FN-695

PARTICLE POPULATION INSIDE A TIGHT STATIONARY BUCKET

K.Y. Ng

Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510

K.M. Fung and S.Y. Lee

Physics Department, Indiana University, Bloomington, IN 47405

(July, 2000)

Abstract

The population of a bunch in a rf bucket is discussed, and the general method is presented. In order that the bunch matches the bucket, the population must be according to the invariant tori inside the bucket.

---

1 INTRODUCTION

To do simulation in the longitudinal phase space, one usually starts with the random population of macro-particles inside a longitudinal bucket, according to some phase-space distribution, for example, bi-Gaussian, elliptical, etc. Such population is easy for a small bunch inside a large bucket, because the invariant tori near the center of the bucket are closely in the form of ellipses. The story is quite different for a large bunch inside a tight bucket. This is because the tori deviate very much from ellipses, especially those in the vicinity of the separatrices.

Consider the normalized Hamiltonian (see Appendix A)

$$H(\phi, p) = \frac{p^2}{2} + (1 - \cos\phi) \;, \tag{1.1}$$

where[†] Φ is the rf phase coordinate ranging from $-\pi$ to π and P is the normalized energy offset or the momentum canonical to Φ. It is easy to see that the stable fixed point (SFP) is at $(\phi, p) = (0, 0)$, while the unstable fixed points (UFP) are at $(\phi, p) = (\pm\pi, 0)$. The separatrix has Hamiltonian value or *energy* value $H = 2$ and forms a bucket with half height $P = 2$. For small ϕ,

$$H(\phi, p) = \frac{p^2}{2} + \frac{\phi^2}{2} \;, \tag{1.2}$$

which describes just simple harmonic motion. In this normalized representation, the tori near the SFP or the center of the bucket are essentially circles. In other words, when the bunch is small enough, the rms rf phase σ_ϕ and the rms momentum σ_p are equal. This is definitely not true when the bunch is large enough to nearly fill up the whole bucket, because the bucket has a half width π but a half height of only 2. As an example, we populate a large bunch with $\sigma_\phi = 0.75$ rad randomly with 20000 macro-particles using bi-Gaussian distribution according to the Hamiltonian in Eq. (1.2). First, we find, as shown in Fig. 1, many particles fall outside the momentum aperture of the bucket, although nearly no particles fall outside the $(-\pi, \pi)$ rf phase limits. One method is to neglect all those particles outside the bucket and consider only those inside. This, however, alters the linear density of the bunch and the rms

---

[†]In this article, we use the convention in probability and statistic that a random variable is represented by a capitalized symbol, while the value it assumes is in the lower case. For example, the energy offset variable is denoted by P and the value it assumes by p.
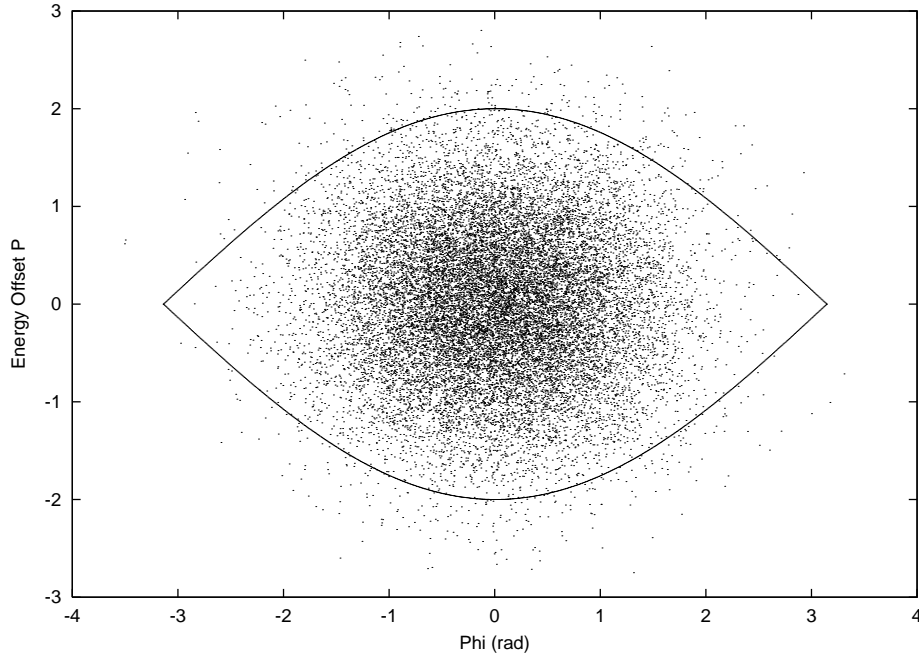
Figure 1: Population of 20000 macro-particles in a bi-Gaussian distribution with rms spread equal 0.75 rad for both the phase and energy, according to Eq. (1.2). Notice that quite a number of particles fall outside the bucket, and also the population does not fit the bucket.

phase spread is not the same as the one originally demanded. Another method is to truncate the bi-Gaussian distribution during the population at the bucket height. This will certainly ensure all particles to be inside the bucket and the rms phase spread will be what we desire. The bunch distribution, however, does not match the bucket. After several synchrotron oscillations, it will filament and lengthen the rms phase spread.

The above undesirable outcomes occur because we have populated the bunch according to ellipses which are not trajectories of the particle motion in the longitudinal phase space. In order that the populated bunch distribution matches the bucket, we must populate according to the invariant tori inside the bucket instead. We will present the correct way of particle population in Sec. 4 after we review some basic theory of random population.

2 RANDOM POPULATION IN ONE DIMENSION

Consider a *random variable X*. When X assumes the value x, it has the *probability density function $f(x)$*. In other words,

$$\Pr\left(x < X \le x + dx\right) = f(x)\,dx \tag{2.1}$$

and

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \ , \tag{2.2}$$

where Pr denotes probability. The above still applies when X has a finite range, because we can always define $f(x)$ to vanish outside that range. The probability of having $X \le x$ is denoted by

$$F(x) = \int_{-\infty}^{x} f(x')dx' \ , \tag{2.3}$$

and is called the *cumulative probability distribution function*. Obviously $F(X)$ is also a random variable, which we can denote by Y. As an illustration, the probability density function and its cumulative probability distribution function for the standard normal are plotted in Figs. 2(a) and 2(b). Here, we are going to revisit the very important theorem of probability integral transform [1].

Theorem: The random variable $Y = F(X)$, where F is a cumulative probability distribution function, is a *uniform random variable* on [0,1].

Proof: Because $F(X)$ is a cumulative probability distribution function, it has the range [0,1]. From Fig. 2, the outcome of $Y \le t$ is the same as the outcome of $X \le x$, where $F(x) = t$. Hence,

$$\Pr(X \le x) = \Pr(Y \le t) = F(x) = t \ . \tag{2.4}$$

This theorem provides a very helpful way to populate a random variable X according to its probability density function $f(x)$. The procedures are as follows:
(1) Compute the corresponding cumulative probability distribution function $F(x)$,
(2) Generate a uniform random number t on [0,1],
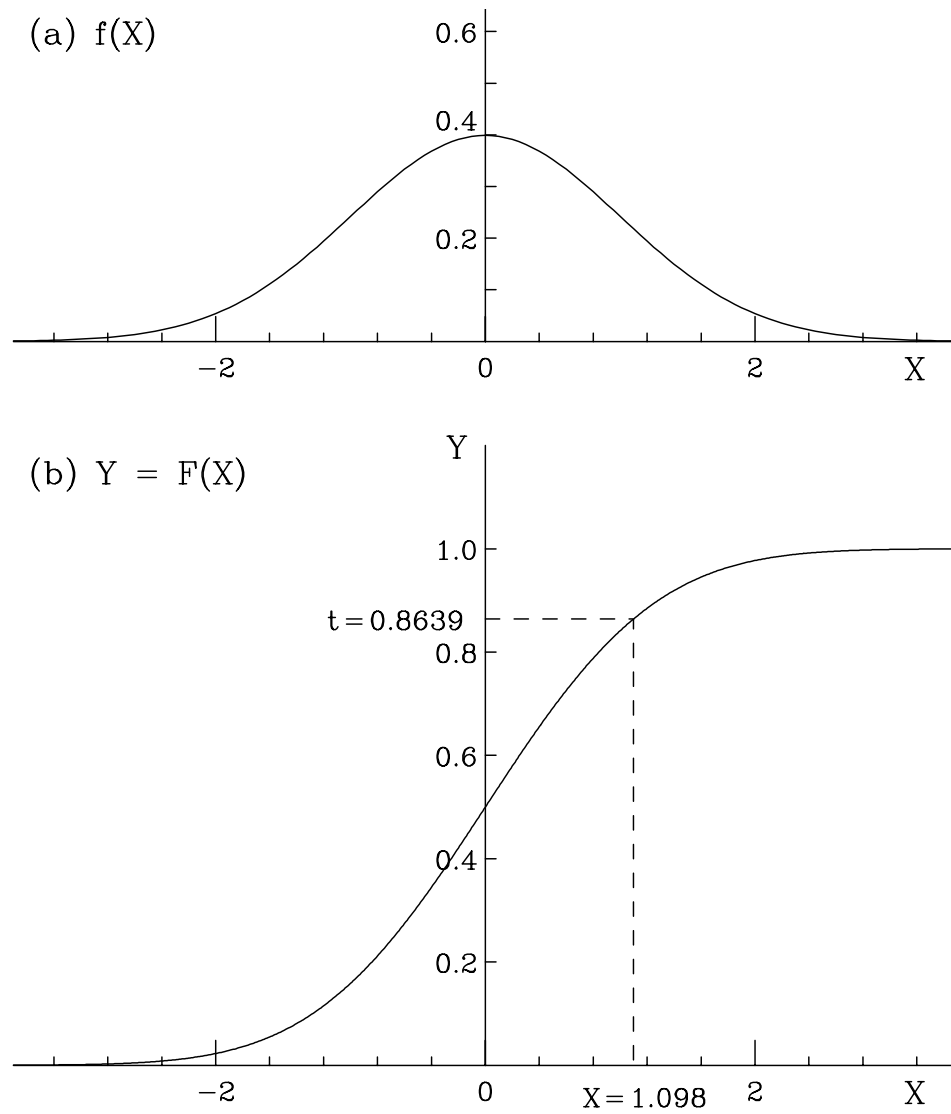(3) From Fig. 2, read off the value of X via $F(x) = t$.

Figure 2: (a) Probability density function $f(x)$ for the standard normal. (b) Cumulative probability distribution function $Y = F(x)$ for the standard normal. For a uniform random number $t = 0.8639$, the random variable X assumes the value $X = 1.098$.

The last step is equivalent to finding the inverse function of $F(x)$, because, for each uniform random number t on [0,1], the value X assumed is

$$x = F^{-1}(t) \ . \tag{2.5}$$

Consider, for example, the Lorentz density function

$$f(x) = \frac{a}{\pi} \frac{1}{x^2 + a^2} \ . \tag{2.6}$$

The cumulative probability distribution function is

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \frac{x}{a} \ , \tag{2.7}$$

and the inverse is

$$F^{-1}(t) = -a \cot(\pi t) \ . \tag{2.8}$$

Thus, the random variable X can be obtained from

$$x = -a \cot(\pi t) \ , \tag{2.9}$$

by generating uniform random numbers t on [0,1]. As another example, the parabolic probability density function

$$f(x) = \begin{cases} \dfrac{3}{4a} \left(1 - \dfrac{x^2}{a^2}\right) & |x| \le a \\ 0 & |x| > a \end{cases} \tag{2.10}$$

has the cumulative distribution function

$$F(x) = \frac{1}{2} + \frac{3x}{4a} - \frac{x^3}{4a^3} \ , \tag{2.11}$$

which is a cubic. For a given uniform random number t on [0,1], compute

$$\theta = \sin^{-1} \sqrt{t} \tag{2.12}$$

in the first quadrant. The inverse of the Eq. (2.11) is given by

$$\begin{cases} x_1 = 2a \cos \dfrac{2\theta}{3} \ , \\ x_2 = 2a \cos \left(\dfrac{2\theta}{3} + \dfrac{2\pi}{3}\right) \ , \\ x_3 = 2a \cos \left(\dfrac{2\theta}{3} - \dfrac{2\pi}{3}\right) \ , \end{cases} \tag{2.13}$$

whichever is in the range $[-a, a]$. In general, only one of the three solutions will be within this range, except for the rare situation that two of the solutions are equal and are equal to $+a$ or $-a$.

Unfortunately, not every cumulative probability distribution function has an inverse in a simple analytic closed form. For the Gaussian density function, the cumulative distribution function can be written in terms of the error function, although the integration cannot be performed analytically. Its inverse, however, cannot be written as an analytic expression.

However, with high-speed computers, this is not a problem at all. There are standard routines to compute Eq. (2.5) numerically. Essentially, these routines compute the percentiles of the density function $f(x)$ using a combination of trapezoidal and Gaussian quadratures to a very high accuracy. A uniform random number t on $[0,1]$ (for example, 0.8643, see Fig. 2) generated will be assigned to one of the percentiles (here, the 86th percentile) and the corresponding value of the random variable X in between two adjacent percentiles (here, the 86th and 87th) is computed using 4-point interpretation (here $X = 1.0980$ for a normal distribution).

3 EXTENSION TO TWO DIMENSIONS

The above procedure of population can be extended to two dimensions. The key is the polar coordinates. In the first example, X and Y are two independent random variables having bi-Gaussian probability density function

$$f(x,y) = \frac{1}{2\pi a_x b_x} \exp\left[-\frac{x^2}{2a_x^2} - \frac{y^2}{2a_y^2}\right] . \tag{3.1}$$

Now go to the polar coordinates (r, θ) with

$$\begin{aligned} x &= a_x r \cos\theta , \\ y &= a_y r \sin\theta . \end{aligned} \tag{3.2}$$

The density function is transformed into

$$f(x,y)\, dx\, dy = \frac{g(r)}{2\pi}\, d\theta\, dr , \tag{3.3}$$

with the radial density function

$$g(r) = r e^{-r^2/2} . \tag{3.4}$$

The polar density is uniform and θ can therefore be integrated out. Now we are in one dimension again. The cumulative radial probability distribution function is

$$G(r) = \int_0^r r' e^{-r'^2/2}\, dr' = 1 - e^{-r^2/2} \,, \tag{3.5}$$

from which the inverse can be derived easily,

$$r(t) = G^{-1}(t) = \sqrt{-2\ln(1-t)} \,. \tag{3.6}$$

Thus, for every set of two uniform random numbers t_1 and t_2 on [0,1], the random variables X and Y assume the values

$$\begin{aligned} x &= a_x r \cos(2\pi t_2) \,, \\ y &= a_y r \sin(2\pi t_2) \,, \end{aligned} \tag{3.7}$$

where

$$r = \sqrt{-2\ln t_1} \,. \tag{3.8}$$

In fact, the bi-Gaussian density function is only *semi-two-dimensional* in the sense that the two random variables X and Y are statistically independent. In other words, the density function is separable, or it is the product of the density functions of X and Y:

$$f(x,y) = \left[\frac{1}{\sqrt{2\pi}a_x} e^{-x^2/(2a_x^2)} \right] \left[\frac{1}{\sqrt{2\pi}a_y} e^{-y^2/(2a_y^2)} \right] \,. \tag{3.9}$$

In that case, polar coordinates do not present any significant advantage, because we can populate X and Y independently. The situation is different for the elliptical distribution

$$f(x,y) = \frac{3}{2\pi a_x a_y} \sqrt{1 - \frac{x^2}{a_x^2} - \frac{y^2}{a_y^2}} \,, \tag{3.10}$$

which gives parabolic density function when projected onto the x-axis or y-axis. Certainly, this density function cannot be written as the product of the density functions of X and Y. So we proceed as before by going to the polar coordinates (r, θ) given by Eq. (3.2). The density function of Eq. (3.10) is transformed into Eq. (3.3) with

$$g(r) = 3r\sqrt{1-r^2} \,, \tag{3.11}$$

which is now one-dimensional. The corresponding cumulative probability distribution function is

$$G(r) = \int_0^r 3r'\sqrt{1-r'^2}\, dr' = 1 - \left(1-r^2\right)^{3/2} \,, \tag{3.12}$$

with the inverse

$$r(t) = G^{-1}(t) = \sqrt{1 - (1-t)^{2/3}} \; . \tag{3.13}$$

Thus, for every set of two uniform random numbers t_1 and t_2 on $[0,1]$, the random variables X and Y assume

$$x = a_x r \cos(2\pi t_2) \; ,$$
$$y = a_y r \sin(2\pi t_2) \; , \tag{3.14}$$

where

$$r = \sqrt{1 - t_1^{2/3}} \; . \tag{3.15}$$

4 POPULATION INSIDE A BUCKET

We return to the Hamiltonian of Eq. (1.1) with the random variables Φ and P. To ensure that the populated bunch fits the bucket, the probability density function must be a function of the Hamiltonian, or

$$f(\phi, p) = f[H(\phi, p)] \; , \tag{4.1}$$

where[‡] the form of the function $f(H)$ is arbitrary. Perform the transformation

$$q = 2 \sin \frac{\phi}{2} \; , \tag{4.2}$$

so that

$$f(H) \, d\phi \, dp = h(q, p) \, dq \, dp \; , \tag{4.3}$$

where

$$h(q, p) = \frac{f(H)}{\sqrt{1 - \dfrac{q^2}{4}}} \; . \tag{4.4}$$

The denominator on the right side arises from the transformation Jacobian. We next adopt the polar coordinates defined by

$$p = r \cos \theta \; ,$$
$$q = r \sin \theta \; . \tag{4.5}$$

[‡]Mathematically, the functions on left and right sides of Eq. (4.1) are different and should be represented by different symbols. However, as physicists, we use the same symbol to represent the two functions which assume the same value.

Now the range of r and the range of θ are independent of each other, although the two random variables are not statistically independent. We can therefore project onto r by integrating over θ and obtain the radial density function $g(r)$,

$$g(r)\,dr = f[H(r)]\,rdr\int_{-\pi}^{\pi}\frac{d\theta}{\sqrt{1-\dfrac{r^2}{4}\sin^2\theta}} = 4r\,f[H(r)]K(\tfrac{r}{2})\,dr \;, \qquad (4.6)$$

where $K(\tfrac{r}{2})$ is the complete elliptical function of the first kind. We now arrive at a density function in one dimension. The cumulative probability distribution function is

$$G(r) = \int_0^r 4r'\,f[H(r')]K(\tfrac{r'}{2})\,dr' \;. \qquad (4.7)$$

Given a uniform random number on [0,1], the value of r can be found as indicated in Sec. 2 and 3. Knowing r, we need to determine θ which does not have a uniform density. In fact, from Eq. (4.6), the polar density function is

$$v(\theta) = \frac{1}{4K(\tfrac{r}{2})}\frac{1}{\sqrt{1-\dfrac{r^2}{4}\sin^2\theta}} \;, \qquad (4.8)$$

the physical meaning of which will be studied in Appendix B. The corresponding cumulative distribution function

$$V(\theta) = \frac{1}{4K(\tfrac{r}{2})}\int_0^\theta \frac{d\theta'}{\sqrt{1-\dfrac{r^2}{4}\sin^2\theta'}} \qquad (4.9)$$

is just the elliptical integral and the inverse can be expressed in terms of one of the Jacobian elliptic functions. Here, we choose the sine-like sn and arrive at

$$\theta(t) = V^{-1}(t) = \sin^{-1}\text{sn}\left[4tK(\tfrac{r}{2}), \tfrac{r}{2}\right] \;, \qquad (4.10)$$

where the arcsine function is considered to be a multi-valued function assigning values to all the 4 quadrants (or $0 \le \theta < 2\pi$). Alternatively, we can also consider

$$\tilde{\theta}(t) = V^{-1}(t) = \left|\sin^{-1}\text{sn}\left[tK(\tfrac{r}{2}), \tfrac{r}{2}\right]\right| \;, \qquad (4.11)$$

so that $0 \le \tilde{\theta} < \tfrac{1}{2}\pi$. The angle θ is then determined by assigning $\tilde{\theta}$ to one of the 4 quadrants according to which quarter t belongs to.

For illustration, we take a bi-Gaussian density function

$$f(p, \phi) = f\left[-\frac{H}{a^2}\right] = A \exp\left[-\frac{1}{2a^2}\left(p^2 + 4\sin^2\frac{\phi}{2}\right)\right] , \qquad (4.12)$$

where A is the normalization constant and a is a parameter, which designates the rms spreads of both the random variables Φ and P when $a \ll 1$. Transforming to the polar coordinates, the density function for r is

$$g(r) = 4ArK(\tfrac{r}{2})\,e^{-r^2/(2a^2)} , \qquad (4.13)$$

where

$$\begin{cases} p = r\cos\theta , \\ \phi = 2\sin^{-1}\dfrac{r\sin\theta}{2} . \end{cases} \qquad (4.14)$$

In other words, given two uniform random numbers t_1 and t_2 on [0,1], the random variables Φ and P assume the values according to Eq. (4.14), where r is a complicated function of t_1 and is obtained numerically as indicated in Sec. 2 via Eq. (4.13), while θ is obtained from t_2 and r via Eq. (4.11).

A number of comments are in order:

(1) The random variables Φ and P are not statistically independent of each other because of the restriction of the bucket boundary. In the normalized Hamiltonian, the restriction is

$$p^2 + 4\sin^2\frac{\phi}{2} = p^2 + q^2 = r^2 \leq 4 . \qquad (4.15)$$

Thus r has the range of [0,2] in the probability density function $g(r)$ and the cumulative probability distribution function $G(r)$. The finite range of r sets a correlation between Q and P, so that the density function cannot be fractorized into the product of the density functions of Q and P. Thus, one cannot address the distributions of Q and P separately. In fact, this is the reason why the problem complicates. Let us examine the difficulties of performing the population using the variables Φ and P directly. We need to integrate over P from $-2\cos\frac{\phi}{2}$ to $+2\cos\frac{\phi}{2}$ to obtain the density function of Φ,

$$f_\phi(\phi) = \sqrt{2\pi}aA\,\mathrm{erf}\left(\frac{\sqrt{2}\cos\frac{\phi}{2}}{a}\right)\exp\left(-\frac{2\sin^2\frac{\phi}{2}}{a^2}\right) , \qquad (4.16)$$

where erf is the error function. After populating numerically the value of $\Phi = \phi$ between $\pm\pi$ according to the density function of Eq. (4.16) with a uniform random

number on [0,1], we can populate the value of P according to the density function

$$
f_p(p) = \begin{cases} \left[\sqrt{2\pi}\,a\,\mathrm{erf}\left(\dfrac{\sqrt{2}\cos\frac{\phi}{2}}{a} \right) \right]^{-1} \exp\left(-\dfrac{p^2}{2a^2} \right) , & |p| \le 2\cos\frac{\phi}{2} , \\ 0 & \text{otherwise} . \end{cases} \tag{4.17}
$$

However, there are two disadvantages. First, notice that for both variables, the inverse of the cumulative distribution functions can not be expressed in the closed form. Worst of all, the distribution for P is different for different values of Φ. This implies that the inverse of a new function has to be searched for every single macro-particle and a lot of computer time will be required. Second, the fact that the integration over p can be expressed in terms of the error function is not typical. In general, a functional expression like Eq. (4.16) is not possible, meaning that the density function for Φ becomes completely numerical and the population of macro-particles will become more tedious and complicated. It is clear that with R and Θ as variables, these two difficulties can be avoided. First, an analytic expression for the density function of R can always be written as Eq. (4.6) independent of the distribution $f(H)$. This is because the integration over θ can always be performed and is, in fact, the same for all distributions $f(H)$. Certainly, an important reason is the fact that the range of θ is r-independent. Second, the distribution of Θ, Eq. (4.9), is also the same for all distributions $f(H)$. Best of all, the inverse of the cumulative distribution of Θ, Eq. (4.10), can be expressed in the closed form, thus eliminating the search of inverse numerically and can speed up the population process.

(2) Usually the linear rms spread σ_ϕ of the bunch is given but not the parameter a, which approaches the rms spread only when the bunch is small. Therefore, a conversion between σ_ϕ and a is necessary. Although we know that

$$
\sigma_\phi^2 = \int \phi^2 f(\phi,p)\, d\phi\, dp = \int \phi^2 A \, \exp\left[-\frac{1}{2a^2}\left(p^2 + 4\sin^2\frac{\phi}{2} \right) \right] d\phi\, dp , \tag{4.18}
$$

unfortunately, the above integral cannot be carried out in the closed form. However, this double integral can be reduced to a single integral by integrating over p giving something similar to Eq. (4.16). This integral can now be computed numerically and the σ_ϕ-a relation is depicted in Fig. 3. Another more general method is to perform the double integral numerically or to populate the bunch for a given a and compute σ_ϕ numerically from the phase locations of the macro-particles. Notice that there is
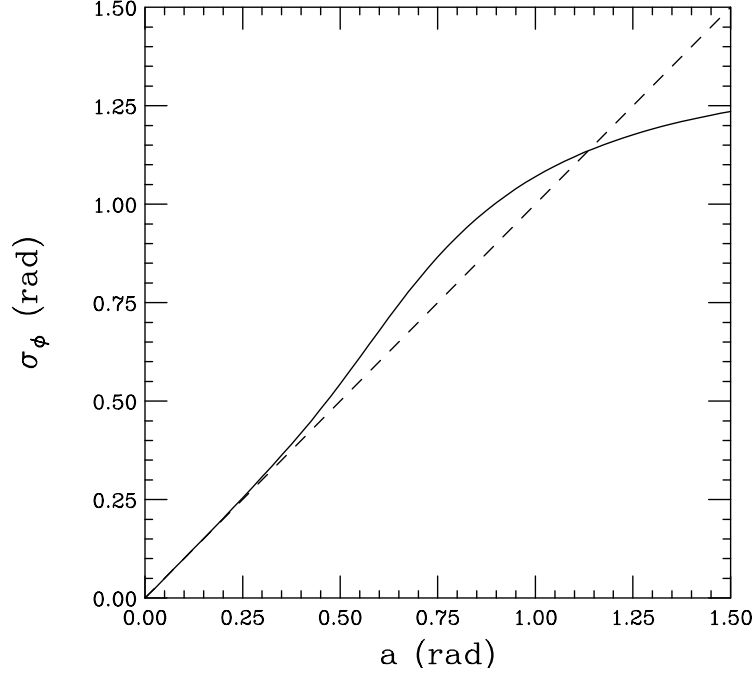
Figure 3: Plot showing the relation between rms phase spread σ_ϕ and the parameter a of a bunch having a probability density function given by Eq. (4.12). The parameter a approaches the rms phase spread of the distribution only when the bunch is small compared with the bucket. The dashed line is at 45° showing the situation if $\sigma_\phi = a$.

an upper bound for σ_ϕ. This is because as $a \to \infty$, the population inside the bucket becomes uniform. It is easy to show that this upper bound is $2\sqrt{\frac{1}{4}\pi^2 - 2} = 1.3673$.

Now, given a σ_ϕ, we can read off the corresponding a either from interpolation or a fitted relationship. Then, the population can be performed accordingly. Figure 4 shows such a population with $\sigma_\phi = 0.75$ rad (or $a = 0.653$) in the normalized coordinates. This is to be compared with the population in Fig. 1. It is clear that no particles ever fall outside the bucket now and the bunch does fit the bucket nicely.
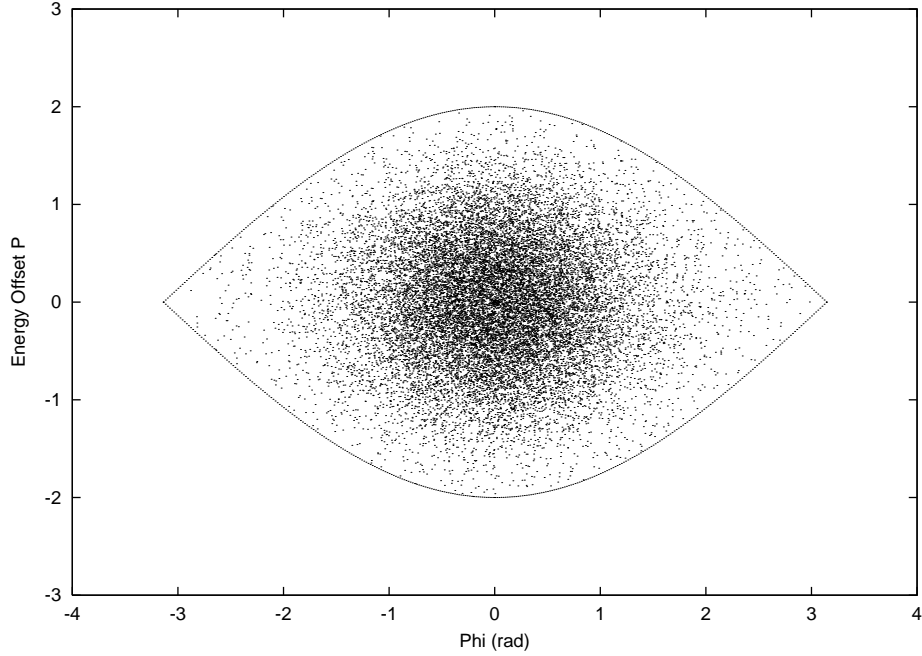
Figure 4: Population of 20000 macro-particles in a bi-Gaussian-like distribution with rms phase spread equal 0.75 rad (or $a = 0.653$), according to probability density of Eq. (4.12). Compared with Fig. 1, the bunch fits the bucket and no particles fall outside the bucket.

APPENDIX

A NORMALIZED HAMILTONIAN

When a particle traverses the rf cavity gap, it sees the rf phase ϕ and has a fractional momentum spread δ. Usually, it takes many revolutions to complete one synchrotron oscillation. It is therefore reasonable to introduce a continuous independent *time* variable ϑ', which advances by 2π per revolution. The equations of motion are

$$\frac{d\phi}{d\vartheta'} = h\eta\delta \ , \tag{A.1}$$

$$\frac{d\delta}{d\vartheta'} = -\frac{eV_{\rm rf}\,{\rm sgn}(\eta)}{2\pi\beta^2 E}\sin\phi \ , \tag{A.2}$$

where $V_{\rm rf}$ is the peak rf voltage, h the rf harmonic, e the particle charge, η the slip parameter, E the energy of the synchronous particle, and β its velocity with respect to

the velocity of light. We have restricted the discussion to the situation of a stationary bucket only.

Introduce a normalized momentum offset p, such that

$$p = \frac{h\eta}{\nu_s}\delta \ , \tag{A.3}$$

where

$$\nu_s = \sqrt{\frac{h|\eta|eV_{\mathrm{rf}}}{2\pi\beta^2 E}} \tag{A.4}$$

is the small-amplitude synchrotron tune. The equations of motion are transformed into

$$\frac{d\phi}{d\vartheta'} = \nu_s p \ , \tag{A.5}$$

$$\frac{dp}{d\vartheta'} = -\nu_s \sin\phi \ . \tag{A.6}$$

When p is considered the momentum conjugate to ϕ, the equations of motion are derivable from the Hamiltonian

$$H = \tfrac{1}{2}\nu_s p^2 + \nu_s(1 - \cos\phi) \ . \tag{A.7}$$

Finally, we obtain the Hamiltonian of Eq. (1.1) by changing the independent variable from ϑ' to $\vartheta = \nu_s\vartheta'$, where ϑ advances by 2π per small-amplitude synchrotron period.

B   DENSITY ALONG TORUS

We wish to have an understanding of the particle density along a torus, which is illustrated as solid in Fig. 5. The location of a particle on that torus is determined by the distance s along the torus from the P-axis to the arrow head. It can also be described by the angle θ defined in Eq. (4.5), which is measured from the P-axis. Here, we will concentrate on the particle density along this particular torus only. An element ds of length along the torus at distance s is related to the element $d\theta$ by

$$ds = \sqrt{dp^2 + d\phi^2} = \sqrt{\frac{1 - \dfrac{r^2}{4}\sin^4\theta}{1 - \dfrac{r^2}{4}\sin^2\theta}}\, r\, d\theta \ , \tag{B.1}$$

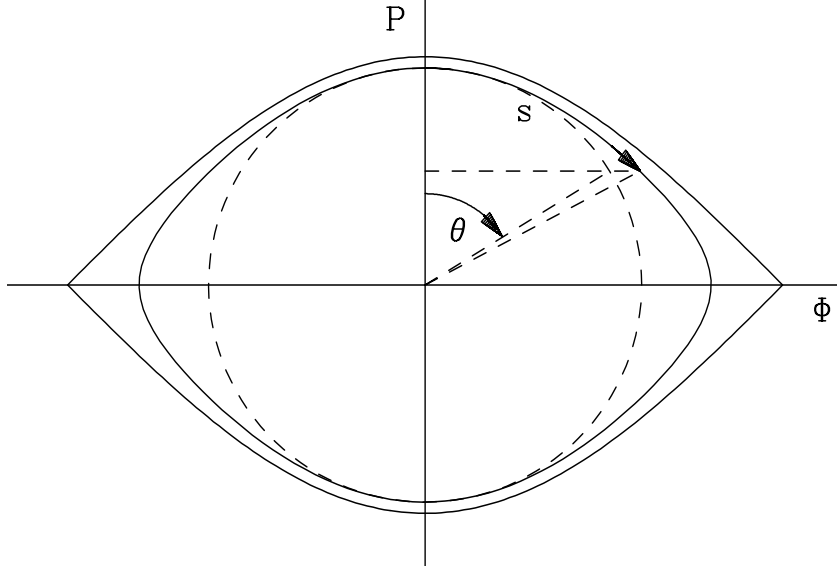where $\tfrac{1}{2}r^2$ is the Hamiltonian value of the torus.

Figure 5: Plot showing the distance s along a torus from P-axis to the arrow head as a function of the angle θ defined in Eq. (4.5) measured from the P-axis. The auxiliary dashed circle of radius r is drawn for the definition of θ.

The number of particles in a length ds along the torus is $\tilde{\rho}(s)ds$. To ensure that the particle density $\tilde{\rho}(s)$ is stationary or does not change with time, we demand

$$\text{div}\left[\tilde{\rho}(s)\vec{v}(s)\right] = 0 \ , \tag{B.2}$$

where $\vec{v}$ is the velocity of particle. This is the velocity at which the particle performs synchrotron rotation and is in the direction along the torus. To be stationary, $\tilde{\rho}(s) \, ds/d\vartheta$ must be s- or θ-independent. In other words, we must have

$$\tilde{\rho}(s) \propto \frac{d\vartheta}{ds} \ . \tag{B.3}$$

Here ϑ is the independent time variable of the Hamiltonian of Eq. (1.1) as defined and derived in Appendix A. Recall that ϑ advances by 2π per synchrotron period for particles on a torus with $r \to 0$. The time for a particle to move a distance ds along the torus, or for a corresponding change $d\theta$ or $d\phi$, can be obtained through the phase equation of the Hamiltonian like Eq. (A.5):

$$d\vartheta = \pm \frac{d\phi}{r\sqrt{1 - \dfrac{4}{r^2}\sin^2\dfrac{\phi}{2}}} \ , \tag{B.4}$$

where the positive sign applies when the rotation is in the first and fourth quadrants of the longitudinal phase space and the negative sign for the other two quadrants. Transforming from ϕ to θ using Eqs. (4.2) and (4.5), we arrive at

$$d\vartheta = \frac{d\theta}{\sqrt{1 - \frac{r^2}{4}\sin^2\theta}} \,. \tag{B.5}$$

Thus, the velocity along the torus is

$$v = \frac{ds}{d\vartheta} = r\sqrt{1 - \frac{r^2}{4}\sin^4\theta} \,. \tag{B.6}$$

Since the velocity along the torus is position dependent, this explains why the particle density cannot be uniform along the torus. Expressing in terms of θ, the density becomes [using $\rho(\theta)\,d\theta = \tilde{\rho}(s)\,ds$]

$$\rho(\theta) = \tilde{\rho}(s)\frac{ds}{d\theta} \propto \frac{d\vartheta}{d\theta} \,, \tag{B.7}$$

which is the same expression in Eq. (4.8).

References

[1] See any book on probability and statistics, for example, J.R. Blum and J.I. Rosenblatt, *Probability and Statistics*, W.B. Saunders, 1972.