the last quadrupole than in the focusing plane. This forces the last two quadrupoles close together and requires a very short last quadrupole, but a long next-to-last quadrupole.

2.3.1 Quadrupole match example

We apply this to the example shown in Fig. 4. The RF quadrupole requires matched $\beta_T = 6.125 \text{ cm}$ and the beam comes from a FODO section, where $\beta_T = 84.6 \text{ cm}$. With final quadrupoles of effective length $L = 8.4 \text{ cm}$, we get $\hat{x} = 1.8 \text{ cm}$, $f = 9.7 \text{ cm}$, with desired acceptance $\epsilon = 0.005 \text{ cm}$.¹ We get

$$\frac{\Delta\epsilon}{\epsilon} \sim \frac{1.8^4}{0.005 \times 9.7^2 \times 8.4} = 2.6. \quad (26)$$

To improve the emittance growth, we shorten the final quadrupole in order to decrease the size of the beam in the next-to-final quadrupole. In this case (Fig. 5), the beam size is only one-half as large, so the result is 1/16 or

$$\frac{\Delta\epsilon}{\epsilon} \sim 0.17. \quad (27)$$

These calculations were verified with COSY- ∞ [6] and gave $\Delta\epsilon/\epsilon \approx 0.1$. Final phase spaces are shown in Fig. 6.

It is important to understand that codes like TRANSPORT and TRACE3D can result in aberrations as large as that of the blue curve in Fig. 6, without any hint that there is a problem.

2.4 Dipole bender

Dipole magnets or electrostatic benders are similar to quadrupoles in that their lowest-order aberrations are insensitive to field falloff shaping and extent. The difference is that the lowest order is quadratic rather than cubic force. See my Snowmass talk [14] for derivations.

For the sector magnet the non-linear kicks are as follows ($L = \rho\theta$, the trajectory length):

$$\Delta x' = -\frac{L}{2\rho^3}y^2 \text{ and } \Delta y' = -\frac{L}{\rho^3}xy. \quad (28)$$

Often for dipoles the length L is not small compared with the focal length, so the thin-lens approximation used in estimating the aberration is not very good. In that case the calculation can simply split the dipole once or twice.

¹Commonly: ‘ $50\pi \text{ mm-mrad}$ ’.

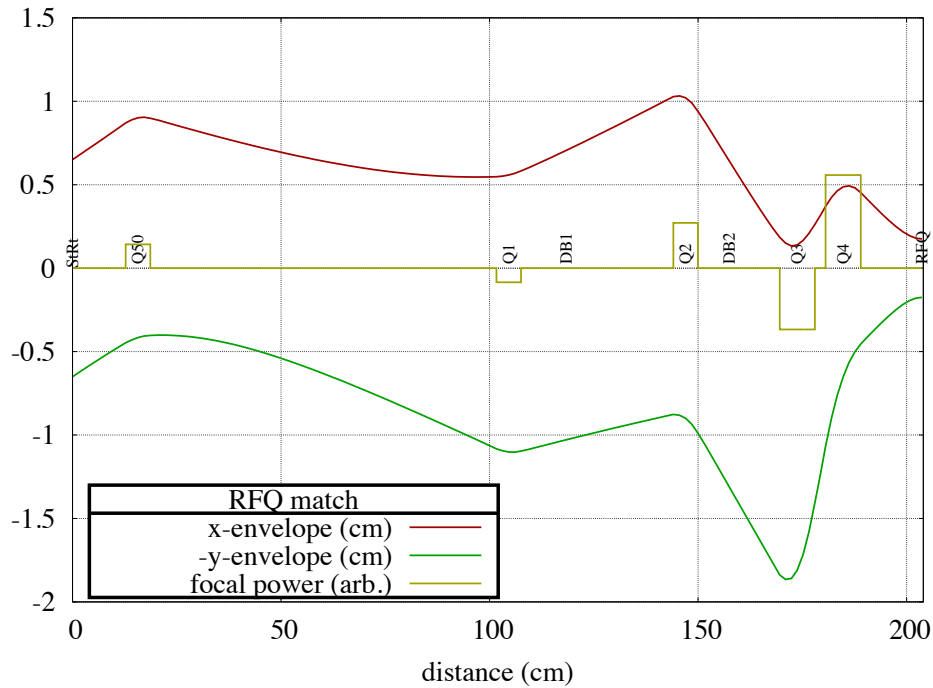


Fig. 4: First-order beam envelopes for match to RF quadrupole. Emittance is $50 \mu\text{m}$. Red (top) is x -envelope, green (bottom) is y -envelope, plotted for clarity as if it is negative. Yellow (middle) is quadrupole strength function. This is for exact first-order match, ignoring the third-order aberrations. Emittance growth is very bad, as can be seen in the blue curve of Fig. 6.

For the electrostatic dipole bender, I have solved the general toroidal case. Here, c is the ratio of electrode curvature in the bend plane to the non-bend plane. This means that $c = 0$ is a cylindrical bend and $c = 1$ is a spherical one:

$$\Delta x' = \frac{L}{\rho^3} \left[\left(-4 + \frac{7}{2}c - c^2 \right) x^2 + \left(-\frac{1}{2}c + c^2 \right) y^2 \right], \quad (29)$$

$$\Delta y' = \frac{L}{\rho^3} (-c + 2c^2) xy. \quad (30)$$

3 Beam transport optimization technique

Emittance growth factors are found for each transport element and combined with the emittance growth due to mismatch, to form a function to be minimized. We simply add (in quadrature) together, and minimize this as we vary such parameters as quadrupole position, length, or strength.

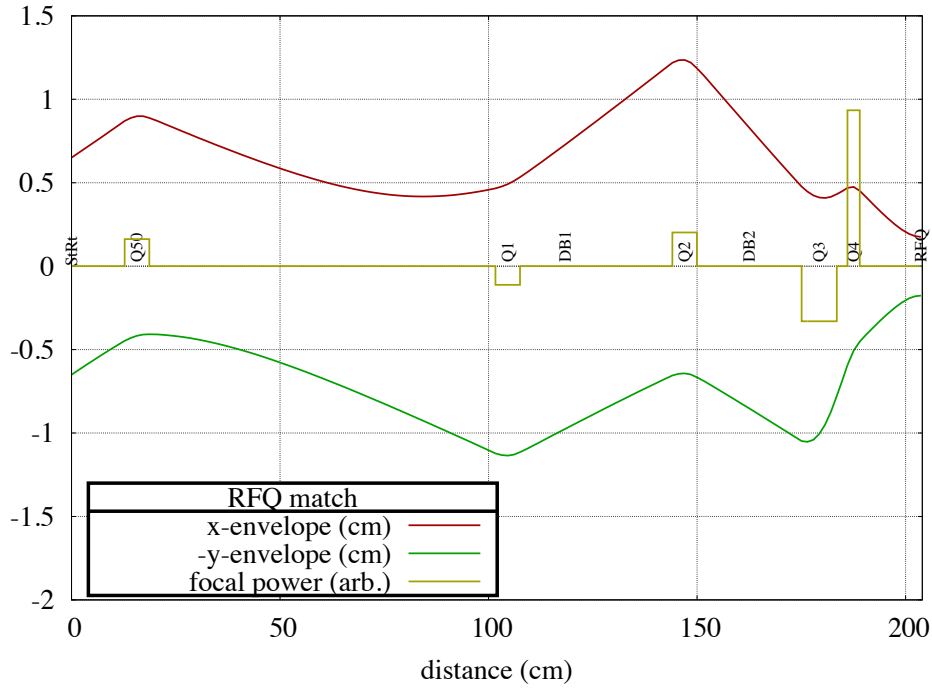


Fig. 5: Optimized RF quadrupole match, taking into account third-order aberrations, according to Eq. 21

3.1 Growth from higher order

Lowest higher-order kicks are used to estimate emittance growth. See Table 1, which is not meant to be exhaustive.

3.2 Growth from mismatch

The mismatch factor is the standard definition from, e.g., Bove *et al.* [15], see Fig. 7. The ratio of area of the ellipse that is similar but encapsulates the mismatched ellipse to the area of the mismatched ellipse is $D + \sqrt{D^2 - 1}$, where $D \equiv (\beta_2\gamma_1 + \beta_1\gamma_2)/2 - \alpha_1\alpha_2$ and α, β, γ are the ‘Twiss’ parameters of the mismatched ellipse (subscript 1) and the matched ellipse (subscript 2). As $D \ll 1$, it is sufficient to summarize this as $\Delta\epsilon/\epsilon \approx \sqrt{2(D - 1)}$.

This does not pretend to be a higher-order calculation; indeed, it is possible (though highly unlikely) that the higher order of one element is compensated by another; this would only show up in an actual higher-order calculation. As stated, the intention is not to find accurate higher-order effects, only to ensure that they are negligible.

3.3 Optimization engine

There are many optimization techniques for designing and tuning beam transport lines. Some are built into the transport codes themselves. Almost all of these work on the basis of reducing

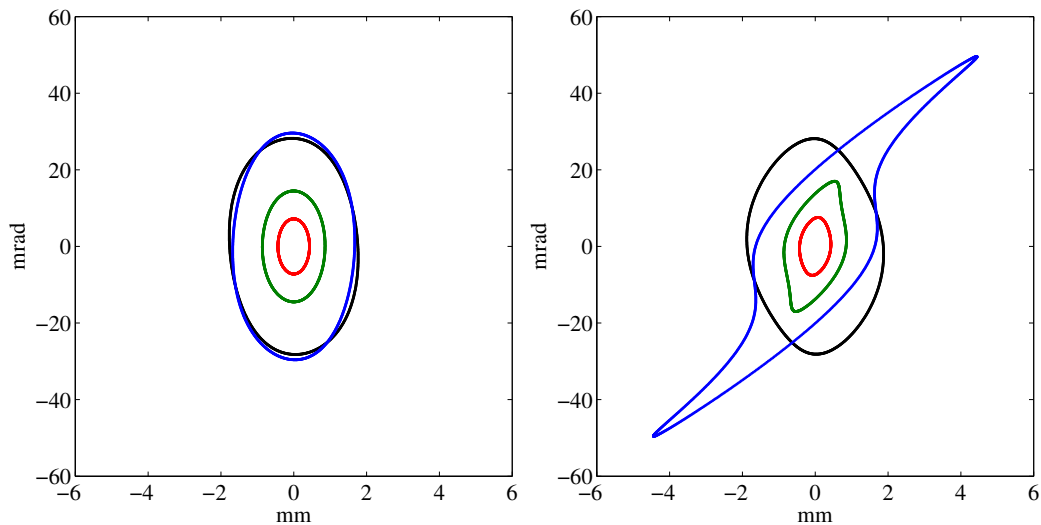


Fig. 6: Emittance contours (i.e. constant action) in phase space for previous envelope figures, calculated from the COSY- ∞ transfer maps. Left, x - P_x ; right, y - P_y . Black is the optimized case for the design acceptance $\epsilon = 50 \mu\text{m}$, where the long next-to-last quadrupole is brought forward by shrinking the final quadrupole, as shown in Fig. 5. From outside in, blue, green, and red are for the non-optimum configuration: blue, matched beam $\epsilon = 50 \mu\text{m}$; green, $\epsilon = 12.5 \mu\text{m}$; red, $\epsilon = 3 \mu\text{m}$.

Table 1: Higher-order kicks for a few chosen common elements

Element	$\Delta x'$	$\Delta y'$
Solenoid	$-\frac{x(x^2+y^2)}{fwL}$	$-\frac{y(x^2+y^2)}{fwL}$
$w = \text{fringe-field length, } f = \text{focal length, } L = \text{eff. length}$		
Einzel lens	$-\frac{x(x^2+y^2)}{\eta faL}$	$-\frac{y(x^2+y^2)}{\eta faL}$
$a = \text{aperture radius, } f = \text{focal length, } L = \text{total length, } 1/2 < \eta < 3/2$		
Magnetic quadrupole	$\frac{-1}{f^2L} (\frac{1}{3}x^3 + xy^2)$	$\frac{-1}{f^2L} (\frac{1}{3}y^3 + yx^2)$
$f = \text{focal length, } L = \text{eff. length}$		
Electric quadrupole	$\frac{-1}{f^2L} (\frac{7}{6}x^3 - \frac{1}{2}xy^2)$	$\frac{-1}{f^2L} (\frac{7}{6}y^3 - \frac{1}{2}yx^2)$
Magnetic sector bend	$-\frac{\theta y^2}{2\rho^2}$	$-\frac{\theta xy}{\rho^2}$
$\theta = \text{bend angle, } \rho = \text{bend radius}$		
Electric cylindrical bend	$-\frac{4\theta x^2}{\rho^2}$	0
Electric spherical bend	$\frac{\theta(-3x^2+y^2)}{2\rho^2}$	$\frac{\theta xy}{\rho^2}$

$$\text{where } D = \frac{1}{2} (\beta_2 \gamma_1 + \gamma_2 \beta_1 - 2\alpha_1 \alpha_2)$$

$$= 1 + \frac{(L_2 - L_1)^2 + (S_2 - S_1)^2}{2L_1 L_2}$$

For meaning of $\alpha, \beta, \gamma, L, S$ see 3.2 .

3.4.5 Three ellipses

- a) Area of ellipse S_3 , similar to S_2 , such that S_3 circumscribes S_1 : (area $S_1 = \text{area } S_2 = S$)

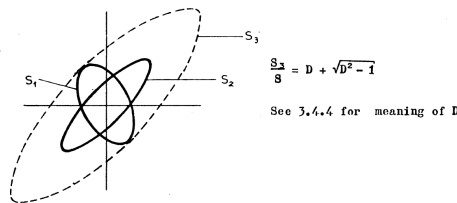


Fig. 7: Excerpt from Bovet *et al.* [15]

an error to zero by finding local derivatives of the error with respect to the parameters. I use a downhill simplex method. It is fast and robust.

It is also easily modified to incorporate simulated annealing for more than three free parameters. I use routines from the book *Numerical Recipes* by Press *et al.* [16]. They give the following description.

Offered a succession of options, a simulated thermodynamic system was assumed to change its configuration from energy E_1 to energy E_2 with probability $p = \exp[(E_2 - E_1)/(kT)]$. Notice that if $E_2 < E_1$, this probability is greater than unity; in such cases the change is arbitrarily assigned a probability $p = 1$, i.e., the system always took such an option. This general scheme, of always taking a downhill step while sometimes taking an uphill step, has come to be known as the Metropolis algorithm. To make use of the Metropolis algorithm for other than thermodynamic systems, one must provide the following elements:

1. A description of possible system configurations.
2. A generator of random changes in the configuration; these changes are the options presented to the system.
3. An objective function E (analogue of energy) whose minimization is the goal of the procedure.
4. A control parameter T (analogue of temperature) and an annealing schedule which tells how it is lowered from high to low values, e.g., after how many random changes in configuration is each downward step in T taken, and how large is that step.

Applied to the design of a transport channel, we can identify these four points as:

1. a set of parameters, such as quadrupole strengths, locations, and their allowed ranges;
2. use the vertices of a ‘simplex’ in the space of all possible parameters, and change these randomly, independently according to T ; for $T = 0$, the technique reduces to the standard ‘simplex’ method of minimization;
3. E = the sum of all possible effects to be minimized, such as mismatch or emittance growth, with their appropriate weights;
4. use T , such that $T = 1$ implies that parameters are varied through their entire ranges, specify N , the number of random changes at each temperature, and $\alpha (< 1)$, the factor with which T is multiplied at each step.

For as few as three parameters, α can be 0.88, with $N = 8$, resulting in only about 100 evaluations of the beamline to achieve 10^{-4} accuracy. But this number of evaluations increases exponentially with the number of parameters. For six parameters, we need ten times slower cooling $\alpha = 0.98$, $N = 50$, and this results in roughly 100 times more evaluations ($10^4 - 10^5$) before convergence is obtained. An empirical relation found from many such runs is that for n variable parameters, we need $N = 2^n$, $\alpha = 1 - \frac{1}{N}$. Still, the calculations are simple so that millions of beamline calculations can be made and an optimum found, for the same amount of computing time as a single million-particle simulation.

3.4 TRANSOPTR: Space charge and other complications

TRANSOPTR [17] was originally written in 1981. Like the original envelope codes TRANSPORT and TRACE3D, it included only elements whose transfer matrices could be written in closed form and were known. This required all elements to be ‘hard-edged’. The extension to other elements and to space charge requires that the envelope code itself numerically integrates through the elements. The exact differential equations of Sacherer [18] were incorporated into TRANSOPTR by Mark deJong [19], but greatly expanded by me to include many other element types [20]. In principle, any linear element can be incorporated, from solenoids and quadrupoles with bell-shaped strength functions, to linear accelerators and bunchers. All that is needed is the Hamiltonian to quadratic order [21].

4 Conclusions

Simple formulae can be found for the lowest-order aberrations and used to estimate emittance growth. A simple strategy for beam transport system design optimization is to add these growths to the growth from mismatch, to form an error function to be minimized. Efficient first-order transport codes (with or without space charge) can be augmented in this way. An example is TRANSOPTR, which uses a minimization engine to vary beamline parameters to find an optimum with negligible emittance growth.

At TRIUMF, many beamlines have been efficiently designed in this way, some with strong space charge effects, and perform as predicted, in spite of never having been simulated with any multiparticle code [22–24].

Acknowledgement

I would like to thank Dr Thomas Planche for his careful reading and helpful comments.

References

- [1] S. Penner, *Rev. Sci. Instrum.* **32** (1961) 150, <https://doi.org/10.1063/1.1717300>.
- [2] K.L. Brown, A first-and second-order matrix theory for the design of beam transport systems and charged particle spectrometers, Tech. Rep. SLAC-75, Stanford Linear Accelerator Center, CA, 1971.
- [3] K.L. Brown *et al.*, Transport: a computer program for designing charged-particle beam-transport systems, Tech. Rep. SLAC-R-91-REV-3, Stanford Linear Accelerator Center, CA, 1980.
- [4] K.R. Crandall and D. Rusthoi, *Trace 3-D Documentation* (Los Alamos National Laboratory, Los Alamos, 1987), <https://doi.org/10.2172/6290515>.
- [5] H. Wollnik *et al.*, GIOS—a program for the design of general ion optical systems, ii, Physikalisches Institut, Justus-Liebig-Universität, Giessen, 1988.
- [6] M. Berz, *Nucl. Instrum. Methods Phys. Res. A* **298** (1990) 473, [https://doi.org/10.1016/0168-9002\(90\)90649-Q](https://doi.org/10.1016/0168-9002(90)90649-Q).
- [7] J. Back *et al.*, Commissioning of the low energy beam transport, MOPEC078, Proc. IPAC, 2010 vol. 10.
- [8] B. Biswas, *Pramana*, **86** (2016) 1299, <https://doi.org/10.1007/s12043-016-1197-7>.
- [9] J.D. Jackson, *Classical Electrodynamics* (John Wiley & Sons, New York, 2007).
- [10] A. Galejs and P. Rose, in *Focusing of Charged Particles*, Ed. A. Septier (Academic Press, Orlando, 1967), vol. ii, pp. 297–326, <https://doi.org/10.1016/B978-0-12-636902-1.50014-7>.
- [11] R. Baartman, Intrinsic third order aberrations in electrostatic and magnetic quadrupoles, Proc. Particle Accelerator Conf., 1997, (IEEE, 1997), vol. 2, p. 1415.
- [12] R. Baartman, Quadrupoles, to third order, *arXiv:1508.01576*, 2015.
- [13] R. Baartman, Emittance convention, Tech. Rep. TRI-BN-15-07, TRIUMF, 2015.
- [14] R. Baartman, End effects of beam transport elements, Snowmass, 2001. <http://lin12.triumf.ca/text/Talks/2001Snowmass/FF.pdf>
- [15] C. Bovet *et al.*, A selection of formulae and data useful for the design of Ag synchrotrons, Tech. Rep. (CERN, Geneva, 1970).
- [16] W.H. Press *et al.*, *Numerical Recipes* (Cambridge University Press, Cambridge, 1996).
- [17] E. Heighway and R. Hutcheon, *Nucl. Instrum. Methods Phys. Res.* **187** (1981) 89, [https://doi.org/10.1016/0029-554X\(81\)90474-2](https://doi.org/10.1016/0029-554X(81)90474-2).
- [18] F.J. Sacherer, *IEEE Trans. Nucl. Sci.* **18** (1971) 1105, <https://doi.org/10.1109/TNS.1971.4326293>.
- [19] M. deJong and E. Heighway, *IEEE Trans. Nucl. Sci.* **30** (1983) 2666, <https://doi.org/10.1109/TNS.1983.4332918>.

- [20] R. Baartman, TRANSOPTR: changes since 1984, Tech. Rep. TRI-BN-16-06, TRIUMF, 2016. http://lin12.triumf.ca/text/design_notes/b2016_06/TRI-BN-16-06_TRANSOPTR.pdf
- [21] R. Baartman, Fast envelope tracking for space charge dominated injectors, 28th Linear Accelerator Conf. (LINAC'16), East Lansing, MI, 25–30 September 2016, pp. 1017–1021 (JACOW, Geneva, 2017).
- [22] T. Planche *et al.*, Commissioning and early operation of the ARIEL e-Linac, 28th Linear Accelerator Conf. (LINAC'16), East Lansing, MI, 25–30 September 2016, pp. 12–16 (JACOW, Geneva, 2017).
- [23] R. Baartman, ISAC LEBT, ISAC and ARIEL: The TRIUMF Radioactive Beam Facilities and the Scientific Program, pp. 69–77 (Springer, 2013), https://doi.org/10.1007/978-94-007-7963-1_7.
- [24] R. Baartman, TRIUMF 300 keV vertical injection line, European Cyclotron Progress Meeting, 2015. <http://lin12.triumf.ca/text/Talks/2015ECPM/ISIS.pdf>