SLAC-PUB-6073
March 1993
(A)

The Beam-Beam Interaction in e^+e^- Storage Rings

Robert H. Siemann*
Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

1 Introduction

Colliders are designed for studying relatively rare, small impact parameter collisions that produce elementary particles. This is not the dominant interaction between the beams, however. That dominant interaction, the beam-beam interaction, is due to the electromagnetic fields of the beams.

The simplest and most pragmatic treatment of the beam-beam interaction is to parametrize the luminosity, L, in terms of the beam-beam tune shift, ξ. The tune shift is the shift in the vertical betatron tune of a small amplitude particle due to the electromagnetic fields of the other beam. Expressed in terms of beam parameters

$$\xi = \frac{r_e}{2\pi} \frac{N\beta_y^*}{\gamma\sigma_y(\sigma_y+\sigma_x)} \tag{1.1}$$

where $r_e = 2.82\times10^{-15}$ m is the classical electron radius, N is the number of particles in a beam bunch, β_y^* is the vertical β-function at the interaction point, γ is the beam energy in units of rest energy, and σ_y and σ_x are the rms vertical and horizontal beam sizes at the collision point. The luminosity

$$L = \frac{1}{4\pi} \frac{N^2 f_c}{\sigma_y \sigma_x} \tag{1.2}$$

(f_c is the collision frequency) can be rewritten in a frequently used parametrization

$$L = \frac{N f_c \gamma \xi (1+\sigma_y/\sigma_x)}{2 r_e \beta_y^*} \quad . \tag{1.3}$$

* This work was supported by the Department of Energy contract DE-AC03-76SF00515.

Paper Presented at the Joint US-CERN School on Particle Accelerators
Frontiers of Particle Beams: Factories with e^+e^- Rings
Benalmadena, Spain, October 29 - November 4, 1992

It is known from experience that there is a soft limit on ξ, $\xi \lesssim 0.05$. The pragmatist chooses a value of ξ and then uses eq. (1.3) to trade-off luminosity, beam current (the total current is $I_T = eNf_c$), σ_y/σ_x, and β_y^*. A "conservative" design might be based on $\xi = 0.03$ while more aggressive ones might use $\xi = 0.05$. This approach has shortcomings primarily because it isn't based on an understanding of the underlying physics of the beam-beam interaction. Other beam parameters are assumed independent of and unaffected by ξ, and ξ is assumed independent of and unaffected by them. In fact, beam-beam performance depends on many other parameters including some of those in eq. (1.3), betatron and synchrotron tunes, bunch length, lattice errors, and radiation damping. Therefore, eq. (1.3) has limited applicability, and extrapolations into new regimes such as those being considered for heavy quark factories - small β_y^*, crossing angles, closely spaced bunches, unequal beam energies, etc - have uncertainties associated with them. Furthermore, a "conservative" choice of ξ may not lead to a conservative overall design. Other beam parameters or accelerator systems such as the RF and vacuum systems may be pushed unnecessarily.

This article is a personal perspective about the physics of the beam-beam interaction. This is an active area of research combining operational experience, experiments, computer models, and theory with the goal being to overcome the shortcomings above. This research hasn't progressed sufficiently to quantitatively explain beam-beam limits, but there are qualitative explanations of many of the features of the beam-beam interaction and clear directions for future developments.

2 Observations

Experimental aspects of the beam-beam interaction are the subject of several articles that give detailed observations.[1-6] Some of these papers synthesize data from several storage rings. This is a difficult task because there are numerous important parameters, and when colliders are compared many of these parameters are different. It is hard to know which of these differences are essential, which are secondary, and even if all the differences have been identified. Rather than repeating this type of quantitative analysis, this section stresses a common feature of the observations - there are two beam-beam limits.

Normally beams are injected into a collider with electrostatic separation at the interaction point. While the beams are still separated they are Gaussian with rms beam sizes at the interaction point of

$$\sigma_{x0} = \sqrt{\beta_x^* \varepsilon_x} \quad ; \quad \sigma_{y0} = \sqrt{\beta_y^* \varepsilon_y} \quad . \tag{2.1}$$

The emittances, ε_x and ε_y, are determined by the magnet lattice and the properties of synchrotron radiation.[7] Most electron storage rings have had $\sigma_{x0} >> \sigma_{y0}$. That is the natural relation between the sizes because: i) the main dipoles bend in the horizontal plane leading to $\varepsilon_x >> \varepsilon_y$, and ii) a quadrupole doublet is the simplest interaction region configuration. If the quadrupole polarities are chosen to give

$\beta_x^* >> \beta_y^*$, the tune shifts in the two planes, ξ_x (given by eq. (1.1) with x and y interchanged) and ξ_y ($\equiv \xi$), can be made roughly equal

$$\frac{\xi_x}{\xi_y} = \frac{\beta_x^*/\sigma_x}{\beta_y^*/\sigma_y} = \sqrt{\frac{\beta_x^* \varepsilon_y}{\beta_y^* \varepsilon_x}} \quad . \tag{2.2}$$

Synchrotron light measurements show that the horizontal size doesn't change significantly when the beams are brought into collision, $\sigma_x = \sigma_{x0}$. However, above some current the vertical beam size becomes dependent on N rather than being determined by synchrotron radiation, $\sigma_y \neq \sigma_{y0}$. Figures 1 and 2 are a compilation of luminosity and tune shift data for colliders when performance has been optimized. There are two distinct regimes. At low currents $L \propto N^2$ (I_T^2) indicating $\sigma_x \sigma_y$ is constant. The data do not show this regime for all colliders because for them it doesn't exist or is below the current range presented. At high currents $L \propto N$, and, since the horizontal size is unchanged, $\sigma_y \propto N$. TRISTAN has a single beam current limit and is the only collider not to reach this regime. The tune shift is derived from the luminosity measurements using eq. (1.3), so the tune shift plots are another way of presenting the same data. They show the tune shift reaching a limit. This is one of the beam-beam limits. It is associated with the beam core, and, as the figures show, the tune shift limit varies from collider to collider.

The second beam-beam limit comes from changes of the beam distribution. The dominant effect is the appearance of "non-Gaussian" tails in the vertical - the number of particles with large vertical betatron amplitude is greater than that of a Gaussian distribution with the measured rms width of the core. The beam-beam interaction increases the population of the tail of the distribution even more than it increases the core size.

Figure 3 shows a beam distribution measurement. The non-Gaussian tails are clear. Systematic studies of the tails have never been made because of the difficulty of the measurements. The luminosity is insensitive to the tails, and synchrotron radiation monitors are plagued by unwanted reflections from the vacuum chamber that dominate the image beyond 2 - 3 σ_y. The best measurement technique is destructive - measuring the lifetime as a collimator is moved toward the center of the beam. Systematic studies are tedious because new beams must be injected after each measurement.

Particles with sufficiently large amplitudes hit the vacuum pipe causing experimental backgrounds and reducing the beam lifetime. The dynamic aperture due to magnet nonlinearities may play a role in that particles may fall outside the dynamic aperture before hitting the physical aperture. The second beam-beam limit has been reached when the lifetime or backgrounds become unacceptable. This beam-beam limit is associated with the beam tails. It isn't parametrized by a value of ξ; ξ is determined by the size of the core and has reached its limiting value below the second beam-beam limit.

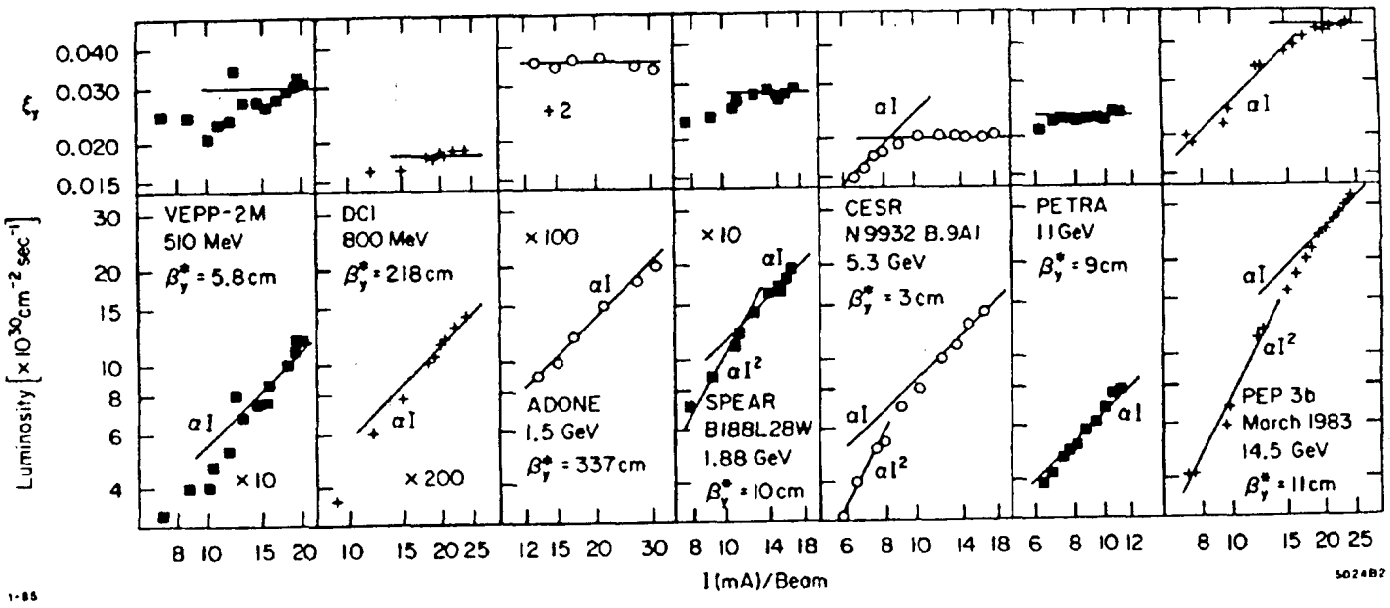

Figure 1: The beam-beam performance of e^+e^- colliders from an article by John Seeman (Ref. 1).

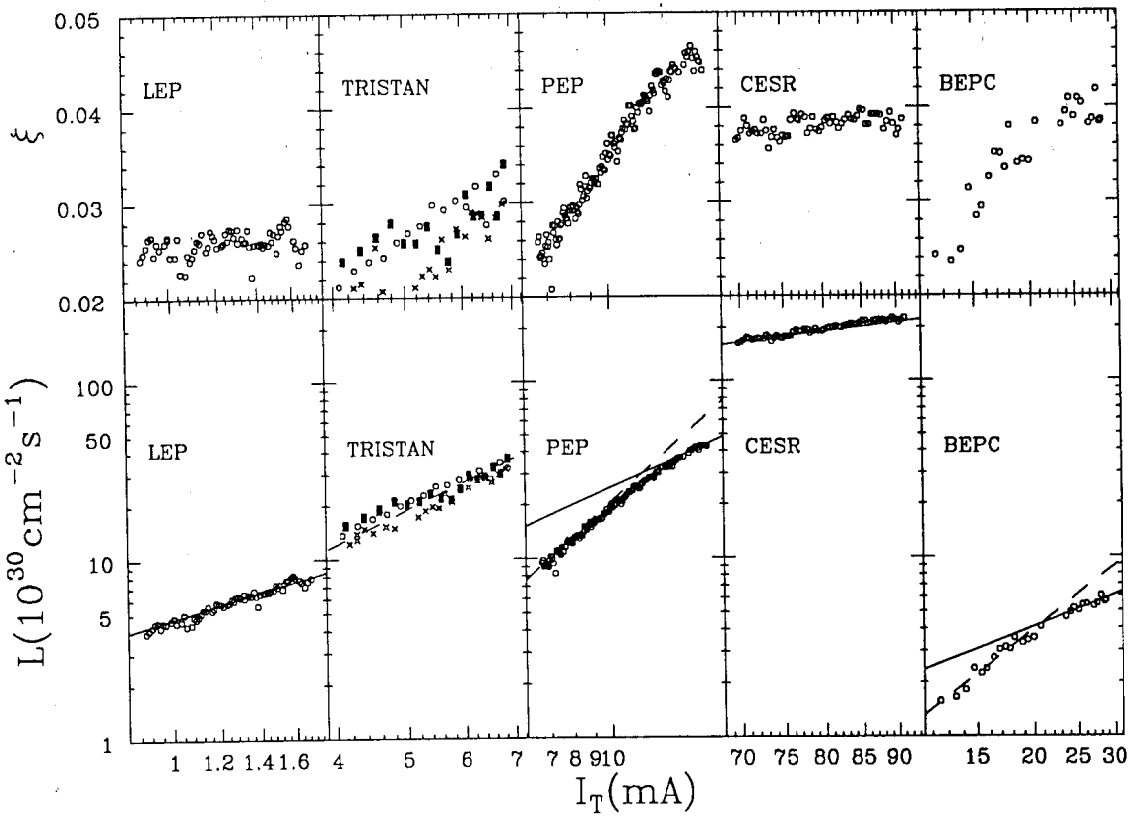

Figure 2: An update of beam-beam performance. The LEP data were provided by J. Jowett, the TRISTAN data by H. Fukuma and Y. Funakoshi, the PEP data by M. Donald, the CESR data by D. Rice, and the BEPC data by C. Zhang. The solid lines are eyeball fits to $L \propto I_T$, and the dashed lines to $L \propto I_T^2$. The different symbols for TRISTAN indicate measurements by different detectors.

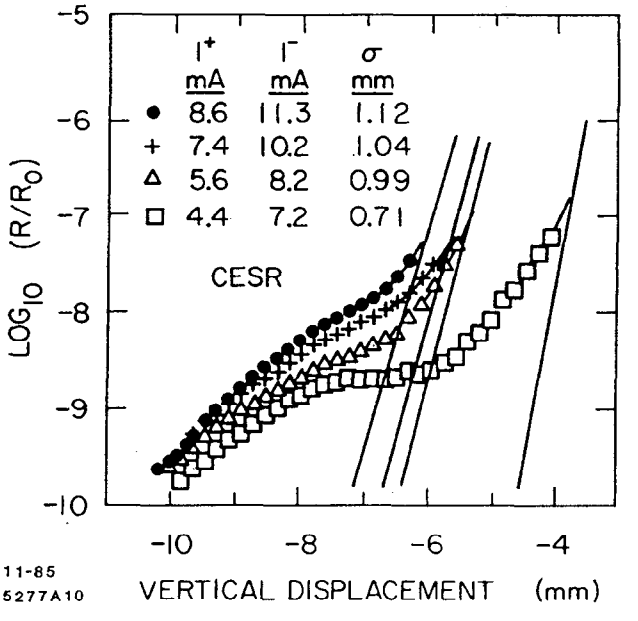

Figure 3: Rate of bremsstrahlung photons produced by a thin Be wire in the vertical tail of the CESR beam. R_0 is the rate at the center of the beam, and the solid lines are extrapolations of the Gaussian core (ref. 8).

3 Underlying Physics

The angles of a relativistic particle passing through an oncoming beam (Figure 4) change by[9]

$$\left.\begin{matrix}\Delta x' \\ \Delta y'\end{matrix}\right\} = \frac{-Nr_e}{\gamma\sigma_x}\sqrt{\frac{2\pi}{1-R^2}}\left.\begin{matrix}\mathrm{Im} \\ \mathrm{Re}\end{matrix}\right\} f_{BB} \quad ; \tag{3.1}$$

$$f_{BB} = W(u+iRv) - \exp(-(1-R^2)(u^2+v^2))\,W(Ru+iv) \quad .$$

In this equation $\sigma_x > \sigma_y$ is assumed, W is the complex error function,[10] $R = \sigma_y/\sigma_x$, and

$$u = \frac{x}{\sigma_x\sqrt{2(1-R^2)}} \quad ; \quad v = \frac{y}{\sigma_y\sqrt{2(1-R^2)}} \quad . \tag{3.2}$$

An example of $\Delta y'$ is plotted in Figure 5. At small y/σ_y, $\Delta y' \propto y$ while at large values it falls like 1/y.

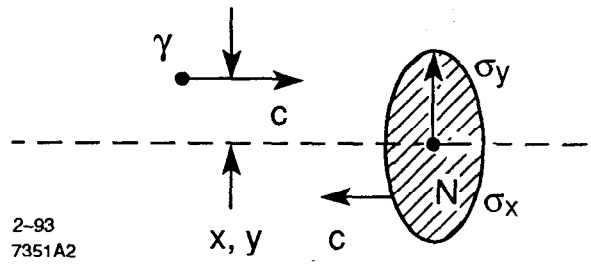

Figure 4: A relativistic particle of energy γ passing through an oncoming thin pancake of N oppositely charged particles. The transverse distributions of the pancake are Gaussian with rms widths σ_x and σ_y. The particle is displaced from the center of the oncoming beam by (x,y).

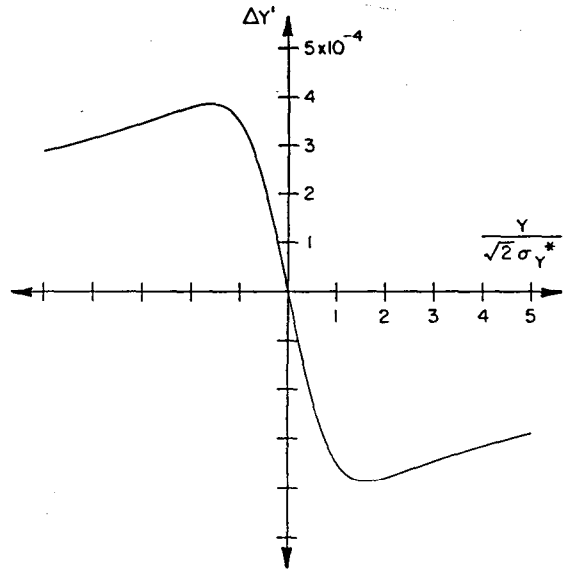

Figure 5: The vertical kick experienced by a typical electron in CESR (ref. 11).

The complex error function can be approximated for small arguments as $W(\zeta) = 1 + 2i\zeta/\pi^{1/2}$ giving

$$\left.\begin{matrix}\Delta x' \\ \Delta y'\end{matrix}\right\} = \frac{-2Nr_e}{\gamma(\sigma_x+\sigma_y)} \left.\begin{matrix}\mathrm{Im} \\ \mathrm{Re}\end{matrix}\right\} \left\{ i\frac{x}{\sigma_x} + \frac{y}{\sigma_y} \right\} . \qquad (3.3)$$

This is a focusing quadrupole in both dimensions. Displacement, betatron amplitude and betatron phase are related by $z = A_z\cos(\psi_z)$ where z = x or y. Small amplitude particles experience this linear focusing for all values of phase, and it changes their tunes by

$$\Delta Q_y = \frac{r_e}{2\pi} \frac{N\beta_y^*}{\gamma\sigma_y(\sigma_y+\sigma_x)} = \xi_y \equiv \xi \ , \tag{3.4}$$

and

$$\Delta Q_x = \frac{r_e}{2\pi} \frac{N\beta_x^*}{\gamma\sigma_x(\sigma_y+\sigma_x)} = \xi_x \ . \tag{3.5}$$

Particles with larger betatron amplitudes are outside the linear region of eq. (3.1) for a range of phases, and the focusing averaged over phase is weaker than for a small amplitude particle. Very large amplitude particles have almost no change in tune due to the beam-beam interaction. The beam has a range of tunes from $\{\Delta Q_x, \Delta Q_y\} = \{\xi_x, \xi_y\}$ at small amplitudes to $\{0, 0\}$ at very large amplitudes. For this reason ξ_x and ξ_y are often, appropriately, called the beam-beam tune spreads. The tune spreads are one consequence of the nonlinearity of the beam-beam interaction.

Sufficiently strong resonances within the tune spreads can lead to beam-beam phenomena like that discussed above. Those resonances can be associated with the magnet lattice, or they can come from the beam-beam interaction itself which produces nonlinear resonances because $\Delta x'$ and $\Delta y'$ are nonlinear functions of x and y. These nonlinear resonances are the second consequence of the nonlinearity of the beam-beam interaction.

So far single particle physics has been discussed, the effects of the fields of one beam on a particle in the second beam. This is *incoherent, weak-strong* physics. The particles in the second beam are independent of each other, and the fields and distribution of the first beam are unaffected by the second beam. In fact, the fields at the collision point are strong, and the beams modify each other's distributions and fields. This leads to multiple particle, *coherent, strong-strong* physics.

The connection between the physics regime and core blowup and the tune shift limit is an open issue. It could be incoherent or coherent physics, and it could be that one or the other is more important in a particular collider. Most theoretical analyses of the tune shift limit concentrate on incoherent physics, and most computer simulations make approximations that inhibit coherent physics. Experiments that would distinguish between them have not been performed, but there are some observations that can be explained only by coherent physics.

The large amplitude particles leading to lifetime and background limitations are rare. These rare particles are independent of each other and cannot affect the distribution of the other beam. Clearly these particles are described by single particle, incoherent physics.

4 The Incoherent Beam-Beam Interaction

4.1 Hamiltonian Formalism

Hamiltonian methods are extremely powerful. They can be used to calculate many aspects of the incoherent beam-beam interaction, and they provide a framework for addressing such topical questions as the effects of synchrotron motion, finite bunch length, unequal beam energies, optical errors, crossing angles, etc.

In the absence of the beam-beam interaction the Hamiltonian for the transverse motion of a single particle is[12]

$$H_0(x,p_x,y,p_y,s) = -eA_s(s) - \left[1+\frac{x}{\rho(s)}\right]\left[p^2-p_x^2-p_y^2\right]^{1/2} . \quad (4.1)$$

In this equation: s is the usual coordinate along the reference orbit, ρ is the bending radius, A_s is the s-component of the vector potential describing the magnet lattice, $p = [E^2/c^2 - m^2c^2]^{1/2}$ is the total momentum, and p_x and p_y are the momenta conjugate to x and y. The ideal solution is to find constants of the motion. This is possible when the lattice is made up of dipoles and quadrupoles. In that case canonical transformations can be used to simplify H_0 to[12]

$$H_0 = \frac{2\pi Q_{x0}}{C} I_x + \frac{2\pi Q_{y0}}{C} I_y \quad (4.2)$$

where C is the accelerator circumference, Q_{x0} and Q_{y0} are the betatron tunes and it has been assumed for simplicity of illustration that there are no skew quadrupoles. The canonical transformation has been used to transform to "action-angle coordinates"; the actions, I_x and I_y, are constants of the motion because their conjugate coordinates, the angles ψ_x and ψ_y, do not appear in the Hamiltonian and $dI_z/ds = -\partial H_0/\partial\psi_z = 0$ (z = x or y).

The "smooth approximation" is a convenient description of transverse motion where the betatron phases advance at constant rates, $d\psi_z/ds = 2\pi Q_{z0}/C$, rather than at the actual instantaneous rates $d\psi_z/ds = 1/\beta_z(s)$. The transverse coordinates in terms of the action-angle coordinates of the Hamiltonian in eq. (4.2) are

$$z = \sqrt{2\beta_z(s) I_z}\cos(\psi_z+\chi_z(s)) \quad (4.3)$$

where

$$\chi_z(s) = \int_0^s \frac{d\zeta}{\beta_z(\zeta)} - Q_{z0}\frac{2\pi s}{C} . \quad (4.4)$$

The angle ψ_z is the betatron phase in the smooth approximation, $d\psi_z/ds = \partial H_0/\partial I_z = 2\pi Q_{z0}/C$, and $\chi_z(s)$ accounts for the difference between the approximate and actual phases. (There are alternate action-angle coordinates where the angle is the betatron phase rather than the phase in the smooth approximation, but using the choice in eqs. (4.2) - (4.4) allows one to account easily for the rapid phase advance near the interaction point where β is small.) Note that χ_z is periodic with period C

$$\chi_z(s+nC) = \chi_z(s) \quad . \tag{4.5}$$

The Hamiltonian including the beam-beam interaction is

$$H(x,p_x,y,p_y,s) = H_0+V_{BB}(x,y,s) \tag{4.6}$$

where V_{BB} is the beam-beam potential. This potential is nonlinear, and it isn't possible to find constants of the motion. A perturbation method must be used. The steps in this method are:

1. Write the Hamiltonian, H, in terms of the action-angle coordinates of the unperturbed Hamiltonian, H_0.
2. Fourier analyze H with respect to s, ψ_x, and ψ_y.
3. Calculate the dependences of tunes on action from the average value of the perturbation.
4. Determine resonance conditions and resonance properties from the slowly varying terms of H.

A sample calculation is performed in the next section to illustrate the techniques and arrive at conclusions relevant to heavy quark factories. It is a generalization of previously published work[13,14] to finite length, flat beams.

4.2 Bunch Length Effects

4.2.1 Write the Hamiltonian, H, in Terms of the Action-Angle Coordinates of the Unperturbed Hamiltonian, H_0

The picture is that given by Figure 4 with the thin pancake replaced by an oncoming bunch with rms length σ_L. The beam-beam potential, assuming only one interaction point, is

$$V_{BB} = \frac{-Nr_e}{\gamma} \frac{2}{\sqrt{2\pi\sigma_L^2}} \sum_{n=-\infty}^{\infty} V_F(x,y,s) \exp\left[-2(s-(nC+c\tau))^2/\sigma_L^2\right] . \tag{4.7}$$

The potential V_F comes from a solution of Poisson's equation[15]

$$V_F = \int_0^{\infty} \frac{dq}{\sqrt{(2\sigma_x^2+q)(2\sigma_y^2+q)}} \exp\left\{-\left[\frac{x^2}{2\sigma_x^2+q} + \frac{y^2}{2\sigma_y^2+q}\right]\right\} ; \tag{4.8}$$

V_F has explicit s dependence because $\sigma_z^2 = \sigma_{z0}^2 + \varepsilon_z s^2/\beta_z^*$ near the collision point. This is important in the vertical dimension, and the calculation is valid for $\sigma_L/\beta_y^* < 1$. The modulation of the collision point due to synchrotron motion of tune Q_s and amplitude $\hat{\tau}$ is

$$\tau = \frac{\hat{\tau}}{2} \cos(2\pi n Q_s) \tag{4.9}$$

where eqs. (4.7) and (4.9) have a factor of one-half associated with them that arises from the relative motion of the particle and the opposing beam.

The transverse coordinates can be rewritten using eq. (4.3) to give

$$V_F(I_x,\psi_x,I_y,\psi_y,s) =$$

$$\int_0^\infty \frac{dq}{\sqrt{(2\sigma_x^2+q)\,(2\sigma_y^2+q)}} \exp\left\{-\left[\frac{2\beta_x I_x \cos^2(\psi_x+\chi_x)}{2\sigma_x^2+q} + \frac{2\beta_y I_y \cos^2(\psi_y+\chi_y)}{2\sigma_y^2+q}\right]\right\} \tag{4.10}$$

$$\simeq \int_0^\infty \frac{dq}{\sqrt{(2\sigma_{x0}^2+q)\,(2\sigma_{y0}^2+q)}} \exp\left\{-\left[\frac{2\beta_x^* I_x \cos^2(\psi_x+\chi_x)}{2\sigma_{x0}^2+q} + \frac{2\beta_y^* I_y \cos^2(\psi_y+\chi_y)}{2\sigma_{y0}^2+q}\right]\right\} . \tag{4.11}$$

The approximation $\sigma_L/\beta_y^* < 1$ was used in going from eq. (4.10) to (4.11), and the only remaining s-dependence in eq. (4.11) is in χ_x and χ_y.

4.2.2 Fourier Analyze H with Respect to s, ψ_x, and ψ_y

This Fourier analysis will show the resonant structure of the beam-beam interaction. The Hamiltonian is periodic in ψ_x and ψ_y with period 2π, so

$$H = H_0(I_x,I_y) - \frac{Nr_e}{\gamma} \sum_{p,r=-\infty}^{\infty} \int_{-\infty}^{\infty} dk\, A_{pr}(I_x,I_y,k) \exp(i(p\psi_x+r\psi_y-ks)) . \tag{4.12}$$

As discussed below coefficient A_{pr} corresponds to resonances with horizontal order |p| and vertical order |r|. It is

$$A_{pr} = \frac{1}{(2\pi)^3} \int_0^{2\pi} d\psi_x \int_0^{2\pi} d\psi_y \int_{-\infty}^{\infty} ds\, \exp(-i(p\psi_x+r\psi_y-ks)) \times \frac{2}{\sqrt{2\pi\sigma_L^2}} \sum_{n=-\infty}^{\infty} V_F \exp\left[-2(s-(nC+c\tau))^2/\sigma_L^2\right] . \tag{4.13}$$

A first result of making the approximation going from eq. (4.10) to eq. (4.11) is that the ψ integrals and the s integral can be factored by making a change of variables $\theta_z = \psi_z + \chi_z$

$$A_{pr}(I_x,I_y,k) = T_{pr}(I_x,I_y)\ \frac{1}{2\pi}\int_{-\infty}^{\infty} ds\ \exp(i(p\chi_x+r\chi_y+ks))$$

$$\times\ \frac{2}{\sqrt{2\pi\sigma_L^2}}\sum_{n=-\infty}^{\infty}\exp\left[-2(s-(nC+c\tau))^2/\sigma_L^2\right]\ ; \qquad (4.14)$$

$$T_{pr} = \frac{1}{(2\pi)^2}\int_0^{2\pi} d\theta_x e^{-ip\theta_x}\int_0^{2\pi} d\theta_y e^{-ir\theta_y}\int_0^{\infty}\frac{dq}{\sqrt{(2\sigma_{x0}^2+q)\,(2\sigma_{y0}^2+q)}}$$

$$\times\ \exp\left\{-\left[\frac{2\beta_x^* I_x\cos^2\theta_x}{2\sigma_{x0}^2+q}+\frac{2\beta_y^* I_y\cos^2\theta_y}{2\sigma_{y0}^2+q}\right]\right\}\ . \qquad (4.15)$$

The integral T_{pr} is zero when either p or r is odd.

The s integral can be performed by i) using the periodicity of χ_z (eq. (4.5)), ii) making a second use of the approximation $\sigma_L/\beta_y^* < 1$ to write $\chi_z(s) = s/\beta_z^* - 2\pi Q_z s/C$, and iii) using an integral from Gradshteyn and Ryzhik[16]

$$A_{pr} = \frac{1}{2\pi}\,T_{pr}(I_x,I_y)\,\exp\left[\frac{-k_{pr}^2\sigma_L^2}{8}\right]$$

$$\times\ \sum_{n=-\infty}^{\infty}\exp\left[\frac{ik_{pr}\hat{\tau}c}{2}\cos(2\pi nQ_s)+iknC\right]\ . \qquad (4.16)$$

The wavenumber k_{pr} is

$$k_{pr} = k + p(1/\beta_x^* - 2\pi Q_{x0}/C) + r(1/\beta_y^* - 2\pi Q_{y0}/C)\ . \qquad (4.17)$$

Finally, using a Bessel function sum[17] and the Poisson sum rule[18]

$$A_{pr} = \frac{1}{C}\,T_{pr}(I_x,I_y)\,\exp\left[\frac{-k_{pr}^2\sigma_L^2}{8}\right]$$

$$\times\ \sum_{m,\,n=-\infty}^{\infty} i^m J_m(k_{pr}\hat{\tau}c/2)\,\delta(kC-2\pi(n-mQ_s)) \qquad (4.18)$$

where J_m is a Bessel function of order m. Sidebands spaced by Q_s have developed around the betatron resonances. The decomposition in eq. (4.18) makes sense only when a few of the sidebands are important.

Substituting back into eq. (4.12)

$$H = H_0(I_x, I_y) - \frac{Nr_e}{C\gamma} \sum_{\substack{m,n,p,\\ r=-\infty}}^{\infty} T_{pr}(I_x, I_y) \exp\left[\frac{-k_{pr}^2 \sigma_L^2}{8}\right] \times i^m J_m(k_{pr}\hat{\tau}c/2) \exp(i(p\psi_x + r\psi_y - 2\pi(n - mQ_s)s/C)) \quad . \tag{4.19}$$

The four-fold sum is the beam-beam perturbation. This all seems terribly messy, but the next steps show how this formality pays off.

4.2.3 Calculate the Dependences of the Tunes on Action from the Average Value of the Perturbation

The average value of the perturbation is given by the term with p = r = m = n = 0. All the other terms are oscillatory. When its phase varies rapidly, the effect of a term on the motion averages to zero quickly. A phase varies slowly if the tunes have special values leading to resonances, or if one or more of the tunes is low. Usually the fractional parts of the betatron tunes are not close to zero, but the synchrotron tune can be low. It is in hadron colliders where the effects of synchrotron motion are averaged over hundreds or thousands of turns. This leads to the possibility of adiabatic behavior where: i) resonance conditions change slowly enough that particles trapped in a resonance stay trapped as resonance conditions change,* and ii) the decomposition in eq. (4.18) isn't the most illuminating approach.

The synchrotron tune in electron colliders is large enough that the synchrotron motion is averaged in tens of turns. The important terms in the perturbation are its average value and a few resonances. The phase advance is $d\psi_z/ds = \partial H/\partial I_z$. The derivative of H_0 gives $2\pi/C$ times the nominal tune, and the derivative of the average value of the perturbation gives the average phase advance from the perturbation. It is

$$\left\langle \frac{d\psi_z}{ds} \right\rangle \left[\equiv \frac{2\pi}{C} Q_z \right] = \left\langle \frac{\partial H}{\partial I_z} \right\rangle = \frac{2\pi}{C} Q_{z0} - \frac{Nr_e}{C\gamma} \frac{\partial T_{00}}{\partial I_z} \quad . \tag{4.20}$$

For example, in the vertical

---

* The synchrotron amplitude and tune enter the criteria for adiabatic motion, but the focus is on Q_s because $\hat{\tau}/\beta^* \sim 1$ in modern electron and hadron colliders. The differences between the "quasilinear" and "adiabatic" regimes are discussed in ref. 19.

$$Q_y = Q_{y0} + \frac{2Nr_e\beta_y^*}{\gamma(2\pi)^3}\int_0^{2\pi} d\theta_x \int_0^{2\pi} d\theta_y \cos^2\theta_y \int_0^{\infty} \frac{dq}{\sqrt{(2\sigma_{x0}^2+q)(2\sigma_{y0}^2+q)^3}}$$

$$\times \exp\left\{-\left[\frac{2\beta_x^* I_x \cos^2\theta_x}{2\sigma_{x0}^2+q} + \frac{2\beta_y^* I_y \cos^2\theta_y}{2\sigma_{y0}^2+q}\right]\right\} . \qquad (4.21)$$

In the limit $I_x, I_y \to 0$,

$$Q_y = Q_{y0} + \frac{Nr_e\beta_y^*}{2\pi\gamma}\int_0^{\infty} \frac{dq}{\sqrt{(2\sigma_{x0}^2+q)(2\sigma_{y0}^2+q)^3}} . \qquad (4.22)$$

Evaluating the integral

$$Q_y = Q_{y0} + \xi_y . \qquad (4.23)$$

Equation (4.21) and the analogous one for Q_x (eq. (4.21) with x and y interchanged) give the dependences of tunes on actions. Usually the integrals have to be done numerically. Figure 6 shows the results of such a calculation.

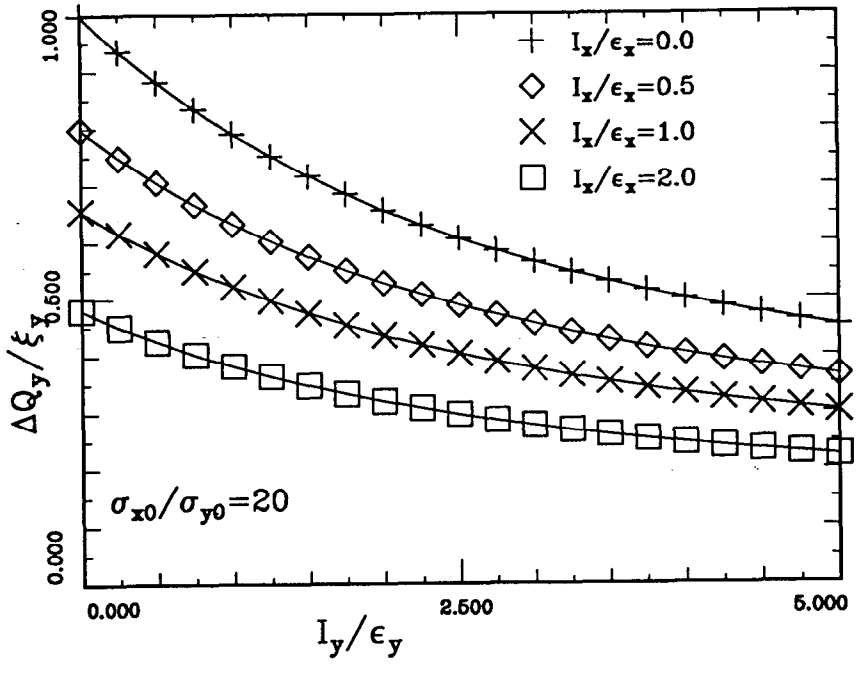

Figure 6: The vertical tune, $Q_y = Q_{y0} + \Delta Q_y$, for different values of the horizontal and vertical actions calculated from eq. (4.21).

4.2.4 Determine Resonance Conditions and Resonance Properties from the Slowly Varying Terms of H

A term in the Hamiltonian has a cumulative effect, resonant build-up, when its phase varies slowly. The resonance condition is

$$\frac{d}{ds}\left[p\psi_x + r\psi_y - 2\pi(n-mQ_s)\frac{s}{C}\right] \tilde{=} 0 \ , \qquad (4.24)$$

or

$$pQ_x + rQ_y + mQ_s = n \ . \qquad (4.25)$$

The tunes are the actual tunes, not the nominal tunes, and eqs. (4.21), the analogous one for Q_x, and (4.25) can be solved for the locus of resonant actions, $\{I_{xR}, I_{yR}\}$. This is illustrated in Figure 7. The resonance order equals |p| + |r| + |m|.

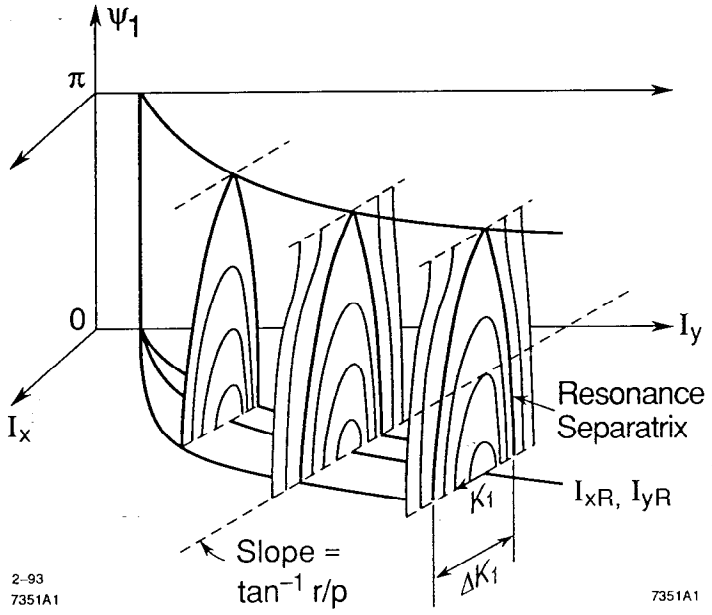

Figure 7: A nonlinear resonance illustrated in a three-dimensional phase space. The axes are the unperturbed actions I_x and I_y and the resonant angle $\psi_1 = p\psi_x + r\psi_y - 2\pi(n-mQ_s)s/C$. The resonance condition, eq. (4.24), is satisfied for $\{I_{xR}, I_{yR}\}$. The resonance action is K_1, and, from eq. (4.29), the resonance oscillation direction is $\tan^{-1}(r/p)$. This figure is adapted from ref. 20.

The Hamiltonian of a single resonance, assuming it is isolated from all the others, is

$$H_{prm} = H_0 - \frac{Nr_e}{C\gamma} T_{00} - \frac{2Nr_e}{C\gamma} T_{pr}(I_x, I_y) \exp\left\{-\frac{1}{2}\left[\frac{p\sigma_L}{2\beta_x^*} + \frac{r\sigma_L}{2\beta_y^*}\right]^2\right\}$$

$$\times J_m\left\{\left[\frac{p}{2\beta_x^*} + \frac{r}{2\beta_y^*}\right]c\hat{\tau}\right\}\cos(p\psi_x + r\psi_y - 2\pi(n-mQ_s)s/C) \ , \tag{4.26}$$

$$= H_0 - \frac{Nr_e}{C\gamma} T_{00} - \frac{2Nr_e}{C\gamma} F_{prm}(I_x, I_y, \hat{\tau})\cos(p\psi_x + \ldots) \ . \tag{4.27}$$

Equation (4.25) has been used, the term with (-p, -r, -m, -n) has been combined with the term for (p, r, m, n), and the initial phase of the last term has been neglected. Usually $\beta_x^* >> \beta_y^* \sim \sigma_L$, and

$$F_{prm}(I_x, I_y, \hat{\tau}) \simeq T_{pr}(I_x, I_y)\exp\left\{-\frac{1}{2}\left[\frac{r\sigma_L}{2\beta_y^*}\right]^2\right\} J_m\left\{\frac{rc\hat{\tau}}{2\beta_y^*}\right\} \ . \tag{4.28}$$

A standard, well-known canonical transformation can be used to go from the action-angle coordinates of the unperturbed Hamiltonian to the action-angle coordinates of a resonance Hamiltonian.[19] Two new pairs of action-angle coordinates result; one pair is K_1 and $\psi_1 = p\psi_x + r\psi_y - 2\pi(n-mQ_s)s/C$ which are the action-angle coordinates of the resonance. The second action is a constant of the motion; this leads to a constraint

$$pI_y - rI_x = \text{constant} \ . \tag{4.29}$$

The resonance is illustrated in Figure 7.

There are linear oscillations about the center of the resonance with

$$\frac{d\psi_1(\hat{\tau})}{ds} = \frac{Nr_e}{C\gamma}\sqrt{\left|2F^R_{prm}(\hat{\tau})\Lambda_{pr}\right|} \tag{4.30}$$

for small values of K_1, and the full width of the resonance separatrix is

$$\Delta K_1 = 4\sqrt{\left|\frac{2F^R_{prm}(\hat{\tau})}{\Lambda_{pr}}\right|} \ . \tag{4.31}$$

The quantity Λ_{pr} is proportional to the rate of change of tune with action

$$\frac{dQ_1}{dK_1} = \frac{C}{2\pi}\frac{d^2\Psi_1}{ds dK_1} = \frac{Nr_e}{2\pi\gamma}\Lambda_{pr}$$

$$= \frac{Nr_e}{2\pi\gamma}\left[p^2\frac{\partial^2 T_{00}}{\partial I_x^2} + r^2\frac{\partial^2 T_{00}}{\partial I_y^2} + 2pr\frac{\partial^2 T_{00}}{\partial I_x \partial I_y}\right]\Bigg|_{I_{xR}, I_{yR}}, \qquad (4.32)$$

and

$$F^R_{prm}(\hat{\tau}) = F_{prm}(I_{xR}, I_{yR}, \hat{\tau}) \ . \qquad (4.33)$$

4.2.5 Discussion

The last section contains the results. The factors determining resonance properties are:

1. The strength of the perturbation, Nr_e/γ. The small amplitude frequency depends on it, but the separatrix size is independent of it. Qualitatively, the resonance potential well stays the same size but gets deeper as the beam-beam strength increases.
2. The rate of change of the resonance tune, $pQ_x + rQ_y + mQ_s$, along the direction of oscillation in the $\{I_x, I_y\}$ plane is proportional to the "detuning", Λ_{pr}.
3. The remainder of the dependence on the resonant actions is given by $T_{pr}(I_{xR}, I_{yR})$. It must be calculated numerically; sample calculations are shown in Figure 8. When p or r is odd, $T_{pr} = 0$. Odd order resonances can be introduced by an offset at the interaction point. When I_{xR}/ε_x and I_{yR}/ε_y are small, T_{pr} is small and the potential well is small and shallow. As p and r increase, T_{pr} decreases, reducing the importance of high order resonances.
4. The form factor $\exp(-\frac{1}{2}(r\sigma_L/2\beta^*_y)^2)$ accounts for the nonlinear force acting over a range of vertical betatron phase. The resultant phase averaging increases with r, the vertical order of the resonance. The horizontal phase does not change over σ_L, and, therefore, p and β^*_x do not enter. It is likely that phase averaging is the mechanism contributing to good CESR performance with $\sigma_L/\beta^*_y \sim 1.1$.[21]
5. There are resonances with m = 0 involving betatron motion only and synchrobetatron resonances with $m \neq 0$ arising from the modulation of the collision point from synchrotron motion. $J_m(rc\hat{\tau}/2\beta^*_y)$ gives the dependence on synchrotron amplitude. The Bessel function $J_m(\zeta)$ has its first maximum at $\zeta \sim m$. The m*th* synchrobetatron resonance is important for $\hat{\tau} \geq 2\beta^*_y m/rc$, and particles with large synchrotron amplitudes have more synchrobetatron resonances.

"Difference" resonances have sign(p) = -sign(r), and, from eq. (4.29), $|p|I_y + |r|I_x$ is a constant. The energy associated with the transverse motion can be transferred between horizontal and vertical motions as long as the above sum remains constant. "Sum" resonances have sign(p) = sign(r) and the constant of motion is

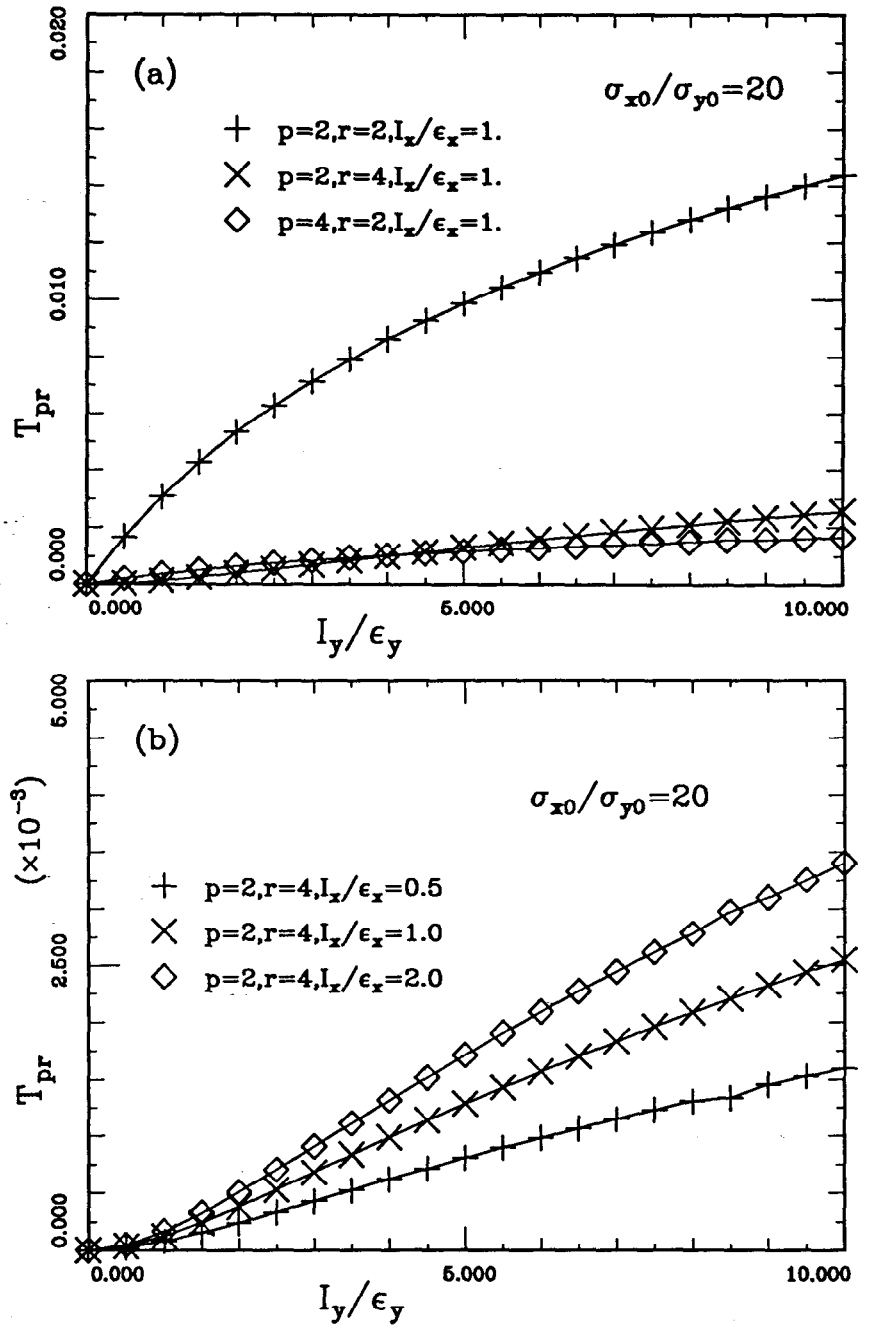

Figure 8: Plots of T_{pr} (eq. (4.15)) for different resonances and different values of the actions.

$|p|I_y - |r|I_x$. There is no restriction on the energy associated with transverse motion. The horizontal and vertical actions can grow without bound, and reduced lifetimes are probably associated with sum resonances. Radiation damping is outside the scope of this Hamiltonian analysis, but the naive expectation is that it limits the actions. In fact, it may enhance the ability of particles to reach large amplitudes through resonance streaming.[22]

The discussion so far has concentrated on single, isolated resonances. They explain the beam-beam performance in hadron colliders[23] and much of the beam core behavior in e^+e^- simulations,* but the interaction between resonances could be important, particularly for lifetime effects. Synchrobetatron sidebands are separated in tune by Q_s and in action by

$$\delta K_1 = \frac{2\pi\gamma Q_s}{Nr_e |\Lambda_{pr}|} \quad . \tag{4.34}$$

Stochastic motion occurs when resonances overlap.[26] The Chirikov criterion is that there is resonance overlap and resulting chaotic motion when

$$\frac{\Delta K_1}{\delta K_1} > \frac{2}{\pi} \quad \text{or} \quad \frac{Nr_e}{\gamma} > \frac{Q_s}{\sqrt{2\left|F^R_{prm}(\hat{\tau})\Lambda_{pr}\right|}} \quad . \tag{4.35}$$

The threshold Nr_e/γ for chaotic motion decreases with Q_s until the adiabatic regime with stable motion is entered. At this point the picture of separated synchrobetatron sidebands is not appropriate, and there is a transition from the quasilinear to the adiabatic regime.[19]

Particles with large synchrotron amplitudes have a number of sidebands, and the resonance overlap criterion suggests a connection between particles that are determining the lifetime and particles with large synchrotron amplitudes. Such connections have been seen in simulations,[27] but, as far as I know, there is no convincing connection between particles with large synchrotron amplitudes and particle losses. An alternative mechanism for reaching large amplitudes involves the interaction between nonlinear resonances and synchrotron radiation. It is the subject of the next section.

4.3 Resonance Streaming[22] and Phase Convection[20]

Nonlinear resonances can combine with the noise and damping from synchrotron radiation to produce non-Gaussian tails and, possibility, explain the reduced lifetime that is the second beam-beam limit. This possibility has motivated general studies of the interaction of nonlinear resonances, noise, and damping,[20,22]

* References 24 and 25 are reviews of beam-beam simulations with complete references.

and the results have been applied to collider lifetimes.[28] Only betatron motion has been considered, and this section has that restriction.

A particle subject to noise and deterministic forces from isolated nonlinear resonances and damping could reach large vertical amplitudes by a variety of routes. The "most probable" route is the one with the weakest net damping. One could imagine starting with a number of particles at the same value of I_y and different values of (I_x, ψ_x, ψ_y) and tracking them backwards in time in a system without noise. They all damp to the origin of phase space, but at different rates. The particle to damp slowest has backtracked along the most probable route to the starting value of I_y.[20] Resonances, through the mechanism of resonance streaming,[22] often determine the most probable route to large vertical amplitudes.

There are two extremes of the relative importance of a resonance versus synchrotron radiation. In one the time for damping and fluctuations to transport a particle across the resonance is short compared to the oscillation period. Resonant build-up cannot occur, and the resonance is not important. The dominant motion in the other extreme is oscillation about the resonance center $\{I_{xR}, I_{yR}\}$. Using eq. (4.25) the slope of the resonance center is

$$\frac{dI_{yR}}{dI_{xR}} = \frac{-\left[p\,\frac{\partial Q_x}{\partial I_x} + r\,\frac{\partial Q_y}{\partial I_x}\right]}{p\,\frac{\partial Q_x}{\partial I_y} + r\,\frac{\partial Q_y}{\partial I_y}} \quad . \tag{4.36}$$

The derivatives are evaluated at $\{I_{xR}, I_{yR}\}$. Tunes decrease with increasing amplitude, so the partial derivatives are all negative. The slope is negative for sum resonances and can be positive or negative for difference resonances. The slopes can be large when the beams are flat and $I_y >> \sigma_{y0}^2$.

The different effects of damping from sum and difference resonances can be understood using Figure 9. Assume for the sake of illustration that the damping is only in the I_x direction. A decrease in I_x changes the centers of the resonance oscillations from A to B. In the case of the sum resonance damping has shifted the oscillation center to a larger vertical action; the damping process has a component along $\{I_{xR}, I_{yR}\}$ that increases the vertical amplitude. This is resonance streaming. It occurs for sum resonances only, as contrasting the left- and right-hand sides of Figure 9 shows, and for flat beams, $\sigma_x/\sigma_y > 15$.[29]

The Fokker-Planck Equation describes the evolution of the phase space density in systems with damping and noise. When the potential is independent of time, there is a stationary solution similar to the Boltzmann distribution. When the potential is time (or s) dependent, as it is for the beam-beam interaction, there is no stationary solution. It is possible to calculate the distribution far from the core and independent of the initial distribution at large times when damping and noise are weaker than the resonances, however.[20] That density is affected strongly by resonances. It is enhanced by resonance streaming with resonances providing the most likely routes to large amplitude. Particles that fall into a sum resonance stream to large amplitudes and then leave the resonance when noise and damping become dominant. Once they

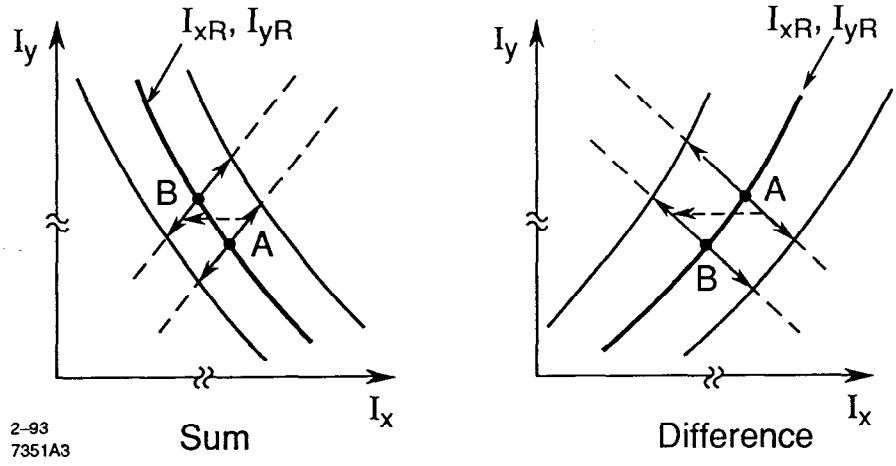

Figure 9: The intersections of resonance "tubes" such as shown in figure 7 with the I_x, I_y plane for sum and difference resonances. The resonance centers, the edges of the separatrices, and the directions of oscillation are shown.

leave the resonance they damp back to the origin of phase space. The result is a circulation of particles in phase space called "phase convection".[20]

The beam-beam interaction does not satisfy the approximations needed for calculating the density distribution far from the core (the beam tails), but a phenomenological model has been used to make estimates.[28] Phase convection and resonance streaming remain as central features of the results. In addition, these estimates show that nonlinearities of the magnet lattice are critical in determining the tails. Parametrizing that nonlinearity as if it were due to single octupole, Gerasimov and Dikansky find that positive lattice nonlinearity, $dQ_{x0}/dI_x > 0$ enhances the tails strongly. This agrees with an experiment at VEPP-4.[3]

4.4 Concluding Remarks

This chapter started with a detailed calculation and ended with qualitative discussions about lifetime limiting mechanisms. The calculation showed the relationship between β_y^* and two longitudinal parameters, σ_L and $\hat{\tau}$. Similar calculations following the same procedure can be used to show the effects of crossing angles, unequal beam energies, phase advance errors between interaction points, etc.

It is difficult to go from the results of these calculations to quantitative statements about beam-beam limits. The value of the calculations is that they give insight into underlying physics, provide a framework for guiding and interpreting experiments and simulations, and identify possible methods for improving performance. Establishing direct, quantitative connections between calculations and performance is one of the themes of ongoing research into the beam-beam interaction. The situation is different for the two beam-beam limits. It isn't clear that the single

particle physics of the incoherent beam-beam interaction is sufficient to explain the behavior of the beam core and the associated tune shift limit. In fact, there is some evidence that it isn't. The role of coherent beam-beam physics needs to be understood.

Beam tails and lifetime limits are clearly single particle physics; they are caused by rare particles far from the core. Three possible mechanisms for particles reaching large amplitudes have been discussed: i) the absence of a restriction on the energy associated with transverse motion for sum resonances, ii) stochastic motion caused by resonance overlap, most likely at large synchrotron amplitudes, and iii) resonance streaming and phase convection. The lifetime limit is a topic ripe for experiments and simulations to evaluate these possibilities.

5 Beam-Beam Simulations

Simulations are an important part of beam-beam research. They are used for making performance estimates of existing or proposed colliders, and they are ideal for performing "experiments" that can be done with a degree of control and a variety of diagnostics that are impossible in real colliders.

Test particles representative of those in the beams are followed for a large number of turns with each turn consisting of transport between interaction points and collisions at the interaction points. Simulation of the arcs almost always includes betatron and synchrotron oscillations, radiation damping and quantum excitations. The techniques are standard.[30] Simulations have included chromaticity, lattice nonlinearities, lattice errors, collective effects, etc depending on the physics under study. References 24 and 25 review much of this work.

The difference between simulations is the treatment of the collisions. This is also the area of recent progress. Simulations are weak-strong when test particles from only one beam, the weak beam, are tracked and the distribution of the opposing strong beam is unaffected by the test particles. This is single particle, incoherent physics. The strong beam is usually Gaussian with the beam-beam impulse given by eq. (3.1). It is necessary to segment the strong beam longitudinally when $\sigma_L \sim \beta_y^*$.[14] This was discovered in a simulation experiment and later understood with a calculation similar to that in section 4.2.

Two beams are tracked in strong-strong simulations, and they modify each other's distributions. The test particles are representative of the beam, and their coordinates are used to determine the distribution and, from that, the electromagnetic fields at the interaction point and the beam-beam impulse. The most common and most straightforward procedure is to calculate the means and rms widths from test particle coordinates, z_k (k = 1,..., K for each beam),

$$\bar{z} = \frac{1}{K} \sum_{k=1}^{K} z_k \ , \qquad (5.1a)$$

$$\sigma_z^2 = \frac{1}{K}\sum_{k=1}^{K}(z_k - \bar{z})^2 \ , \qquad (5.1b)$$

and use these in eq. (3.1) which gives the beam-beam impulse for a Gaussian distribution. This was considered a reasonable approximation because beam distributions remained roughly Gaussian in strong-strong simulations.

A number of effects that are important in operating colliders are seen in strong-strong simulations and are outside the scope of weak-strong simulations. First and foremost is blow-up of the vertical beam size leading to a tune shift limit and luminosity proportional to the total beam current. Actual colliders have been modeled and the tune shift limits from operation and simulation compared. The results are mixed. In general, agreement is found for well-established operating points, but the prediction of new, better operating points is poor. The well-established points have many hours of operator tuning invested in them. This tuning has gradually improved luminosity, presumably through the elimination of small errors that combine with the beam-beam interaction to determine performance. On the other hand, there is rarely enough accelerator studies time to tune extensively at exploratory operating points. It is impossible to know in any detail what errors are removed with tuning and include them in a model. At best one could select errors randomly and simulate an ensemble of colliders to determine the range of possible performance. My conclusion is that i) errors have been tuned out at well-established operating points, the tune shift limit there is due to the beam-beam interaction alone, and, therefore, it can be explained by simulations; and ii) either there has been insufficient tuning at exploratory points or additional physics must be included in simulations. The assumption of the fields from a Gaussian distribution is one possibility that has been investigated recently.

A second effect seen in operation and in strong-strong simulations is the "flip-flop" effect where the two beams have substantially different vertical sizes. This is a hysteretic effect with small differences determining which beam is larger. It is difficult to reproduce actual performance in simulations because of the importance of small differences.

The third common effect is coherent centroid motion which is routinely observed in operations. There are two modes: the "0-mode" where the beam centroids are in phase at the collision point and the "π-mode" where they are 180^o out of phase. These oscillations have limited amplitudes and are helpful as diagnostics for measuring ξ_x and ξ_y since the differences between the π- and 0-mode tunes are[31]

$$\frac{\Delta Q_x}{\xi_x} = \Lambda(r) = 1.330 - 0.370r + 0.279r^2$$

$$\frac{\Delta Q_y}{\xi_y} = (1 - r)\Lambda(r) \qquad (5.2)$$

where $r = \sigma_y/(\sigma_x+\sigma_y)$. The coefficients result from coherent oscillations modifying the charge distribution, and they cannot be reproduced exactly in simulations that restrict the fields to be those of a Gaussian beam.

The increase in vertical beam size and the flip-flop are strong-strong, multiple particle effects, but nonlinear motion of individual particles could account for changes in distributions without adding additional physics. However, coherent centroid oscillations cannot be explained within the framework of the incoherent beam-beam interaction.

The space charge compensation experiments at DCI provide a second, strong piece of evidence that coherent beam-beam effects exist.[32] Those experiments indicate that coherent shape oscillations lead to a tune shift limit. This is in sharp contrast to the harmless nature of coherent centroid oscillations. These experiments had four beams, an electron beam and a positron beam going in one direction colliding with an electron and a positron beam going in the opposite direction. The estimate was that the beam-beam potential was reduced by a factor of ten, and yet there was no striking improvement in performance. The tune shift limit was set by ξ rather than its residual compensated value.

Is there any evidence of coherent shape oscillations in the more normal situation of two beams colliding? There is no experimental evidence. Seeing such oscillations requires imaging a beam on a single turn. Appropriate instruments have become available only recently,[33] and they have not been used in storage ring colliders. There is evidence for coherent oscillations in strong-strong simulations where the beam-beam impulse is calculated for a general distribution rather than using the expression for a Gaussian beam.[34]

Using the means and rms widths, eq. (5.1), together with eq. (3.1) for the impulse may not be a reasonable approximation. Strong-strong simulations are a relaxation calculation; the beam distributions and fields must be consistent with each other. Restricting the fields to those of a Gaussian beam indirectly restricts the beams to remain Gaussian. Relaxing that restriction has a good and a bad effect: it introduces new physics, but it makes the simulation sensitive to noise. Combining eqs. (5.1) and (3.1) is relatively noise free because only a few properties of the beam are extracted from the test particle coordinates. Statistical techniques are needed to distinguish noise from real effects when a general expression for the beam-beam impulse is wanted. An adaptive, least-squares fitting procedure has been developed for nearly round beams, $\sigma_x = \sigma_y$.[35] Coherent shape oscillations are found using this procedure. Figure 10 is an example showing a coherent beam-beam resonance at $Q_{x0} = Q_{y0} = 5/6$. The beam shapes vary turn-by-turn with the extreme being one beam with a dense core and the other with a hollow core (Figure 11). It is impossible to represent such beam shapes with a Gaussian, and it isn't surprising that such turn-by-turn variations in sizes and shapes are not seen when fields from a Gaussian beam are used.

Whether coherent effects are important for flat beams awaits experiments and development of a beam-beam algorithm for flat beams.

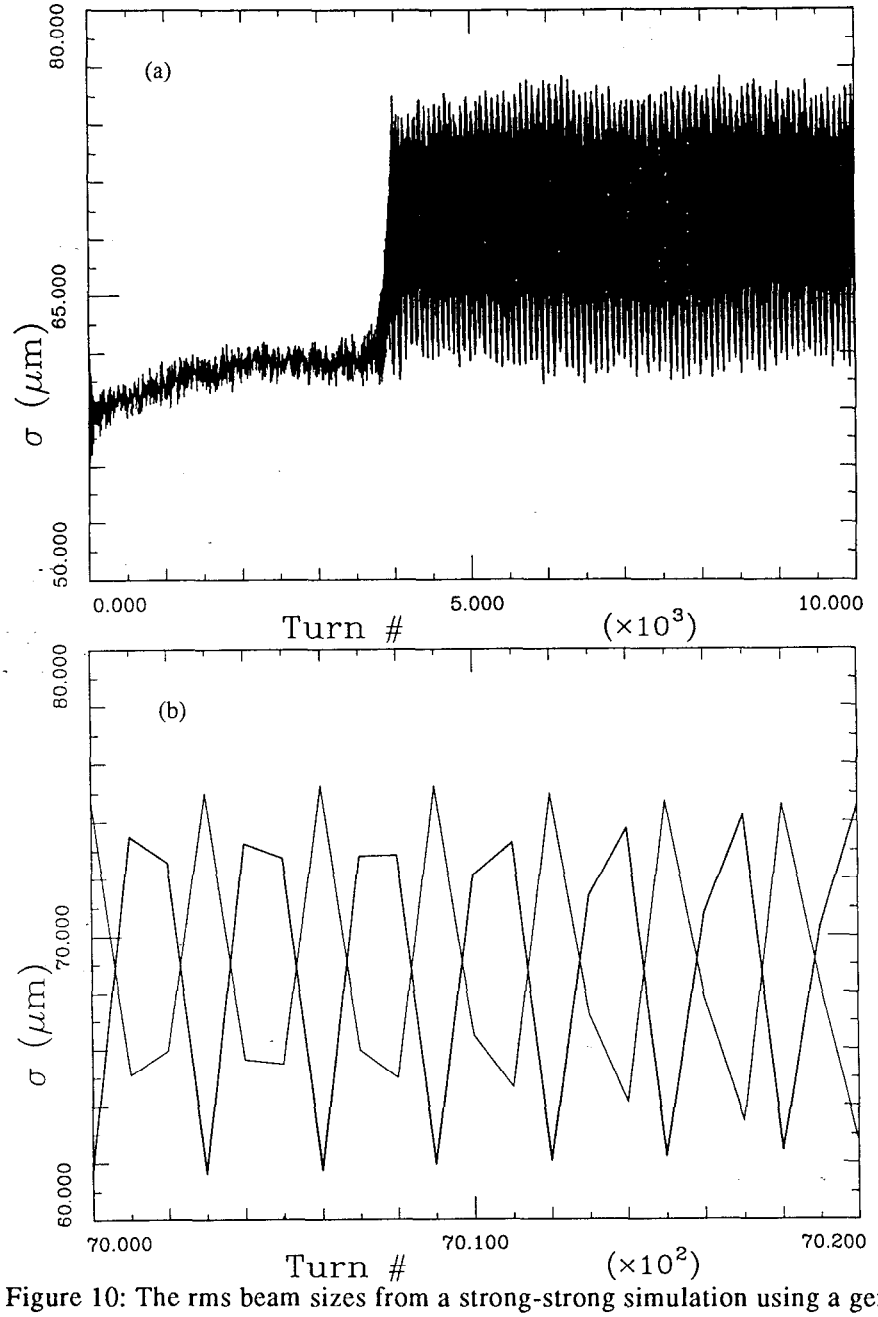

Figure 10: The rms beam sizes from a strong-strong simulation using a general expression for the beam-beam impulse. The parameters of the simulation are : $Q_{x0} = Q_{y0} = 0.79$, $\sigma_{x0} = \sigma_{y0} = 55\ \mu m$, $\xi = 0.10$, and fractional energy loss per turn = 1×10^{-3}. a) shows the onset of the instability for one of the beams, and b) shows that the size variations of the two beams. They are anticorrelated and repeat every three turns. (From ref. 34)

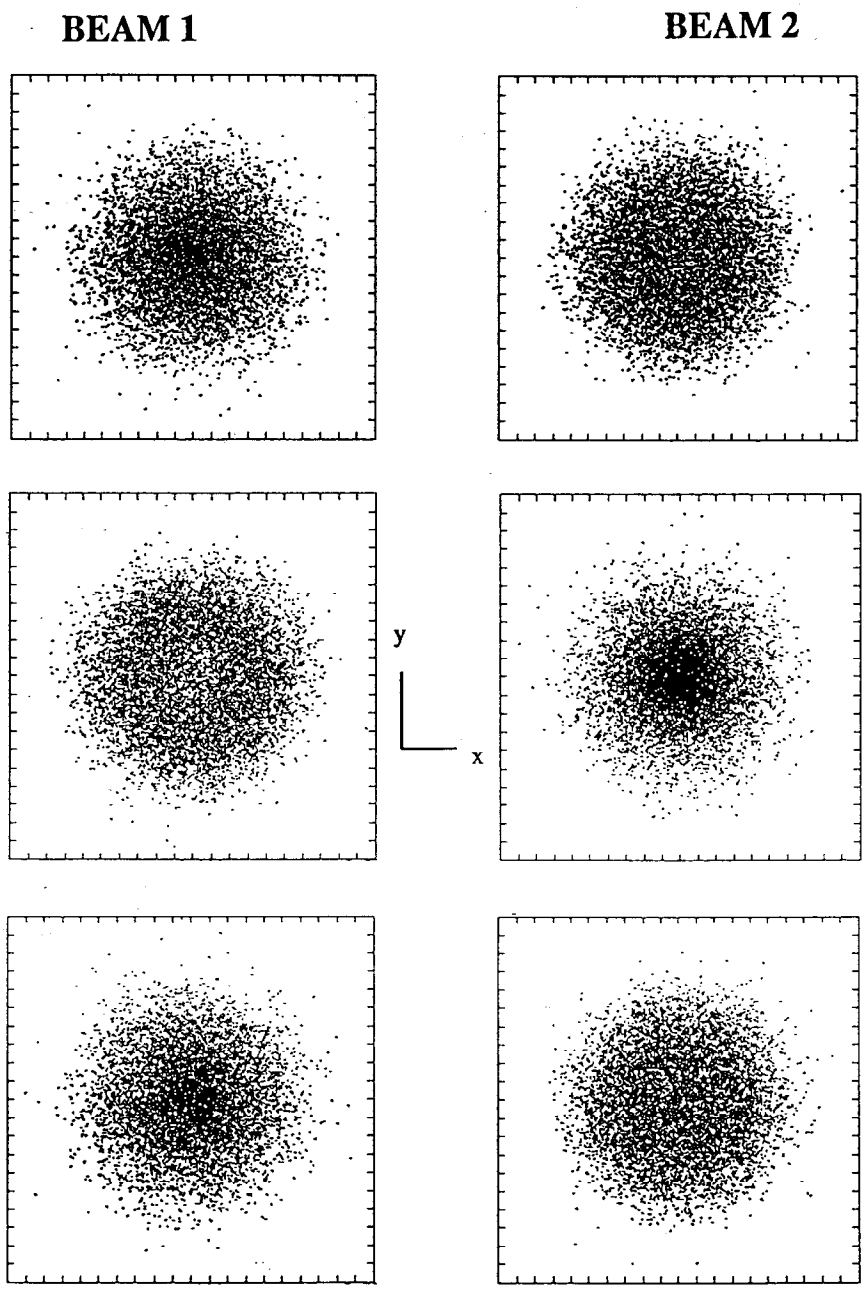

Figure 11: Scatter plots showing the beams on three successive turns for $Q_{x0} = Q_{y0} = -0.80$.

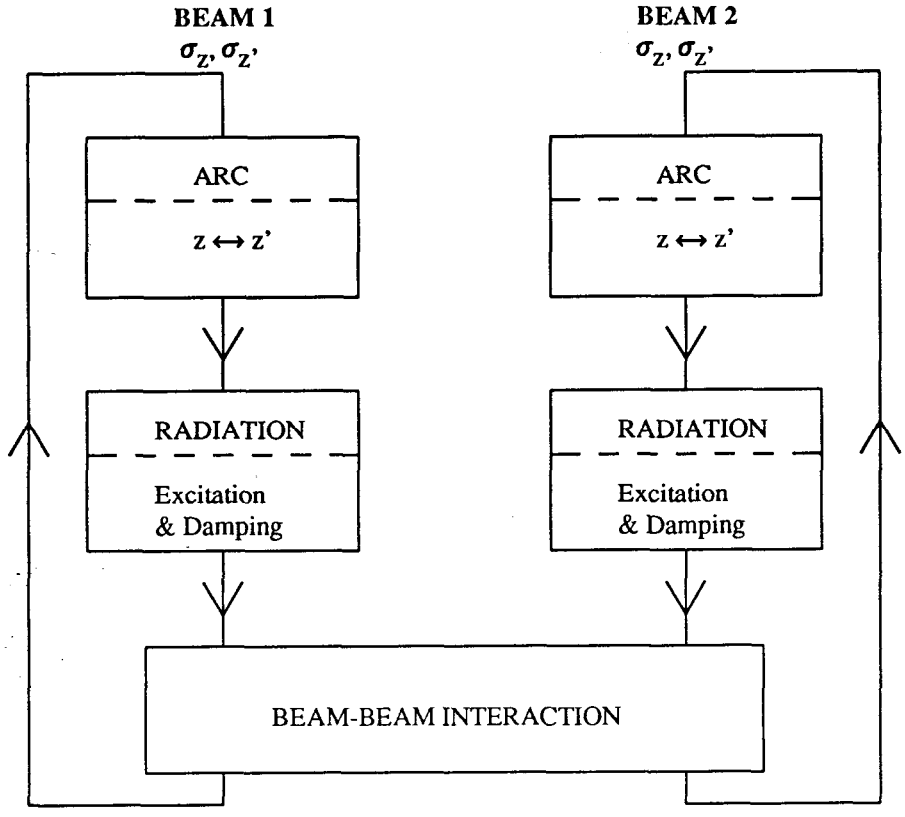

Figure 12: Schematic of an iterated map calculation of coherent beam-beam interactions.

6 The Coherent Beam-Beam Interaction

6.1 Iterated Maps of Moments

There are two different classes of coherent beam-beam interaction theories. One is based on iteration of a one-turn map for the beam moments and the other on solutions of the Vlasov equation.

The iterated map is shown schematically in Figure 12. The rms widths and angular spreads of the beams, the second moments of the distributions, are mapped for a single turn. The first elements of the map are the arcs transporting the beams between collision points. Tune dependent linear combinations of the moments at the beginning of the arc give values at the end. The second step accounts for radiation. The moments are reduced by a fractional amount for damping, and random terms are added to model quantum excitations.

The beams collide in the third step. Two different procedures have been used. One of them is to calculate the increase in the angular spread of each beam due to a Gaussian beam with the width of the other beam.[36] The angular distributions after the collision of two Gaussian beams are not Gaussian because of the nonlinearities of the beam-beam impulse, and this procedure does not account for that. This is similar to the approximation made in strong-strong simulations that use the fields of a Gaussian beam. Consideration has been given to including higher moments in the map.[37] The second procedure is to use a linear beam-beam interaction, essentially eq. (3.3).[38] The beams remain Gaussian because the interaction is linear, but the beam-beam impulse is an approximation.

The one-turn map is iterated by taking the moments after the beam-beam interaction as the inputs for the next turn. Stable solutions are found after many iterations. They include: i) equal beam sizes at low beam-beam strengths, ii) flip-flop solutions with one beam larger than the other, and iii) period-n solutions where the beam sizes follow a pattern that repeats every n turns. These results are qualitatively similar to effects seen in operations and simulations, but details such as tune shift limits, phase space distributions, and tune dependence do not agree. Whether this can be resolved with further developments remains to be determined.

6.2 Solutions of the Vlasov Equation

The Vlasov equation is usually used in accelerator physics to analyze single beam stability. It gives the evolution of phase space density, $\Phi(\vec{x}, \vec{p}, s)$,

$$\frac{d\Phi}{ds} = \frac{\partial\Phi}{\partial s} + \frac{\partial\Phi}{\partial\vec{x}} \cdot \frac{d\vec{x}}{ds} + \frac{\partial\Phi}{\partial\vec{p}} \cdot \frac{d\vec{p}}{ds} = 0 \quad . \tag{6.1}$$

The products are vector dot products. The forces acting on a particle come from external sources and from other particles. As a consequence the derivatives $d\vec{x}/ds$ and $d\vec{p}/ds$ can depend on Φ and the Vlasov equation is quadratic in Φ. It can be linearized by adding a small perturbation ϕ to the equilibrium distribution, Φ_0,

$$\Phi(\vec{x},\vec{p},s) = \Phi_0(\vec{x},\vec{p},s) + \phi(\vec{x},\vec{p},s) \tag{6.2}$$

When $\phi << \Phi_0$, terms of order ϕ^2 can be ignored leaving an equation linear in ϕ. Instabilities are unstable perturbations that are sought by analyzing the resultant equation for growing solutions. The characteristics of unstable perturbations are determined, but it isn't possible to tell whether these instabilities grow forever or are limited by some nonlinearity. That is outside the approximation used to linearize the Vlasov equation.

The Vlasov equation has been linearized and solved by two pairs of authors, Dikansky and Pestrikov[39] and Chao and Ruth,[40] for the one dimensional beam-beam interaction. There are two examples of the one dimensional beam-beam interaction: i) very flat beams, $\sigma_x >> \sigma_y$, where the vertical beam-beam force depends on y only, and ii) round beams, $\sigma_x = \sigma_y$, where only the radial coordinate matters. Dikansky and Pestrikov have analyzed both cases and found similar results. Chao and Ruth did

their calculation for flat beams. I am more familiar with their work, and, for that reason only, the next section follows it.

6.3 Coherent Instabilities in the One-Dimensional Beam-Beam Interaction[40]

The Vlasov equation when only the vertical coordinate is considered is

$$\frac{d\Phi_k}{ds} = \frac{\partial\Phi_k}{\partial s} + y'\frac{\partial\Phi_k}{\partial y} + y''\frac{\partial\Phi_k}{\partial y'} = 0 \ . \tag{6.3}$$

The index k = 1,2 indicates that there is an equation for each beam. The distributions Φ_k are normalized to unity

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Phi_k(y,y')\,dy\,dy' = 1 \ , \tag{6.4}$$

and their projections onto the y-axis are

$$\rho_k(y) = \int_{-\infty}^{\infty}\Phi_k(y,y')\,dy' \ . \tag{6.5}$$

The second derivative, y'', depends on the storage ring lattice and the beam-beam interaction which, in turn, depends on the distribution of the other beam. The beam-beam impulse for a particle in beam 1 comes from Gauss' Law. It is

$$\begin{aligned}\Delta y' &= \frac{-4\pi N r_e}{\gamma L_x}\left[\int_{-\infty}^{y}\rho_2(\zeta)\,d\zeta - \int_{y}^{\infty}\rho_2(\zeta)\,d\zeta\right] \\ &= \frac{-4\pi N r_e}{\gamma L_x}\int_{-\infty}^{\infty}\rho_2(\zeta)\,\Theta(y-\zeta)\,d\zeta\end{aligned} \tag{6.6}$$

where

$$\Theta(\zeta) = \begin{cases} 1 & \zeta > 0 \\ -1 & \zeta < 0 \end{cases} , \tag{6.7}$$

and L_x is the horizontal width of the beam. A Gaussian beam has $L_x = (8\pi)^{1/2}\sigma_x$ near x = 0.

The equilibrium distribution is the same for the two beams, $\Phi_1 = \Phi_2 = \Phi_0$, and it satisfies

$$\frac{\partial\Phi_0}{\partial s} + y'\frac{\partial\Phi_0}{\partial y} - F(s,y)\frac{\partial\Phi_0}{\partial y'} = 0 \ . \tag{6.8}$$

with

$$F(s,y) = K(s)y + \frac{4\pi N r_e}{\gamma L_x} \sum_{n=-\infty}^{\infty} \delta(s-nC) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi_0(\zeta,y')\Theta(y-\zeta)\,dy'\,d\zeta \quad . \tag{6.9}$$

The first term in F gives the focusing of the magnet lattice, and the second term comes from the beam-beam interaction. This equation is quadratic in Φ_0, and solving it is difficult. However, a solution is not required, and that simplifies the situation. There is more on this below.

Following eq. (6.2), eq. (6.3) is linearized by substituting $\Phi_k = \Phi_0 + \phi_k$. The approximate equation for ϕ_1 is

$$\frac{\partial \phi_1}{\partial s} + y'\frac{\partial \phi_1}{\partial y} - F(s,y)\frac{\partial \phi_1}{\partial y'} - \frac{\partial \Phi_0}{\partial y'} \frac{4\pi N r_e}{\gamma L_x} \sum_{n=-\infty}^{\infty} \delta(s-nC) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_2(\zeta,y')\Theta(y-\zeta)\,dy'\,d\zeta \tilde{=} 0. \tag{6.10}$$

Terms involving only the equilibrium distribution make no contribution because of eq. (6.8), and one term proportional to

$$\frac{\partial \phi_1}{\partial y'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_2(\zeta,y')\Theta(y-\zeta)\,dy'\,d\zeta \tag{6.11}$$

has been neglected. There is a similar equation for ϕ_2 that is coupled to eq. (6.10). They can be uncoupled by introducing

$$\phi_{\pm} = \phi_1 \pm \phi_2 \tag{6.12}$$

which satisfy two independent equations

$$\frac{\partial \phi_{\pm}}{\partial s} + y'\frac{\partial \phi_{\pm}}{\partial y} - F(s,y)\frac{\partial \phi_{\pm}}{\partial y'} \mp \frac{\partial \Phi_0}{\partial y'} \frac{4\pi N r_e}{\gamma L_x} \sum_{n=-\infty}^{\infty} \delta(s-nC) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\pm}(\zeta,y')\Theta(y-\zeta)\,dy'\,d\zeta \tilde{=} 0. \tag{6.13}$$

The equilibrium distribution enters in three ways. An approximation is used to simplify the calculation making it possible to characterize instabilities but making it inconsistent and of limited quantitative value - three different Φ_0's are used, one for each way it enters. First, Φ_0 must satisfy eq. (6.8) to eliminate the terms involving only Φ_0 from the linearized Vlasov equation. Finding that Φ_0 is analogous to determining the longitudinal distribution of a single beam taking account of potential well distortion. It isn't necessary to know this Φ_0 for characterizing instabilities, but it does affect thresholds.[41]

Second, Φ_0 gives the contribution of the beam-beam interaction to the focusing through F(s,y), eq. (6.9). Assuming a stable equilibrium distribution exists and is such that it produces a linear focusing force, F(s,y) becomes

$$F(s,y) = F(s)y \ . \tag{6.14}$$

Since F is linear in the displacement just like a quadrupole lattice, a β-function can be found and action-angle coordinates, I and ψ, exist. The β-function accounts for the focusing of Φ_0 as well as the magnet lattice. The action is a constant of the motion, dI/ds = 0, and Φ_0 is a function of I only, $\Phi_0(I,\psi) = \Phi_0(I)$. Equation (6.13) can be rewritten in terms of the action-angle coordinates using these facts plus i) β is a minimum at the interaction point so that $\beta' = d\beta/ds = 0$ there, ii) $d\psi/ds = 1/\beta$, iii) y'' = - F(s)y, and iv) the chain rule. It becomes

$$\frac{\partial \phi_{\pm}}{\partial s} + \frac{1}{\beta}\frac{\partial \phi_{\pm}}{\partial \psi} \pm \frac{d\Phi_0}{dI}\sqrt{2I\beta}\sin\psi \times$$

$$\frac{4\pi N r_e}{\gamma L_x}\sum_{n=-\infty}^{\infty}\delta(s-nC)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_{\pm}(\zeta,y')\Theta(y-\zeta)\,dy'\,d\zeta = 0. \tag{6.15}$$

There is a rough analogy in transverse single beam stability calculations to the simplification of eq. (6.14); it is the assumption that the deflecting fields leading to the transverse impedance are linear in displacement.

The third appearance of Φ_0 is as a weighting factor $d\Phi_0/dI$ for the beam-beam contributions of $\phi_{\pm}$. The "water-bag" model with constant phase space density out to a boundary is used

$$\Phi_0(I) = \frac{1}{\pi\varepsilon} H(\varepsilon/2 - I) \tag{6.16}$$

where ε is the vertical emittance,

$$H(\zeta) = \begin{cases} 1 & \zeta < 1 \\ 0 & \zeta > 1 \end{cases}, \tag{6.17}$$

and the normalization is

$$2\pi\int_0^{\infty}\Phi_0(I)\,dI = 1 \ . \tag{6.18}$$

With this distribution the last term in eq. (6.15) is a delta-function, and the perturbation is localized to I = ε/2.

Fourier analyzing $\phi_{\pm}$

$$\phi_{\pm} = \delta(I-\varepsilon/2)\sum_{m=-\infty}^{\infty} g_m^{\pm}(s)\,e^{im\psi} \ . \tag{6.19}$$

Substituting into eq. (6.15)

$$\frac{dg_m^{\pm}}{ds} + \frac{im}{\beta} g_m^{\pm} \mp \frac{2Nr_e}{\pi\gamma L_x}\sqrt{\frac{\beta}{\varepsilon}} \sum_{n=-\infty}^{\infty} \delta(s-nC) \sum_{k=-\infty}^{\infty} M_{mk} g_k^{\pm} = 0 \quad . \tag{6.20}$$

The matrix element M_{mk} of matrix **M** is

$$M_{mk} = \int_0^{2\pi} d\psi \, \sin\psi e^{-im\psi} \int_0^{2\pi} d\psi' e^{ik\psi'} \Theta(\cos\psi - \cos\psi') \tag{6.21}$$

$$= \begin{cases} \dfrac{-32im}{\left[(m+k)^2-1\right]\left[(m-k)^2-1\right]} & m+k \text{ even} \\ 0 & m+k \text{ odd} \end{cases} \quad .$$

Equation (6.20) gives the evolution of the perturbation. Follow it for one turn assuming only one interaction point. The Fourier components are independent between collisions, and g_m, the perturbation with phase space periodicity m, advances at -m times the betatron phase advance;

$$g_m(s) = g_m(0)\exp\left[-im\int_0^s \frac{ds}{\beta}\right] \quad . \tag{6.22}$$

The total betatron phase advance for a complete turn is $2\pi Q$. The betatron tune, Q, includes the focusing from Φ_0. Let $g_m(0)$ denote the value just after the interaction point and $g_m(C)$ the value one turn later taking into account all but the last term in eq. (6.20)

$$g_m(C) = g_m(0)\exp(-i2\pi mQ) = R_{mm} g_m(0) \quad . \tag{6.23}$$

This defines the elements of a diagonal matrix **R**. The perturbations are coupled to each other at the interaction point. The change during a collision is

$$\Delta g_m^{\pm} = \mp \frac{2Nr_e}{\pi\gamma L_x}\sqrt{\frac{\beta}{\varepsilon}} \sum_{k=-\infty}^{\infty} M_{mk} g_k^{\pm}(C) \quad . \tag{6.24}$$

The one-turn transformation of the vector $\vec{g} = (\ldots g_2, g_1, g_0, g_{-1}, g_{-2}, \ldots)^T$ is

$$\vec{g}^{+} = \left[\mathbf{I} \mp \frac{2Nr_e}{\pi\gamma L_x}\sqrt{\frac{\beta}{\varepsilon}}\,\mathbf{M}\right]\mathbf{R}\vec{g}^{+} = \mathbf{T}\vec{g}^{+} \quad . \tag{6.25}$$

There is a coherent beam-beam instability when one of the eigenvalues of **T** has an absolute value greater than unity; this occurs when $|\mathrm{Tr}(\mathbf{T})| > 2$. Different Fourier components are unstable at different tunes. Consider only g_m and g_{-m} to learn the instability condition. For these two components

$$\mathbf{T} = \begin{bmatrix} 1 \pm i\alpha_m & i\alpha_m \\ -i\alpha_m & 1 \mp i\alpha_m \end{bmatrix} \begin{bmatrix} \exp(-im2\pi Q) & 0 \\ 0 & \exp(im2\pi Q) \end{bmatrix}$$

$$= \begin{bmatrix} (1 \pm i\alpha_m)\exp(-im2\pi Q) & i\alpha_m \exp(im2\pi Q) \\ -i\alpha_m \exp(-im2\pi Q) & (1 \mp i\alpha_m)\exp(im2\pi Q) \end{bmatrix} \tag{6.26}$$

where

$$\alpha_m = \frac{32m}{4m^2-1}\left[\frac{2Nr_e}{\pi\gamma L_x}\sqrt{\frac{\beta}{\varepsilon}}\right] \quad . \tag{6.27}$$

The motion is stable if

$$\left|\frac{1}{2}\mathrm{Tr}(\mathbf{T})\right| = \left|\cos(2\pi Qm) \pm \alpha_m \sin(2\pi Qm)\right| < 1 \quad . \tag{6.28}$$

Instabilities occur near Q = n/2m. In terms of $\delta Q = Q - n/2m$ eq. (6.28) is

$$\left|\cos(2\pi m\delta Q) \pm \alpha_m \sin(2\pi m\delta Q)\right| < 1 \quad . \tag{6.29}$$

Chao and Ruth perform more detailed calculations including several values of lml, multiple interaction regions, and multiple bunches, but the important results have been obtained above. These are:

1. There are coherent beam-beam resonances for Q = n/2m corresponding to perturbations with phase space periodicity m. The resonances are even order only.
2. At resonance the betatron tune of the lattice is less than n/2m because of the focusing from the equilibrium distribution, Φ_0. Treating the beam-beam interaction as a thin quadrupole producing a tune shift ξ, the lattice and betatron tunes are related by

$$\cos(2\pi Q) = \cos(2\pi Q_0) - 2\pi\xi\sin(2\pi Q_0) \tag{6.30}$$

which for the water-bag model implies

$$\xi = \frac{4Nr_e}{\pi L_x \gamma\sqrt{\varepsilon/\beta}} \quad . \tag{6.31}$$

3. From eq. (6.29) ϕ_- is unstable for $\delta Q < 0$. Since $\phi_- = \phi_1 - \phi_2$, the beams are anticorrelated when ϕ_- is unstable. The beams are correlated when ϕ_+ is unstable which occurs when $\delta Q > 0$.
4. The full width of the resonance is

$$\Delta Q = \frac{32\xi}{\pi(4m^2-1)} \quad . \tag{6.32}$$

The resonances become narrower as m increases.

A number of approximations have been made, and Landau and radiation damping have not been included. They should determine the important resonances.

Simulations are the appropriate way to judge the validity of this calculation and to make quantitative predictions.

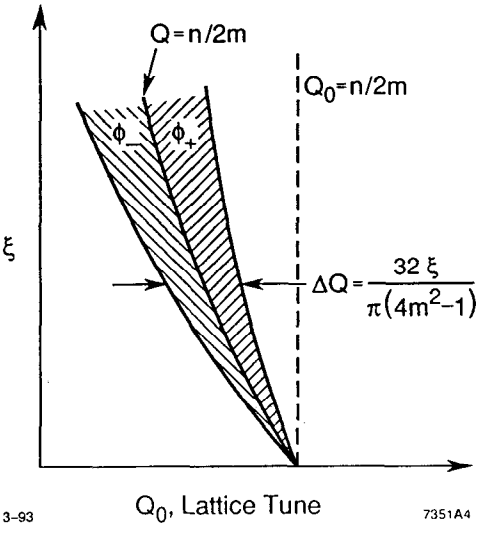

Figure 13: The results of the coherent beam-beam calculation of Chao and Ruth (ref. 40).

6.4 Coherent Beam-Beam Simulations, Revisited

The coherent resonance shown in Figures 10 and 11 agrees with all but one of the results of the Vlasov equation solution. This particular resonance is Q = 5/6, and the beams are anticorrelated as expected if ϕ_- was unstable. The resonance has a stopband with a width about a factor of two smaller than eq. (6.32) (see Figure 14). The only substantive disagreement is that an instability with the beams correlated has never been seen. It could be that ϕ_+ is unstable initially, but the limiting behavior has different characteristics.

Other resonances have been searched for. Resonances with Q = n/2 and Q = n/4 are seen even when the fields of a Gaussian beam are used, and there is a coherent resonance at Q = 7/8. There are no odd order resonances. The behavior at Q = 4/6 is close to that at Q = 5/6 suggesting that both are sixth order resonances.

The simulation shows several features that are outside the Vlasov equation calculation. Nonlinearities limit the instability, and there are damping effects. The stopband does not extend to $\xi = 0$, but the width shrinks to zero at $\xi \sim 0.06$, presumably as a result of Landau damping. The instabilities are sensitive to radiation damping. The sixth order resonance is present when the fractional energy loss per turn is as large as 10^{-3}, but it must be reduced to 10^{-5} to see the eighth order resonance. The former is large compared to the energy loss in heavy quark factories

while the latter is comparable to the energy loss in DCI where coherent phenomena are thought to have limited the improvement from space charge compensation.

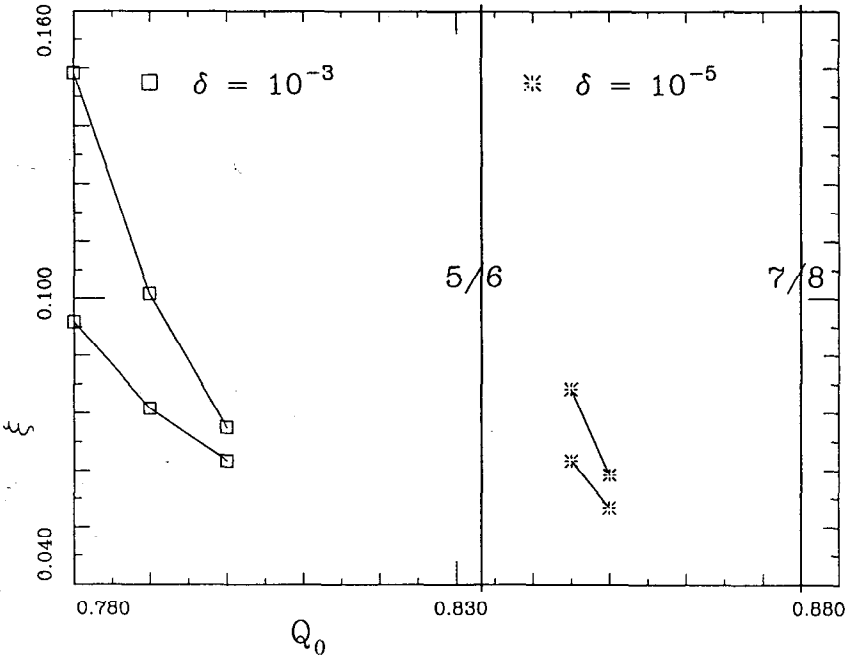

Figure 14: Onset and offset values of ξ as a function of Q_0 for sixth-order (squares) and eight-order (asterisks) coherent resonances. In each case the region of coherent motion is between the lines. δ is the fractional energy loss per turn. (From ref. 34).

6.5 Concluding Remarks

The coherent beam-beam interaction has many of the qualitative features of the tune shift limit. Although the simulation suggests that the tune shift limit associated with it is higher than that achieved in operating colliders, the simulation was done for the special case of a nearly round beam with essentially a one-dimensional beam-beam interaction. A field calculating algorithm for flat beams has proven more difficult to develop, but work is continuing there. Perhaps the tune shift limit will be lower for flat beams and a two dimensional beam-beam interaction.

The most direct evidence about the role of the coherent beam-beam interaction should come from colliders operating at the beam-beam limit, however. The signature is clear: beam distributions changing turn-by-turn with the changes of the two beams correlated. Such measurements are possible and should be performed.

7 Acknowledgements and Dedication

I have thought about the beam-beam interaction for many years and have learned what I know about it from many people. A few have been especially important to me. Either I have worked with them for extended periods or gained particular insight from their work. They are Jeff Tennyson, Gerry Jackson, Srinivas Krishnagopal, Andrei Gerasimov, Dave Rice, Joel Le Duff, Felix Izrailev, and Alex Chao. It has been a pleasure to work with and learn from them.

Jeff Tennyson passed away last year while in the prime of his career. I would like to dedicate this article to him.

8 Citations

1. John T. Seeman, Nonlinear Dynamics Aspects of Particle Accelerators, Springer-Verlag, Berlin, edited by J. M. Jowett, S. Turner and M. Month, 121 (1986).
2. John T. Seeman, Proceedings of the 12th International Conference on High-Energy Accelerators, Fermilab, Batavia, IL, edited by F. T. Cole and R. Donaldson, 212 (1983).
3. A. B. Temnykh, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 5 (1989).
4. A. Piwinski, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 12 (1989).
5. D. H. Rice, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 17 (1989).
6. P. M. Ivanov *et al*, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 26 (1989).
7. Matthew Sands, The Physics of Electron Storage Rings - An Introduction, SLAC-121 (Nov, 1970).
8. G. Decker and R. Talman, IEEE Trans Nucl Sci **NS-30**, 2188 (1983).
9. M. Bassetti and G. Erskine, CERN-ISR-TH/80-06 (1980).
10. W. Gautschi, Handbook of Mathematical Functions, Nat. Bureau of Standards, Washington, Ninth printing, edited by M. Abramowitz and I. A. Stegun, 295(1970).
11. G. P. Jackson, PhD Thesis, Cornell Univ (1988).
12. Ronald D. Ruth, AIP Conf Proc **153**, 150 (1987).
13. F. M. Izrailev and I. B. Vasserman, Proc of the 7th All Union Conference on Charged Particle Accelerators, Dubna, USSR, 288 (1980).
14. S. Krishnagopal and R. Siemann, Phys Rev **D41**, 2312 (1990).
15. S. Kheifets, PETRA Note 119, DESY (Jan, 1976).
16. I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products (Academic Press, New York, Fourth Edition, 1965) eq. 3.323.2.

17. I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products (Academic Press, New York, Fourth Edition, 1965) eq. 8.511.4.
18. P. M. Morse and H. Feshbach, Methods of Theoretical Physics (McGraw-Hill Book Co, New York, 1953) p. 483.
19. A. L. Gerasimov, F. M. Izrailev and J. L. Tennyson, AIP Conf Proc **153**, 474 (1987).
20. A. L. Gerasimov, Physics Letters A **135**, 92 (1989)
A. L. Gerasimov, Physica **41D**, 89 (1990).
21. David H. Rice, Part Accel **31**, 107 (1990).
22. Jeffrey Tennyson, Physica **5D**, 123 (1982).
23. L. R. Evans, AIP Conf Proc **127**, 243 (1985).
24. S. Myers, Nonlinear Dynamics Aspects of Particle Accelerators (Berlin: Springer-Verlag, 1986, edited by J. M. Jowett, S. Turner & M. Month), p. 176.
25. R. H. Siemann, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 110 (1989).
26. B. V. Chirikov, Physics Reports **52**, 263 (1979).
27. A. B. Temnykh, private communication.
28. A. L. Gerasimov and N. S. Dikansky, Nucl. Instr. Meth. **A292**, 209 (1990).
A. L. Gerasimov and N. S. Dikansky, Nucl. Instr. Meth. **A292**, 221 (1990).
A. L. Gerasimov and N. S. Dikansky, Nucl. Instr. Meth. **A292**, 233 (1990).
29. J. L. Tennyson, private communication.
30. R. H. Siemann, AIP Conf Proc **127**, 368 (1985).
31. K. Yokoya *et al*, KEK Preprint 89-14 (1989).
32. J. Le Duff *et al*, Proc of 11th Inter Conf on High-Energy Accel, 707 (1980).
J. Le Duff and M. P. Level, "Experiences Faisceau-Faisceau sur DCI", LAL/RT/80-03 (Orsay, 1980).
33. M. Minty *et al*, SLAC-PUB-5993 (1992).
34. S. Krishnagopal and R. Siemann, Phys Rev Lett **67**, 2461 (1991).
35. S. Krishnagopal and R. Siemann, LBL-31094, SLAC/AP-90 (1991).
36. Kohji Hirata, Phys Rev Lett **58** 25; **58**, 1798 (E) (1987).
Kohji Hirata, Phys Rev **D37**, 1307 (1988).
37. Kohji Hirata, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 46 (1989).
38. M. A. Furman, K. Y. Ng, A. W. Chao, SSC-174 (1988).
M. Furman, Third Advanced ICFA Beam Dynamics Workshop, INP, Novosibirsk, edited by I. Koop and G. Tumaikin, 52 (1989).
39. N. S. Dikansky and D. V. Pestrikov, Part Accel **12**, 27 (1982).
40. A. Chao and R. Ruth, Part Accel **16**, 201 (1985).
41. Katsunobu Oide and Kaoru Yokoya, KEK Preprint 90-10 (1990).