Amplitude Dependence of Time of Flight and its Connection to Chromaticity

J. Scott Berg

Submitted to Nuclear Instruments and Methods in Physics Research A

July 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.
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.
Amplitude Dependence of Time of Flight and its Connection to Chromaticity

J. Scott Berg

Brookhaven National Laboratory; Building 901A; P.O. Box 5000; Upton, NY 11973-5000

Abstract

In general, the time of flight of a charged particle in an accelerator will depend on the particle’s transverse amplitude. This effect can become important in machines with large transverse emittances, such as muon accelerators. We will describe the effect, its physical origin, and the circumstances where it becomes important. We will then demonstrate that the effect is directly related to chromaticity. We will describe the effect on the longitudinal dynamics in various circumstances, including linacs and fixed field alternating gradient accelerators (FFAGs). We will describe methods for correcting the effect, particularly in FFAGs.

Key words: synchro-betatron coupling, longitudinal, transverse, chromaticity, FFAG, accelerator

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

1. Introduction

In 2003, Palmer [1] demonstrated that in a linear non-scaling FFAG, there was a dependence of the time of flight on the transverse amplitude. Since the effect was small compared to both the RF period and the variation in the time of flight itself, it was hoped that the effect would not be too significant. However, at the 2005 FFAG workshop at KEK, Machida [2] demonstrated that particles with large transverse amplitude were not accelerated in a linear non-scaling FFAG. He also showed the same time-of-flight dependence on the transverse amplitude.

In this paper, we will first demonstrate that the time-of-flight dependence on transverse amplitude is directly related to the chromaticity (transverse oscillation frequency dependence on energy) of the machine. While this is a fairly straightforward consequence of Hamiltonian dynamics, it has not generally been of great interest in accelerator studies because
- Transverse oscillation amplitudes tend to be small when compared to longitudinal oscillation amplitudes.
- Chromaticity is corrected to near zero in most circular machines.
- Longitudinal (synchrotron) oscillations will switch the places of late and early arriving particles, reducing the impact of the late arrival of a particle with large transverse amplitude.

This effect becomes interesting for muon machines, in particular, because
- Transverse emittances (areas in phase space) are starting to become comparable to longitudinal emittances: after cooling for a neutrino factory, for instance, they the transverse emittance is about 20% of the longitudinal.
- Many of the systems, such as the initial accelerating linac and the FFAGs for later acceleration,
cannot have their chromaticity corrected.
- Those same systems do not have significant synchrotron oscillations.

Past studies have primarily looked at the effect of the longitudinal motion on the transverse, in the context of synchro-betatron coupling and resonances [3,4]. This effect has been studied before, in the context of ionization cooling studies [5–7]. It manifested itself in a correlation between the transverse amplitude and the energy, as will be explained later in this paper. This paper connects the phenomenon with chromaticity, and applies the theory to a wider range of systems, where its effects are even stronger. In particular, the effects in non-scaling FFAGs are thoroughly examined.

2. Physical and Mathematical Description

In its simplest form, when magnets are linear, the time of flight depends on transverse amplitude because a trajectory oscillating about the closed orbit has a longer length than the closed orbit itself, as can be seen in Fig. 1. This arises from the nonzero angles that the oscillations make with respect to the closed orbit. For small displacements about the reference orbit, the length added to the orbit is proportional to the square of the orbit’s displacement. This makes the time of flight increase with increasing transverse amplitude.

To compute the effect in general, begin with the Hamiltonian for the accelerator without RF cavities,

\[ H(z, E, s) \]

where \( z \) is a four-dimensional vector of the transverse phase space coordinates, \( E \) is the energy, and \( s \) is the arc length along a reference curve that defines the coordinate system (and is the independent variable). We find the energy-dependent closed orbit for the transverse variables \( z_0(E, s) \), which is a solution for \( z \) of Hamilton’s equations of motion and, for a closed ring (or periodic lattice) where the length of the reference curve for a turn/period is \( L \), is a periodic function of \( s \) with period \( L \). We then compute a \( 4 \times 4 \) matrix \( A(E, s) \) and a function \( t_0(E, s) \), which are periodic in \( s \), such that after the transformation from \( z \) and \( t \) (the time) to \( \tilde{z} \) and \( \tilde{t} \) defined by

\[ z = A(E, s)\tilde{z} + z_0(E, s) \]

\[ t = \tilde{t} + t_0(E, s) + \frac{1}{2} \tilde{z}^T A^T S(\partial_E A)\tilde{z} \]

\[ + \tilde{z}^T A^T S \partial_E z_0 + \frac{1}{2} z_0^T S \partial_E z_0, \]

the Hamiltonian becomes

\[ \frac{2\pi \nu(E) \cdot J_n}{L} + H_T(E) + O(J_n^{3/2}). \] (3)

where \( J_n \) is the two-dimensional vector of action variables, defined by \( J_n,1 = (\tilde{z}_1^2 + \tilde{z}_2^2)/2, J_n,2 = (\tilde{z}_3^2 + \tilde{z}_4^2)/2, \) and \( S \) is the 4-dimensional symplectic metric

\[ S = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \] (4)

I have introduced the notation \( \partial_E \), which is the partial derivative of the following function with respect to \( E \). The subscript \( n \) is an indication that I am not using the usual action variables with dimensions of length, but instead I have not divided the Hamiltonian and the momenta by the total momentum or a reference momentum as is usually done, so the dimensions are \( J_n \) are energy-time. \( J_n \) should be nearly invariant as one accelerates. \( \nu \) is the tune as generally defined.

Note that the difference between \( t \) and \( \tilde{t} \) is a periodic function of \( s \). Thus, any time advance from one turn/period to the next will be governed by the Hamiltonian (3). From Hamilton’s equations of motion

\[ \frac{d\tilde{t}}{ds} = -\partial_E H_T - \frac{2\pi(\partial_E \nu) \cdot J_n}{L} + O(J_n^{3/2}). \] (5)

This is the fundamental theoretical result of this paper: that the time of flight depends on the transverse amplitude, and the lowest order dependence is directly related to the tune variation with energy, also known as the chromaticity.

The result is qualitatively consistent with the original description of the phenomenon: namely that increased orbit lengths for particles with large transverse amplitudes increases the time of flight. Since for a lattice with no chromaticity correction, the tune decreases with increasing energy, Eq. (5) indicates that the time of flight will increase with increasing transverse amplitude, consistent with the expected behavior. Furthermore, Eq. (5) is linear in \( J_n \), and therefore quadratic in the transverse displacement, again as expected geometrically.

But interestingly, correction of chromaticity is able to reduce or eliminate the dependence of time of flight on transverse amplitude. Sextupoles are introduced into a machine at points with nonzero dispersion to correct chromaticity [8–10]. One wants
to increase the tune with increasing energy to correct the chromaticity. For this argument, assume a FODO cell with a positive dispersion. To increase the horizontal chromaticity, one needs a focusing strength at the focusing quadrupole (primarily) which increases with momentum. This requires a vertical field whose derivative increases with increasing position, since the dispersion is positive. This implies a sextupole which has a vertical field that is parabolic in the midplane and positive. Now, consider oscillations in that field. Since the field is quadratic, the average field seen by a particle undergoing a nonzero amplitude horizontal oscillation is positive. This acts like an increase in the average dipole field, resulting in an average radius of the orbit which is reduced. This reduces the length of the orbit, and thus the time it takes to move along the orbit. The average field coming from the sextupole is proportional to the square of the maximum horizontal displacement in the oscillation. Thus, correcting horizontal chromaticity results in a reduced time of flight for large transverse amplitudes due to the chromatic correction sextupoles, and consistent with Eq. (5).

3. Applications to Accelerator Systems

We now examine the effect of the amplitude dependence of the time of flight on various accelerator subsystems, all of which appear in most muon acceleration systems. We will consider motion in a stationary RF bucket, acceleration in a linac, and acceleration in a non-scaling FFAG. We will obtain some rough numerical results for muon systems with reasonable parameters to give an idea of the significance of the effect.

3.1. Stationary RF Bucket

For a stationary RF bucket, one first wants to find the fixed point about which one oscillates. This can be done using Eq. (5), by finding the solution where \( \frac{d\bar{t}}{ds} = \frac{h}{(f_{RF}L)} \), where \( f_{RF} \) is the RF frequency, and \( h \) is the harmonic number. Say that \( E_0 \) is the energy where \( \partial \nu H_T = \frac{h}{(f_{RF}L)} \). Expanding \( \partial \nu H_T \) to first order near \( E_0 \), one finds that for small \( J_n \), the fixed point energy is

\[
E_0 - \frac{2\pi(\partial \nu) \cdot J_n}{L \partial^2 \nu H_T},
\]

where all the derivatives are evaluated at \( E_0 \). In terms of the usual quantities, this is

\[
E_0 + \frac{2\pi \xi \cdot J_n}{T_0 \eta},
\]

where \( T_0 \) is the time of flight along the closed orbit at energy \( E_0 \), \( \xi \) is the chromaticity, defined such that the tune at energies near \( E_0 \) is given by \( \nu + \xi(\Delta p/p) \), with \( p \) being the total momentum, and \( \eta \)
is the frequency slip factor. A similar result holds for a non-stationary bucket.

It seems curious that the effect of the time of flight depending on amplitude is to have the fixed point energy change with amplitude. The reason is that the particles always oscillate about a point which is synchronized with the RF frequency. Thus, the fixed point comes about by adjusting the energy so as to adjust the time of flight for particles with a finite transverse amplitude so that they are synchronized with the RF. Therefore, for a group of particles arriving at the same time and energy at the zero-amplitude fixed point, the particles with large transverse amplitudes will start to oscillate about the large-amplitude fixed point, effectively increasing their longitudinal emittance. The entire RF energy change with amplitude. The reason is that depending on amplitude is to have the fixed point for a non-stationary bucket.

Let’s take the example of a muon cooling lattice [11]. The chromaticity is about 0.24, and \( \eta = -1/\gamma^2 \), with \( \gamma \) being the ratio of the total energy to the muon’s rest mass energy. The cell length is about 75 cm. The maximum normalized transverse amplitude at the beginning of the channel is about 60 mm (which is \( 2J_n/mc \) where \( m \) is the muon mass and \( c \) is the speed of light). The resulting closed orbit shift at that amplitude and a reference momentum of 220 MeV/c is about 31 MeV. Considering that the full energy spread of the beam extracted from the cooling channel is about \pm 46 MeV/c, this is significant. Those who have studied ionization cooling have been aware of this effect [5–7], and have found it important to correctly place an amplitude-energy correlation into a beam before it enters a cooling channel.

3.2. Linac

Consider a linac where the particles are relativistic enough that synchrotron oscillations can be neglected. This is true for muons once they have reached a few hundred MeV. Say that the tune per cell in a linac is adjusted to be independent of the momentum as the beam is accelerated (if it is not, see the next subsection). If one assumes a constant accelerating gradient, then a particle with nonzero transverse amplitude arrives at a time differing from that of a zero amplitude particle by

\[
-\frac{2\pi}{\Delta E} \ln \left( \frac{p_f}{p_i} \right) \xi \cdot J_n, \tag{8}
\]

where \( p_i \) is the initial momentum and \( p_f \) is the final momentum (for the entire length of linac), and \( \Delta E \) is the energy gain per cell, assuming that \( \xi \) is the chromaticity for that same cell.

Taking, for example, the last section of the low energy acceleration linac from [11], which accelerates muons from around 520 MeV to 1500 MeV, considering particles with a normalized transverse amplitude of 30 mm, and assuming a chromaticity of -0.25, the large transverse amplitude particles are behind the zero transverse amplitude particles by 0.4 ns. For 201.25 MHz RF, this is about 30° of phase. This is comparable to the bunch length that we are accelerating. Thus, one should take these high amplitude particles into account when designing the linac, to try to ensure that their energy gain is not so different from the low amplitude particles. This will be complicated, however, by the desire to ensure that the all the particles in the longitudinal distribution also gain roughly the same energy.

Increasing the energy gain per cell reduces the time of flight variation with transverse amplitude, essentially because the number of cells the particles traverse is inversely proportional to the energy gain per cell, and the time of flight variation is essentially proportional to the number of cells traversed.

3.3. FFAGs

FFAGs behave in many ways like a linac, but there are some important differences. First, because an FFAG is a multiturn machine, the tune of each cell is not in general (except for scaling FFAGs) adjusted to be the same, so Eq. (8) is not quite correct. Secondly, the RF dynamics can be more complicated [12–15].

The time of flight difference for large amplitude particles can be computed, assuming a uniform accelerating gradient, to be

\[
-\frac{2\pi (\Delta \nu) \cdot J_n}{\Delta E}, \tag{9}
\]

where \( \Delta E \) is the energy gain per cell, and \( \Delta \nu \) is the change in the tune per cell from the initial to the final energy. Just as in a linac, the effect is reduced if the amount of acceleration per cell is increased. The effect can also be improved by reducing the range of tunes over the acceleration range. One in principle has much greater control over the tunes in an FFAG than in a linac, since there is bending and therefore the opportunity to correct chromaticity. In fact, the scaling FFAG [16] allows one to in principle elimi-
nate the tune variation with energy entirely, eliminating the time of flight variation with transverse amplitude to lowest order. For various reasons not discussed here, however, there are reasons to consider FFAGs where the tune does vary with energy, in particular linear non-scaling FFAGs [17,18].

Let’s consider the example of a 10–20 GeV linear non-scaling FFAG for muons, with a transverse normalized amplitude of 30 mm, an energy gain of 12.75 MeV per cell, and a tune range of 0.21. This results a 0.55 ns time slip over the acceleration cycle for a high amplitude particle when compared to a zero amplitude one. For 201.25 MHz RF, this corresponds to a 40° phase slip.

In addition, FFAGs tend to use multiple stages to perform their acceleration. Thus, a second stage will tend to increase the time difference even more. If one could perform half a synchrotron oscillation between stages, then in principle the low and high amplitude particles would switch places and would come back together at the end of two stages. However, constructing such a synchrotron oscillation would require an additional ring (or arc) with RF voltage comparable to that required for a single FFAG stage, and thus becomes extremely costly.

One could attempt to correct the effect by creating a positive chromaticity in the transfer line between the FFAG stages. One can compute the required chromaticity for the transfer line using Eqs. (5) and (9):

$$\xi = -\frac{\beta^2 E}{\Delta E} \Delta \nu,$$

where $\beta$ is the reference velocity divided by the speed of light, $E$ is the reference energy, $\Delta E$ and $\Delta \nu$ are as they were in Eq. (9), and $\xi$ is the chromaticity for the entire beam line. The difficulty is the ratio $E/\Delta E$, which is approximately the number of turns in the FFAG times the number of cells. The transfer line cannot expect to create a chromaticity per cell that is larger than $-\Delta \nu$, and thus the number of cells in the transfer line would need to be extremely large for this to work.

3.3.1. Chromaticity Correction of the FFAG

Of course, one could attempt to correct the chromaticity in the FFAG itself. However, in doing so, one quickly runs up against what makes a linear non-scaling FFAG work well: it is able to tolerate the tune variation despite passing through low-order resonances because

- the magnets are highly linear, so any resonances are driven very weakly,
- acceleration is rapid, so resonances are passed through very quickly, and
- The lattice is highly symmetric, consisting entirely of short, simple, identical cells.

Adding nonlinearities has the potential to violate the first of these conditions. In fact, any non-scaling lattice with a significant amount of nonlinearity thus far constructed has failed to have sufficient dynamic aperture, at least for acceleration of muon beams that are only modestly cooled [19–21]. Nonetheless, one might hope that the effect can at least be reduced by adding a modest amount of chromaticity, while simultaneously keeping a sufficiently large dynamic aperture (with purely linear magnets, the dynamic aperture is generally much larger than necessary, even with the large dynamic aperture required).

Figure 2 shows the zero-amplitude tune coverage over the energy range of a non-scaling FFAG for various levels of chromaticity correction. A lattice is first constructed which is below the $3\nu_x = 1$ line, and whose tune is equidistant from the $3\nu_x = 1$ and $\nu_x - \nu_y = 0$ lines at the low energy end and the $\nu_x - 2\nu_y = 0$ line at the high energy end. This
avoids the resonances that would be directly driven by the sextupole once we add it, as well as avoiding the linear coupling resonance. Higher horizontal tunes are generally preferred for linear non-scaling lattices [22], so one is making a sacrifice by staying below the $3\nu_x = 1$ resonance, but adding sextupoles will strongly drive that resonance [23,11], so it is necessary.

We then add sextupole and adjust the dipole and quadrupole fields in the lattice such that the tunes at the low and high energies are given by

$$
\nu_{lo}(x) = (1 - x/2)\nu_{lo}(0) + (x/2)\nu_{hi}(0)
$$

$$
\nu_{hi}(x) = (x/2)\nu_{lo}(0) + (1 - x/2)\nu_{hi}(0),
$$

where $x$ is the fraction by which the chromaticity is corrected. Figure 2 shows the tune coverage for values of $x$ up to 0.5. Note that it is really the difference between the low and high energy tunes that are being corrected, not the chromaticity at a given energy. Doing so would have required adding higher order multipoles, which would have then driven resonances (such as $\nu_x = 1/4$) directly that we have not avoided, leading to potential losses of dynamic aperture.

Tracking by Machida [24] indicates that the dynamic aperture is acceptable up to $x$ around 0.3. One does have particle losses when crossing the $4\nu_x = 1$ resonance, but the lattice parameters can be adjusted to stay just below that. For higher amounts of chromaticity correction, the dynamic aperture is reduced at all energies, so avoiding resonances will not improve the situation. If a higher level of chromaticity correction were possible, one could even think about running at a large horizontal tune by staying entirely above the $3\nu_x = 1$ line, but the limited level of chromaticity correction that seems to be possible precludes that. One cannot create multiple sextupole families to try to improve the dynamic aperture since every cell must be identical (to the extent possible) to avoid introducing even more resonances.

3.3.2. FFAG Phase Space Dynamics

Further complications arise when one considers the longitudinal phase space dynamics of a linear non-scaling FFAG. The time of flight in a linear non-scaling FFAG can be made to be a nearly symmetric parabola [25] when the particles are highly relativistic. This leads to particles moving through a serpentine channel through phase space, as shown in Fig. 3 [12–15]. Figure 4 shows what the time of flight as a function of energy looks like for both zero amplitude particles and particles with a larger transverse amplitude. The RF frequency (or the cell length) is adjusted so that particles with a certain time of flight are synchronized with the RF. Particles can only be accelerated over the desired energy range for a limited range of values for this synchronization time, which is indicated for the zero transverse amplitude particles in Fig. 4 [15]. In fact, the volume of phase space transmitted is better for certain values for that synchronization time than others. Since particles with a nonzero transverse amplitude have a different time of flight as a function of energy, the optimal synchronization time will be different for large amplitude particles than for zero
amplitude particles. In fact, it may be that there is no value for the synchronization time for which both low amplitude and high amplitude particles will be accelerated.

Another way to look at the problem is to examine the longitudinal phase space. Figure 5 shows the longitudinal phase space through which particles of low amplitude will be accelerated, and the phase space through which particles of high amplitude will be accelerated, for one sample set of parameters. Notice the limited overlap at low energies. Thus, only a very limited range of phases in this example can be accelerated, much more limited than the range of phases one would have expected by only looking at low amplitude particles. If the zero amplitude particles alone were used to determine the initial conditions, many of the high amplitude particles would not be accelerated at all, which is precisely what is seen in [2]. Furthermore, since the initial conditions that overlap cannot be centered in the channel at either amplitude, there will be much greater longitudinal distortion of the bunch than one would have if one were centered in the channel and only had particles with small transverse amplitudes. However, this is ameliorated by the fact that in most cases, one only needs a small longitudinal acceptance for particles with large transverse amplitudes.

Finally, as discussed before, particles with different amplitudes clearly arrive at different times, even though they start out at the same point in longitudinal phase space. This can be seen from Fig. 5. Particles that start out in the overlapping region at low energy will arrive with a phase close to zero if they have low transverse amplitudes, but will arrive near a phase of $\pi/2$ if they had high transverse amplitudes, since the particles move along lines that are more or less parallel to the left and right separatrices. In fact, due to differing times (really numbers of cells) to traverse the paths, they will also probably arrive with different energies as well.

3.3.3. Addressing the Problem in FFAGs

To address the amplitude dependence of the time of flight in non-scaling FFAGs, one should probably apply a number of methods, each of which individually will make the problem somewhat better, and hopefully together will make the machine perform acceptably:

- Introduce sextupole components into the magnets to partially correct the chromaticity.
- Create positive chromaticity in the transfer line. This is less likely to give the dynamic aperture problems found in the FFAG since one is no longer required to make every cell identical, since one only need handle the energy spread of the beam in the transfer line.
- Introduce second or third harmonic RF into the FFAG. This will reduce the energy spread in the particles that comes from the different longitudinal phase space dynamics that particles with different transverse amplitudes have.
- Increase the amount of voltage per cell. This helps both because the time of flight variation with transverse amplitude is reduced (Eq. (9)), but also because the phase space channel for the particles gets larger, increasing the overlapping area for different transverse amplitudes [15]. Optimized machine designs often left empty cells because doing so reduced the magnet aperture and therefore the cost [26]. It appears more important to correct the problem described here. One could even use two-cell cavities rather than single cell to improve the behavior further, but this could potentially increase costs significantly.

4. Conclusions

We have demonstrated that the lowest order dependence of the time of flight on transverse amplitude is directly related to the chromaticity of a machine. We have used that fact to quickly estimate its effect in several accelerator systems. In particular, we have shown why it is important in non-scaling FFAGs, and have explored possible methods of addressing the problem.
References


