Fermi National Accelerator Laboratory

TM-1475

Primer on Beam Dynamics in Synchrotrons*

L. Teng
Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

August 1987

*To be published in *Physics of High Energy Colliders*, edited by A. Chao, H.T. Edwards, and M. Month

Operated by Universities Research Association Inc. under contract with the United States Department of Energy

PRIMER ON BEAM DYNAMICS IN SYNCHROTRONS

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

1 INTENSITY-INDEPENDENT DYNAMICS

We discuss first the factors controlling the beam dynamics when the beam intensity, hence the self-field generated by the beam, is negligible. In this case the motions of the particles in the beam are independent and we have the so-called single-particle dynamics. The single-particle dynamics is clearly controlled only by external electromagnetic fields and external physical barriers such as beam collimators and vacuum chamber walls. The transverse motion is controlled principally by the magnetic field, and the longitudinal motion is controlled by the radiofrequency electric field. These motions are discussed in more detail below.

1.1 Linear Transverse Motion

The closed orbit, hence the overall geometry of the beam, is determined by the dipole field on the orbit. The closed orbit is generally unique for a given particle momentum. The dipole field is usually designed to be vertical on the closed orbit and to have a high sector-periodicity. This gives a planar closed orbit with a sector-periodic geometry. Since particles travel in a narrow beam the effects of the magnetic field are most simply discussed by expanding the field in powers of the transverse coordinates x(horizontal) and y(vertical) measured from the closed orbit. The coefficients defining these multipole fields are given by

$$B_y + iB_x = B_o \Sigma_n (b_n + ia_n)(x + iy)^n \qquad (1)$$

where all quantities are functions of the distance s along the closed orbit, B_o is the vertical field on orbit ($x = y = 0$), and b_n and a_n are respectively the normal and skew $2(n + 1)$-pole coefficients. So defined $b_o = 1$ and $a_o = 0$ for error-free field. The quadrupole fields specified by b_1 and a_1 produce linear focusing actions. The particles are guided by the quadrupole field to oscillate stably (betatron oscillations) about the closed orbit. Generally the quadrupole field is designed such that the skew component $a_1 = 0$ and the normal component b_1 has the same high orbital sector-periodicity. The horizontal and vertical betatron oscillations are then uncoupled, each given by

$$x(\text{or } y) \propto \sqrt{\beta}\, e^{\pm i\int ds/\beta} \qquad (2)$$

where $\beta = \beta(s)$ having the sector-periodicity is called the amplitude function. The number of oscillations per revolution or the tune, ν, is defined by

$$\nu = \frac{1}{2\pi}\int_0^{2\pi R} ds/\beta \tag{3}$$

where $2\pi R$ is the circumference of the closed orbit

Errors in the dipole field cause closed orbit distortions. Closed orbit distortions are undesirable because they reduce the effective aperture of the ring. Vertical dipole field errors are produced by construction errors of the dipole magnets and horizontal misalignments of the quadrupole magnets; horizontal dipole field errors are produced by roll errors of the dipole magnets and vertical misalignments of the quadrupole magnets. Closed orbit distortions can be corrected by realigning the magnets or by using steering dipoles.

Errors in the quadrupole field cause distortions in the amplitude function β. Amplitude distortions also reduce the effective aperture of the ring but their magnitude is generally smaller than that due to closed orbit distortions and correction is seldom necessary.

Both the closed orbit distortion and the amplitude distortion blow up on resonances. It is easy to see that if the tune is an integer, m, orbits with oscillation are also closed, and hence the closed orbit is no longer unique. Any dipole error with harmonic m will drive an arbitrarily large closed orbit distortion. It is less transparent but equally suggestive that if the oscillation is closed in two revolutions (2ν = integer = m) quadrupole errors with harmonic m will drive an arbitrarily large amplitude. Since driving error fields are unavoidable, in either the case of integer or half-integer resonance the motion becomes unstable. The half-integer resonance has, in fact, a finite width $\Delta\nu$ within which the motion is always unstable. This "stopband" width is given by

$$\Delta\nu = m^{th} \text{ harmonic amplitude of } \left(\frac{\beta}{4}\frac{b_1}{\rho}\right)$$

$$= c_2 \text{ defined by Eq. (11) below} \tag{4}$$

where $\rho = \rho(s)$ is the radius of the closed orbit. The coupled linear resonances $\nu_x \pm \nu_y$ = integer = m are driven by the mth harmonic of the skew quadrupole field given by a_1. The sum resonance $\nu_x + \nu_y$ = m also has a finite width stopband.

1.2 Momentum Effects

A ring magnet lattice is capable of confining particles over a limited range of momentum, each particle traveling about the closed orbit corresponding to its own momentum. For planar orbits the orbits of different momenta are separated horizontally. The orbit displacement per relative momentum increment, $\Delta p/p$, is called the dispersion function D, which has, of course, also the sector-periodicity and is given by

$$D \equiv \frac{\Delta x}{\Delta p/p} = \frac{1}{2\nu \sin \pi\nu} \int_s^{s+2\pi R} \frac{1}{\rho} \cos[\nu\pi + \phi(s) - \phi(r)]dr \qquad (5)$$

where $\phi(s) = \int ds/\beta$ is the betatron oscillation phase. The relative circumference increment per $\Delta p/p$ is called the momentum compaction factor α and is given by

$$\alpha = \frac{\Delta R/R}{\Delta p/p} = \frac{1}{2\pi R} \int_o^{2\pi R} \frac{D}{\rho} ds \qquad (6)$$

The amplitude function is also dependent on momentum. The relative change in the amplitude function β per $\Delta p/p$ is given by

$$\frac{\Delta\beta/\beta}{\Delta p/p} = \frac{1}{2\sin 2\pi\nu} \int_s^{s+2\pi R} \frac{b_1}{\rho} \beta \cos 2[\nu\pi + \phi(s) - \phi(r)]dr \ . \qquad (7)$$

Integrating around the closed orbit gives the momentum dependence of the tune. The tune change per $\Delta p/p$ is called the chromaticity, ξ, and is given by

$$\xi \equiv \frac{\Delta\nu}{\Delta p/p} = -\frac{1}{2\pi} \int_o^{2\pi R} \frac{\Delta\beta/\beta}{\Delta p/p} \frac{ds}{\beta} = -\frac{1}{4\pi} \int_o^{2\pi R} \frac{b_1}{\rho} \beta ds \ . \qquad (8)$$

When the momentum deviation $\Delta p/p$ is too large, the dispersion may cause the particle to strike the horizontal aperture or the chromaticity may run the tune onto resonance values. Thus both the dispersion and the chromaticity act to define the momentum aperture of the ring.

1.3 Nonlinear Transverse Motion

Nonlinear fields are introduced either deliberately or inadvertently through errors and beam-beam interactions in colliders. Sextupole field is introduced at places where the dispersion function D is large to modify or compensate for the natural chromaticity. Dispersion puts orbits with different momenta at different horizontal positions in the sextupole field, hence under the actions of different quadrupole fields. This introduces an additional momentum dependence of the tune which may be adjusted to compensate for the natural chromaticity. To modify the chromaticities independently in both transverse planes we need two sets of sextupoles placed at locations with greatly different ratios β_y/β_x and hence having very different effects on the horizontal and the vertical chromaticities.

Octupoles must be introduced when half-integer is employed for slow extraction.

In any case, the dynamics of the beam particles is, in fact, always nonlinear. With nonlinearities, the tunes ν_x and ν_y are amplitude dependent. As the amplitudes grow the tunes will encounter a succession of resonances. The 4-dimensional phase space is thus crisscrossed with intersecting surfaces in the shape of tori on which the motion has resonant tune values. All resonances, linear and nonlinear, may be summed up in the formula

$$j\nu_x \pm k\nu_y = m \tag{9}$$

where j, k, and m are positive integers or zero. Each resonance is excited by the mth harmonic of the 2n-pole field where $n \equiv j + k$ is the "order" of the resonance. The linear integer and half-integer resonances thus have orders 1 and 2 respectively. Nonlinear resonances are of orders ≥ 3. Those with $j = n, n - 2, n - 4, \cdots$ are excited by the normal field b_{n-1} and those with $j = n-1, n - 3, \cdots$ are excited by the skew field a_{n-1}. For example, those excited by norm$\nu_x \pm 2\nu_y = m$, and $2\nu_x \pm \nu_y = m$, $3\nu_y = m$.

To understand or visualize the features of nonlinear motions it $\sqrt{\beta}\, x$ and $\sqrt{\beta}\, x' - (\beta'/2)(x/\sqrt{\beta})$ (prime means d/ds). Consider the turn-to-turn mapping curve of an oscillation with small amplitude tune ν_o not on a resonance. In the normalized coordinates the mapping curve of a small linear oscillation is just a circle. For large oscillation, as the tune approaches a resonance, say $n\nu = m$, the mapping curve takes on the shape of a regular n-sided polygon with rounded corners. All stable phase points are contained in the central stable region, which is an area bounded by an n-cornered figure formed by the separatrices of the resonance. The corners are the unstable fixed points. To first approximation the radial distance of the unstable fixed points from the origin is given by

$$r_n = \left(\frac{\left| \nu_o - \frac{m}{n} \right|}{C_n} \right)^{\frac{1}{n-2}} \tag{10}$$

where C_n is the resonance driving harmonic amplitude given by

$$C_n \cos\left(m \frac{s}{R} - \text{"phase"}\right) = \frac{\beta_x^{\frac{n}{2}}}{2^n} \frac{b_{n-1}}{\rho} . \tag{11}$$

To this lowest order of approximation the separatrices can also be expressed in simple analytical forms.

This stable region defines the dynamic aperture. Outside this region the motion is at least locally unstable. The stable area is somewhat smaller than $\pi(r_n)^2$ and goes to zero as ν_o approaches the resonant value m/n. The tune deviation $|\nu_o - m/n|$ for which the stable area is just enough to contain the beam, namely just equal to the emittance of the beam (see Section 2 below) is defined as the half-width of the resonance. In practical cases this first-order

single degree-of-freedom picture always gives resonance widths much smaller than measured values, indicating that more precise two degrees-of-freedom computations are needed. Clearly the dynamic aperture has to be larger than the beam, but, depending on the magnitude of the driving term and the separation of the small amplitude tune from resonance, the dynamic aperture may be smaller than the physical aperture. In this case the stable particle motion is limited by the dynamic aperture.

Nonlinear motions in two coupled degrees of freedom have the same behavior but are more complicated and more difficult to visualize. In the 4-dimensional phase space the mapping points of oscillations not on resonances lie on closed 2-dimensional tori. The projections of these mapping points onto 2-dimensional phase planes corresponding to each degree of freedom cover broad bands which encircle the origins. The projection points seem to scatter all over the bands. The scatter makes the emittance appear larger and have a fuzzier boundary. This makes the beam loss versus tune curves show broad valleys at resonances instead of narrow gulches.

The shapes of the tori and of the stability boundaries, and hence the dynamic aperture, can all be derived analytically to any arbitrary order. To the lowest order the effect of the nonlinear term is to introduce distortion functions on the linear amplitude and phase. However, the algebra involved is rather complicated especially when many high-order nonlinear terms are present. It is easier to use a straightforward tracking program to compute the dynamic aperture numerically. This is the favored approach at present.

The dynamic aperture was investigated long ago and understood in connection with sector-focusing cyclotrons and fixed-field alternating-gradient accelerators, in which the fields are extremely nonlinear and the physical apertures (at least the horizontal aperture) are essentially nonexistent.

Another complication of the nonlinear dynamics is the existence of the stochastic regime of solutions. These solutions generally appear in stochastic layers near separatrices and unstable fixed points, where many high-order resonances overlap. In the projections of motion near the stability limit, sometimes higher-order resonance loops do show up. These stochastic layers make the boundaries of the central stable region fuzzy, but they are fairly narrow and do not sensibly affect the definition of the dynamic aperture. The stochastic regime motion plays a more major role in determining the limitations for the beam-beam interaction in colliders. This will be discussed in Section 2 below.

Distinct from the dynamic aperture is the physical aperture defined by beam collimators or beam pipe walls. The physical aperture is, of course, much more definite, easier to understand, and simpler to calculate.

1.4 Longitudinal Motion

The longitudinal motion of the particles is controlled mainly by the radio frequency electric field. The motion is intrinsically

nonlinear and can be approximated as linear only when the amplitude is small. The nonlinearity is generally that of a sinusoidal electric field. The coordinate is the longitudinal displacement, z, from the synchronous position. The stable region in the (z, z' ≡ dz/ds) phase plane is bounded by a single separatrix passing through a single unstable fixed point. The stable region is shaped like a tear drop and is called the rf bucket. Similar to the transverse motion, the small, approximately linear oscillation can be written as

$$z \propto \sqrt{\beta_z}\, e^{\pm i\int ds/\beta_z} \tag{12}$$

where

$$\beta_z = \sqrt{\frac{2\pi}{h}\,\frac{E}{eV_o}\,\frac{\beta^2R^2}{\eta\cos\phi_s}}\ ,$$

$E = mc^2\gamma$ = total energy

h = harmonic number

V_o = peak rf voltage per turn

$\eta = \frac{d\omega/\omega}{dp/p}$ = revolution frequency dispersion

ϕ_s = synchronous phase angle

except now β_z is an adiabatic constant and the motion is sinusoidal. The number of oscillations per turn, namely the longitudinal tune ν_z is given by

$$\nu_z = \frac{1}{2\pi}\int_0^{2\pi R}\frac{ds}{\beta_z} = \frac{R}{\beta_z} \tag{13}$$

The longitudinal focusing is generally rather weak, and ν_z is very small, approaching 0.1 only for high energy, high repetition rate synchrotrons. Thus, there are no pure longitudinal resonances. The longitudinal oscillation (also called synchrotron oscillation or phase oscillation) does, however, contribute to transverse resonances through coupling to the horizontal oscillation by the orbit dispersion. The lowest order coupling term in the Hamiltonian is proportional to

$$(xD' + x'D)\ z, \qquad \text{prime} = d/ds\ . \tag{14}$$

The longitudinal oscillation therefore contributes to the side-bands

$$\nu_x \pm \ell\nu_z, \qquad \ell = \text{integer}\ , \tag{15}$$

of the horizontal oscillation and creates the overall resonant conditions

$$j\nu_x \pm k\nu_y \pm \ell\nu_z = m \ . \tag{16}$$

This so-called synchro-betatron coupling, and hence the resonance strengths, become progressively weaker at higher ℓ values. The coupling can be eliminated altogether either by placing the rf cavities in zero dispersion ($D = D' = 0$) straight sections or by judiciously distributing the cavities around the ring lattice such that all the coupling terms [Eq. (14)] add up to zero.

The longitudinal oscillation remains fairly linear within the central half of the area of the bucket. The motion becomes strongly nonlinear only when it gets close to the wall of the bucket (the separatrix). As for all nonlinear motions, the longitudinal motion becomes stochastic within a stochastic layer next to the separatrix and the unstable fixed point, but the stochastic layer is usually very thin.

2 INTENSITY DEPENDENT DYNAMICS

2.1 Emittance

Particles in a beam bunch oscillate about the synchronous closed orbit (plotted as the origin of the phase space) and populate a central volume of the phase space. The coordinate variables of the 6-dimensional phase space are simply x, y, and z as defined in Section 1 above and the commonly used momentum variables are listed in Table I, where p_s is the synchronous momentum.

Table I

Coordinate variables	Momentum variables		
	I	II	III
x	p_x	$x' \equiv p_x/p_s$	$\beta\gamma x' = p_x/mc$
y	p_z	$y' \equiv p_y/p_s$	$\beta\gamma y' = p_z/mc$
z	$p_z(=p-p_s)$	$z' \equiv \eta p_z/p_s$	$\beta\gamma z' = \eta p_z/mc$

The independent variable is either time t or the distance s along the closed orbit. Set I gives the proper conjugate momentum variables. With these variables the 6-dimensional volume of the phase space has the unit $(\text{eV sec})^3$ and is an invariant of the motion. With variables of set II the 6-dimensional phase volume has the simpler unit m^3 but shrinks as p_s^{-3}. The set III variables provide both an invariant phase volume and the simple unit.

The 2-dimensional area formed by projecting the 6-dimensional phase volume, which is populated by particle phase points, on the phase plane of one specific degree of freedom is called the

emittance in that degree of freedom. This is not a well-defined concept, because the density distribution in the populated phase generally fuzzy. Two extreme density distributions in the 2-dimensional projection are usually considered, uniform and bi-Gaussian.

If the density is uniform inside an area with sharp boundaries,the emittance is simply the bounded area.

For a linear lattice the "closed" boundary shape (namely one that has the sector periodicity) is an ellipse. The extent of the ellipse along the coordinate variable x, say, is just the x-width of the beam. Denoting the half-width by a_x, we can write the area of the x phase-ellipse or the x emittance as

$$\epsilon_x = \begin{cases} \dfrac{\pi a_x^2}{\beta_x} & \begin{pmatrix}\text{un-normalized} \\ \text{in x, x' plane}\end{pmatrix} \\ \\ \beta\gamma \dfrac{\pi a_x^2}{\beta_x} & \begin{pmatrix}\text{normalized} \\ \text{in x, } \beta\gamma\text{x' plane}\end{pmatrix} . \end{cases} \tag{17}$$

If the density is bi-Gaussian in x and x', and the rms beam half-width is σ_x, the emittance is usually defined as

$$\epsilon_x = \begin{cases} \dfrac{6\pi\sigma_x^2}{\beta_x} & \text{(un-normalized)} \\ \\ \beta\gamma \dfrac{6\pi\sigma_x^2}{\beta_x} & \text{(normalized)} \end{cases} \tag{18}$$

which contains 95% of the beam.

The density function of a real beam is never this simple, especially with nonlinear fields present. One has the choice of using either an iso-density curve or an ellipse that contains, say, 95% of the beam, to define the emittance. The latter is more practical because in all likelihood the beam will be further transported in a linear periodic lattice, and the phase points inside the iso-density curve will be smeared out to fill an ellipse.

The choice of 95% is arbitrary. The CERN convention is to use $4\pi\sigma^2/\beta$ for the Gaussian distribution. Such an emittance contains 86.5% of the beam.

2.2 Static Beam-Field Effects

Assuming the beam is stable and the density distribution in the beam bunch is in a steady state, one can calculate the effect of the electromagnetic field produced by the beam (beam-field or self-field) on the motion of individual particles in the beam. In the transverse plane the effect is a detuning of the betatron

oscillations. It is convenient to consider the effect as resulting from two different contributions, the space-charge contribution and the image-charge contribution.

The transverse space-charge (and current) force has an energy dependence of $1 - \beta^2 = 1/\gamma^2$, the result of the cancellation between the electric defocusing force (factor 1) and the magnetic focusing force (factor β^2). The tune shift $\delta\nu$ contains, in addition, the energy factor $1/\beta^2\gamma$, where β^2 arises from the tune being expressed in terms of the angular velocity, and γ arises from the relativistic mass increase. The dependence of the force on the transverse coordinates is related crucially to the particle density distribution in the beam. If the density distribution is uniform inside an elliptical beam, the transverse force is linear up to the edge of the beam, and the tune depression is independent of amplitude for oscillations inside the beam. The x and y tune depressions are given by

$$\begin{cases} \delta\nu_x = -\dfrac{r_p}{\beta^2\gamma^3}\,\dfrac{1}{\epsilon_x}\displaystyle\int \dfrac{\lambda ds}{1+\dfrac{a_y}{a_x}}, & \epsilon_x \equiv \dfrac{\pi a_x^2}{\beta_x} \\[2ex] \delta\nu_y = -\dfrac{r_p}{\beta^2\gamma^3}\,\dfrac{1}{\epsilon_y}\displaystyle\int \dfrac{\lambda ds}{1+\dfrac{a_x}{a_y}}, & \epsilon_y \equiv \dfrac{\pi a_y^2}{\beta_y} \end{cases} \tag{19}$$

where $r_p \equiv e^2/mc^2 = 1.535\text{x}10^{-18}$ m = classical radius of proton, $a_x(s)$ and $a_y(s)$ are the semi-axes of the beam cross-section ellipse, and $\lambda = \lambda(s)$ is the local linear particle density.

As expected, the tune shift is larger in the direction of the minor axis of the ellipse. This simple but rather unrealistic distribution is called the Kapchinsky-Vladimirsky distribution. In the 4-dimensional transverse phase space (x, x', y, y') this corresponds to a δ-function distribution on a 4-dimensional ellipsoidal shell.

If the density distribution in the elliptical beam is bi-Gaussian in x and y, we get a spread in the tune depressions. The depressions are greatest for the smallest amplitude oscillations in the dense core of the beam and are

$$\begin{cases} \delta\nu_{x_{max}} = -\dfrac{3r_p}{\beta^2\gamma^3}\,\dfrac{1}{\epsilon_x}\displaystyle\int \dfrac{\lambda ds}{1+\dfrac{\sigma_y}{\sigma_x}}, & \epsilon_x \equiv \dfrac{6\pi\sigma_x^2}{\beta_x} \\[2ex] \delta\nu_{y_{max}} = -\dfrac{3r_p}{\beta^2\gamma^3}\,\dfrac{1}{\epsilon_y}\displaystyle\int \dfrac{\lambda ds}{1+\dfrac{\sigma_x}{\sigma_y}}, & \epsilon_y \equiv \dfrac{6\pi\sigma_y^2}{\beta_y} \end{cases} \tag{20}$$

where σ_x and σ_y are the standard deviations of the Gaussian distributions or the rms half-widths of the beam. With the usual definition of $\epsilon = 6\pi\sigma^2/\beta$ for the Gaussian distribution, the maximum tune shift is 3 times that of the uniform distribution.

Neither of these distributions is realistic, but this discussion shows convincingly that the realistic space-charge tune depression has a spread from a value approaching zero for the largest amplitude oscillations to a value approaching but likely no greater than that given by Eq. (20) for the smallest amplitude oscillations.

The image charge (and current) force does not contain the electric-magnetic cancellation factor γ^{-2} and therefore tends to dominate at high energies. To first order it depends only on the linear density of the beam and the cross-sectional dimensions of the imaging beam pipe and magnet poles, and not on the cross-section of the beam. Blowing up the transverse dimensions of the beam reduces the space-charge tune depression but not the image-charge term. To reduce the image-charge term, one has to enlarge the beam pipe. Also, since the image charge (and current) is external to the beam, the effects of its field on the beam are opposite in the two transverse planes as necessitated by the Laplace equation.

The electric image tune shifts are

$$\begin{cases} \delta\nu_x = c_1 \dfrac{r_p}{\beta^2\gamma} \dfrac{2\pi R\lambda}{\langle \pi h^2/\beta_x \rangle} \\[2ex] \delta\nu_y = - c_1 \dfrac{r_p}{\beta^2\gamma} \dfrac{2\pi R\lambda}{\langle \pi h^2/\beta_y \rangle} \end{cases} \tag{21}$$

where 2h = vertical separation of the assumed electric imaging surfaces, c_1 = numerical factor depending on the shape of the imaging surface (= $\pi^2/48$ for parallel planes), and $\langle\rangle$ = averaging around the ring.

The magnetic image tune shifts are

$$\begin{cases} \delta\nu_x = - c_1 \dfrac{r_p}{\gamma} \dfrac{2\pi R\lambda - N}{\langle \pi h^2/\beta_x \rangle} + c_2 \dfrac{r_p}{\gamma} \dfrac{N}{\langle \pi g^2/\beta_x \rangle} \\[2ex] \delta\nu_y = c_1 \dfrac{r_p}{\gamma} \dfrac{2\pi R\lambda - N}{\langle \pi h^2/\beta_y \rangle} - c_2 \dfrac{r_p}{\gamma} \dfrac{N}{\langle \pi g^2/\beta_y \rangle} \end{cases} \tag{22}$$

where 2g = vertical separation of the assumed dc magnetic imaging surfaces, $N = \int\lambda ds$ = number of particles in ring, and c_2 = numerical factor depending on the shape of the surface (= $\pi^2/24$ for parallel planes).

In these expressions the first terms are the ac magnetic image terms and the second terms are the dc magnetic terms. The shape factors c_1 and c_2 have been computed for other than the parallel plane geometry.

We have assumed that the image forces are instantaneously in phase with the beam bunch. Since the bunch is moving and since the imaging vacuum pipe, magnet poles, etc., are all electromagnetically active elements, the image force can have out-of-phase components and can therefore induce oscillations in the beam. This will be discussed further below.

For the longitudinal beam-field force, as a simple approximation, we assume the beam bunch to be a line charge with linear density $\lambda(z)$ along the center line of a beam-pipe that has a capacitance per unit length C. For a circular conducting beampipe of radius b and a circular beam of radius a, $C = [1 + 2\ell n(b/a)]^{-1}$ is a reasonable approximation. The charge distribution then produces a voltage distribution $e\lambda/C$ and a longitudinal field $E_z = -(e/C)d\lambda/dz$. If λ is parabolic, say

$$\lambda(z) = \frac{3}{4}\frac{N}{a_z}\left(1-\frac{z^2}{a_z^2}\right) \tag{23}$$

with $2a_z$ = bunch length and N = number of particles in bunch, we have

$$E_z = \frac{3}{2}\frac{eN}{Ca_z^3}z\ , \tag{24}$$

namely a linear force directed away from the midpoint of the beam bunch (z = 0). The response of the particles is defocusing below transition ($\eta > 0$) and focusing above transition ($\eta < 0$). Thus, the particle behaves as though it has a negative longitudinal mass above transition. The consequences of the longitudinal beam-field force and the negative mass effect on transition crossing will be discussed in further detail later.

In reality the linear density distribution is likely to be more complicated than simple parabolic, and the varying transverse size of the beam will make the longitudinal self-field force dependent also on x and y. The above descriptions of both the transverse and the longitudinal effects are oversimplifications that help create a physical understanding of the basic processes involved and nevertheless give quantitatively reasonable and approximate estimates.

2.3 Coherent Instabilities

The particle beam traveling in an accelerator is surrounded by and coupled to a great number of electromagnetically active elements each of which can be represented electromagnetically by a complex impedance. These include, e.g., the resistive beampipe wall, discontinuities or structures formed in the beampipe such as bellows

and rf cavities, apparatus inserted inside the pipe such as kicker magnets, beam position monitors, etc. The bunched beam current I is rich in harmonic content and induces a voltage per turn IZ where Z is the total impedance of all the electromagnetic elements in the ring which are coupled to the beam. This voltage acts back on the beam particles with a force $U + iV \equiv eIZ$. If the action on the particles is a positive feedback, and if the motions of the particles stay coherent for a long enough time that the positive feedback can be considered as acting on the beam as a whole, a coherent instability in the beam will result and the beam may be lost. Depending on the length of the decay time of this wakefield, a beam bunch may feel its own wakefield on the next turn around and become unstable; this is called the self-excited or turn-to-turn instability. Or the wakefield of one beam bunch may be felt by the succeeding bunches and induce the coupled-bunch or bunch-to-bunch instability. Rather complete analyses have been made of the behaviors of the different modes and the onset thresholds of these instabilities. Fortunately, in most practical cases the wakefield of one beam bunch is effectively attenuated before the next bunch arrives, and these instabilities are not excited. Even in the case when they are excited, these stabilities are easily cured or damped either by electronic feedback or by Landau damping ("decoherencing" motions of individual bunches).

We are therefore left with only the intra-bunch or single-bunch instabilities. The frequencies of these single-bunch instabilities are too high for damping by available electronics and the effect of Landau damping is limited in magnitude. Together with some incoherent effects discussed below, the single-bunch coherent instability usually imposes ultimate limitations on the beam current. To increase the beam current that can be accelerated we must either reduce the impedance or reduce the coherence time (i.e. increase the Landau damping).

Wakefields with long decay times are generally induced by the high-Q parasitic modes of the accelerating rf cavities and can usually be eliminated by damping out these modes in the cavities. Or we can "decoherence" the motions of the different bunches. Longitudinally we can make the synchrotron oscillation frequencies of the bunches different by adding a cavity operating, e.g., at the harmonic number h + 1. Transversely the betatron tunes of different bunches can be made different by using a radiofrequency quadrupole that imposes different quadrupole fields on different bunches. For proton synchrotrons it is generally sufficient just to damp out the harmful parasitic modes in the cavities.

The longitudinal single-bunch instability, commonly known as the microwave instability, induces very short longitudinal lumping of the particles at microwave frequencies within a beam bunch. This instability is stabilized by Landau damping through a spread in the revolution frequency due to the momentum spread. The threshold of the instability expressed as the maximum allowed longitudinal impedance Z_ℓ for given beam current and momentum spread is

$$\frac{|Z_\ell|}{n} \le F_\ell \frac{E}{eI} \beta^2 |\eta| \left(\frac{\Delta p}{p}\right)^2 \tag{25}$$

where n = (instability frequency)/(revolution frequency) = mode number, I = peak current in bunch, $\Delta p/p$ = peak momentum spread (FWHM) in bunch, and F_ℓ = form factor of order unity.

The attainable value of $|Z_\ell|/n$ has a practical lower limit, but with proper care a value of ~ 1 Ω can easily be attained.

The transverse single-bunch instability is also known as the high-mode head-tail instability. The primary excitation mechanism is the following. The field generated by the transverse oscillation of the head of the beam bunch acts on particles in the tail. Because of the phase difference between the head and the tail produced through a non-zero chromaticity by the momentum swing during synchrotron oscillation, this excitation force has the necessary out-of-phase component to induce instability. The particles in the head and the tail are continually interchanged by synchrotron oscillation. Thus, the instability of the whole beam bunch is self-regenerative. This instability can be "cured" mainly by Landau damping arising from a spread in betatron tune. In principle, it can also be damped by a spread in the synchrotron oscillation frequency which produces a mixing of particles along the length of the bunch. But in practice it is difficult to attain sufficiently rapid longitudinal mixing. The threshold of the instability expressed as an upper limit for the transverse impedance per unit length, Z_t, is

$$|Z_t| \le \pi F_t \frac{E}{eI} \frac{\beta^3 \nu}{R} \Delta\nu \tag{26}$$

where F_t is another form factor of order unity and $\Delta\nu$ is the betatron tune spread.

Both the momentum spread $\Delta p/p$ in Eq. (25) and the tune spread $\Delta\nu$ in Eq. (26) are limited by resonance. The limitation is stronger for colliders in which, because of the long storage time requirement, much higher resonances have to be avoided. Thus, we must reduce the impedances to the minimum.

Again, the contributions to the impedances Z_ℓ and Z_t have two sources, that from the space charge/current and that from the image charge/current on the beampipe wall. The image contributions to Z_ℓ and Z_t are related to each other. For a circular beampipe of radius b the simple approximate relation is

$$Z_t \simeq \frac{2R}{\beta b^2} \frac{Z_\ell}{n} . \tag{27}$$

As stated above, a practical lower limit for $|Z_\ell|/n$ is about 1 Ω, a value more or less independent of the size of the ring. This relation then shows that Z_t is larger for higher energy machines for which R is larger and b is smaller. Therefore, one expects transverse instability to be more troublesome for higher energy machines. A great deal of effort has been devoted to computing the

impedances for special geometries of the beam and the chamber wall. But the eventual conclusion must be based on measurements.

2.4 Transition Problems

Several special problems are caused by longitudinal beam self-field forces in crossing transition. Below transition the revolution frequency dispersion

$$\eta \equiv \frac{d\omega/\omega}{dp/p} = 1/\gamma^2 - 1/\gamma^2_t$$

(where γ_t is the transition energy in units of mc^2), is positive (velocity increasing faster with momentum than orbit length) and a particle responds to longitudinal force as though it has a positive mass, i.e. accelerates in the direction of the force. Above transition $\eta < 0$ and a particle responds as though it has a negative mass, i.e. accelerates in the direction opposite to the force. This reversal of response is usually taken care of by making a phase jump in the accelerating rf field from a positive slope (converging force) to a negative slope (diverging force). Together with the change in sign of the "mass," this phase jump keeps the effect always focusing.

When the longitudinal self-field force becomes comparable to the force due to the external rf field, the following problems arise: (1) Unlike the rf force, the self-field force cannot be reversed in sign at transition and remains diverging both before and after; thus it subtracts from the rf force below transition and adds to it above transition, causing mismatch in the force constant and hence a blowup in the longitudinal emittance. (2) Near transition η is sensibly zero, and there can be no Landau damping to stabilize the beam against the longitudinal microwave instability.

Both problems can be resolved or at least alleviated by employing the transition jump scheme. In this scheme fast pulsed quadrupoles are installed in the ring to jump the orbit length dispersion, $(\gamma_t)^{-2}$, at transition crossing so that η is changed abruptly from a non-zero positive value to an appropriate non-zero negative value. With proper adjustments, matching can be reestablished even in the presence of the beam-field force, and, since $|\eta|$ is never zero or even small, Landau damping is always present to damp the microwave instability. In addition to taking care of these longitudinal problems, one must remember to switch the chromaticity ξ from a negative value below transition to a positive value above transition to keep the transverse head-tail instability under control.

All these transition problems, if not properly resolved, although they may not cause direct beam loss, will invariably blow up the longitudinal emittance and perhaps also the transverse emittance if the head-tail instability is not appropriately damped.

2.4 Incoherent High Intensity Effect: Intrabeam Scattering

In the rest frame of a beam bunch the particles are confined in a potential well in all three degrees of freedom. In addition, the particles interact via Coulomb scattering. This intrabeam scattering can be expected to cause noticeable growth in the 6-dimensional emittance if the particle density is sufficiently high. The total growth rate is given by

$$\frac{1}{\tau} = \frac{\pi^2}{\gamma} \frac{c r_p^2 N(\log)}{\Gamma} \langle H(\lambda_1,\lambda_2,\lambda_3)\rangle \tag{28}$$

where $r_p = e^2/mc^2 = 1.535\times10^{-18}$m = classical proton radius, N = number of protons in bunch, log = Coulomb logarithm $\simeq 20$, $\Gamma \equiv (2\pi\beta\gamma)^3(\sigma_x^2/\beta_x)(\sigma_y^2/\beta_y)(\sigma_z^2/\beta_z)$ = invariant 6-dimensional phase volume occupied by beam in Gaussian distributions, and the expression $\langle H(\lambda_1, \lambda_2, \lambda_3)\rangle$ is a dimensionless and homogeneous "momentum shape factor". The quantities $(\lambda_1)^{-1/2}$, $(\lambda_2)^{-1/2}$, and $(\lambda_3)^{-1/2}$ measure the principal axes of the momentum ellipsoid of the beam bunch, and $\langle\rangle$ denotes averaging around the ring. The function H equals zero if $\lambda_1 = \lambda_2 = \lambda_3$, namely if the momentum spread is isotropic. This shows that the effect of the intrabeam scattering is to equipartition the momentum spread among all three degrees of freedom. On the other hand, in an alternating gradient lattice the λ's cannot be equal everywhere. Hence the emittance will always grow.

In the general case, the function H can be expressed in terms of elliptic integrals, but in the special case when $\lambda_1 > \lambda_2 = \lambda_3$, namely when the momentum distribution is an oblate circular ellipsoid with the short axis along the 1 direction, one can write

$$H = \frac{2(\lambda_1 + 2\lambda_2)}{\sqrt{\lambda_2(\lambda_1 - \lambda_2)}} \sin^{-1}\sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1}} - 6 \ . \tag{29}$$

Formulas, slightly more complicated, exist also for $1/\tau_x$, $1/\tau_y$, and $1/\tau_z$, namely the individual growth rates for the emittances in each degree of freedom.

For the Fermilab accelerators operating at the present intensity, intrabeam scattering has not been much of a problem.

2.6 Luminosity Issues

Colliding beams give the possibility of reaching high center-of-mass energies. On the other hand the luminosity is naturally lower than that attainable with a single beam on a fixed target. To maximize the luminosity, the beams are focused hard to a tiny spot at the point of collision. This unfortunately increases the electromagnetic forces between the beams. These forces are extremely nonlinear and act to disrupt or at least blow up the beam so that the beam lifetime is reduced.

For the head-on collision of two circular beam bunches, each of radius a and having N particles, the integrated luminosity is just

$$L_b = \frac{N^2}{\pi a^2} = \beta\gamma \frac{N^2}{\beta^* \epsilon_n} \tag{30}$$

where β^* = amplitude function β at point of collision, and $\epsilon_n = \beta\gamma\ \pi a^2/\beta^*$ = normalized emittance.

For given energy and particle number, to maximum L we must minimize ϵ_n (low emittance beam) and β^* (low β^* obtained by an insertion of strong quadrupoles). If the time interval between bunches in each beam is τ_b, the luminosity is

$$L = \frac{L_b}{\tau_b} = \beta\gamma \frac{N^2}{\beta^* \epsilon_n \tau_b} \ . \tag{31}$$

The disruptive effects of the nonlinear beam-beam forces are difficult to calculate analytically. But for a given density distribution in the beam, i.e. a given mix of nonlinear forces, the effects in each transverse degree of freedom can be measured in terms of only one parameter, the linear tune-shift. We demonstrate this for the case of one transverse degree of freedom, say x. To begin with, neglecting the effect of the beam pipe, the electric potential of the second beam as seen by a particle in the first beam can be written as

$$V(x,s) = eNf\left(\frac{x^2}{a^2},\ s\right) \tag{32}$$

where N is the number of particles in the bunch and f is a properly normalized function. It is clear that f is an even function of x and that x should be scaled by the half-width a of the beam. The effect of the beam-beam force alone on a particle in the first beam is given by

$$m\gamma \frac{d^2x}{dt^2} = -\ (1 + \beta^2)e \frac{\partial V}{\partial x} = -(1 + \beta^2)e^2 N \frac{\partial f}{\partial x}$$

where the factor $(1 + \beta^2)$ arises from the reinforcement (because the beams are going in opposite directions) of the electric and the magnetic forces. With s as the independent variable this equation becomes

$$\frac{d^2x}{ds^2} = -\ \frac{1 + \beta^2}{\beta^2\gamma}\ r_p N \frac{\partial f}{\partial x}$$

This shows that the total Hamiltonian of the motion could be written as

$$H(x,p;s) = \frac{1}{2}\ (p^2 + Kx^2) + \frac{1 + \beta^2}{\beta^2\gamma}\ r_p N f\ , \tag{33}$$

We make the usual canonical transformation to the angle-action variables ϕ and J by

$$\begin{cases} x = \sqrt{2\beta_x J}\cos\phi \\ p = -\sqrt{\dfrac{2J}{\beta_x}}\left(\sin\phi - \dfrac{\beta_x'}{2}\cos\phi\right) \end{cases} \tag{34}$$

and obtain the new Hamiltonian

$$k(\phi,J;s) = \nu J + \frac{1+\beta^2}{\beta^2\gamma}\, rpNf\left[\frac{2\beta xJ}{a^2}\cos^2\phi, s\right] . \tag{35}$$

We then define a scaled action variable

$$I \equiv \frac{2\beta_x}{a^2} J .$$

Keeping in mind that $a^2 \alpha \beta_x$, and hence β_x/a^2 is independent of s, we can write the canonical equations of K as follows

$$\begin{cases} \dfrac{d\phi}{ds} = \dfrac{\partial K}{\partial J} = \nu - 4\pi(\delta\nu_b)\dfrac{\partial f}{\partial I} \\ \dfrac{dI}{ds} = \dfrac{2\beta_x}{a^2}\dfrac{dJ}{ds} = -\dfrac{2\beta_x}{a^2}\dfrac{\partial K}{\partial\phi} = 4\pi(\delta\nu_b)\dfrac{\partial f}{\partial\phi} \end{cases} \tag{36}$$

where the beam-beam tune shift

$$\delta\nu_b \equiv \frac{1}{2\pi}\,\frac{1+\beta^2}{\beta^2\gamma}\,\frac{r_p N\beta_x}{a^2} = -\frac{1}{2}\,\frac{1+\beta^2}{\beta}\,\frac{r_p N}{\epsilon_n} \tag{37}$$

has the form of usual space-charge tune shift except that the factor $1 - \beta^2$ is changed to $1 + \beta^2$. In Eq. (36) $f = f(I\cos^2\phi, s)$ and the nonlinearity in $I\cos^2\phi$ is derived from the original nonlinearity in x^2/a^2. Equation (36) shows that the motion is characterized only by the linear tune ν and the beam-beam tune shift $\delta\nu_b$.

We can now write a phenomenological expression for the beam-beam blowup rate,

$$\frac{1}{T} \propto \frac{1}{\sqrt{\tau_c}}\, G\left(\frac{\delta\nu_b}{\delta\nu_c}\right) \tag{38}$$

where τ_c = time interval between collisions of a beam bunch with others, and $\delta\nu_c$ = critical tune shift.

The blowup effects from collisions with different bunches are not coherent, hence the dependence on τ_c is taken to be $\sqrt{\tau_c}$. G is a function that rises sharply for $\delta\nu_b > \delta\nu_c$. Experiences at the $\bar{p}pS$ indicates a critical tune shift value of $\delta\nu_c \sim 0.003$.

To summarize, the disruptive effects of the nonlinear beam-beam forces can be measured by the linear beam-beam tune shift $\delta\nu_b$, and

experience shows that, to obtain reasonable lifetime for the colliding beams, $\delta\nu_b$ should not be >0.003.

In addition to the nonlinear effects, the simple linear tune shift per revolution is, as before, limited by resonances. To avoid resonances up to, say, the 7th order, we must have

$$N_c \delta\nu_b \lesssim 0.02 \tag{39}$$

where N_c is, for a given bunch, the number of collisions with other bunches in one revolution. If $\delta\nu_b \sim 0.002$, say, this imposes an upper limit on N_c of about 10. For pp colliders, or $\bar{p}p$ colliders with beam separators, the beam bunches collide only at the interacting points (IP's) for experiments, and N_c is equal to the number of IP's. A value of 10 for N_c is quite acceptable. But for $\bar{p}p$ colliders without beam separators, all bunches of beam 1 collide with all bunches of beam 2. A value of $N_c = 10$ will severely limit the number of bunches per beam.

An entirely different limitation on the integrated luminosity per bunch crossing, L_b, is imposed by the resolving power of the detector for the events produced. The total cross-sections for pp or $\bar{p}p$ collisions are of the order of 100 mbarn = 10^{-25} cm^2 at TeV energies. It is difficult to resolve more than, say, 2 events during the collision of two beam bunches. Thus L_b should be of the order of 10^{25} cm^{-2} and no greater, and to obtain a luminosity of $L = 10^{33}$ cm^{-2} sec^{-1} we need $\tau_b \sim 10^{-8}$ sec or $\sim$ 10 nsec. This imposes rather stringent demands on the geometry of the interacting point. The bunches must collide more or less head-on (otherwise the luminosity will be reduced) at the IP, and 5 nsec (1.5 m) away on either side the bunches must be separated, at least by as much as their width. These demands can indeed be met, but only with difficulty.

The strategy of getting $L_b \sim 10^{25}$ cm^{-2} with the lowest particle number N is also clear. For this we want to use the lowest possible emittance ϵ_n and the lowest possible β^*. We should then check that Eq. (37) gives a beam-beam tune shift $\delta\nu_b$ smaller than $\delta\nu_c \sim 0.003$.

APPENDIX A SPACE-CHARGE TUNE SHIFT

A.1 Circular Beams

Case I Uniform density distribution

The electric field is radial and is a function only of the radial coordinate r. As long as one stays inside the beam ($r < a$) the field is given by

$$\frac{1}{r}\ \frac{d(rE_r)}{dr} = -4\pi e\rho = -\frac{4e\lambda}{a^2} \tag{A-1}$$

or

$$E_r = -\frac{2e\lambda}{a^2}\, r \qquad (r<a) \tag{A-2}$$

where a is the radius and $\lambda(s)$ is the linear density of the beam. The force f_r is then

$$f_r = \frac{eE_r}{\gamma^2} = -\frac{2e^2\lambda}{\gamma^2 a^2}\, r \tag{A-3}$$

where the γ^2 factor arises from the cancellation between the electric and the magnetic forces. Thus, the tune shift is

$$\delta\nu_r = \frac{1}{4\pi}\int \frac{\beta_r}{pc\beta}\frac{df_r}{dr}\, ds = -\frac{r_p}{2}\frac{1}{\beta^2\gamma^3}\frac{N}{\epsilon_r} \tag{A-4}$$

where $r_p = \frac{e^2}{mc^2} = 1.535\text{x}10^{-18}$ m = classical proton radius,

$\epsilon_r \equiv \frac{\pi a^2}{\beta_r}$ = emittance,

$N = \int \lambda ds$ = total number of particles,

and r denotes either x or y.

<u>Case II</u> Gaussian density distribution

In this case we write

$$\rho = \frac{\lambda}{2\pi\sigma_r^2}\, e^{-\frac{r^2}{2\sigma_r^2}} \tag{A-5}$$

and obtain

$$\delta\nu_r = -3r_p\ \frac{1}{\beta^2\gamma^3}\ \frac{1}{\epsilon_r}\int \lambda\left\{ e^{-\frac{r^2}{2\sigma_r^2}} - \frac{\sigma_r^2}{r^2}\left(1 - e^{-\frac{r^2}{2\sigma_r^2}}\right)\right\} ds \tag{A-6}$$

where, again, r denotes either x or y. The maximum tune shift is obtained for oscillations with vanishing amplitude or as $r \to 0$. This gives

$$\delta\nu_{r\ max} = -\frac{3r_p}{2}\frac{1}{\beta^2\gamma^3}\frac{N}{\epsilon_r} \tag{A-7}$$

where $$\epsilon_r \equiv \frac{6\pi\sigma_r^2}{\beta_r} = 95\%\ \text{emittance}\ .$$

A.2 Elliptical Beams

Case I Uniform density distribution

With uniform density and an elliptical beam boundary with semi-axes a and b, it is easy to see that a potential function

$$V(x,y) = \frac{2\lambda}{a+b}\left[\frac{x^2}{a} + \frac{y^2}{b}\right] \tag{A-8}$$

satisfies the Poisson equation

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 4\pi\frac{\lambda}{\pi ab} = 4\pi\rho \tag{A-9}$$

and the boundary condition that the elliptical beam boundary should be an equipotential. Thus the electric fields are

$$\begin{cases} E_x = -\dfrac{\partial V}{\partial x} = -\dfrac{4\lambda}{a(a+b)}\,x \\[2ex] E_y = -\dfrac{\partial V}{\partial y} = -\dfrac{4\lambda}{b(a+b)}\,y \end{cases} \tag{A-10}$$

and the tune shifts are

$$\begin{cases} \delta\nu_x = -\dfrac{r_p}{\beta^2\gamma^3}\dfrac{1}{\epsilon_x}\displaystyle\int\frac{\lambda ds}{1+\frac{b}{a}}\ , & \epsilon_x \equiv \dfrac{\pi a^2}{\beta_x} \\[2ex] \delta\nu_y = -\dfrac{r_p}{\beta^2\gamma^3}\dfrac{1}{\epsilon_y}\displaystyle\int\frac{\lambda ds}{1+\frac{a}{b}}\ , & \epsilon_y \equiv \dfrac{\pi b^2}{\beta_y} \end{cases} \tag{A-11}$$

Case II Bi-Gaussian density distribution

The density distribution is

$$\rho(x,y) = \frac{\lambda}{2\pi ab}\exp\left[-\frac{x^2}{2a^2} - \frac{y^2}{2b^2}\right] \tag{A-12}$$

where a and b are, now, the standard deviations. The potential function for such a density distribution is

$$V(x,y) = e\lambda \int_0^\infty dt \, \frac{1 - \exp\left[-\frac{x^2}{2(a^2+t)} - \frac{y^2}{2(b^2+t)}\right]}{\sqrt{(a^2+t)(b^2+t)}} \, . \tag{A-13}$$

To get the x tune shift, e.g., we have

$$E_x = -\frac{\partial V}{\partial x} = -e\lambda x \int_0^\infty dt \, \frac{\exp\left[-\frac{x^2}{2(a^2+t)} - \frac{y^2}{2(b^2+t)}\right]}{(a^2+t)\sqrt{(a^2+t)\,(b^2+t)}} \, . \tag{A-14}$$

The x tune shift is then

$$\delta\nu_x = \frac{1}{4\pi}\frac{r_p}{\beta^2\gamma^3} \int \frac{1}{e}\,\beta_x \frac{\partial E_x}{\partial x}\, ds$$

$$= -\frac{1}{4\pi}\frac{r_p}{\beta^2\gamma^3} \int \lambda\beta_x ds \int_0^\infty dt \, \frac{\left(1 - \frac{x^2}{a^2+t}\right)\exp\left[-\frac{x^2}{2(a^2+t)} - \frac{y^2}{2(b^2+t)}\right]}{(a^2+t)\sqrt{(a^2+t)(b^2+t)}} \, . \tag{A-15}$$

The maximum tune shift is for vanishingly small oscillations corresponding to x, y → 0. This gives

$$\delta\nu_{x\ max} = -\frac{1}{4\pi}\frac{r_p}{\beta^2\gamma^3} \int \lambda\beta_x ds \int_0^\infty \frac{dt}{(a^2+t)\sqrt{(a^2+t)(b^2+t)}}$$

$$= -\frac{3r_p}{\beta^2\gamma^3}\,\frac{1}{\epsilon_x}\int \frac{\lambda ds}{1+\frac{b}{a}} \, , \qquad \epsilon_x \equiv \frac{6\pi a^2}{\beta_x} \, . \tag{A-16}$$

Similarly we have

$$\delta\nu_y = -\frac{1}{4\pi}\frac{r_p}{\beta^2\gamma^3} \int \lambda\beta_y ds \int_0^\infty dt \, \frac{\left(1 - \frac{y^2}{b^2+t}\right)\exp\left[-\frac{x^2}{2(a^2+t)} - \frac{y^2}{2(b^2+t)}\right]}{(b^2+t)\sqrt{(a^2+t)\,(b^2+t)}} \, . \tag{A-17}$$

and

$$\delta\nu_{y\ max} = -\frac{3r_p}{\beta^2\gamma^3}\frac{1}{\epsilon_y}\int\frac{\lambda ds}{1+\frac{a}{b}}, \qquad \epsilon_y \equiv \frac{6\pi b^2}{\beta_y}. \tag{A-18}$$

APPENDIX B FORMULAS RELATED TO LONGITUDINAL OSCILLATION (SYNCHROTRON OSCILLATION, PHASE OSCILLATION)

In Section 2, for clarity we used a unified notation for all three degrees of freedom. Here we are under no such constraint. The simplest and most directly obvious starting equations are

$$\begin{cases} \frac{d\phi}{dt} = \omega_o - h\omega \\ \frac{dp}{dt} = \frac{eV_o}{2\pi R}\sin\phi \end{cases} \tag{B-1}$$

where ϕ = rf phase as seen by the particle,
$\omega = c\beta/R$ = particle revolution frequency,
h = harmonic number,
p = particle momentum, and
ω_o, V_o = frequency and peak voltage of rf.

To put the equations in canonical form we define the variable W, remembering that $d\phi = (h/R)dz$, by

$$dW \equiv \frac{R}{h}dp = \frac{dE}{h\omega}, \qquad E = mc^2\gamma = \text{total energy}$$

and write the equations as

$$\begin{cases} \frac{d\phi}{dt} = \omega_o - h\omega = \frac{\partial H}{\partial W} \\ \frac{dW}{dt} = \frac{eV_o}{2\pi h}\sin\phi = -\frac{\partial H}{\partial\phi} \end{cases} \tag{B-2}$$

with the Hamiltonian

$$H(\phi,W;t) = \omega_o W - E(W) + \frac{eV_o}{2\pi h}\cos\phi .$$

We define the synchronous values (subscript s) by

$$p_s \equiv (e/c)B\rho \quad \text{(synchronous momentum)}$$

where B = B(t) = guide magnetic field and ρ = constant bending radius, are all given parameters. This then defines the synchronous phase ϕ_s by

$$\frac{dp_s}{dt} = \frac{eV_o}{2\pi R}\sin\phi_s .$$

Expanding the first of Eq. (B-2) to the first-order term in $w \equiv W - W_s$ gives

$$\begin{cases} \dfrac{d\phi}{dt} = (\omega_o - h\omega_s) - \dfrac{h^2\omega_s}{Rp_s}\,\eta w = -\dfrac{h^2\eta}{mR^2\gamma_s}\,w = \dfrac{\partial K}{\partial w} \\[2ex] \dfrac{dw}{dt} \quad \dfrac{eV_o}{2\pi h}\,(\sin\phi - \sin\phi_s) = -\dfrac{\partial K}{\partial \phi} \end{cases} \tag{B-3}$$

where

$$\eta \equiv \frac{d\omega/\omega}{dp/p} = \gamma^{-2} - \gamma_t^{-2} = \text{revolution frequency dispersion,}$$

$$\gamma_t^{-2} \equiv \frac{dR/R}{dp/p} = \text{orbit length dispersion, and } \gamma_t = \text{transition-gamma,}$$

and where the rf is tuned such that $\omega_o = h\omega_s$. The new Hamiltonian is

$$K(\phi,w;t) = -\frac{h^2\eta}{mR^2\gamma_s}\,\frac{w^2}{2} + \frac{eV_o}{2\pi h}\,(\cos\phi + \phi\sin\phi_s)\ .$$

The adiabatic phase trajectories are given by K = constant. The separatrix or the "bucket boundary" is the trajectory that passes through the single unstable fixed point at

$$w = 0, \qquad \phi = \pi - \phi_s \equiv \phi_2$$

and is given by the equation

$$w = \pm\,\frac{R}{hc}\sqrt{\frac{1}{\pi h}\,\frac{eV_oE_s}{\eta}}\,\left[\cos\phi + \cos\phi_s + \phi\sin\phi_s - (\pi - \phi)\sin\phi_s\right]^{1/2}\ . \tag{B-4}$$

This gives the following dimensions for the rf bucket:

<u>Horizontal extent</u>: ϕ_1 to ϕ_2. Both are given by the solutions of

$$\cos\phi + \cos\phi_s + \phi\sin\phi_s - (\pi - \phi_s)\sin\phi_s = 0\ . \tag{B-5}$$

Then $\phi_2 = \pi - \phi_s$ is an obvious solution. The other solution ϕ_1 is on the opposite side of ϕ_s as ϕ_2 and must be obtained numerically.

<u>Vertical extent</u>: $2w_{max}$. It is easy to see that w is maximum at $\phi = \phi_s$. Thus, we get

$$w_{max} = \frac{RE_s}{hc}\sqrt{\frac{2}{\pi h}\,\frac{eV_o}{E_s}\,\frac{1}{\eta}}\ \beta(\phi_s) \tag{B-6}$$

with

$$\beta(\phi_s) \equiv \sqrt{\cos\phi_s + (\phi_s - \tfrac{\pi}{2})\sin\phi_s} \tag{B-7}$$

so defined that $\beta(\phi_s = 0) = 1$. The other more physical momentum variables introduced in section 2 are related to w by

$$\begin{cases} p_{z\ max} = \frac{h}{R} w_{max} = \frac{E_s}{c} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \frac{1}{\eta}}\ \beta(\phi_s) \\ z'_{max} = \frac{\eta h}{p_s R} w_{max} = \frac{1}{\beta_s} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \eta}\ \beta(\phi_s) \\ \beta\gamma z'_{max} = \frac{\eta h}{mcR} w_{max} = \gamma_s \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \eta}\ \beta(\phi_s) \end{cases} \tag{B-8}$$

<u>Area of bucket</u>: This is given in the basic units of (ϕ, w) by

$$A(\phi, w) = 8 \frac{RE_s}{hc} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \frac{1}{\eta}}\ \alpha(\phi_s) \tag{B-9}$$

where

$$\alpha(\phi_s) \equiv \frac{1}{4\sqrt{2}} \int_{\phi_1}^{\phi_2} \left[\cos\phi + \cos\phi_s + \phi\sin\phi_s - (\pi - \phi_s)\sin\phi_s\right]^{1/2} d\phi \tag{B-10}$$

so defined that $\alpha(\phi_s = 0) = 1$. In the other more physical variables the bucket area is given by

$$\begin{cases} A(z, p_z) = 8 \frac{RE_s}{hc} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \frac{1}{\eta}}\ \alpha(\phi_s) = 8 \frac{R}{h} p_{z\ max} \frac{\alpha(\phi_s)}{\beta(\phi_s)} \\ A(z, z') = 8 \frac{R}{h\beta_s} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \eta}\ \alpha(\phi_s) = 8 \frac{R}{h} z'_{max} \frac{\alpha(\phi_s)}{\beta(\phi_s)} \\ A(z, \beta\gamma z') = 8 \frac{R\gamma_s}{h} \sqrt{\frac{2}{\pi h} \frac{eV_o}{E_s} \eta}\ \alpha(\phi_s) = 8 \frac{R}{h} (\beta\gamma z')_{max} \frac{\alpha(\phi_s)}{\beta(\phi_s)} \end{cases} \tag{B-11}$$

The functions $\phi_1(\phi_s)$, $\alpha(\phi_s)$, and $\beta(\phi_s)$ are all given in tabulated form in CERN Report CERN/MPS-SI/Int.DL/70/4.

APPENDIX C FORMULAS RELATED TO TRANSITION CROSSING

We start with Eqs. (B-3) of Appendix B. For small oscillations we expand to linear terms in $\psi \equiv \phi - \phi_s$ (ϕ_s = constant). We also assume a linear increase of energy in time, namely

$$\dot{\gamma} = \frac{c\beta_t}{2\pi R} \frac{eV_o \sin \phi_s}{mc^2} = \text{constant} \tag{C-1}$$

where $c\beta_t$ is the particle velocity at transition. This gives

$$-\eta = \frac{1}{\gamma_t^2} - \frac{1}{\gamma^2} \simeq \frac{2\dot{\gamma}}{\gamma_t^3} t \qquad \text{(transition at } t = 0)$$

and

$$\begin{cases} \dfrac{d\psi}{dt} = -\dfrac{h^2\eta}{mR^2\gamma_s} w \simeq \left(\dfrac{hc}{R}\right)^2 \left(\dfrac{2\dot{\gamma}}{\gamma_t^4}\right) \dfrac{t}{E_o} w \equiv at\, w \\ \\ \dfrac{dw}{dt} = \dfrac{eV_o \cos\phi_s}{2\pi h} \psi \equiv b \cot\phi_s\, \psi \end{cases} \tag{C-2}$$

where we have defined

$$a = 2 \frac{\omega_\infty^2}{E_o} \frac{\dot{\gamma}}{\gamma_t^4} > 0 \ , \ E_o = mc^2 = \text{rest energy},$$

$$\omega_\infty = \frac{hc}{R} = \text{rf frequency at } \infty \text{ energy, and}$$

$$b = \frac{c\beta_t}{2\pi R} \frac{eV_o \sin\phi_s}{mc^2} \frac{R}{hc} \frac{mc^2}{\beta_t} = \frac{E_o}{\omega_\infty} \frac{\dot{\gamma}}{\beta_t} > 0 \ .$$

We also define a scaled time variable $x \equiv t/T$ where T is related to the synchrotron oscillation frequency Ω by

$$\Omega^2 = -abt \cot\phi_s \equiv \frac{|t|}{T^3} \qquad \begin{cases} t < 0 \ , & \phi_s < \frac{\pi}{2} \\ t > 0 \ , & \phi_s > \frac{\pi}{2} \end{cases}$$

or

$$T^{-3} = ab|\cot\phi_s| \qquad \text{and } (\Omega T)^2 = |t|/T \ . \tag{C-3}$$

So defined T is the time away from transition when the synchrotron oscillation phase has advanced one radian and is hence a measure of the "width" of the transition.

To solve Eq. (C-2) we make the transformations

$$y = \frac{2}{3}\, x^{3/2} = \frac{2}{3}\left(\frac{t}{T}\right)^{3/2} \quad \text{and} \quad \psi = x\Phi \ ,$$

and obtain for Φ the equation

$$\frac{d^2\Phi}{dy^2} + \frac{1}{y}\,\frac{d\Phi}{dy} - \left[1 + \frac{(2/3)^2}{y^2}\right]\Phi = 0 \ . \tag{C-4}$$

This is the Bessel equation giving the solution

$$\begin{cases} \psi = Ax\left(\cos\chi\ J_{2/3} + \sin\chi\ N_{2/3}\right) \\ aT^2x^{1/2}w - x^{-3/2}\psi = Ax\left(\cos\chi\ J'_{2/3} + \sin\chi\ N'_{2/3}\right) \end{cases} \tag{C-5}$$

where A and χ are the "amplitude" and "phase" constants. The phase plane trajectory or equivalently the boundary of the phase area covered by the beam is obtained by eliminating χ in Eqs. (C-5). This gives

$$\gamma_\psi \psi^2 + 2\alpha_\psi \psi w + \beta_\psi w^2 = \frac{\epsilon_\psi}{\pi} \tag{C-6}$$

where

$$\begin{cases} \gamma_\psi \equiv \dfrac{1}{aT^2x^{1/2}}\left(J^2_{-1/3} + N^2_{-1/3}\right)\left(J_{2/3}N_{-1/3} - N_{2/3}J_{-1/3}\right)^{-1} \\ \alpha_\psi \equiv -\left(J_{2/3}J_{-1/3} + N_{2/3}N_{-1/3}\right)\left(J_{2/3}N_{-1/3} - N_{2/3}J_{-1/3}\right)^{-1} \\ \beta_\psi \equiv aT^2x^{1/2}\left(J^2_{2/3} + N^2_{2/3}\right)\left(J_{2/3}N_{-1/3} - N_{2/3}J_{-1/3}\right)^{-1} \end{cases} \tag{C-7}$$

and the "amplitude" A is related to the phase space area or longitudinal emittance ϵ_ψ by

$$A^2 = \left(\frac{\epsilon_\psi}{\pi}\right)\left(\frac{aT^2}{x^{3/2}}\right)\left(J_{2/3}N_{-1/3} - N_{2/3}J_{-1/3}\right)^{-1} \tag{C-8}$$

It is interesting that:

<u>At large y</u> ($y \to \infty$, away from transition)

$$\begin{cases} J_\nu \to \sqrt{\dfrac{2}{\pi y}}\ \cos\left(y - \dfrac{\pi\nu}{2} - \dfrac{\pi}{4}\right) \\ N_\nu \to \sqrt{\dfrac{2}{\pi y}}\ \sin\left(y - \dfrac{\pi\nu}{2} - \dfrac{\pi}{4}\right) \end{cases}$$

and Eq. (C-6) becomes

$$\left(aT^2x^{1/2}\right)^{-1}\psi^2 + \left(aT^2x^{1/2}\right)w^2 = \frac{\epsilon_\psi}{\pi} \tag{C-9}$$

which is an upright ellipse with semi-axes

$$\begin{cases} \hat{\psi} = \sqrt{\dfrac{\epsilon_\psi}{\pi}}\ \sqrt{aT^2}\ x^{1/4} \equiv \theta_o x^{1/4} \equiv \theta \\[2ex] \hat{w} = \sqrt{\dfrac{\epsilon_\psi}{\pi}}\ \dfrac{1}{\sqrt{aT^2}}\ x^{-1/4} = \dfrac{\epsilon_\psi}{\pi}\ \dfrac{1}{\theta} \end{cases} \tag{C-10}$$

and Eq. (C-8) becomes

$$A^2 = \frac{\epsilon_\psi}{3}\ aT^2\ . \tag{C-11}$$

<u>At small y</u> ($y \to 0$, at transition)

$$J_{2/3} \to \frac{3}{2\ \Gamma(2/3)}\left(\frac{y}{2}\right)^{2/3}, \qquad J_{-1/3} \to \frac{1}{\Gamma(2/3)}\left(\frac{y}{2}\right)^{-1/3}$$

$$N_{2/3} \to -\frac{2}{\sqrt{3}\ \Gamma(1/3)}\left(\frac{y}{2}\right)^{-2/3}, \quad N_{-1/3} \to -\frac{1}{\sqrt{3}\ \Gamma(2/3)}\left(\frac{y}{2}\right)^{-1/3}$$

and Eq. (C-6) becomes

$$\frac{4/3}{kaT^2}\ \psi^2 + 2\ \frac{1}{\sqrt{3}}\ \psi w + kaT^2w^2 = \frac{\epsilon_\psi}{\pi}\ . \tag{C-12}$$

The ellipse is slightly tilted. The maximum extents of the ellipse are given by the usual relations to be

$$\begin{cases} \hat{\psi}_t = \sqrt{k}\ \sqrt{\dfrac{\epsilon_\psi}{\pi}}\ \sqrt{aT^2} \equiv \sqrt{k}\ \theta_o \equiv \theta_t \\[2ex] \hat{w}_t = \dfrac{2}{\sqrt{3k}}\ \sqrt{\dfrac{\epsilon_\psi}{\pi}}\ \dfrac{1}{\sqrt{aT^2}} = \dfrac{2}{\sqrt{3}}\ \dfrac{\epsilon_\psi}{\pi}\ \dfrac{1}{\theta_t} \end{cases} \tag{C-13}$$

where $$k \equiv \frac{3^{1/3}}{\pi} \left[\Gamma(2/3)\right]^2 = 0.842$$

and $$\theta_o \equiv \sqrt{\frac{\epsilon_\psi}{\pi}} \sqrt{aT^2} \tag{C-14}$$

will be used for scaling below. Eq. (C-8) again becomes Eq. (C-11).

The envelope equation of ψ derived from Eq. (C-2) is

$$\frac{d}{dx}\left(\frac{1}{x}\frac{dY}{dx}\right) - \delta Y = \frac{x}{Y^3} \tag{C-15}$$

where $$x = \frac{t}{T} = \text{scaled time} ,$$

$$\delta \equiv \frac{\cot \phi_s}{|\cot \phi_s|} = \begin{cases} +1 & \phi_s < \frac{\pi}{2} \\ -1 & \phi_s > \frac{\pi}{2} \end{cases}$$

and $$Y = \frac{\theta}{\theta_o} = \text{scaled envelope of } \psi$$

$$(\theta = \hat{\psi} = \text{envelope of } \psi) .$$

We are now ready to investigate the two unpleasant features of transition crossing.

C.1 Microwave Instability

The stability condition (with $F_\ell = 1$) is given by Eq. (25):

$$\frac{|Z_\ell|}{n} < \frac{E}{eI} \beta^2 \, |\eta| \left(\frac{\Delta p}{p}\right)^2 . \tag{C-16}$$

Near transition, we have approximately

$$|\eta| \simeq \frac{2\dot{\gamma}}{\gamma_t^3} |t|$$

and

$$\frac{\Delta p}{p} \simeq \frac{2\,\frac{h}{R}\,\hat{w}_t}{mc\beta_t\gamma_t} = \sqrt{\frac{8}{3\pi k}} \sqrt{\frac{\epsilon_\psi}{E_o\dot{\gamma}T^2}} \; \frac{\gamma_t}{\beta_t}$$

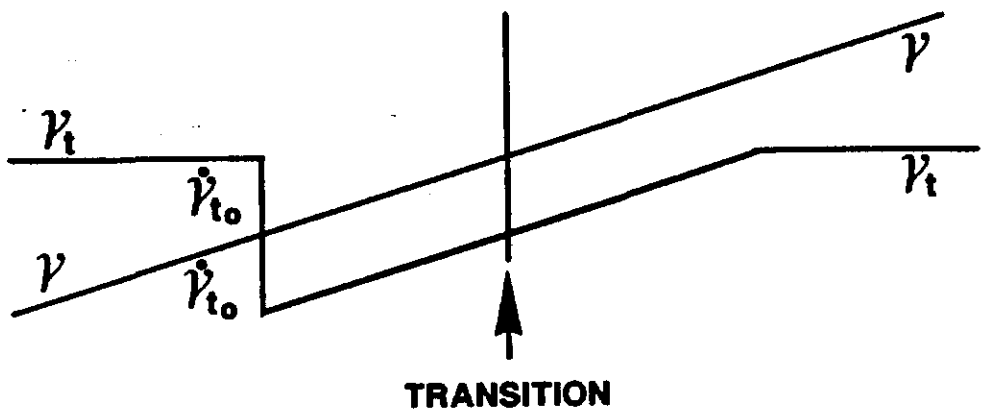
which, when substituted in Eq. (C-16), give

$$|t| > \frac{3\pi k}{16} \frac{eI|Z_\ell|/n}{\epsilon_\psi} T^2 \equiv t_o . \tag{C-17}$$

Thus the beam is unstable against microwave instability from $t = (-t_o \text{ to } t_o)$ or for

$$\Delta\gamma \equiv \gamma - \gamma_t = (-\dot{\gamma} t_o \text{ to } \dot{\gamma} t_o) \ . \qquad \text{(C-18)}$$

One can calculate the blowup factor e^S where S is an integral from $-t_o$ to t_o. The blowup can be avoided by making a γ_t jump of $2\dot{\gamma}\, t_o$ as shown below



C.2 Space-Charge Mismatch

We assume a parabolic longitudinal distribution with bunch half-length

$$\hat{z} = \frac{R}{h}\hat{\phi} = \frac{R}{h}\theta$$

as follows:

$$\lambda(z) = \frac{3}{4}\frac{N}{\hat{z}}\left[1 - \frac{z^2}{\hat{z}^2}\right]$$

where λ is the linear density and N is the number of particles in the bunch. Then the longitudinal electric field is

$$E_z = -eg\frac{d\lambda}{dz} = \frac{3}{2}\frac{egN}{\hat{z}^3}z \qquad \text{(C-20)}$$

with $$g \simeq 1 + 2\ln\left(\frac{\text{beam pipe radius}}{\text{beam radius}}\right)$$

and

$$\left(\frac{dw}{dt}\right)_{s.c.} = \frac{3}{2}\frac{r_p}{c}\frac{E_o\omega_\infty}{\gamma^2}\frac{gN}{\theta^3}\phi \equiv \frac{G}{\gamma^2}\frac{\phi}{\theta^3} \qquad \text{(C-21)}$$

where, as defined,

$$G \equiv \frac{3}{2}\frac{r_p}{c}E_o\omega_\infty gN$$

and $\qquad r_p \equiv \frac{e^2}{mc^2} = 1.535\text{x}10^{-18}\text{m}$ = classical proton radius.

The synchrotron oscillation Eqs. (C-2) now becomes

$$\begin{cases} \frac{d\phi}{dt} = \text{at } w \\ \frac{dw}{dt} = \left(b \cot\phi_s + \frac{G}{\gamma^2\theta^3}\right)\phi \ . \end{cases} \tag{C-22}$$

The mismatch arises from the fact that the sign of $\cot\phi_s$ is switched from positive at $t < 0$ (before transition) to negative at $t > 0$ (after transition) by shifting the synchronous phase from ϕ_s to $\pi - \phi_s$, but there is no way to switch the sign of the space-charge term $G/(\gamma^2\theta^3)$. Hence the synchrotron oscillation frequency is effectively shifted from

$$\Omega^2 = a\left(b \cot\phi_s + \frac{G}{\gamma^2\theta^3}\right)(-t) \qquad \text{before transition, } t < 0$$

to $$\Omega^2 = a\left(b\left|\cot\phi_s\right| - \frac{G}{\gamma^2\theta^3}\right)(t) \qquad \text{after transition, } t > 0 \ .$$

The degree of mismatch is exhibited by the Sørenssen parameter η_{oo}, which is defined by

$$\eta_{oo} = \frac{\frac{G}{\gamma_t^2\theta_o^3}}{b\left|\cot\phi_s\right|} = \frac{aG}{\gamma_t^2}\left(\frac{T}{\theta_o}\right)^3 \equiv \frac{K}{\gamma_t^2} \ . \tag{C-23}$$

(Sørenssen used $\theta_t = \sqrt{k}\,\theta_o = 0.02\theta_o$ instead of θ_o. This of course makes little difference in the discussion.) If $\eta_{oo} << 1$ the space-charge mismatch is negligible. Otherwise matching can be reestablished by making γ_t and/or ϕ_s appropriate functions of time.

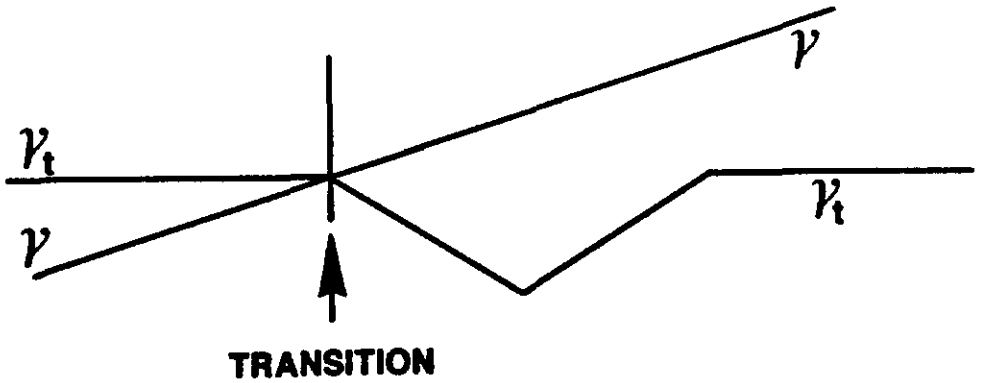
With space charge, the scaled envelope Eq. (C-15) becomes

$$\frac{d}{dx}\left(\frac{1}{x}\frac{dY}{dx}\right) - \left(\delta + \frac{K}{\gamma^2}\frac{1}{Y^3}\right)Y = \frac{x}{Y^3} \tag{C-24}$$

where, as defined in Eq. (C-23)

$$K \equiv aG\left(\frac{T}{\theta_o}\right)^3 \ .$$

Many different γ_t or ϕ_s jump schemes are possible. One such γ_t jump that will reestablish matching is as follows. Some ϕ_s jump schemes also do very well in rematching. But ϕ_s jump cannot avoid blowup due to microwave instability.



If one wants both to reestablish matching and to cure microwave instability one must employ γ_t jump or both γ_t and ϕ_s jump. The optimal design of the jump(s) is best done by using multi-particle numerical simulation.