



**BNL-74966-2006-JA-R1
MUC-PUB-ACCELERATION-326**

***Minimizing Longitudinal Distortion in a Nearly
Isochronous Linear Nonscaling Fixed-Field
Alternating Gradient Accelerator***

J. Scott Berg

Submitted to Physical Review Special Topics – Accelerators and Beams

March 2006

Physics Department/Bldg. 901A

Brookhaven National Laboratory

P.O. Box 5000

Upton, NY 11973-5000

www.bnl.gov

Notice: This manuscript has been authored by employees of Brookhaven Science Associates, LLC under Contract No. DE-AC02-98CH10886 with the U.S. Department of Energy. The publisher by accepting the manuscript for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

This preprint is intended for publication in a journal or proceedings. Since changes may be made before publication, it may not be cited or reproduced without the author's permission.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Minimizing longitudinal distortion in a nearly isochronous linear non-scaling fixed-field alternating gradient accelerator

J. Scott Berg*

Brookhaven National Laboratory; Building 901A; PO Box 5000; Upton, NY 11973-5000

(Dated: 7 March 2006)

Linear non-scaling FFAGs (Fixed-Field Alternating Gradient accelerators) are machines that use linear magnets to achieve an extremely large energy acceptance (generally a factor of 2 or more). This paper examines the longitudinal dynamics in such a machine, focusing on the longitudinal acceptance, the phase space area that is transmitted without excessive distortion. The paper shows how to compute the distortion in two ways: computing the emittance growth, and computing the distortion of an initial ellipse from an elliptical shape. The paper will describe a model for the longitudinal dynamics in a linear non-scaling FFAG, show how to compute the longitudinal distortion in such a machine using a Dragt-Finn factorization, examine the accuracy of the calculation, and describe how longitudinal acceptance can interact with other performance criteria for an FFAG.

PACS numbers: 29.27.Bd, 41.85.-p, 45.10.Hj

Keywords: FFAG; longitudinal dynamics; Dragt-Finn factorization

I. INTRODUCTION

Fixed field alternating gradient (FFAG) accelerators are machines which accelerate over a large range of energy (generally a factor of 2 or more) without varying the magnet fields. Since the magnet fields don't vary during acceleration, acceleration is potentially very rapid. FFAGs were first studied and built in the 1950s [1–3], but little further work was done on these machines until very recently. In the last few years, two FFAGs have been built in Japan [4–8], and FFAGs have been proposed there for several other projects.

All of the FFAGs built until now have been what are today called “scaling” FFAGs. They are called scaling because of an important property of the machine: after applying a particular energy-dependent linear transformation to the phase space variables, the dynamics of the machine become completely independent of energy. In particular, the tunes and momentum compaction are independent of energy, and the energy-dependent closed orbits are geometrically similar. The energy-independent tune and the energy independence of the phase space (except for an energy-dependent linear transformation) allows one to choose a good working tune, as in a synchrotron, and one will then avoid resonances over the entire acceleration process.

There are of course difficulties with scaling FFAGs. The time of flight depends strongly on energy, even when the velocity is nearly the speed of light, so one must either vary the RF frequency to match the time of flight, which limits the acceleration rate, or one must fix the RF frequency and accelerate so quickly that the bunch does not leave the RF crest. The large energy range and the machine's nonzero dispersion require large apertures.

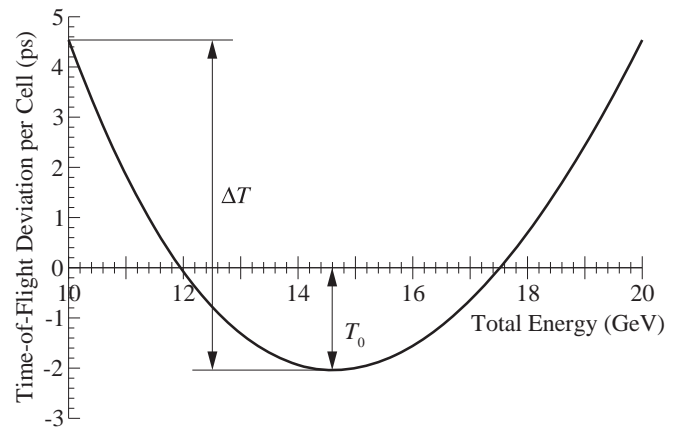


FIG. 1: Time of flight as a function of energy in a linear non-scaling FFAG.

Finally, the highly nonlinear magnets required for a scaling FFAG can potentially limit the dynamic apertures for those machines.

A new type of FFAG, the linear non-scaling FFAG [9, 10], was proposed to improve the performance of FFAGs with respect to these difficulties, particularly for muon acceleration. Muons must be accelerated very rapidly to avoid decays, yet one desires to make as many passes through the RF as possible to use the very expensive RF cavities more efficiently. The rate of acceleration prevents the RF frequency from being varied, and thus the range in time-of-flight must be minimized. Linear non-scaling FFAGs for muon acceleration applications make the machine as isochronous as possible by making the machine isochronous somewhere close to the middle of the energy range of the machine (as in Fig. 1). In a scaling FFAG, the momentum compaction is a nonzero constant, leading to a larger time-of-flight range than for a comparable non-scaling FFAG. Muon accelerators also require an exceptionally large dynamic aperture, so to

*Electronic address: jsberg@bnl.gov; URL: <http://pubweb.bnl.gov/people/jsberg/>

improve that, the linear non-scaling FFAGs use only linear magnets, as opposed to the highly nonlinear magnets required for a scaling FFAG. Finally, the non-scaling nature of the FFAG means that orbits will not be geometrically similar as in the scaling FFAG. This extra freedom can be used to make the apertures smaller in some of the magnets.

As can be seen in Fig. 1, the time of flight in a linear non-scaling FFAG will be well approximated by a parabola. As can be determined from the shift of the minimum from the center in that figure, cubic order terms only contribute around 10% to the time of flight within the energy range. The characteristics of the motion in phase space are determined qualitatively by this parabolic shape, and the underlying causes of the phase space distortion examined here are present when the time of flight is approximated by a parabola.

Instead of using linear magnets, one can use nonlinear magnets and make a non-scaling FFAG lattice almost perfectly isochronous [11, 12]. Higher order effects (such as the time of flight dependence on transverse amplitude) and errors will inevitably make the isochronicity imperfect, but one can certainly reduce the time-of-flight range substantially from what can be achieved in a linear non-scaling FFAG. While this may come at the cost of dynamic aperture (due to nonlinearities) and magnet cost, such machine require consideration. This paper will more directly address the case of linear non-scaling FFAGs, where the time of flight as a function of energy is nearly parabolic; the results from this paper can be useful in comparing the machine performance of linear non-scaling FFAGs to these other types of non-scaling FFAGs.

This paper describes the longitudinal dynamics in a machine where the time of flight is exactly a parabolic function of energy, with the minimum of the parabola at the center of the energy range, and the RF voltage is a sinusoidal function of the phase. In the model, the RF voltage and the time-of-flight advance are both distributed uniformly around the ring. This system has been examined previously [13, 14]. A primary constraint on the design parameters for these FFAGs will be the amount of longitudinal phase distortion one can tolerate. This paper will quantify this distortion and describe a method for computing it for this system.

II. HAMILTONIAN IN SCALED VARIABLES

The FFAG ring is assumed to consist entirely of identical cells, all of which contain an RF cavity. I take the further step of ignoring the longitudinal variation within the cell, making a continuous approximation of the system. The time-of-flight τ is approximated here to be a parabolic function of energy E :

$$\frac{d\tau}{ds} = \Delta T \left(\frac{2E - E_i - E_f}{\Delta E} \right)^2 - T_0, \quad (1)$$

where the machine is designed to accelerate from E_i to E_f , $\Delta E = E_f - E_i$, ΔT is the difference between the time of flight per unit length at E_i (or E_f) and the central energy, s is the distance along a reference curve which defines the coordinate system, and T_0 is chosen so that $d\tau/ds$ is zero when the phase of a particle in the RF cavities does not change from one cell to the next. The energy gain in the cavities is sinusoidal:

$$\frac{dE}{ds} = V \cos(\omega\tau), \quad (2)$$

where V is the average energy gain per unit length for on-crest acceleration in a cell, and ω is the angular RF frequency. The choice of placing the minimum of the parabola at $(E_i + E_f)/2$, and the fact that there is no phase in Eq. (2), is based on symmetry.

As an example of what these parameters might be like for a real machine, consider a muon FFAG design: the machine would accelerate from 10 to 20 GeV, using 201.25 MHz superconducting RF. One design [15] has 91 4.7 m long combined-function doublet cells. With a gradient of 10 MV/m in the cavities, V is about 1.5 MeV/m. ΔT is about 1.4 ps/m in this lattice.

One can perform the change of variables

$$x = \omega\tau \quad p = \frac{E - E_i}{\Delta E} \quad u = s\omega\Delta T \quad (3)$$

to get the new equations of motion

$$\frac{dx}{du} = (2p - 1)^2 - b \quad \frac{dp}{du} = a \cos x \quad (4)$$

where

$$a = \frac{V}{\omega\Delta T\Delta E} \quad b = \frac{T_0}{\Delta T}. \quad (5)$$

Note that b is a relatively simple quantity to adjust in a machine design: a small change of frequency, phase relationship between cavities, or cell length can generally make b any desired value. a is generally more costly to vary: adjusting it requires that the amount of RF voltage be changed, that the number of cells be changed (to change ΔT [16]), or that the cell length (to change ΔT [16]), RF frequency, or energy range be changed significantly. The muon accelerator parameters described above came from a lattice designed for $a = 1/12$.

In these scaled variables, the goal of the machine is to accelerate from $p = 0$ to $p = 1$. This scaled system is governed by the Hamiltonian

$$\frac{1}{6}(2p - 1)^3 - \frac{b}{2}(2p - 1) - a \sin x. \quad (6)$$

Henceforth, I will work with this scaled Hamiltonian.

When $b \geq 0$, this Hamiltonian has unstable fixed points at $x = \pm\pi/2$ and $p = (1 \mp \sqrt{b})/2$. The value of the Hamiltonian on the corresponding separatrices is

$$\pm \frac{1}{3}b^{3/2} \mp a. \quad (7)$$

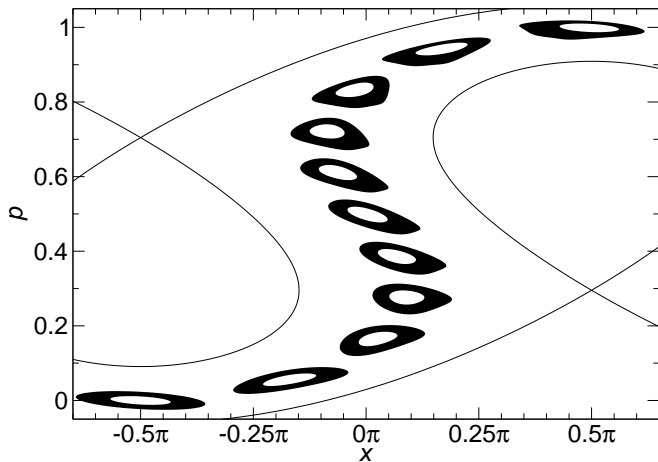


FIG. 2: Phase space in normalized variables, showing separatrices delimiting the phase space and an elliptical ring in phase space at 11 points in the acceleration process from $p = 0$ to $p = 1$.

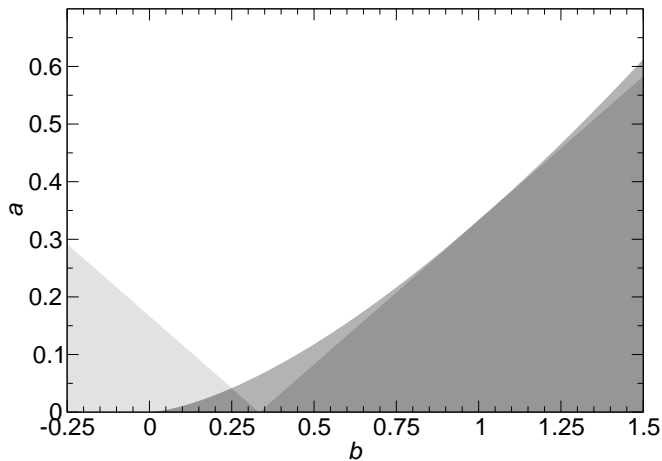


FIG. 3: Region of a - b parameter space for which particles can be accelerated from $p = 0$ to $p = 1$, shown in white. Particles will not accelerate from $p = 0$ to $p = 1$ because of condition (9) (light grey) or condition (8) (medium grey) or both (darker grey).

The separatrices delineate regions of phase space which cannot be crossed. Figure 2 shows the separatrices dividing the phase space. From the Hamiltonian values on the separatrices and an examination of the phase space it can be seen that there will be no region of phase space connecting $p = 0$ to $p = 1$ if $a < (1/3)b^{3/2}$ if $b \geq 0$; if $b < 0$, there are no separatrices. Thus, to accelerate from $p = 0$ to $p = 1$

$$a > \frac{1}{3}b^{3/2} \quad (8)$$

if $b \geq 0$; there is no restriction from the separatrices if $b < 0$. Furthermore, there must be at least one trajectory

crossing both $p = 0$ and $p = 1$. By symmetry, this means that the trajectory passing through $x = 0$ and $p = 1/2$ must pass through $p = 0$ and $p = 1$. It must be possible for the Hamiltonian to be zero when $p = 0$ and $p = 1$. Since $|\sin x| \leq 1$,

$$a > \left| \frac{1}{6} - \frac{b}{2} \right|. \quad (9)$$

The combination of these restrictions leads to

$$\frac{1}{3} - 2a < b < (3a)^{2/3}. \quad (10)$$

Note that as a result, there is an absolute restriction that $a > 1/24$. The region of the a - b parameter space that permits acceleration from $p = 0$ to $p = 1$ is shown in Fig. 3.

III. LIE ALGEBRAIC FORMULATION

To analyze the phase space transmission in this system, I will write the map for evolution in u in Lie Algebraic form, using the Dragt-Finn factorization [17, 18]. In this paper, I will analyze the map about the central trajectory, passing through $x = 0$ and $p = 1/2$. It is convenient to start with the map from $x = 0$, $p = 1/2$ forward in u . Its Lie factorization can be written as

$$\mathcal{M}_{\text{half}} = e^{-:g_1:} e{:f_5:} e{:f_4:} e{:f_3:} e{:f_2:} e{:f_1:}. \quad (11)$$

$g_1 = x/2$ is the operator that translates to the initial conditions $(x, p) = (0, 1/2)$. The remaining operators will be computed through numerical integration as described in [18]. The map from beginning to end is just

$$\mathcal{M}_{\text{full}} = \mathcal{C} \mathcal{M}_{\text{half}}^{-1} \mathcal{C}^{-1} \mathcal{M}_{\text{half}}, \quad (12)$$

where

$$\mathcal{C} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} -x \\ 1-p \end{pmatrix}. \quad (13)$$

Note that the operator order follows the Lie Algebraic convention of first to last being left to right. Writing \mathcal{C} in Lie algebraic notation,

$$\mathcal{C} = \mathcal{F} e{:x:} = e^{-:x:} \mathcal{F}, \quad (14)$$

where \mathcal{F} is the reflection operator

$$\mathcal{F} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} -x \\ -p \end{pmatrix}. \quad (15)$$

Note that \mathcal{F}^2 is the identity operator, as is \mathcal{C}^2 . The combined map can thus be written

$$\begin{aligned}
\mathcal{M}_{\text{full}} &= e^{-:x:} \mathcal{F} e^{-:f_1:} e^{-:f_2:} e^{-:f_3:} e^{-:f_4:} e^{-:f_5:} e^{:g_1:} \mathcal{F} e^{:x:} e^{-:g_1:} e^{:f_5:} e^{:f_4:} e^{:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} \mathcal{F} e^{-:f_2:} e^{-:f_3:} e^{-:f_4:} e^{-:f_5:} \mathcal{F} e^{:x-2g_1:} e^{:f_5:} e^{:f_4:} e^{:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} e^{-:f_2:} \mathcal{F} e^{-:f_3:} e^{-:f_4:} \mathcal{F} e^{2:f_5:} e^{:f_4:} e^{:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} e^{-:f_2:} e^{:f_3:} e^{-:f_4:} e^{2:f_5:} e^{:f_4:} e^{:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} e^{-:f_2:} e^{:f_3:} \exp(2:e^{-:f_4:} f_5:) e^{:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} e^{-:f_2:} \exp(2:e^{:f_3:} e^{-:f_4:} f_5:) e^{2:f_3:} e^{:f_2:} e^{:f_1:} \\
&= e^{:f_1-x:} \exp(2:e^{-:f_2:} e^{:f_3:} e^{-:f_4:} f_5:) \exp(2:e^{-:f_2:} f_3:) e^{:f_1:} \\
&\approx e^{:f_1-x:} \exp(2:e^{-:f_2:} f_5:) \exp(2:e^{-:f_2:} f_3:) e^{:f_1:}.
\end{aligned}$$

I use the transformation rule for similarity transformations on Lie operators, the fact that \mathcal{F} leaves even order homogeneous polynomials invariant, and changes the sign of odd order homogeneous polynomials, that first-order Lie exponentials can be combined by adding their exponents, and that $g_1 = x/2$. The final approximation is appropriate since the factorization was truncated at fifth order already, and including the effects of f_3 and f_4 on f_5 would result in terms higher order than fifth. Note that the linear part of the map is the identity and that any f_4 terms disappear from the final map. Note that if the time of flight were not perfectly parabolic or the RF waveform were asymmetric, the linear part would not be the identity (which will not matter for these results) and there would be a nonzero f_4 term.

A. Quantities to Analyze

To give optimal performance of a machine, we will want to minimize some quantity which is related to the deviation of the machine from linearity. Different types of machines will require the minimization of different quantities. For colliders, the r.m.s. energy spread and bunch length at collision are the important quantities. Thus, the machine should minimize the emittance growth for that case. For a neutrino factory, however, the machine is generally designed to transmit a certain phase space volume and not much more. In that case, one wishes to minimize the growth of the boundary of an elliptical

phase space volume.

One wants to find the effect of the FFAG on an elliptical distribution. A translation has no effect on the distribution shape, so one only needs to analyze the effect of

$$\mathcal{M} = \exp(2:e^{-:f_2:} f_5:) \exp(2:e^{-:f_2:} f_3:) \quad (16)$$

on an ellipse. If $\mathcal{A} = e^{:a_2:}$ is a linear transformation that transforms a circle to the desired ellipse, then one can instead analyze the effect of

$$\begin{aligned}
\mathcal{M}_C &= \mathcal{A} \mathcal{M} \mathcal{A}^{-1} \\
&= e^{:a_2:} \exp(2:e^{-:f_2:} f_5:) \exp(2:e^{-:f_2:} f_3:) e^{-:a_2:} \\
&= \exp(2:e^{:a_2:} e^{-:f_2:} f_5:) \exp(2:e^{:a_2:} e^{-:f_2:} f_3:) \quad (17)
\end{aligned}$$

on a circularly symmetric distribution in phase space. A linear map transforms a homogeneous polynomial to a homogeneous polynomial of the same order. Thus

$$\mathcal{M}_C = e^{:g_5:} e^{:g_3:} \quad (18)$$

$$g_m = 2e^{:a_2:} e^{-:f_2:} f_m = \sum_{k=0}^m g_{mk} x^{m-k} p^k \quad (19)$$

In most cases, I will try to find the optimal ellipse orientation (i.e., I will try to choose \mathcal{A}) so as to minimize the quantity of interest.

I will need to compute the effect of \mathcal{M}_C on a phase space vector to third order; this is

$$\mathcal{M}_C \begin{pmatrix} x \\ p \end{pmatrix} \equiv \mathcal{M}_C \mathbf{z}_1 = \mathbf{z}_1 + \mathbf{z}_2 + \mathbf{z}_3 \quad (20)$$

$$\mathbf{z}_2 = [g_3, \mathbf{z}_1] = \begin{pmatrix} -g_{31}x^2 - 2g_{32}xp - 3g_{33}p^2 \\ 3g_{30}x^2 + 2g_{31}xp + g_{32}p^2 \end{pmatrix} \quad (21)$$

$$\mathbf{z}_3 = \frac{1}{2} [g_3, [g_3, \mathbf{z}_1]] = \begin{pmatrix} (g_{31}^2 - 3g_{30}g_{32})x^3 + (g_{31}g_{32} - 9g_{30}g_{33})x^2p + (g_{32}^2 - 3g_{31}g_{33})xp^2 \\ (g_{31}^2 - 3g_{30}g_{32})x^2p + (g_{31}g_{32} - 9g_{30}g_{33})xp^2 + (g_{32}^2 - 3g_{31}g_{33})p^3 \end{pmatrix}. \quad (22)$$

There are also fourth order terms, but they are not needed for the initial computations, so I will delay considering them. Because of the presence of the transformation \mathcal{A} , I will examine transformations of circles and circular distributions under \mathcal{M}_C .

To optimize the design, I will vary the transformation \mathcal{A} (which has two free parameters) to minimize the quantity of interest for a given a and b . I will then find the b which minimizes the quantity for a given a (since, as noted above, it costs very little to change b), producing a plot of the quantity versus a .

1. Emittance Growth

To compute the emittance, we assign

$$x = \sqrt{2J} \cos \theta \quad p = -\sqrt{2J} \sin \theta \quad (23)$$

and average over an arbitrary distribution in J and a uniform distribution in θ . The emittance is defined to be the square root of the determinant of the covariance matrix $\langle z z^T \rangle - \langle z \rangle \langle z \rangle^T$ ($\langle f \rangle$ is the average of the quantity f over the distribution, and z is the phase space vector). For $z = z_1$, the emittance is $\langle J \rangle$. For $z = \mathcal{M}_C z_1$, the result to lowest order is

$$\langle J \rangle + \frac{1}{2} \langle 2z_1^T z_3 + z_2^T z_2 \rangle - \frac{1}{2} \langle z_2 \rangle^T \langle z_2 \rangle. \quad (24)$$

The fourth order moments needed are

$$\langle x^4 \rangle = \langle p^4 \rangle = \frac{3}{2} \langle J^2 \rangle \quad \langle x^2 p^2 \rangle = \frac{1}{2} \langle J^2 \rangle \quad (25)$$

$$\langle x^3 p \rangle = \langle x p^3 \rangle = 0. \quad (26)$$

Thus, the emittance growth is

$$\begin{aligned} & \frac{3}{4} \langle J^2 \rangle (9g_{30}^2 + 5g_{31}^2 + 5g_{32}^2 + 9g_{33}^2 - 6g_{30}g_{32} - 6g_{31}g_{33}) \\ & - \frac{1}{2} \langle J \rangle^2 [(g_{31} + 3g_{33})^2 + (g_{32} + 3g_{30})^2]. \end{aligned} \quad (27)$$

This is only necessarily nonnegative when $\langle J^2 \rangle \geq (4/3)\langle J \rangle^2$. Mathematically, one must have $\langle J^2 \rangle \geq \langle J \rangle^2$. For the Gaussian distribution $e^{-J/\epsilon}/\epsilon$, $\langle J \rangle = \epsilon$ and $\langle J^2 \rangle = 2\epsilon^2$. At the opposite extreme, for an ‘‘airbag’’ distribution $\delta(J - \epsilon)$, $\langle J \rangle = \epsilon$ and $\langle J^2 \rangle = \epsilon^2$. A ‘‘waterbag’’ distribution (uniform phase space density) has $\langle J^2 \rangle = (4/3)\langle J \rangle^2$.

For the analysis, I will restrict the discussion to $\langle J^2 \rangle > (4/3)\langle J \rangle^2$. The minimum emittance growth as a function of a is shown in Fig. 4. For a close to its minimum value of $1/24$, the emittance growth is proportional to $(a - 1/24)^{-2}$. Larger values of $\langle J^2 \rangle / \langle J \rangle^2$ give larger values for the emittance growth. For a given values of $\langle J^2 \rangle / \langle J \rangle^2$, the emittance growth is proportional to $\langle J \rangle^2$.

The optimal b is shown in Fig. 5. The reason for the sudden change near 0.41 is shown in Fig. 6. For small b , the minimum emittance growth occurs at the smallest

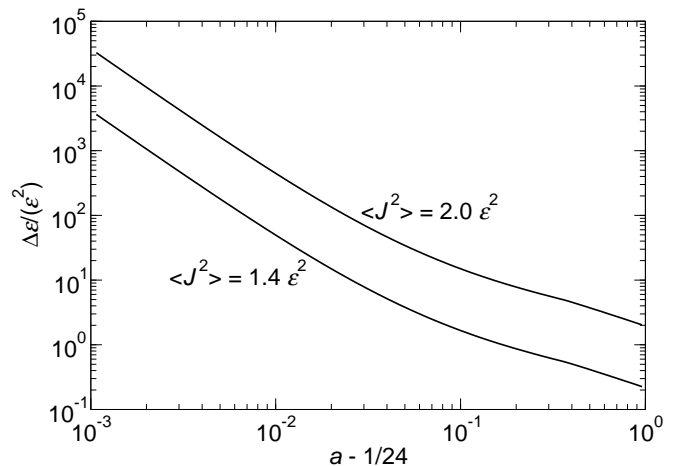


FIG. 4: Emittance growth, minimized over b and \mathcal{A} . ϵ is $\langle J \rangle$.

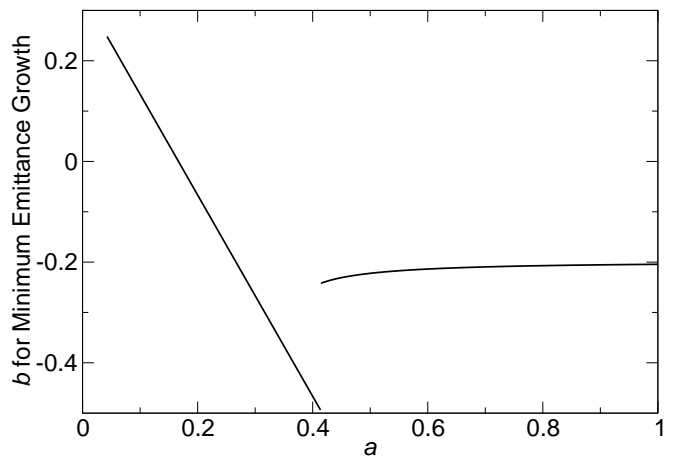


FIG. 5: Value for b which gives the minimum emittance growth.

allowed b . The optimal b is independent of the ratio $\langle J^2 \rangle / \langle J \rangle^2$.

If the time of flight were not perfectly parabolic (as in the real case shown in Fig. 1), there would potentially be relative corrections of the order of the relative size of the difference from a parabola, due to the distortion of the orbit about which the expansion is performed and changes to the derivatives about that. The addition of f_4 terms to the map will give no contribution to the emittance growth. Specific quantitative results will be left to a subsequent paper.

2. Boundary Distortion

If instead we wish to minimize the distortion of a circle due to \mathcal{M}_C , first compute the change in the radius of the

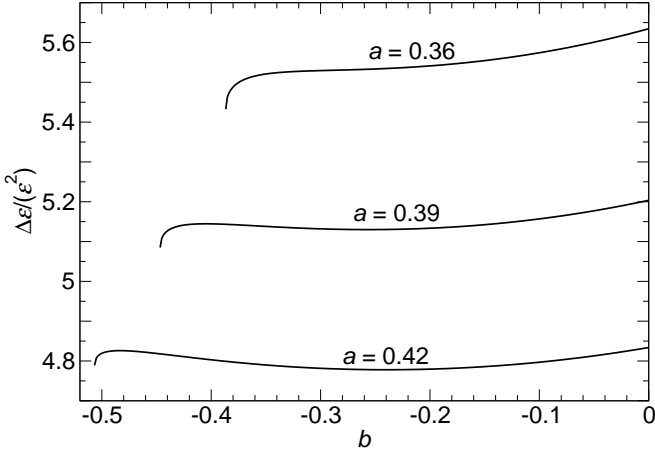


FIG. 6: Emittance growth as a function of b for various values for a .

circle under \mathcal{M}_C to lowest order:

$$\Delta r^2 = 2\Delta J = 2\mathbf{z}_1^T \mathbf{z}_2 = -2g_{31}x^3 + (6g_{30} - 4g_{32})x^2p - (6g_{33} - 4g_{31})xp^2 + 2g_{32}p^3. \quad (28)$$

Replacing x and p by J and θ using Eq. (24), this becomes

$$-4\sqrt{2J^3}[g_{31}\cos^3\theta + (3g_{30} - 2g_{32})\cos^2\theta\sin\theta + (3g_{33} - 2g_{31})\cos\theta\sin^2\theta + g_{32}\sin^3\theta]. \quad (29)$$

To find the extrema of this, take the derivative with respect to θ :

$$-4\sqrt{2J^3}[(3g_{30} - 2g_{32})\cos^3\theta + (6g_{33} - 7g_{31})\cos^2\theta\sin\theta - (6g_{30} - 7g_{32})\cos\theta\sin^2\theta - (3g_{33} - 2g_{31})\sin^3\theta]. \quad (30)$$

To find the zeros of this, divide by $\cos^3\theta$ or $\sin^3\theta$, and you will get a third order polynomial in $\tan\theta$ or $\cot\theta$ respectively. The zeros of this can be found analytically, and the resulting zeros used to find the extrema of Eq. (29). Note that the distortion of the radius of the circle is proportional to the square of the radius of the circle.

Figure 7 shows the minimum ellipse distortion as a function of a . The value of $\Delta J/(2J)^{3/2}$ is proportional to $(a - 1/24)^{-1}$ for a close to $1/24$.

The b which gives this minimum distortion is identical to the b which gives the minimum emittance growth that was shown in Fig. 5.

The comments in the subsection on emittance growth regarding a time of flight which is not perfectly parabolic also apply to the boundary distortion calculation.

3. Boundary Distortion with Linear Transform

If one is interested only in the distortion of the outer boundary of the ellipse, one need not require that an ellipse with nonzero size have the same center, orientation,

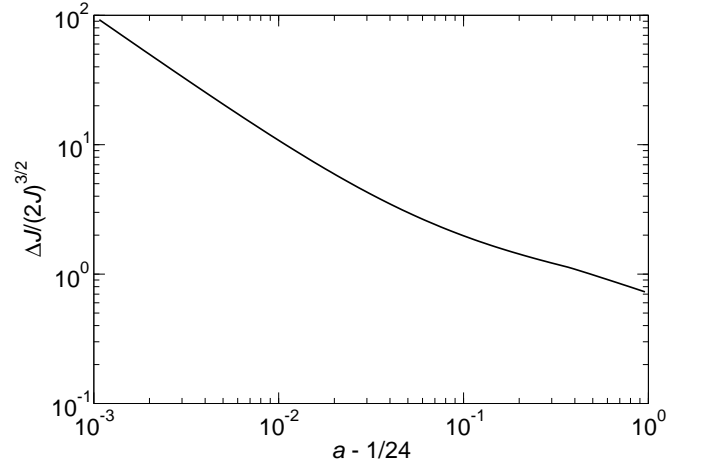


FIG. 7: Distortion of the ellipse, minimized over b and \mathcal{A} .

and aspect ratio as an infinitesimally small ellipse. We thus want to look at the quantity

$$[\mathbf{z} - \mathbf{z}_0(J)]^T B(J)[\mathbf{z} - \mathbf{z}_0(J)] - 2J, \quad (31)$$

and minimize it to lowest order in J , while simultaneously finding $\mathbf{z}_0(J)$ and $B(J)$, varying both the transformation \mathcal{A} as well as $\mathbf{z}_0(J)$ and $B(J)$. Note that $\mathbf{z}_0(0) = \mathbf{0}$ and $B(0) = I$, the identity matrix. Substituting $\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2$, $\mathbf{z}_0(J) = \mathbf{z}_{02}J$, and keeping terms to third order, Eq. (31) becomes

$$(2J)^{3/2} \left(\frac{3}{2}(g_{33} - g_{31})\cos 3\theta - \frac{3}{2}(g_{30} - g_{32})\sin 3\theta - \frac{1}{2}(3g_{33} + g_{31})\cos\theta - \frac{1}{2}(3g_{30} + g_{32})\sin\theta \right) - (2J)^{3/2} \mathbf{z}_{02}^T \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix}. \quad (32)$$

Thus, the $\cos\theta$ and $\sin\theta$ terms can be eliminated if

$$\mathbf{z}_{02} = \frac{1}{2} \begin{bmatrix} -3g_{33} - g_{31} \\ 3g_{30} + g_{32} \end{bmatrix}. \quad (33)$$

The $\cos 3\theta$ and $\sin 3\theta$ terms can be eliminated by the freedom in choosing \mathcal{A} (there are 2 free parameters there). One should find an \mathcal{A} such that $g_{32} = g_{30}$ and $g_{31} = g_{33}$. The result is that all the third order terms will be eliminated, and

$$\mathbf{z}_{02} = 2 \begin{bmatrix} -g_{33} \\ g_{30} \end{bmatrix} \quad (34)$$

If we consider the fourth order terms, and write $B(J) = I + B_2J$, then the fourth order terms are

$$2J^2 [\cos\theta \quad -\sin\theta] B_2 \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} + 8J^2 [(g_{33}^2 - g_{30}^2)\cos 2\theta + 2g_{30}g_{33}\sin 2\theta]. \quad (35)$$

Thus, if

$$B_2 = 4 \begin{bmatrix} g_{30}^2 - g_{33}^2 & 2g_{30}g_{33} \\ 2g_{30}g_{33} & g_{33}^2 - g_{30}^2 \end{bmatrix}, \quad (36)$$

the fourth order terms are eliminated. Note that if the problem did not have the symmetry it does, there would be g_4 terms which would lead to nonzero $\cos 4\theta$ and $\sin 4\theta$ terms, and the fourth order terms could not be removed. Since the parabola is not perfectly symmetric, one might argue that it is important to include those fourth order terms from the broken symmetry. However, the symmetry breaking is small in the FFAG designs under consid-

eration, and for the large phase space we wish to transmit, the magnitude of the coefficients of the fourth order terms may be less than the magnitude of the coefficients of the fifth order terms by more than a factor of the magnitude of the phase space variables. If this is true, the symmetric approximation is still important. This paper will discuss only the symmetric case, and the asymmetric case will be addressed in a subsequent paper.

To determine the magnitude of the ellipse distortion, we will need to determine the change in the square of the radius to fifth order in the phase space variables. This requires computing $\mathcal{M}_C \mathbf{z}_1$ to fourth order, which is

$$\begin{aligned} & \frac{1}{6} [g_3, [g_3, [g_3, \mathbf{z}_1]]] + [g_5, \mathbf{z}_1] = \\ & \left(\begin{aligned} & -(5g_{30}^2 + g_{33}^2)g_{33}x^4 + 8(g_{30}^2 - g_{33}^2)g_{30}x^3p + 6(3g_{30}^2 - g_{33}^2)g_{33}x^2p + 16g_{30}g_{33}^2xp^3 + (3g_{33}^2 - g_{30}^2)g_{33}p^4 \\ & (g_{33}^2 - 3g_{30}^2)g_{30}x^4 - 16g_{30}^2g_{33}x^3p + 6(g_{30}^2 - 3g_{33}^2)g_{30}x^2p^2 + 8(g_{30}^2 - g_{33}^2)g_{33}xp^3 + (g_{30}^2 + 5g_{33}^2)g_{30}p^4 \end{aligned} \right) \\ & + \left(\begin{aligned} & -g_{51}x^4 - 2g_{52}x^3p - 3g_{53}x^2p^2 - 4g_{54}xp^3 - 5g_{55}p^4 \\ & 5g_{50}x^4 + 4g_{51}x^3p + 3g_{52}x^2p^2 + 2g_{53}xp^3 + g_{54}p^4 \end{aligned} \right), \quad (37) \end{aligned}$$

where I made use of the fact that $g_{30} = g_{32}$ and $g_{31} = g_{33}$. Including a fourth order term $\mathbf{z}_{40}J^2$ in $\mathbf{z}_0(J)$, the fifth order terms in Eq. (31) will be

$$\begin{aligned} \Delta r^2 = & 2(2J)^{5/2}(r_{50} \cos^5 \theta + r_{51} \cos^4 \theta \sin \theta \\ & + r_{52} \cos^3 \theta \sin^2 \theta + r_{53} \cos^2 \theta \sin^3 \theta + r_{54} \cos \theta \sin^4 \theta \\ & + r_{55} \sin^5 \theta) - \sqrt{8J^5} \mathbf{z}_{04}^T \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix}, \quad (38) \end{aligned}$$

where

$$r_{50} = g_{33}(3g_{30}^2 - g_{33}^2) - g_{51} \quad (39)$$

$$r_{51} = -3g_{30}(g_{30}^2 - 3g_{33}^2) + 2g_{52} - 5g_{50} \quad (40)$$

$$r_{52} = -2g_{33}(3g_{30}^2 - g_{33}^2) - 3g_{53} + 4g_{51} \quad (41)$$

$$r_{53} = -2g_{30}(g_{30}^2 - 3g_{33}^2) - 3g_{52} + 4g_{54} \quad (42)$$

$$r_{54} = -3g_{33}(3g_{30}^2 - g_{33}^2) + 2g_{53} - 5g_{55} \quad (43)$$

$$r_{55} = g_{30}(g_{30}^2 - 3g_{33}^2) - g_{54} \quad (44)$$

Rewrite this as a sum of terms linear in $\cos n\theta$ and $\sin n\theta$, and choose \mathbf{z}_{04} so as to eliminate the $\cos \theta$ and $\sin \theta$ terms:

$$\mathbf{z}_{04} = \frac{1}{2} \begin{pmatrix} -g_{51} - g_{53} - 5g_{55} \\ g_{54} + g_{52} + 5g_{50} \end{pmatrix}. \quad (45)$$

Now, finding the maximum value of Δr^2 is straightforward. Take the derivative of the fifth order terms in Δr^2 with respect to θ , divide by $\cos^5 \theta$ or $\sin^5 \theta$, and find the zeros of the polynomial in $\tan \theta$ or $\cot \theta$. Substitute the resulting θ values into fifth order terms in Δr^2 to get maximum value of Δr^2 .

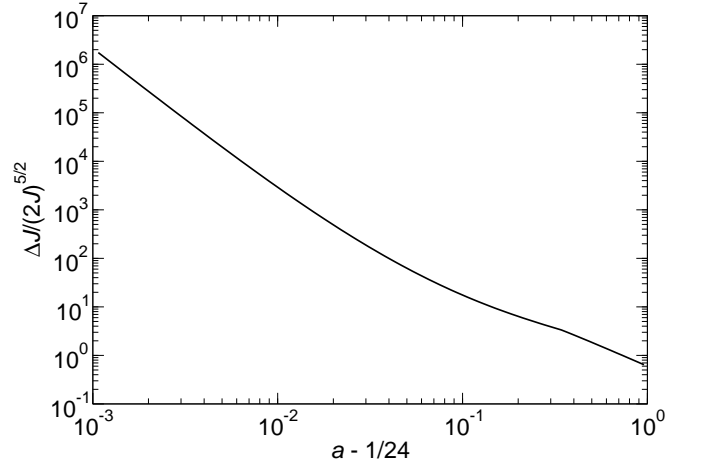


FIG. 8: Distortion of the ellipse, allowing the center and ellipse shape to change with J , minimized over b and \mathcal{A} .

The resulting minimum distortion is shown in Fig. 8, and the b corresponding to that minimum distortion is shown in Fig. 9. The distortion of the radius is proportional to the fourth power of the radius (it was proportional to the square of the radius if the J -dependent shift and ellipse shape change were not included) and for small a , is approximately proportional to $(a - 1/24)^{-3}$.

Depending on how the ellipse is viewed physically, the nonzero $\mathbf{z}_0(J)$ may imply that one is not accelerating all the way from $p = 0$ to $p = 1$. In that case, one can adjust a after the calculation to correct for this.

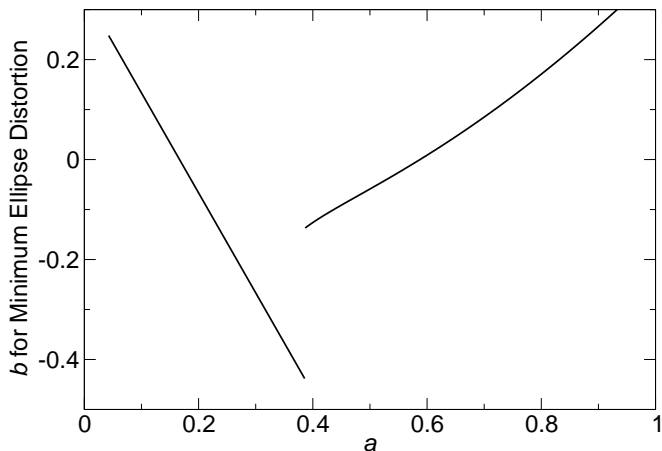


FIG. 9: Value for b which gives the minimum ellipse distortion when the center and ellipse shape are allowed to change with J .

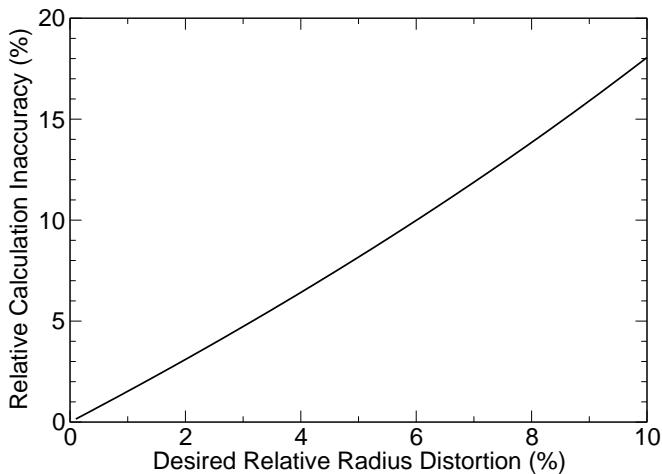


FIG. 10: Relative inaccuracy in the calculation of ΔJ vs. the desired value of $\Delta J/J$, for the boundary distortion without the amplitude-dependent linear transformation.

IV. ACCURACY OF THE CALCULATION

Since this calculation only keeps the lowest-order terms relevant to the desired result, there must be an inaccuracy in the estimate of the ellipse distortion and the emittance growth.

For the ellipse boundary distortion without the amplitude-dependent linear transformation, Fig. 10 shows the relative inaccuracy in the calculation. The calculation is done by choosing a value for a , choosing b and an ellipse orientation to give the minimum value for ΔJ , tracking 1001 points on the ellipse (spaced equally on a circle then transformed to the ellipse), calculating the actual value of ΔJ , and taking the maximum value. ΔJ is calculated from the tracking data based on the expected ellipse (i.e., I did not try to find an ellipse that best fit the tracking data). The calculation was done for a from

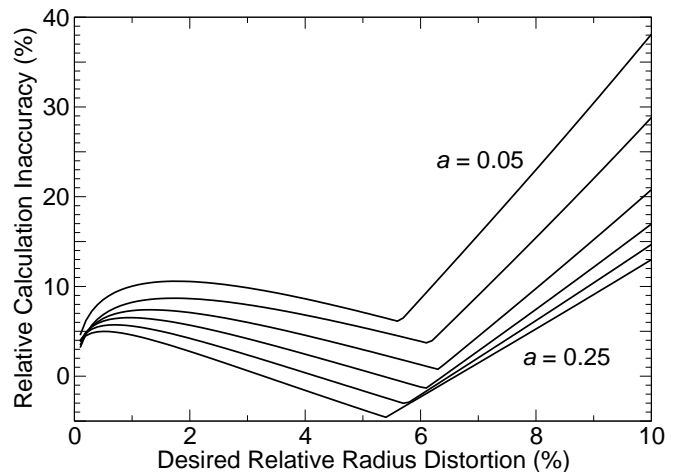


FIG. 11: Relative inaccuracy in the calculation of ΔJ vs. the desired value of $\Delta J/J$, for the boundary distortion with the amplitude-dependent linear transformation. Different curves are for different values of a , uniformly spaced from 0.05 to 0.25.

0.05 to 0.15, and the result was nearly independent of a , so only a single curve is shown in Fig. 10.

Figure 11 shows the result for the same calculation but for the calculation of the ellipse boundary distortion with the amplitude-dependent linear transformation. In this case, the results do depend on a , but the inaccuracies are qualitatively similar for different a . The sharp rise near zero is expected from the fact that $\Delta J/J \propto (2J)^{3/2}$, and the relative correction should be proportional to $(2J)^{1/2}$ (this also explains the linear behavior in Fig. 10). Despite the initial rapid rise, the inaccuracy levels off quickly and stays around 10% or lower (lower for larger a) until the desired inaccuracy gets to around 5–6%; at that point, the inaccuracy begins to increase very rapidly (less rapidly for larger a). As one is willing to tolerate a larger ellipse distortion, the higher order terms begin to become more important, until a point where they cause a rapid reduction in the accuracy of the estimate made here.

For emittance growth, the accuracy of the calculation is shown in in Fig. 12. The emittance growth is computed by integrating 10^5 particles distributed uniformly in θ and in J according to the distribution [19]

$$\frac{7}{9\epsilon} \left(1 - \frac{2J}{9\epsilon}\right)^{5/2}. \quad (46)$$

This distribution function leads to a one-dimensional distribution (in either x or p) which has no particles beyond 3σ . The inaccuracy in the calculation quickly becomes extremely large. This is due to the large-amplitude tails in the distribution, which are more strongly affected by the higher-order terms and contribute disproportionately to the emittance growth, as can be seen in Fig. 13. To reduce this effect, one can apply the following iterative procedure to the final distribution: find the second-order covariance matrix; remove all particles outside of a $k\sigma$ el-

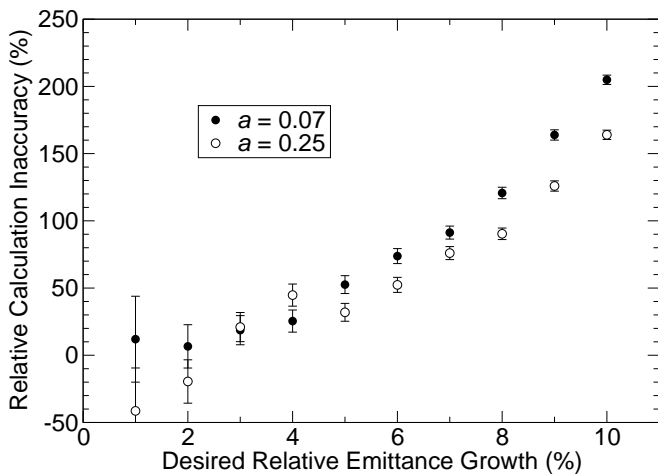


FIG. 12: Relative inaccuracy in the calculation of the emittance growth $\Delta\epsilon$ vs. the desired value of $\Delta\epsilon/\epsilon$, for two different values of a . Error bars show 1σ uncertainties due to the finite number of particles integrated.

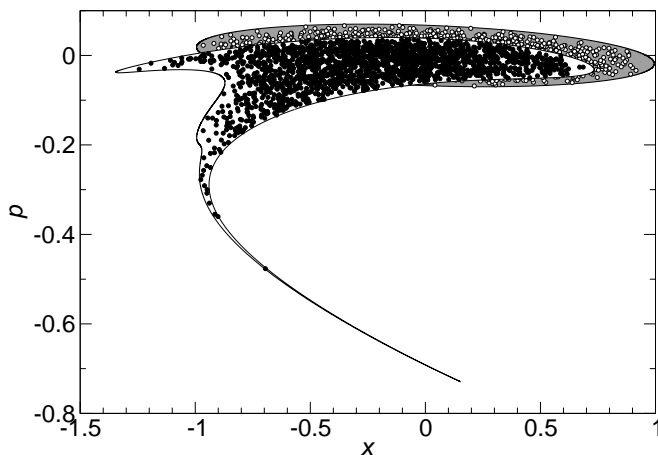


FIG. 13: Initial (white circles) and final (black circles) distributions, where the desired emittance growth is 10%. The grey ellipse denotes the 3σ boundary, and that same boundary for the final distribution is shown surrounding the black circles.

lipse, where k is a given constant; repeat these two steps until no particles are cut. The results of this cutting procedure are shown in Fig. 14. The value of k has an extremely strong effect on the result, and there does not seem to be a clear way to choose the best value for k . The poor accuracy of the emittance growth calculation without the cutting procedure, the lack of a good choice for the cutting amplitude, and the fact that for some distributions, the emittance growth calculation leads to a reduction in the calculated emittance, leads one to conclude that emittance growth does not seem to be a good criterion for the optimization of a machine design, except possibly when only very small emittance growths are desired (in the range of 1–2%) and for Gaussian or nearly Gaussian distributions.

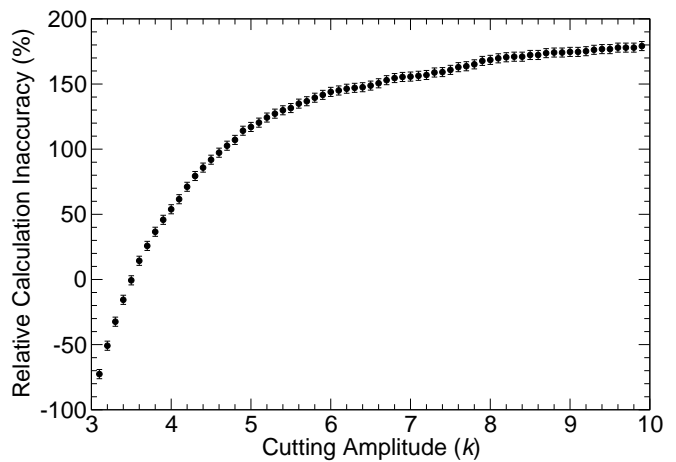


FIG. 14: Relative inaccuracy in the calculation of the emittance growth $\Delta\epsilon$ vs. cutting amplitude k , for $a = 0.10$ and a desired emittance growth of 10%. Error bars show 1σ uncertainties due to the finite number of particles.

V. OTHER CRITERIA FOR OPTIMIZATION

In real machines, there may be criteria other than just the longitudinal distortion that determine the optimal parameters to use.

For example, for small values of a , the minimum longitudinal distortion occurs when b takes on its minimum value, $1/3 - 2a$. One can compute the change in u (and thus the total arc length covered) in accelerating from $p = 0$ to $p = 1$ as a function of a and b ; it is the integral

$$\int_0^1 \frac{6 dp}{\sqrt{36a^2 - (2p-1)^2} [(2p-1)^2 - 3b]^2}. \quad (47)$$

This integral can be done in terms of an elliptic integral; it is

$$\frac{3}{(3(q_0 - b)(q_0 - 3b)^3)^{1/4}} F(\phi(1, q_0, b), m(q_0, b)) \quad (48)$$

where F is the incomplete Jacobi elliptic integral of the first kind, and

$$q_0 = [(3a + \sqrt{(3a)^2 - b^3})^{1/3} + (3a - \sqrt{(3a)^2 - b^3})^{1/3}]^2 \quad (49)$$

$$\phi(q, q_0, b) = \cos^{-1} \left(\frac{(q_0 - q)\sqrt{q_0 - 3b} - q\sqrt{3(q_0 - b)}}{(q_0 - q)\sqrt{q_0 - 3b} + q\sqrt{3(q_0 - b)}} \right) \quad (50)$$

$$m(q_0, b) = \frac{1}{2} - \frac{3}{4} \frac{q_0^2 - 6q_0b + 6b^2}{\sqrt{3(q_0 - b)(q_0 - 3b)^3}}. \quad (51)$$

Note that in the equation for q_0 , one should take the real solution of the cube root. A more restricted version of this result appears in [20].

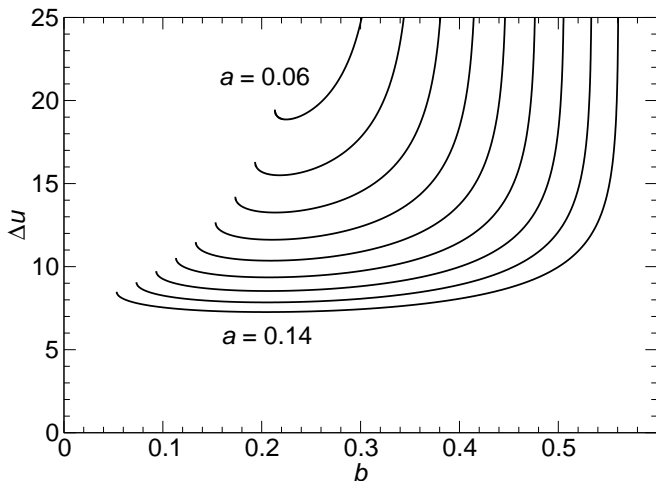


FIG. 15: Change in u as a function of b , for a evenly spaced between 0.06 and 0.14.

The change in u is shown in Fig. 15 as a function of b for various a . It is clear from that figure that the smallest change in u , and therefore the smallest time in the FFAG, does not occur when b is at its minimum value. For muon machines, where muon decay must be minimized, or for machine's where the machines repetition rate must be maximized, one may not want to use the minimum b , even though that does give the smallest longitudinal distortion for a given a . Using the machine parameters referred to in section II, the decay loss at the minimum b is about 8.3%, whereas choosing the minimum Δu results in a decay loss of about 7.7%.

Instead, one may choose a different figure of merit, which combines the longitudinal distortion with some other quantity. For instance, one could combine the decay loss with the longitudinal distortion by imagining that any area outside of the ellipse is loss. If there is only one harmonic in θ in the ellipse distortion, and higher-order terms in $\Delta J/(2J)$ are ignored, then the fractional loss can be written as

$$\frac{\Delta N_{\text{decay}}}{N} + \frac{2}{\pi} \frac{\Delta J}{2J}.$$

For this merit factor, for a 10–20 GeV muon FFAG, 201.25 MHz RF with an average gradient of 1.5 MV/m, and a normalized longitudinal acceptance (transmitted ellipse area divided by π) of 150 mm, Fig. 16 shows the b which minimizes the above fractional loss, and Fig. 17 shows the minimum loss as a function of a . I use the ellipse distortion with the amplitude-dependent shift of the center to compute this. For sufficiently small a , the minimum b is still optimal, since the distortion “losses” far exceed the decay losses. As a gets larger, however, the decay losses become more important, and the optimal b is no longer the minimum, as can be seen from Fig. 15.

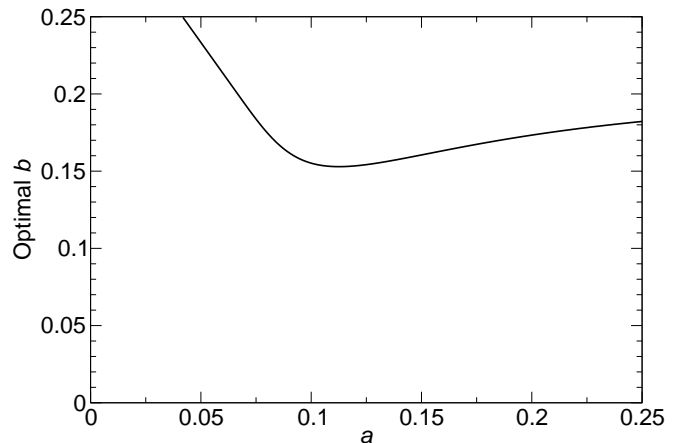


FIG. 16: Value of b which minimizes the fractional loss for the muon FFAG described in the text.

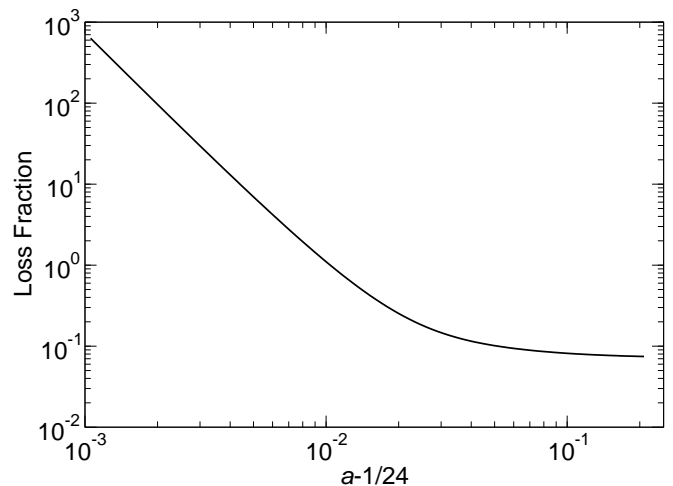


FIG. 17: Fractional loss as a function of a for the muon FFAG described in the text.

VI. CONCLUSIONS

I have described a method for quantifying the longitudinal distortion of a phase space ellipse in a dynamical system. A Dragt-Finn expansion is used to facilitate the lowest-order minimization of the distortion. The method is applied to linear non-scaling FFAGs, where longitudinal phase space distortion is an important design consideration. It yields a procedure for choosing important FFAG design parameters. Using emittance growth to characterize the distortion in this system appears to be problematic, whereas distortion of an ellipse seems to work well.

Finally, this method has broader applicability to single-pass systems, by changing the similarity transformation on the map by \mathcal{A} with separate initial and final transformations.

ACKNOWLEDGMENTS

This work has been supported by the United States Department of Energy, Contract No. DE-AC02-98CH10886.

-
- [1] K. R. Symon, D. W. Kerst, L. W. Jones, L. J. Laslett, and K. M. Terwilliger, *Phys. Rev.* **103**, 1837 (1956).
- [2] F. T. Cole, R. O. Haxby, L. W. Jones, C. H. Pruett, and K. M. Terwilliger, *Rev. Sci. Instrum.* **28**, 403 (1957).
- [3] D. W. Kerst, E. A. Day, H. J. Hausman, R. O. Haxby, L. J. Laslett, F. E. Mills, T. Ohkawa, F. L. Peterson, E. M. Rowe, A. M. Sessler, et al., *Rev. Sci. Instrum.* **31**, 1076 (1960).
- [4] M. Aiba, K. Koba, Y. Mori, R. Muramatsu, C. Ohmori, I. Sakai, Y. Sato, A. Takagi, R. Ueno, T. Yokoi, et al., in *Proceedings of EPAC 2000, Vienna, Austria* (2000), pp. 581–583.
- [5] T. Adachi, M. Aiba, K. Koba, S. Machida, Y. Mori, A. Mutoh, J. Nakano, C. Ohmori, I. Sakai, Y. Sato, et al., in *Proceedings of the 2001 Particle Accelerator Conference, Chicago*, edited by P. Lucas and S. Webber (IEEE, Piscataway, NJ, 2001), pp. 3254–3256.
- [6] M. Aiba, K. Koba, S. Machida, Y. Mori, A. Mutoh, J. Nakano, C. Ohmori, Izumi Sakai, Y. Sato, M. Sugaya, et al., in [21], pp. 1028–1030.
- [7] S. Machida, Y. Mori, A. Muto, J. Nakano, C. Ohmori, I. Sakai, Y. Sato, A. Takagi, T. Uesugi, A. Yamazaki, et al., in [22], pp. 3452–3454.
- [8] S. Machida, Y. Mori, A. Muto, J. Nakano, C. Omori, I. Sakai, Y. Sato, A. Takagi, T. Yokoi, M. Yoshii, et al., in *Proceedings of EPAC 2004, Lucerne, Switzerland* (EPAC, European Physical Society Accelerator Group, 2004), pp. 2643–2645.
- [9] C. Johnstone, W. Wan, and A. Garren, in *Proceedings of the 1999 Particle Accelerator Conference*, edited by A. Luccio and W. MacKay (IEEE, Piscataway, NJ, 1999), p. 3068.
- [10] F. E. Mills and C. Johnstone, in the transparency book for the 4th International Conference on Physics Potential & Development of $\mu^+ \mu^-$ colliders, San Francisco, CA (1999), UCLA, Los Angeles, CA, pp. 693–698.
- [11] G. H. Rees, *Isochronous, FFAG rings with insertions for rapid muon or electron acceleration* (2005), presentation at the 7th International Workshop on Neutrino Factories and Superbeams, Laboratori Nazionali di Frascati (Rome), June 21–26, 2005.
- [12] G. H. Rees, *Design of an isochronous FFAG ring for acceleration of muons* (2005), unpublished note.
- [13] J. S. Berg, in [21], pp. 1124–1126.
- [14] S. Koscielniak and C. Johnstone, in [22], pp. 1831–1833.
- [15] J. S. Berg, in *The International Workshop on FFAG Accelerators, October 13–16, 2004, KEK, Tsukuba, Japan*, edited by S. Machida, Y. Mori, and T. Yokoi (2005), pp. 1–8, preprint MUC-CONF-ACCELERATION-309, available from <http://www-mucool.fnal.gov/notes/notes.html>, URL http://hadron.kek.jp/FFAG/FFAG04_HP/.
- [16] J. S. Berg and C. Johnstone, in [22], pp. 2216–2218.
- [17] A. J. Dragt and J. M. Finn, *J. Math. Phys.* **17**, 2215 (1976).
- [18] A. J. Dragt and E. Forest, *J. Math. Phys.* **24**, 2734 (1983).
- [19] J. S. Berg and F. Ruggiero, in *Proceedings of the 1997 Particle Accelerator Conference*, edited by M. Comyn, M. K. Craddock, M. Reiser, and J. Thompson (IEEE, Piscataway, NJ, 1998), pp. 1712–1714.
- [20] S. Koscielniak, Tech. Rep. TRI-DN-05-12, TRIUMF, Vancouver, B.C., Canada (2005).
- [21] *Proceedings of EPAC 2002, Paris, France* (EPAC, European Physical Society Interdivisional Group on Accelerators, CERN, 2002).
- [22] J. Chew, P. Lucas, and S. Webber, eds., *Proceedings of the 2003 Particle Accelerator Conference* (IEEE, Piscataway, NJ, 2003).