Fermilab-Conf-83/17-THY
January, 1983

THE EFFECTS OF BEAM-BEAM COLLISIONS ON STORAGE-RING PERFORMANCE--A PEDAGOGICAL REVIEW

Jonathan F. Schonfeld
Fermi National Accelerator Laboratory
P.O. Box 500
Batavia, Illinois 60510

Invited lectures presented at the 1982 Summer School on High-Energy Particle Accelerators, SLAC, August 2-13, 1982.

THE EFFECTS OF BEAM-BEAM COLLISIONS ON STORAGE-RING PERFORMANCE--A PEDAGOGICAL REVIEW

Jonathan F. Schonfeld
Fermi National Accelerator Laboratory, Batavia, Illinois 60510

TABLE OF CONTENTS

THE EFFECTS OF BEAM-BEAM COLLISIONS ON STORAGE-RING PERFORMANCE--A PEDAGOGICAL REVIEW

Jonathan F. Schonfeld
Fermi National Accelerator Laboratory, Batavia, Illinois 60510

I. GENERAL INTRODUCTION

The effects of beam-beam collisions have received publicity in the last few years because of the disappointing performance of the latest generation of e^+e^- colliding-beam storage rings. PEP, PETRA and CESR were each designed for a luminosity of $10^{32}cm^{-2}sec^{-1}$. PETRA operates at about a factor of five below this;[1] CESR operates at about a factor of ten below.[1] The luminosity of PEP was for a long time comparable to that of CESR, but it has recently been increased[108] to about 3×10^{31} $cm^{-2}sec^{-1}$. Some of the measures (specifically, changes in betatron tunes) taken to produce this increase were suggested by the results of a computer simulation;[106] nevertheless, the reasons for this improvement in performance are still not understood intuitively.

Luminosity is defined as the number of interactions per second per interaction region, per unit of interaction cross section. Thus it is a direct indicator of the event rate that a single elementary-particle experiment can observe. The lower the luminosity, the longer it takes an experiment to accumulate enough data for a good result.

To get a feel for what these luminosity shortfalls mean in practical terms, consider the following illustrative figures, taken from a recent report:[2] At present, the integral of CESR luminosity over one year (times two for the number of instrumented interaction regions) is about 100 pb^{-1} ($\sim 10^{38}cm^{-2}$). A few times 100 pb^{-1} is needed to accomplish a basic program of measuring the B meson mass and width, and the rates of the dominant transitions between the upsilon and its known excited states. 1000 pb^{-1} or so would be needed to obtain information on processes like $B\text{-}\bar{B}$ mixing, η_b production and decay, and $\tau\bar{\tau}$ production in the upsilon region. 10,000 pb^{-1} or so would be needed if one sought information concerning rare but important processes such as CP violation in the $B\bar{B}$ system.

It is by now widely recognized that design estimates and real performance can differ so because design estimates have been little more than very optimistic guesses.[3] There is at present no reliable deductive or semiphenomenological theory of the way in which beam-beam collisions degrade luminosity in storage rings. This is an urgent open problem in accelerator physics.

The staff of CESR is presently contemplating modifications of, or variations on, that facility that could increase luminosity by factors of ten or even of one hundred.[2] It must be stressed that this will not necessarily improve agreement between design theory and machine performance.

This paper presents a survey of the experimental and theoretical literature on colliding-beam effects in both leptonic and hadronic

storage rings. For the most part, this literature is rather technical and, to the novice, both obscurely written and hard to locate. Although there have already been several symposia on the subject,[I,II] as well as a number of reviews for specialists[3,35,101-104,112] there has up till now been no unified and pedagogical exposition. The present work represents an attempt to fill this gap. Needless to say, there are undoubtedly sources that have been overlooked or misrepresented here; to their authors I apologize in advance.

The reader of this review is assumed to be familiar with the basic facts of linear strong-focusing single-beam storage-ring theory;[4] but no prior familiarity with any other details of colliding-beam physics is presupposed.

Our material is grouped into four major areas: observational phenomenology, computer simulation, mathematical background, and theoretical models.

Almost all the material in this review should be familiar to experts. There is, however, one exception: The second theoretical calculation discussed in detail in Chapter V--a critique of a mechanism proposed by Tennyson[5] to explain the beam-beam effect at SPEAR--is original.

II. EXPERIMENTAL PHENOMENOLOGY

1. Introduction

For orientation, here is a list of the major colliding-beam storage-ring types, together with specific examples (past, present, and approved or proposed), as well as a few words of general description:

1. Electron-positron single ring--ACO (France), ADONE (Italy), DORIS (W. Germany), the VEPP series (USSR), CEA (US, terminated), SPEAR (US), CESR (US), PEP (US), PETRA (W. Germany), LEP (CERN, under construction). Counter-rotating bunched e^+e^- beams. Head-on collisions.

2. Electron-positron, electron-electron, or positron- positron intersecting rings--Stanford-Princeton project (US, terminated, e^-e^-), VEP-1 (USSR, terminated, e^-e^-), DORIS, DCI (France). Two rings, intersecting tangentially (S-P, VEP-1, and DCI), or in several locations at a nonzero angle. Bunched beams.

3. Proton-proton intersecting rings--ISR (CERN), ISABELLE (US; status uncertain). Two laced rings, intersecting in several locations at a nonzero angle. Unbunched ("coasting") beams.

4. Proton-antiproton single ring--SPS collider (CERN), Fermilab Tevatron I (US, under construction); UNK (USSR, proposed). Counter-rotating bunched beams. Head-on collisions.

For more detailed specifications of these facilities, the reader can consult Refs. 6 and 7.

These four types all differ in fundamental ways, and these differences are reflected both in observed machine behavior, and in the mathematical ideas that have traditionally been applied in

theoretical studies. These differences will be pointed out as we go along.

Here is a schematic list of the colliding-beam-related problems that are typically encountered in storage ring operation:

Particle loss at injection
Particle loss (fast or slow) after injection
Increase in range of tunes for which beam cannot stably be stored
Expansion of beam size
Impairment of luminosity
Low maximum storable current
Extreme sensitivity of beam characteristics to machine parameters.

The last four of the these problems are most serious at electron-positron facilities. There have been attempts to induce comparable effects in hadronic storage rings (in order to probe the limits of such machines), but only under extreme and somewhat artificial conditions.

In the next two sections we shall describe the beam-beam problems encountered in normal operation of leptonic and of hadronic storage rings, respectively. In Sec. 4 we describe colliding-beam behavior observed in exploratory ISR studies done several years ago in connection with the design of the SPS proton-antiproton collider.

2. Colliding-beam effects in e^+e^- storage rings

The issue of greatest concern--by far--to users of electron-positron storage rings is the discrepancy between the luminosity expected and the luminosity actually achieved. The issue is actually composed of two subissues, because there are really two major design expectations: Luminosity does not grow as rapidly with beam current as expected; and the maximum storable current is lower than expected.

In this section, we explain the assumptions underlying the design expectations, and survey experimental tests of these assumptions.

I covered much of this material in a lecture presented at last year's Summer School.[3] For the discussion here, that presentation has been reorganized somewhat, rephrased, and updated. Some of the information in Ref. 3 (most notably in the exercises) not carried over to the present discussion may be of interest to the reader who wishes to go more deeply into this subject.

As will be explained later, proton-proton/antiproton colliders do not frustrate these assumptions to the same extent. Thus, when hadronic storage rings are discussed in the next section, the emphasis will be on different phenomena.

Operators of the largest electron-positron storage rings report[8] great difficulty in adjusting these machines for optimum performance, primarily because the behavior of colliding e^+e^- beams is extremely

sensitive to small changes in machine parameters, and the effects of such changes are often not reproducible. Moreover, operators are constrained to perform this optimization with a very small number of adjustments, for the following reasons: It takes a long time (as long as twenty minutes) to measure the luminosity that is to be optimized, because of the smallness of electron/positron interaction cross-sections at high energy; while the beams only last for a few hours because of the degrading effects of synchrotron radiation and gas scattering.

Although extreme sensitivity of luminosity to machine parameters is an object of some attention in the literature on computer simulations (see Chapter III), and has been noted in published reports on ISR exploratory studies (see this chapter, Sec. 4), very little, if anything, has been written about it in the context of routine storage ring operation. Accordingly, this topic will not be discussed further in this section.

We shall not be discussing the effects of beam-beam collisions on electron and positron spin polarization, although recent studies at PETRA and SPEAR indicate that polarization is very difficult to maintain in colliding-beam storage rings. The reader is referred to Ref. 9 for details.

a. Luminosity vs. current

Luminosity in e^+e^- storage rings, where the collisions are head-on, is quite generally defined by

$$L = fB \left(N_1 \frac{A}{A_1}\right) \left(N_2 \frac{A}{A_2}\right) \frac{1}{A} = (fBN_1N_2) \frac{A}{A_1A_2}, \qquad (1)$$

where f is the frequency at which a beam particle circumnavigates the ring; B is the number of bunches per beam; N_1 and N_2 are the numbers of particles per bunch in the two beams; A_1 and A_2 are the cross-sectional (transverse) areas of the two beams at an interaction point; and A is the transverse area in which the two beams overlap. By definition, A is necessarily less than or equal to the smaller of A_1 and A_2. These are effective areas, defined by weighted averages when the beam distributions are not uniform. The meaning of the formula should be clear: $N_1(A/A_1)$ is the fraction of particles in a bunch of beam one that actually collide with beam two; similarly for $N_2(A/A_2)$. So $[N_1(A/A_1)]$ $[N_2(A/A_2)]$ is the number of particle collisions per bunch collision, per unit probability of interaction. Bf is the number of bunch-bunch collisions per unit time at a single interaction point. 1/A is the probability of a single particle-particle collision, per unit cross-section. Typically $N_1=N_2(\equiv N)$.

The conventional design estimate of luminosity is based on the assumptions that $A_1=A_2=A$, and that A is equal to A_0, the current-independent natural electron beam area at an interaction point. A_0 is calculated in a standard way[4] from the parameters that

characterize synchrotron radiation and linear focusing; perturbations such as an opposing beam, interactions between particles in the same beam, beam-wall interactions, etc., are neglected. (In proton or antiproton storage rings, where synchrotron radiation is very weak, beam shapes and sizes are determined by the injection system.)

The mathematical formula expressing the implication of the first of these nonperturbation assumptions for luminosity is

$$L = I^2 \cdot (e_0^2 B f A)^{-1}, \tag{2}$$

where N has been rewritten in terms of the more directly measurable I, the total electric current per beam. (e_0 is the magnitude of the electron's electric charge.) Together with the second nonperturbation assumption ($A=A_0$), Eq. (2) implies that luminosity is expected to be directly proportional to current squared, with a proportionality constant determined entirely by simple properties of single-particle motion in the storage ring.

Empirically, expression (2) with $A=A_0$ is observed to describe the current dependence of luminosity only for low currents. At any energy, at every e^+e^- storage ring, there is a turnover current beyond which L grows more slowly than I^2. Exactly how much more slowly is not the same for all machines under all conditions, as we shall see.

Correspondingly, the nonperturbation assumption is empirically observed to describe the areas A_1, A_2 and A only for low currents. For large currents, either A_1 or A_2 (or both) exceeds A_0. This results in the general expression (1) for luminosity falling short of formula (2) because

$$\frac{A}{A_1 A_2} \leq \frac{\min(A_1,A_2)}{A_1 A_2} = \frac{1}{\max(A_1,A_2)} < \frac{1}{A_0} .$$

This phenomenon is commonly referred to as "beam blowup." Typically, blowup is substantial only in the vertical direction. The natural horizontal width of an electron or positron beam is almost never significantly changed by the presence of an opposing beam.

These trends will now be illustrated with data taken at various storage rings.

Let us begin with measurements of luminosity vs. current. Figure 1 represents data taken at SPEAR,[10] one bunch per beam. At the largest current (about 8mA), luminosity falls short of the quadratic curve, extrapolated from small current, by a factor of about three. Figure 2 shows data taken recently under various operating conditions at CESR,[11] at about 5 GeV, and with one bunch per beam. The luminosities at the highest currents shown are between 1.5 and 3 times less than would have been expected from quadratic extrapolation from small current. Note that these luminosity vs. current curves turn over from quadratic to linear at large I. Consider, by contrast, Fig. 3, which shows data from PEP,[12] one curve

corresponding to one bunch per beam, the other corresponding to three bunches per beam.* The three-bunch curve increases like $I^{1.6}$, even at small current. Figure 4 shows data taken recently at PETRA,[13] with two bunches per beam. The y-axis corresponds to specific luminosity, defined as luminosity per bunch divided by the square of current per bunch, i.e., $(L/2)/(I/2)^2 = 2L/I^2$. The x-axis corresponds to current per bunch, I/2. If the nonperturbation assumption were correct, the data points for the three energies shown would lie on three horizontal lines. Instead, each set of measured specific luminosities declines with current. The largest decline--from lowest to highest current--is 30% at 7 GeV, followed by slightly less than 30% at 11 GeV, and about 8% at 17 GeV.

This last figure exemplifies a general trend: Colliding beams perturb one another less (at a given machine) as energy increases. We will see other examples of this phenomenon elsewhere in this review. One cannot account for this weakening trend solely in terms of the well-known tendency of electromagnetic scattering to weaken with increasing energy.[14] Colliding beam effects in storage rings are cumulative results of many beam-beam encounters; it is necessary to understand why the weakening of individual encounters is not compensated somehow by prolonged repetition. This is often explained by naively identifying the synchrotron damping time with some sort of dynamical memory, and observing that the damping time decreases as energy increases[4] ($\propto 1/E^3$). However, it is not really obvious that this is the right way to look at things.

The CESR and PETRA data presented here were taken after "mini-beta insertions" had been installed. Mini-beta insertions are magnet configurations that lower the scale of vertical beta functions at interactions points from tens of centimeters (the standard when these machines were designed) to centimeters. In the case of CESR,[11] the reduction was from 11 cm to 4 cm; at PETRA,[13] it was from 20 cm to about 9 cm. These reductions increase luminosity because they increase focusing at interaction points, and therefore they decrease all the beam areas A_1, A_2 and A in formula (1), roughly in identical proportions. I mention this here in order to stress that these enhancements in the overall scale of luminosity (a factor of between two and three in each case[11,13]) have not significantly shrunken the range of currents within which the L vs. I curves are noticeably non-quadratic. The nonperturbation assumption appears to fail just as badly with mini-beta insertions as without them. (For a discussion of ultimate limits on the reduceability of interaction-point betas, the reader is referred to Chapter III, Subsection 2b.)

We turn now to direct measurements of beam-blowup. The figures discussed below were all obtained (directly or indirectly) from photographs of synchrotron radiation. Such a photograph is made by

*This data was taken before the modifications that have recently increased PEP's peak luminosity by a factor of about three. At present, PEP luminosity at peak current is about 60% of the value that would follow from the nonperturbation assumptions.[109]

pointing a television camera, or some other type of light monitor, along a tangent to a storage ring. Because synchrotron light is (to a good approximation) emitted tangent to the path of a circulating particle, the output from such a monitor gives a representation of the transverse profile of a beam, light intensity being proportional to beam density.

(It should be noted that the blowups shown in the following two figures do not necessarily coincide exactly with what one would see if one observed the beams at an interaction region, where luminosity is measured. Beam profiles cannot in general be photographed at interaction points because little synchrotron radiation is emitted there--synchrotron radiation requires curved or accelerated motion, whereas the presence of collision detectors precludes the placement of bending magnets or RF cavities at interaction points.)

Figure 5 shows such television images of electron and positron beams under various conditions at SPEAR[16] (energy and current not specified). "Flip flop" refers to an effect that has so far been well documented only at SPEAR:[17] When the beams are in collision, the extent to which they blow up can be controlled by adjusting the relative phase between electromagnetic oscillations in two RF cavities: Move the phase one way and the positron bunch shrinks while the electron bunch expands; move it the other way and the opposite happens. The caption "flip-flop balanced" indicates that the RF cavities are phased so that the two beams are blown up equally. (As it turns out, SPEAR luminosity is always highest when the flip-flop is balanced. The data in Fig. 1, and in Fig. 7, below, were taken under balanced conditions.[10]) "With flip-flop effect" means that the RF cavities are not phased for balance. In either case one sees clearly that at least one beam blows up when two beams collide.

Figures 6a and 6b show intensity profiles of synchrotron light observed at ADONE[18] at two currents, one below the threshold for blowup and the other above. The energy is near 1 GeV. The eight peaks are to be read as follows: Let, for example, $\rho(x,y)$ be the distribution function of electrons at the low current, in the plane perpendicular to the axis of a camera looking at the electron beam. (x and y refer to horizontal and vertical coordinates.) Then the lower left-hand peak in Fig. 6a is proportional to $\int\rho(x,y)dy$, while the lower right-hand part is proportional to $\int\rho(x,y)dx$. The other six peaks are interpreted similarly. The blowup of the electron beam at the higher current is evident in the widening of the lower right-hand peak in Fig. 6b relative to the upper right-hand peak in Fig. 6b, and to the two right-hand peaks in Fig. 6a.

Note, as mentioned earlier, that both in Figs. 5 and 6, beam blowup is apparent only in the vertical dimension.

Although convenient because human eyes can see it easily, beam size of course only crudely indicates how particle distributions are altered when two beams are brought into collision. One can begin to form a more complete picture with the help of some measurements made at SPEAR:[19] Recall that in a linearly focusing e^+ or e^- single beam storage ring, bunch distributions are typically nearly Gaussian.[4] SPEAR results[19] indicate that at high enough current, bunch

distributions in an e^+e^- collider have tails that are much longer than would be the case if the distributions were Gaussian. It will be useful to bear this in mind when we discuss computer simulations in Part III.

b. Maximum current; tuneshift

The maximum luminosity that can be obtained at a given energy is determined in part by the luminosity vs. current curve, and in part by the maximum current that can be stored. Maximum current per beam when there are two beams colliding in a storage ring is typically much less than maximum current at the same energy when only a single beam is circulating. For example, PETRA, at 7 GeV per circulating particle, can hold up to 20 mA when operated as a single-bunch single-beam storage ring, but only up to about 5 mA per bunch when operated as a collider with two bunches per beam;[20] CESR at 5 GeV and one bunch per beam can hold up to 30 mA when operated with a single beam, but only up to 10-15 mA per beam when operated as a collider.[11]

The working definition of maximum current varies from storage ring to storage ring. At SPEAR,[21] PEP[21] and PETRA,[20] maximum current is the point beyond which background in experimental detectors exceeds some pre-specified maximum tolerable level. The apparent arbitrariness in this criterion turns out to be of minor importance, since the rise in background as current approaches maximum is typically very sharp.[21] Other laboratories have applied criteria based on beam lifetime.[18,22,23]

The conventional procedure by which designers have traditionally predicted the maximum colliding current that can be accomodated in e^+e^- storage rings is based on two general hypotheses: (i) The destabilizing strength of the force that one beam exerts on the other beam is completely specified by the value of a single dimensionless parameter, the linear vertical beam-beam tuneshift per interaction (or simply "tuneshift," when no confusion is likely) felt by the second beam. The tuneshift will be defined shortly. The current in a beam that is assumed to have a Gaussian charge distribution is easily expressed in terms of the tuneshift it exerts, the beam energy, and the vertical beta function and the beam's dimensions at an interaction point, through a formula that we shall present later. (One can also define a horizontal tuneshift, although vertical tuneshift has traditionally been regarded as a much more important index of beam stability because, as we have seen, beams in e^+e^- colliders typically blow up vertically much more than they do horizontally. I am not aware of a more rigorous way of justifying this point of view.) (ii) A beam reaches maximum current when the tuneshift exerted on it by the other beam attains some maximum value that can be determined easily by a simple rule of thumb.

To calculate maximum storable current in the conventional way, one substitutes the maximum tuneshift dictated by the rule of thumb, together with the beam energy, the expected vertical beta function at an interaction point, and the unperturbed (see footnote next page) interaction-point beam dimensions, into the formula just mentioned.

The problematical parts of this procedure are the use of the unperturbed beam dimensions (as we have already seen) and the simple rule of thumb used to predict maximum tuneshift. At least three different versions of the rule of thumb have at various times been used seriously by storage ring designers to forecast maximum tuneshift. We will describe these variants, and compare them with the performance of the machines to which they were originally applied, after we have defined the tuneshift below. Sadly, the rule of thumb--in any of its versions--has had little real success.

In the future, if no convincing and flexible analytical theory of beam-beam effects emerges, it is likely that most important design predictions will be made on the basis of computer simulation.

We now turn to the definition of horizontal and vertical beam-beam tuneshifts. In what follows, we shall restrict our attention to storage rings in steady state. Thus we may assume that the explicit time-dependence of the non-noisy part of the force felt by any beam particle is periodic, the period being the time it takes a beam particle to circle the ring once. It is possible to formulate definitions under more general circumstances, but this will not be necessary for our purposes.

Horizontal and vertical tuneshifts felt by a beam in a storage ring are given by the following recipe: Isolate the non-noisy and non-dissipative forces--due to focusing magnets and to the opposing beam-- acting on a test particle in the beam in question. Represent these forces as series' in powers of betatron and synchrotron coordinates and velocities, and then discard all but the linear terms. Evaluate the horizontal and vertical (we are neglecting x-y coupling) Courant-Snyder tunes characterizing oscillatory motion in the presence of such linear, time-periodic forces. Evaluate these tunes again, this time omitting the forces due to the opposing beam. Form two differences: between the two horizontal tunes just evaluated, and between the two vertical tunes just evaluated. The horizontal and vertical beam-beam tuneshifts per interaction felt by the beam in question are given by these two differences, divided by C, where C is the number of collision regions. (In the case of single-ring bunched-beam colliders, C is equal to twice the number (B) of bunches per beam.)

One divides by C for this reason: Storage rings are often designed so that--ignoring inaccuracies in construction and control--the C regions between the C interaction points are identical. Thus a storage ring's natural periodicity is often naively expected to be C^{-1} times its circumference. Instabilities in dynamical systems typically arise when a rational number of internal oscillations take place during a repeat period of the system's

*Throughout this paper, we use the word "unperturbed" to label a quantity that is calculated ignoring interactions between particles in different beams, between particles in the same beam, and between real particles and image charges. For example: A_0, defined earlier, is the unperturbed beam area. (This footnote refers to preceding page.)

time-dependent parameters. In the case of storage rings, we see that the number of horizontal or vertical betatron oscillations per machine period is the horizontal or vertical tune divided by C. This is why it is considered desirable to include division by C in the definition of the tuneshifts.

(To be complete, we must point out that there is a growing number of indications, primarily from computer simulation, but also from real experience with proton-proton and proton-antiproton colliders, that small, irregular departures from ideal storage-ring symmetries (for example, the C^{-1}-periodicity mentioned above) play a major role in the observed behavior of colliding beams. This will be discussed more fully later in this chapter and in Chapter III.)

Here is how the tuneshift rule of thumb has evolved during the history of storage ring design:

Version no. 1: The earliest expectation[24] was that the tuneshift could get as large as the difference between C^{-1} times the vertical tune evaluated neglecting the opposing beam (i.e., C^{-1} times the unperturbed vertical tune), and the nearest integer or half-integer (assuming C^{-1}-periodicity). In other words, beams were expected to be stable as long as the linear approximation to the colliding system did not become linearly resonant.

Version no. 2: The earliest e^+e^- storage rings were all characterized by limiting tuneshifts near .025, much smaller than the limits expected on the basis of Rule no. 1.[25] It might have been natural at that point to try a tuneshift rule based on nonlinear resonance, but to my knowledge this was not done (although it was proposed[24]). In the absences of theoretical guidance, designers adopted the assumption that--for reasons unknown--all e^+e^- storage rings must be characterized by the same maximum tuneshift, which, from earlier experience, was conjectured to be about .025. This was the figure used in the SPEAR design.[26]

Version no. 3: As it turned out, at high energy the tuneshift at SPEAR can approach .06. (The maximum tuneshift at ACO was also .06.) Thus, in the mid-70's one expected that a storage ring of much higher energy could routinely reach .06 in the middle of its range. .06 was thus the maximum tuneshift assumed in the CESR,[27] PEP[28] and PETRA[20] designs.

Subsequent experience has shown this estimate to be much too favorable. The highest tuneshift observed so far at PETRA[13] is close to .04; at CESR[11] the highest tuneshift seen so far is near .02. The highest tuneshift observed at PEP[12] before recent changes[108] in magnet placement and beta functions was near .02; since the changes, PEP has achieved tuneshifts as high as about .05. The original LEP design[29] followed Rule 3 in assuming a maximum achievable tuneshift of .06, but on the basis of CESR, PEP and PETRA performance, this has since[30] been reduced to .03. This guess is likely to be replaced before long by expectations derived from computer simulations like the one recently developed for the LEP project by S. Myers.[31] (Myers' model and its consequences will be discussed in Chapter III.)

We see, in sum, that tuneshifts have turned out to be much more difficult to predict than was originally conjectured. Nevertheless,

perhaps for lack of anything better, the tuneshift has become one of the persistent themes in the literature on storage rings. With this in mind, let us proceed to develop the phenomenology of the beam-beam tuneshift in greater detail.

In what follows we make a number of standard simplifying assumptions: The distribution of particles in a beam at an interaction point, in the transverse plane, is assumed to be Gaussian, centered at the origin of the betatron oscillations of the opposing beam; the disturbance that a test particle suffers when it passes a bunch from the opposing beam is assumed to be due entirely to the Coulombic interaction; the whole of this disturbance is assumed to be felt instantaneously.

Let us comment briefly on these assumptions:

-- We have already mentioned empirical indications that particle distributions in colliders are not Gaussian far from the beam centers. However, this does not invalidate the above Gaussian assumption a priori, since the linear approximations required by the recipe defining the tuneshift are, strictly speaking, meaningful only close to beam centers.

-- In reality, beam distributions are only approximately centered, not perfectly centered, as is assumed here. For a discussion of some effects specifically associated with off-center collisions, see the next section, as well as Section 2 of Chapter V.

-- The assumption that the beam-beam interaction is entirely Coulombic means that one neglects inelastic (i.e., radiative) electromagnetic scattering, as well as hard processes (e.g., nuclear scattering in the case of proton/antiproton colliders). Inelastic electromagnetic scattering cannot be ignored in colliders of extremely high energy,[32] but in existing storage rings it can be safely neglected (for an estimate, see Ref. 3, Problem 5). Hard processes can be ignored because they are generally of much shorter range than the interparticle spacing in a beam. Thus, a beam particle participates in hard processes only occasionally, while it responds to the combined Coulombic force due to all the particles in every passing bunch in the opposing beam.

-- Strictly speaking, the time elapsed during an encounter between a test particle and a bunch from the opposing beam is (in an ultrarelativistic storage ring) half the bunch length (see Fig. 17) divided by the speed of light. This is always very much shorter than the time between encounters. It is also usually much shorter than the time scale on which a beam particle sees variations in the storage ring structure (magnetic fields, etc.) near the interaction region; although these time scales can be comparable if either the vertical or horizontal beta functions at an interaction point becomes less than the bunch length. The beta functions are not that small in existing storage rings. (This point will be explored further in Chapter III. For a discussion of the manner in which rapid variation in beta functions can alter tuneshifts, see Ref. 111.)

Under these assumptions, the tuneshift due to a beam in a single-ring bunched-beam collider satisfies the following equation:

$$\cos 2\pi\,(\mu_y+\Delta\mu_y) = \cos 2\pi\mu_y - 2\pi\xi_y \sin 2\pi\mu_y. \tag{3}$$

(For a derivation, see Ref. 3.) In this equation, $\Delta\mu_y$ is the tuneshift, and μ_y is C^{-1} times the unperturbed vertical storage ring tune. The parameter ξ_y is defined, for oppositely charged beams colliding head on (i.e., e^+e^- or $\bar{p}p$), by

$$\xi_y \equiv \frac{e_0}{2\pi Bf} \quad \frac{\beta_y^* I}{E\sigma_y^*(\sigma_x^*+\sigma_y^*)}\,, \tag{4}$$

where β_y^* is the vertical beta function at an interaction point; σ_x^* and σ_y^* are the horizontal and vertical half-widths of the beam distributions at an interaction point; E is the energy of a beam particle; and all other symbols have been defined before.

When collisions are not head on, Eq. (4) must be modified. For example, the formula to be used in the case of the ISR, or of ISABELLE, is[7]

$$\xi_y = -\frac{e_0}{\sqrt{2\pi c}} \quad \frac{\beta_y^* I}{E\sigma_y^* \tan(\alpha/2)}\,, \tag{4a}$$

where α is the angle at which the beams cross. (Strictly speaking--since such machines involve two rings each--E and β_y^* refer to one ring of a collider, while I and σ_y^* refer to the other.) Note the signs of (4) and (4a): Tuneshifts are positive when the colliding beams have charges of opposite sign; they are negative when the colliding beams have charges of the same sign.

ξ_y is frequently referred to as the "vertical space-charge parameter." For typical storage ring currents, ξ_y is quite small. In particular, for electron-positron storage rings operated under optimal conditions, ξ_y is generally a few times 10^{-2}. At the ISR and the SPS, ξ_y is typically much smaller--between a few times 10^{-4} and a few times 10^{-3}; it is commonly supposed that this is why colliding beam effects are much less pronounced in the ISR and the SPS, than in e^+e^- storage rings.

Thus, for most purposes, it is permissible to reduce Eq. (3) to

$$\Delta\mu_y \simeq \xi_y, \tag{5}$$

as long as μ_y is not too close to an integer or a half-integer. (In ACO and in ADONE, with μ_y close to an integer, $\Delta\mu_y$ and ξ_y could differ considerably:[11,12] In ACO, the largest value of $\Delta\mu_y$ was about .06, while the largest value of ξ_y was about .04; in ADONE, the largest values of $\Delta\mu_y$ and ξ_y were .03 and .08, respectively.) A similar formula applies to $\Delta\mu_x$.

(Note that for $\xi_y > 0$, Eq. (3) cannot be solved for real $\Delta\mu_y$ when μ_y lies within any interval of the form $([n/2 - (1/2\pi)\,\mathrm{Arccot}\,(2\pi\xi_y)],\ n/2)$, where n is an arbitrary integer. [For $\xi_y < 0$, as is the case in proton-proton colliders, the forbidden intervals have the form $(n/2,\ [n/2 - (1/2\pi)\,\mathrm{Arccot}\,(2\pi\xi_y)])$.] For such μ_y, the storage ring with colliding beams, in the linear idealization, is forced to have a tune with a nonzero imaginary part. This means that small-amplitude oscillations of a stored particle about its bunch center, in the presence of the opposing beam, run away exponentially in time. From this, one concludes that e^+e^- or $\bar{p}p$ colliding-beam storage rings should not be operated with single-beam tunes (per C) that lie just below an integer or just below a half-integer. Proton-proton storage rings should not be operated with single-beam tunes that lie just above an integer or a half-integer.[38]

A more sophisticated picture of linear instability leads one to conclude that e^+e^- and $\bar{p}p$ colliding beams should not be operated with single-beam tunes--not divided by C--that lie just below an integer.[38] The modes of oscillation that are expected to become unstable when $C \times (\mu_y + \Delta\mu_y)$ is an integer that does not equal 0, $\pm C/2$, $\pm C$, etc., are coherent--in such a mode, all the particles in each bunch oscillate in phase with one another, i.e., each bunch oscillates as a rigid body. By contrast, the instabilities that set in when $\mu_y + \Delta\mu_y$ is an integer or half-integer, as discussed in the preceding paragraph, are incoherent--the unstable modes are single-particle modes, i.e., the instabilities are not accompanied by special phase relationships between the particles involved. Although signatures of beam-beam induced coherent oscillation have been observed quite clearly in some computer studies,[40] coherent motion has been difficult to see in the behavior of real storage rings without resonant excitation by some external stimulus. Attempts to associate beam-beam collisions at the ISR with the existence of significant numbers of stored protons oscillating in phase are reported in Refs. 41 and 42.)

From (4) and (5), one obtains

$$I = \frac{2\pi B f}{e_0} \cdot \frac{E\sigma_y^*(\sigma_x^* + \sigma_y^*)}{\beta_y^*} \cdot \Delta\mu_y \quad . \tag{6}$$

This is the formula for beam current that was mentioned earlier, when we introduced the conventional procedure for predicting maximum colliding current. In applications to high-energy e^+e^- storage rings, the factor $\sigma_y^*(\sigma_x^* + \sigma_y^*)$ is often replaced by $\sigma_y^*\sigma_x^*$, because σ_y^*/σ_x^* is usually much less than one. The product $\sigma_y^*\sigma_x^*$ is, in turn, often rewritten as $A^*/4\pi$, where A^* is the effective area filled by the Gaussian distribution (the factor 4π accounts for the part of the distribution lying beyond the half-widths). With these changes, Eq. (6) becomes

$$I = \frac{Bf}{2e_0} \cdot \frac{EA^*}{\beta_y^*} \cdot \Delta\mu_y \ . \qquad (6a)$$

Formula (6a) is especially suited to calculations describing the collision of two equivalent beams (equal E's, equal A*'s, equal $\Delta\mu_y$'s). In particular, in this case, substitution of (6c) into (2) (setting $A = A^*$) leads to the following simple relation between luminosity and tuneshift

$$L = \frac{Bf}{4e_0^4} \cdot \frac{E^2A^*}{\beta_y^{*2}} \cdot (\Delta\mu_y)^2 \ . \qquad (7)$$

The design expectations for maximum storable current at CESR, PEP and PETRA were all obtained from Eq. (6a), using $\Delta\mu_y = .06$ (together with expected values of E, β_y^*, and of unperturbed A^* (i.e., of A_0)). Maximum achievable luminosities were predicted from Eq. (7), using the same parameter identifications.

In order to form a sense of how these predictions fail, let us briefly consider them in detail for the case of PEP, prior to the modifications described in Ref. 108. (I choose PEP because its electron and positron beams happen to have roughly equal dimensions even when blown up,[33] so that Eq. (7) is applicable, even when A^* is not equal to A_0): At PEP, until recently, the maximum $\Delta\mu_y$ at the design energy of about 15 GeV was about .02, one third the design value.[12] Furthermore, the beam area A^* exceeded A_0 by about 50% at maximum current.[12] Thus, according to (6a), the overestimate of maximum current due to failure of the second nonperturbation assumption, and to failure of the tuneshift rule of thumb used in design, was $2/3 \times 3$, or a factor of two. Similarly, the overestimate of maximum luminosity due to the same causes was $2/3 \times 3^2$, or a factor of six. (Differences between predicted and achieved maximum current and maximum luminosity also receive contributions from discrepancies between the design values of storage ring magnetic lattice parameters and the lattice specifications achieved in practice, because β_y^* and A_0 depend on the lattice. We are not concerned with such contributions in this paper, as they are due primarily to single-beam problems.)

The easiest way to measure tuneshift is this: Measure current (using an inductive pickup enclosing the beam pipe) and luminosity (observing the rate of some well-understood hard scattering process), and then compute tuneshift from the formula

$$\Delta\mu_y = (2e_0^3)\frac{\beta_y^*}{E} \cdot \frac{L}{I} \ . \qquad (8)$$

This relation is obtained by equating the left and right-hand sides of Eqs. (6a) and (7). The virtue of (8) is that it does not refer to A^*, which cannot easily be observed at a densely instrumented interaction point.

It is instructive to consider the results of some maximum-tuneshift measurements of this kind: Figure 7 shows maximum achievable tuneshift vs. energy as measured at SPEAR.[10] Figure 8 shows tuneshift vs. current per bunch (two bunches per beam) for three energies, as measured at PETRA.[13] The maximum achievable tuneshifts in the latter case are the values of ξ_y at which the ξ_y vs. I_b curves reach plateaus; no plateau is seen in the curve at 17 GeV because at that energy it turns out that collision-unrelated factors destabilize the PETRA beams before the plateau can be reached.

These figures show very clearly that maximum tuneshift can depend significantly on energy at a given colliding-beam storage ring. This is important to bear in mind, because our discussion up to this point may have given the impression--and most of the literature on colliding beam theory definitely gives the impression--that one maximum tuneshift characterizes a storage ring for all energies at which it is designed to operate. This is simply not true.

Note, incidentally, that in each of these figures the maximum tuneshift rises with energy. This is another example of the general tendency, mentioned earlier, of colliding beams to perturb one another less (at a given machine) as energy increases.

Figure 7 is perhaps the more interesting of these two graphs because it shows clearly the existence of two regimes in which the storage ring behaves in two qualitatively different ways: For energy less than about 2 GeV, the maximum SPEAR tuneshift appears to grow as a power of energy ($\simeq E^{2.4}$), while for energy greater than 2 GeV, the maximum tuneshift appears to saturate at a value between .05 and .06.

Of these two regimes, the one at low energy is so far the more thoroughly studied phenomenologically: Wiedemann[16] and Cornacchia[19] have shown that other functions characteristic of SPEAR can be fit to power laws for $E<2$ GeV. Such functions include maximum luminosity vs. energy; and also luminosity vs. current at fixed energy, or vs. energy at fixed current, for currents that exceed the energy-dependent turnover point beyond which L grows more slowly than I^2. Wiedemann[34] has also discovered a simple parametrization of SPEAR data in this regime of low energy and high current that appears successfully to describe the operating characteristics of other storage rings in similar regimes. His universal formula is

$$L \simeq (4\times10^{33}) \cdot \sqrt{f} \cdot E^{1.4} \cdot (I/B)^{1.5}, \qquad (9)$$

where f and B are as defined following Eq. (1). The plot in Fig. 9 shows how close the formula comes to experimental data taken at a number of different storage rings. (Note that SPEAR data taken at high energy and/or low current is not shown in this figure.) A

discussion of power law fits applied by other workers to the behavior of other storage rings can be found in Ref. 35.

Before the SPS was operating routinely as a proton-antiproton collider, there were attempts to use e^+e^- data of the kind shown in Fig. 7 in order to predict, by extrapolation, the maximum tuneshift that could be achieved in a $\bar{p}p$ storage ring. To carry out such an extrapolation, one had commonly to assume that there always exists some (typically low) energy at which a given e^+e^- storage ring behaves as a simple rescaling of the high-energy $\bar{p}p$ storage ring being designed or proposed.

A careful analysis[36] reveals that the parameters of two colliding-beam storage rings must satisfy a number of relations before one can properly consider the two colliders to be rescaled versions of one another. In carrying out tuneshift extrapolations, it was commonly assumed[15] that only one of these relations--specifically, equality of the number of bunch collisions in a radiation damping time--is sufficient to guarantee that two storage rings are equivalent up to overall scale.

Let us determine the energy at which SPEAR is equivalent, according to this simplified criterion, to the 250 GeV SPS collider: The number of bunch collisions in a damping time τ is equal to $2B^2f\tau$. τ depends on storage-ring circumference ℓ_0, particle mass m, and particle energy E as[4], roughly, (constant) $\cdot$ $(m^4\ell_0^2/E^3)$; f depends on these parameters as (constant) $\cdot$ $1/\ell_0$. Thus, assuming equal numbers of bunches in the two storage rings, SPEAR and the SPS are equivalent according to the criterion above when

$$E_{SPEAR} = E_{SPS} \left(\frac{m_{SPEAR}}{m_{SPS}}\right)^{4/3} \left(\frac{\ell_{0\,SPEAR}}{\ell_{0\,SPS}}\right)^{1/3}$$

$$\simeq (250\ \text{GeV}) \cdot (1836)^{-4/3} \cdot (29.5)^{-1/3}$$

$$\simeq .04\ \text{GeV} . \qquad (10)$$

If one extrapolates the low-energy part of the curve shown in Fig. 7 down to this energy, one predicts for the SPS a maximum tuneshift of about 5×10^{-6}. This turns out, with hindsight, to be a bad prediction: Tuneshifts at the SPS collider are routinely as high as a few parts per thousand.[37] Thus, one should be wary of such extrapolations, and of the simple assumptions that go into them.

One of the most serious weaknesses in such an extrapolation is the following: Electron-positron storage rings, from whose behavior one wishes to extrapolate, have radiation damping times that vary between a few milliseconds at very high energy and a few seconds at very low energy. Thus, a great deal of transient behavior has died away by the time measurements are made. By contrast, proton-antiproton storage rings, to which one wishes to extrapolate, have radiation damping times of several months, so that the transients that are not seen in observations of e^+e^- colliders are

necessarily fully present for any practical $\bar{p}p$ measurements made under normal circumstances. In other words: If there are characteristics of $\bar{p}p$ (or pp) storage rings, operated under normal conditions, that can be predicted by extrapolation from e^+e^- machines, operated under normal conditions, then one does not have enough time to see such characteristics set in. An interesting example of an extrapolation that goes the other way--from pp to e^+e^---and involves measurements made under abnormal conditions, will be mentioned in Sec. 4.

One quantity often mentioned in conjunction with the (vertical or horizontal) tuneshift is the (vertical or horizontal) tunespread. Here is a brief schematic definition: Recall first the general form of the vertical betatron oscillation of a particle in a storage ring, when nonlinearities in the forces due to magnets, and to the opposing beam, are neglected:

$$y(s) = \sqrt{2I_y\beta_y(s)}\ \cos\left(\phi_y + \int_o^s \frac{ds'}{\beta_y(s')}\right) . \qquad (11)$$

s is the product of time and the speed of light, and I_y (nonnegative) and ϕ_y are artibrary independent constants of integration; β_y depends neither on I_y, nor on ϕ_y. In a sense that will be made precise in Chapter IV, the leading effect of the nonlinearities (due primarily to the opposing beam) neglected in (11) is that the form (11) is replaced by

$$y(s) = \sqrt{2I_y\beta_y(s)}\ \cos\left(2\pi\left(\frac{s}{\ell_o/C}\right)\delta\mu_y(I_x,I_y) + \phi_y + \int_o^s \frac{ds'}{\beta_y(s')}\right), \qquad (12)$$

where ℓ_o is the storage ring circumference, and $\delta\mu_y$ is some function of I_x and I_y. In particular, in the presence of nonlinearities, the tune of vertical betatron oscillations--in (12) given, per C, by $\mu_y + \delta\mu_y(I_x,I_y)$--depends, in a first approximation, on the amplitude of oscillation. The vertical tunespread is the difference between the largest and smallest values taken by this amplitude-dependent tune (i.e., by $\delta\mu_y$, since μ_y is independent of amplitude). Horizontal tunespread is defined in a similar way.

(It may be of interest to note that a nonzero tunespread can contribute to beam stability. In particular, certain resonant instabilities, to which a single beam can be subject when its tune is very narrowly determined by the storage-ring hardware, can become less severe under the influence of beam-beam collisions, because they can spread out the range of tunes in the beams beyond the instability bandwidths. Such an effect has been observed at ACO.[39])

In general, tuneshifts and tunespreads all have the same order of magnitude (in other words, the tuneshift sets the scale for both linear and nonlinear beam-beam perturbations). This rough equality can be quite useful. It tells us, for example, that one cannot use compensating magnets to cancel the nonlinear forces that cause beam blowup and limit storable colliding current, because tunespreads as large as the beam-beam tuneshifts typical of e^+e^- storage rings are two orders of magnitude greater than can be practically reproduced by arrays of magnetic lenses.[39] A more promising method for cancelling these destabilizing forces is the subject of the next subsection.

c. The DCI project

To the best of my knowledge, the French DCI project[39] represents the only significant attempt to develop a means of cancelling the nonlinear forces that cause beam blowup and limit current in electron-positron colliding-beam storage rings.

DCI consists of two e^+e^- storage rings (peak energy 1 GeV) with two regions of tangency, as shown in Fig. 10. In one ring, electrons travel clockwise and positrons travel counterclockwise; in the other ring, these directions of circulation are reversed. In the interaction regions, bunches from the two oppositely-charged counterclockwise beams can (when the timing is right) combine to form electrically neutral bunches that collide with electrically neutral combinations of positively and negatively charged clockwise bunches. The more complete the neutralization, the more complete the cancellation of Coulombic beam-beam kicks; and therefore, presumably, the smaller the blowup and the higher the maximum colliding current.

(This scheme makes the experimentalist's job harder, because he or she must be able to distinguish between electron-positron, electron-electron, and positron-positron collisions; but this would be a small price to pay for a large increase in luminosity.)

The success of this project has been limited. On the one hand, beam blowup has, under certain conditions, been suppressed; on the other hand, no increase of maximum colliding current has been achieved.

The results of four experiments performed on DCI[43] are shown in Fig. 11. In one experiment ("e^+e^- lower ring") equal electron and positron currents were circulated in one ring, with the other ring vacant; the second experiment ("e^+e^- upper ring") was identical to the first, except that the two rings were interchanged. In the third experiment ("3 beams"), one ring contained a large current of one charge and a very small trace current of the opposite charge, while the other ring contained only a single beam, of current equal in magnitude to that of the strongly populated beam in the first ring, but of opposite charge. In the fourth experiment ("4 beams"), four beams of equal current were circulated at the same time. Neither beam could be neutralized in cases one and two; in case three, the high current beams could in principle neutralize one another; in case four, all beams could in principle be neutralized.

The abscissa of Fig. 11 refers to the current in any one of the two or four beams in the first, second or fourth case; in case three, the abscissa refers to the current in either of the two strongly populated beams. The ordinate of Fig. 11 refers to the cross-sectional area (measured at some fixed storage ring reference point) of the trace beam in the third case; the beam whose area is given by the ordinate in the other cases is not specified; in all cases, the area is normalized to the low-current area of a single isolated beam.

In the first two cases, a strong current-dependent beam blowup is clearly evident in Fig. 11. In the other two cases, neutralization has the desired effect of suppressing blowup. However, this success is offset in the realistic "4 beams" case by a low maximum current.

3. Colliding-beam effects in pp and $\bar{p}p$ storage rings

a. Introduction

The beam-beam effects that dominate the literature on routine operation of e^+e^- storage rings do not figure in corresponding discussions of the ISR and the SPS: Under normal conditions, ISR and SPS beams do not generally blow up; and the ISR or SPS current per beam that can be stored with beams in collision is not generally smaller than the current that can be stored when only a single beam is present. (Observations made under abnormal conditions are discussed in Sec. 4.)

For this reason, the literature on routine operation of the ISR and the SPS has never had a dominant theme. Although miscellaneous effects of beam collisions are easily detectable in normal ISR and SPS performance, these effects are generally too small, or too slow, to justify the kind of singleminded attention that is typically paid to blowup and to tuneshift shortfall in reports on electron-positron storage rings.

In this section, we discuss four such small or slow effects that have been publicly documented, two at the ISR and two at the SPS. Where appropriate, we shall try to assess the importance of what has been seen.

Three of these four otherwise diverse effects have this in common: They occur most strongly when the orbital parameters of some beam particles come close to satisfying resonance conditions. The form of the most general such condition is

$$n_x \nu_x + n_y \nu_y + \sum_i n_i f_i / f_o = n \ , \qquad (13)$$

or equivalently (redefining n)

$$n_x\nu_x + n_y\nu_y + \sum_i n_i\left(\frac{f_i - f_o}{f_o}\right) = n \quad . \tag{14}$$

In these equations, the n's are integers, ν_x and ν_y are the storage ring tunes (not divided by C), f_o is the revolution frequency of such a resonant particle, and the f_i are frequencies of any periodic beam perturbations that are not synchronized with f_o. Such perturbations usually arise from coupling to longitudinal oscillations, and from the interaction of beam particles whose orbital frequencies differ slightly because of slight differences in energy. Thus, neglecting longitudinal coupling (as is often done) f_o and the f_i are typically very close; for most purposes (but not for all--see below) one does not distinguish between them, and therefore the summation term in (13) or (14) is ignored.

The resonant relations that have actually been correlated with beam-beam phenomena at the ISR or SPS involve n_x's, n_y's, and n_i's that are not larger in absolute value than about ten. Such correlations would not be meaningful if a much wider range of integers were involved, since any pair of tunes can be closely approximated by a ν_x and a ν_y satisfying a relation of the form (13) or (14), provided one uses sufficiently large integer coefficients.

The existence of discrete, often unstable, phenomena that are associated with small numbers of resonance conditions is a typical characteristic of systems with weak nonlinearities. When nonlinearities are strong, it is often difficult to associate observed effects with specific resonances. For these reasons, one often identifies ISR and SPS tuneshifts/tunespreads ($0(10^{-3})$) with weak nonlinearity, and e^+e^- tuneshifts/tunespreads ($0(10^{-2})$) with strong nonlinearity.

The term $n_x\nu_x$ is usually omitted in discussions of the ISR, because the ISR beam-beam interaction hardly couples to horizontal oscillations. The reason for this is as follows: Recall that the ISR beams both lie in the horizontal plane, are both continuous, and cross at an angle. Thus a particle in one beam sees the other beam, at a crossing point, as a charge distribution that is nearly uniform in the horizontal coordinate (see Fig. 12), and therefore no kick that a particle gets from the opposing beam has an appreciable horizontal component.

By convention, one usually writes resonance conditions using integers n_x, n_y, $\{n_i\}$, and n that have no common divisor. However, as we shall see in Chapter IV, there is a mathematical sense in which one might distinguish between a resonance indexed by (n_x, n_y, $\{n_i\}$, n) and one indexed by (mn_x, mn_y, $\{mn_i\}$, mn), with m an integer not equal to one, zero, or minus one. In the discussions that follow, we shall assume that all resonances encountered are described by integers (n_x, n_y, $\{n_i\}$, n) that have no common divisor, even when interpreted according to the mathematical formalism to which we have just alluded. When the integers n_x and n_y are small (one or two),

this might lead to meaningful ambiguities. However, in this section we shall be discussing the effects of resonances for which either n_x or n_y is so large that it would be difficult to impute the same effect to a resonance indexed by mn_x and mn_y, with $|m| \geq 2$. The effects associated with a resonance usually diminish rapidly as the order $(|n_x|+|n_y|)$ of the resonance grows.

b. Colliding-beam effects in the ISR

The effects to be discussed here are: loss of current (over and above the single beam loss rate) while two coasting ISR beams collide; and loss of current (over and above the single beam loss rate) while the ISR rings are being filled. The single beam loss rate is a few parts per million per minute;[44] it is primarily due to hard nuclear collisions with residual gas molecules that have not been removed by the storage ring vacuum pumps.

In the context of the preceding discussion, the fundamental difference between these two effects is this: In the former case, all the forces on a circulating particle are synchronous with its revolution frequency, and therefore there are no f_i terms in the corresponding resonance conditions (14). In the latter case, circulating particles experience asynchronous perturbations, and therefore there are nontrivial f_i terms in the corresponding resonance conditions.

Current loss from coasting beams in collision: Guignard[45] reports that although the loss rate for ISR coasting beams in collision is usually comparable to the rate of single-beam current loss, colliding loss rates as high as 20-60 ppm per minute have been observed when the storage rings are operated in a standard configuration that places the vertical tunes near a fifth order ($n_y = 5$) resonance.

This particular rate enhancement turns out to be one of the few concrete numbers in colliding beam physics that has been plausibly explained by a theoretical calculation. This calculation is described in Chapter V.

Even without a detailed calculation, one can immediately draw an important conclusion when beam loss is enhanced near an odd-order vertical resonance: Contrary to a common idealization, the beam distributions cannot be exactly symmetrical about the horizontal plane of the design orbit. (Indeed, in the case reported by Guignard, the centers of the colliding beams were vertically misaligned by about .2 mm--about one-tenth the vertical size of a beam.) The reason: In a machine with perfect vertical symmetry, one expects (as will be explained in Chapter IV) to see enhancement of current loss only near vertical resonances of even order. We shall use a similar argument to draw a similar conclusion from one of the SPS effects described in the next subsection.

Current loss during filling: According to Gourber et al.,[46] a persistent enhancement of current loss during storage ring filling was at one time a problem at the ISR. (The precise magnitude of the problem is difficult to abstract from the published reports--although

measurements of current loss are tabulated, the scales are suppressed. Beam blowup during filling is also mentioned, but in this case no quantitative measurement--with or without a scale--is cited.) The remedy that eliminated this effect was suggested by a qualitative theory that we explain below.

Before we proceed, it should be pointed out that the ISR beams do not actually intersect while the storage rings are being filled. During filling, magnetic separators[47] keep a vertical distance of about 8 mm between beam centers at crossing points,[46] while, as mentioned earlier, the vertical thickness of a beam is only about 2 mm. Thus, loss enhancement during ISR filling differs from the other phenomena discussed in this paper in that it arises from a long range Coulombic interaction between the two beams.

Enhancement of current loss during ISR filling is understood[46] to be a consequence of the beam bunching that necessarily accompanies the filling process. Let us briefly explain why filling--or "stacking"--a beam of uniform longitudinal distribution entails some bunching in intermediate stages of the process. We will then explain how this bunching provokes beam losses.

A storage ring is filled with a coasting beam in the following way:[48] An injecting accelerator deposits some particles into a storage ring orbit. The radius of the orbit is determined by the particles' energy, which is in turn determined by the characteristics of the injector. In order to make room for the next batch of injected particles, the batch already in the storage ring must be moved to a different radius. This is done by changing its energy, i.e., by accelerating it slightly, using RF cavities. After the next batch of particles is injected, acceleration takes place again, to make room for another batch, and so on. The bunching mentioned above is maintained by this periodic acceleration, as dictated in the usual way by the theory of phase stability in circular accelerators.[4]

Because of this bunching, a beam that is in the process of being filled has a complicated shape. The circling of the beam causes details in this shape to propagate around the storage ring--a feature at a given radius circles the storage ring at (roughly speaking) the orbital frequency of particles stored at that radius. As there is a range of orbital radii represented in the beam, there is a range of circulating frequencies represented in the beam's shape.

With this in mind, consider now a test particle in the opposing beam, assumed to be already filled. It feels the beam that is being stacked as a superposition of periodic perturbations whose frequencies are given not only by the frequency with which the test particle itself circles its storage ring but also by the frequencies with which details in the stacking beam's shape circle the stacking beam's storage ring. In view of the foregoing paragraph, these frequencies are spread with nonzero bandwidth about the test particle's orbital frequency. Enhanced current loss is thought to occur during stacking because[46] some of these frequencies can serve as f_1's to the test particle's f_0 in destabilizing resonances of the form[7] (14). Such resonances in this context are called "two-beam overlap-knockout" resonances. When filling is completed and the RF

cavities are turned off, details of the beam distributions wash themselves out, only f_o is left to determine the periodicity of the beam-beam force, the overlap-knockout resonances disappear, and current loss settles down.

Direct--although somewhat artificial--support for this picture comes from the following type of experiment: Two low-current beams of nearly equal, narrowly defined energies are loaded into the ISR rings. One beam is allowed to coast (unbunched); the other beam--necessarily bunched--is accelerated slowly but steadily. The current in the coasting beam is monitored throughout the acceleration cycle, and then graphed as a function of the varying fractional difference between the energies of the two beams. Since the difference in the two beams' revolution frequencies is proportional to the fractional difference between their energies (as long as that difference is small), one derives, from the picture above, the following qualitative expectation for this graph: It should show strong variation* at discrete values of $\Delta E/E$ satisfying

$$n_y \nu_y + n_1 \zeta \left(\frac{\Delta E}{E}\right) = n \,, \qquad (15)$$

where the n's are integers, as before, and ζ is some numerical constant.

One such graph[46] is reproduced here as Fig. 13. As expected, the slope is amplified at discrete values of the relative energy difference. The labels at these points indicate n_y and n_1 values that have been assigned by the experimenters. "First order" means $n_y = 1$, "second order" means $n_y = 2$. The same integer n is associated with all points having the same n_y.

It must be borne in mind that such experiments exaggerate the effect that overlap-knockout resonances have on current loss during routine filling. In routine filling the beam energies and frequencies are not so narrowly defined, and therefore the onset of a resonant instability is not so distinct.

The remedy for enhanced current loss and beam blowup during ISR filling turned out to be a simple modification of the RF system. This modification lengthened the bunches that are formed during acceleration. Loosely speaking, this smooths out the features in the shape of the beam that is being stacked, thereby reducing the asynchronous components of the perturbation felt by a particle in the opposing beam.

c. Colliding-beam effects in the SPS

We describe briefly two proton-antiproton effects that were highlighted in a recent report on the SPS by Evans and Gareyte:[49] Large variations of beam lifetimes that result from small variations in tune; and slow asymmetric particle losses that result in

*A quantitative theoretical prediction of the strength and shape of this variation is not available.

equalization of the emittances of two initially unequal-emittance colliding beams.

Tune-dependence of beam lifetimes: Tune variations as small as about .01 have been observed to result in large variations in beam lifetime. A sample measurement of antiproton beam lifetime, as a function of horizontal SPS tune, is shown in Fig. 14. This measurement was performed with one bunch per beam, and with $\nu_x - \nu_y$ held fixed at about .02. One sees that the lifetime drops by a factor of about three between $\nu_x = .71$ and $\nu_x = .72$.

It should be noted that although the minimum lifetime--about two hours--reached by the curve in Fig. 14 is comparable to typical beam lifetimes (about five hours) in electron-positron storage rings, it is much shorter than is desirable at the SPS. Ideally, beam lifetime should exceed the time needed for accumulating and collimating a replacement antiproton bunch. At the SPS, this replacement time is about twelve hours. The largest lifetimes observed at the SPS are in the range 40-50 hours.

The experimenters have indicated, by the markings "7th order" and "10th order," that the curve in Fig. 14 shows structure at the resonances $7\nu_x = 187$ and $10\nu_x = 267$ ($\nu_x = 26.714$ and $\nu_x = 26.700$, respectively). As in the preceding subsection, one can immediately draw two important conclusions from these features: First, either one of the two bunch distributions is not left-right symmetric when viewed along its beam axis, or the horizontal alignment of the two bunches is not perfect when they collide (or both). Second, the two regions between the two interactions are not completely identical, contrary to one of the expectations articulated in Sec. 2b.

The reasons for these conclusions are as follows: In the behavior of a storage ring with left-right symmetric beams that collide horizontally centered, one expects to see tune-dependent structure only near horizontal resonances of even order; the resonance $7\nu_x = 187$ violates this condition. In the behavior of a storage ring built of two identical sections, one expects to see tune-dependent structure only near resonances $n_x\nu_x + \nu_y\nu_y = n$ with n even; this condition is violated by $7\nu_x = 187$ and $10\nu_x = 267$. The mathematical basis for these symmetry expectations will be explained in Chapter IV.

Emittance equalization: At injection, the antiproton beam in the SPS has mean emittances approximately 50% greater than those of the proton beam (i.e., transverse dimensions approximately 25% greater than those of the proton beam). In experiments involving two proton bunches and one antiproton bunch ($\sim 10^{11}$ particles per proton bunch; $\sim 10^9$-10^{10} particles per antiproton bunch[110]), it was observed that over a period of about seven hours, antiprotons having emittances significantly larger than the mean proton emittances were lost (absorbed by the pipe wall) faster than their counterpart protons, and faster than could be explained by encounters with residual gas left behind by the vacuum system. After about seven hours the transverse dimensions of the antiproton beam matched those of the proton beam. Thereafter (for the remaining ~40 hours of storage), beam dimensions remained matched, and particle loss proceeded at a

rate adequately accounted for by gas scattering. The role--if there is one--of resonances in this phenomenon is not at all as apparent as the roles played by resonances in the other ISR and SPS effects discussed in this section.

4. Nonstandard ISR studies

In this section, we discuss several experimental studies carried out at the ISR under nonstandard operating conditions. Some of these studies were in part conducted to help designers determine in advance how $\bar{p}p$ colliders would perform, and what their limitations would be. In each study, an artifice was employed for the purpose of reproducing the operating conditions anticipated in a head-on bunched-beam collider. Every nonstandard study was also in part--if not entirely--conducted for the purpose of accumulating experimental information about the consequences of large nonlinearity. In each, some artifice was employed with the aim of substantially increasing the tuneshift felt by one of the beams.

A priori, an ISR beam is a poor model of a beam in a head-on collider like the SPS, mainly because it is not bunched, and because the localized target (the opposing beam) in its path resembles Fig. 12b, rather than Fig. 12a. The ISR has on separate occasions been modified in two different ways to compensate for one or the other--unfortunately not for both--of these deficiencies: Experimenters have bunched one ISR beam, leaving the other continuous. (Because bunching substantially reduces beam current, the continuous beam in such an arrangement is hardly perturbed; thus the test beam is necessarily the bunched one.) In this case, bunching is the characteristic being simulated. Experimenters have also placed a system of current-carrying bars (a "nonlinear lens") so that the bulk of its magnetic field lies in the path of an unbunched ISR beam (the other beam is shut down). In such a study, the nonlinear lens is introduced to simulate the shape of the target in a head-on collider.

In order to enhance the tuneshift felt by an ISR beam (bunched or unbunched) against which a second high-current beam is circulated, experimenters have on separate occasions tried decreasing its energy and increasing the value of its beta function at interaction points, in either case leaving the opposing beam's characteristics unchanged. The rationale for these measures is provided by Eq. (4a), according to which the tuneshift felt by the test beam is expected to vary in direct proportion to β^*_y, and in inverse proportion to E. In nonlinear lens studies, tuneshift is increased by raising the current in the current-carrying bars.

Energy-reduction experiments[50,51] (with and without bunching) have been inconclusive, because apparent single-beam instabilities turn out to limit the extent to which energy can be reduced practically. For this reason, I shall not comment further on studies involving energy reduction.

Zotter[52] used the high-beta technique on continuous beams to obtain tuneshifts between .005 and .02 in absolute value. At each

tuneshift, he measured the lifetime of the test beam. His data is reproduced in Fig. 15. According to the figure, the dependence of lifetime on tuneshift is approximately exponential, through two orders of magnitude in lifetime. As far as I know, this has not been theoretically explained.

Zotter pointed out[52] that his data, extrapolated to a tuneshift of .06--SPEAR'S maximum--predicts a lifetime of a few milliseconds--the order of magnitude of SPEAR's synchrotron radiation damping time. This coincidence is probably not significant, considering that at CESR's maximum tuneshift, about .02, Zotter's data shows an ISR lifetime of about one hour, six orders of magnitude longer than CESR's damping time.

Hofman et al.[51] used the high-beta technique on a bunched ISR beam and obtained tuneshifts as high as about .0035 in absolute value, comparable to current SPS values. Their main finding at this tuneshift concerned the effects of beam misalignment: They found that when the colliding beams were well-centered, test-beam lifetime was about 30 hours; but when the beam centers were separated vertically at crossing points by about half a beam thickness, test beam lifetime declined to about 1.5 hours. This is reminiscent of the SPS asymmetry effects mentioned in the preceding section.

Keil and LeRoy[53] conducted a nonlinear lens study at the ISR and obtained tuneshifts as high as .1 in absolute value. At each tuneshift, they measured the beam decay rate for a variety of storage ring tunes. They found that if, at high lens tuneshift, a machine tune (horizontal or vertical) was changed by as little as a few percent of the tunespread, the corresponding change in beam loss rate could be a few orders of magnitude greater than that which a similar tune change could cause with the nonlinear lens turned off. Keil and LeRoy give the following example: With the lens inactive they found that a tune change (horizontal or vertical) of .0025 changed the loss rate by at most a few times 10^{-7} sec^{-1} (i.e., the order of magnitude associated with beam-gas scattering); but with a lens tuneshift of about -.05, tune variations of about .0025 changed the beam loss rate by as much as about 10^{-4} sec^{-1}. As far as I know, this extreme sensitivity has not been explained theoretically.

Their data is represented schematically in Fig. 16. Each bar indicates the range of loss rates observed on several occasions at the corresponding tuneshift. The crosses indicate arithmetic means of highest and lowest loss rates. Notice that the loss rates in this figure appear to saturate at tuneshifts above about .05 in absolute value. On the basis of a computer simulation, Keil and LeRoy interpreted this saturation as a symptom of "large-scale stochasticity"--a type of behavior phenomenologically similar to random motion, characteristic of strongly nonlinear systems. As far as I know, this interpretation has not been pursued more deeply. We shall say more about stochasticity in Chapter III, Subsection 3b, and in Chapter IV, Subsection 2c.

A nonlinear lens study was later undertaken at the SPS.[54] Since this work was not nearly as extensive or systematic as in its ISR counterpart, we shall not discuss it further.

III. NUMERICAL PHENOMENOLOGY

As computing technology has developed[55]--and confidence in analytical theory has deteriorated--more and more effort has been invested in computer simulation of colliding-beam effects in storage rings. This kind of work serves a number of purposes: testing and refining mathematical models of beam-beam phenomena; determining in advance how proposed colliders will perform; exploring ways to enhance the performance of existing storage rings.

All simulations of which I am aware have in common the same basic structure. This structure is described in Sec. 1, below, after which we discuss specific results from recent simulations of electron-positron storage rings (Sec. 2), and from recent simulations of proton-antiproton and proton-proton storage rings (Sec. 3).

1. Basic structure of simulations

As commonly understood at present, a storage ring simulation is a computer program that applies long sequences of a few types of linear and nonlinear transformations to the initial phase space coordinates of a small number of mathematical test particles, and then tabulates statistical properties of the resulting set of final phase space coordinates. Each type of transformation idealizes the effects of a different physical process. The processes which have so far been included in beam-beam simulations are: passage through magnets, passage through RF cavities, radiation damping, radiation noise, and beam-beam encounters. Potentially important processes which, to my knowledge, have so far not appeared in simulations that involve colliding beams include: interactions between particles in the same beam, and interactions with image charges induced in the vacuum pipe.

Although the parameters defining these transformations might be varied during the course of a simulation (see below), the program always applies them in a fixed, periodically repeating order. For example: magnet transport, RF transport, damping, noise, beam-beam encounter, magnet transport, RF transport, etc.... Separation of these effects into discrete sequential transformations is an idealization that one makes for convenience only. In reality, damping and noise, for example, carry on at the same time that particles travel through banks of magnets. In general, one transformation cycle is equivalent either to one revolution of a test particle around the storage ring, or to one passage of a test particle through a single interaction region and a single inter-collision storage-ring arc.

A mathematical test particle is not always to be literally interpreted as a model of a single beam particle. A storage ring beam, for example, contains as many as a few times 10^{11} particles per bunch, while simulations involving more than a few hundred test particles are impractically time-consuming. If N is the number of particles (charge $\pm e_0$, mass m_0) in the real beam being modelled, and M is the number of test particles whose orbits one can afford to

digitally compute, then each test particle is defined to have charge $e = \pm(N/M)e$ and mass $m = (N/M)m$.

The total number of transformation cycles in a simulation depends on the type of storage ring being simulated. If the beam consists of electrons and positrons, then the number of transformation cycles is usually chosen so that the length of the simulation is equivalent, in real time, to a few (unperturbed) radiation damping times. (The unperturbed damping time is generally assumed--rightly or wrongly--to set the scale for the decay of transients in colliding beam distributions.) This number is a few thousands for large colliders such as CESR, PEP, PETRA and LEP, and a few tens of thousands for small colliders such as SPEAR. (For this reason, large e^+e^- storage rings are easier to simulate than small ones.[31]) If the beams consist of protons and/or antiprotons, then the number of transformation cycles is generally set equal to the largest number that one can afford to program, since radiation damping is so slow in proton/antiproton colliders.

The mainframe CPU time corresponding to these numbers of transformation cycles can range from a few minutes per model beam, for some electron-positron simulations,[56] to a few hundred hours per model beam for some proton-antiproton simulations.[73]

One drawback of limiting the number of e^+e^- transformations to the equivalent of a few damping times is that one is thereby prevented from being able to derive values of maximum storable current from the outputs of simulations. The reason is this:[56] In real storage ring operation, maximum current is determined by the maximum tolerable beam loss rate, which is usually about a part per thousand per damping time. This is too small a loss to observe in a statistically significant way in a model beam containing the usual number--a few hundred or less--of test particles.

Let us now consider in detail the different types of transformations that are applied to the phase-space coordinates of such test particles in the course of a typical simulation.

Passage through magnets: The effects of magnets encountered during a single transit between two adjacent interaction regions are generally modelled by a linear transformation of the following form:

$$\begin{pmatrix} y \\ y' \end{pmatrix} \rightarrow \begin{pmatrix} \cos 2\pi\mu_y & \beta_y^* \sin 2\pi\mu_y \\ -(1/\beta_y^*) \sin 2\pi\mu_y & \cos 2\pi\mu_y \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} \qquad (16)$$

(similarly for x). The input to this transformation consists of the betatron coordinates and conjugate momenta of a test particle immediately before passage through the magnets; the output consists of the coordinates and conjugate momenta just after passage through the magnets. As in Chapter II, β_x^* and β_y^* are the beta function values at either interaction point (one assumes the same β's at all interaction points); μ_x and μ_y are the amounts by which the betatron

phases advance between the two interaction points. Magnet nonlinearities are generally ignored.

Programmers sometimes use μ's that vary during the course of a simulation. Two types of variation are common: One type of variation is realized by terms in μ_x and μ_y that take the same pair of values every time the test paticles pass between the same pair of interaction regions, but which take different pairs of values when the particles pass between different pairs of interaction regions. Such terms are typically employed to model storage ring irregularities. The other type of variation is realized by terms in μ_x and μ_y that depend on the amplitude and phase of a test particle's longitudinal displacement with respect to the center of its bunch as it leaves an interaction region. (This type of term gives rise to μ's that can be different for different test particles.) One rationale for such a term is the combined effect of longitudinal (or synchrotron or energy) oscillation and chromaticity ($E\partial\nu/\partial E$). Another rationale is this: A particle that oscillates longitudinally is sometimes near the head of its bunch, and sometimes near the tail. Thus, sometimes it encounters the opposing beam early, and sometimes late. Consequently, the distance it travels between interactions varies--sometimes greater than average, sometimes less. Because this distance is the interval over which one integrates $1/\beta_x$ and $1/\beta_y$ in order to obtain betatron phase advances, these advances must also vary.

Programmers generally translate this into mathematical terms as follows: Let ℓ be the longitudinal distance by which a test particle leads the center of its bunch. During a collision, this particle passes the center of an opposing bunch when the particle is a distance $\ell/2$ (see Fig. 17) past the center of the interaction region (where both bunch centers coincide). Thus, if $\Delta\ell$ is the change in ℓ during one transformation cycle, then the distance travelled by the test particle between collisions exceeds the average by $\Delta\ell/2$. Accordingly, the consequent changes in μ_x and μ_y can be approximated by $1/2\pi\ (\Delta\ell/2\beta_x^*)$ and $1/2\pi\ (\Delta\ell/2\beta_y^*)$.

(It should be noted that this approximation neglects the variation of β_x and β_y near an interaction point. β_x^* and β_y^* are both beta-function minima. In general,[4] if s measures longitudinal distance from the point where β_y reaches its minimum β_y^*, then for small s

$$\beta_y \approx \beta_y^* \left(1 + (s/\beta_y^*)^2\right), \tag{17}$$

and similarly for β_x. This is a negligible effect when β_x^* and β_y^* are much larger than the characteristic value of s--one-quarter the bunch length; but it could be a potentially important source of parameter variation in simulation transformations if either of the β^*'s becomes comparable to the bunch length.)

Several of the studies that we discuss in the next two sections indicate that model beams become significantly less stable when the parameters defining storage ring transformations are made to vary

during the course of a simulation. This is in accord with a rule of thumb often articulated by analysts of nonlinear systems: Suppose one is given two nonlinear systems, the first having more degrees of freedom than the second, or having an externally modulated parameter that the second does not have; then the first is inherently less stable than the second, because (loosely speaking) it's behavior involves a larger number of frequencies, and therefore there are more ways for it to be resonant (or near resonant). A more detailed exposition of this viewpoint can be found in Ref. 5.

Damping of betatron oscillations, per interaction, is sometimes (e.g. see Ref. 57) modelled by a transformation of the form

$$\begin{pmatrix} x \\ x' \end{pmatrix} \longrightarrow \begin{pmatrix} x \\ (1 - \frac{2}{C}\frac{T}{\tau_x})\, x' \end{pmatrix} \tag{18}$$

(and similarly for y). τ_x is the horizontal damping time. T is the (average) time it takes for an ultrarelativistic beam particle to travel around the ring--i.e., the ring circumference divided by the speed of light. Use of this transformation is almost equivalent to assuming that all damping takes place at a single point: When $\tau_x = \tau_y$, Eq. (18) (together with its vertical counterpart) is an exact expression for the effect that a single infinitely thin RF cavity has on transverse phase space.[4] A different transformation is used in Ref. 31:

$$\begin{pmatrix} x \\ x' \end{pmatrix} \longrightarrow e^{-T/C\tau_x} \begin{pmatrix} x \\ x' \end{pmatrix} \tag{19}$$

(and similarly for y). Use of this transformation is equivalent to assuming that damping is smoothly distributed along the region between interactions, and that $T/C\tau_x << \mu_x$ (i.e., that oscillation is much more rapid than damping).

Longitudinal motion: RF cavities and damping: A published report on a simulation almost never explicitly displays the transformation employed to model the effects that RF cavities and intercavity radiative energy loss have on the coordinates of longitudinal phase space. It is most probably one time-step in a simple discretization of the system of differential equations that one finds in the standard textbook discussions[4] of damped synchrotron oscillation:

$$\frac{d\epsilon}{dt} = \frac{V(\tau)-U(\epsilon)}{T}$$

$$\frac{d\tau}{dt} = -\alpha\frac{\epsilon}{E}\,, \tag{20}$$

where: ε is the difference between a test particle's energy and the energy, E, of a stable orbit synchronous with the RF system; τ is the time by which a test particle leads the nearest synchronous orbit (τ and ℓ are related by $\ell = c\tau$, where c is the speed of light); $V(\tau)$ is the net energy gained from all RF cavities upon one storage ring revolution, as a function of the test particle's timing; α is the momentum compaction factor; $U(\varepsilon)$ is the average energy lost to synchrotron radiation in the course of one revolution; and T is the revolution time. In most programs, $V(\tau) - U(\varepsilon)$ is approximated by a linear function of τ and ε. The simulation described in Ref. 31 uses a fully sinusoidal expression for V.

In a proton-proton or proton-antiproton simulation, the program typically assigns the same time-dependent values of ε and τ to all test particles; in other words, all test particles are made to undergo synchrotron oscillations in phase, and with identical amplitude. This is done to save computing time, so that the number of transformation cycles can be made as large as possible. However, such a procedure makes simulation results difficult to interpret, for this reason: In a real storage ring, different particles undergo longitudinal oscillations with different amplitudes, and with different timing relative to orbital revolution. Thus, an effect which is conspicuous in a fixed amplitude, fixed-phase simulation may characterize only a small fraction of particles in a real collider.

The typical electron-positron simulation does not suffer from this problem, because in every transformation cycle the mapping that models quantum noise (see below) adds a different random vector to the longitudinal coordinates of each different test particle.

Quantum noise is generally modelled by a transformation of the form

$$\begin{pmatrix} x \\ x' \end{pmatrix} \longrightarrow \begin{pmatrix} x + \delta_x \\ x' + \delta_{x'} \end{pmatrix} . \qquad (21)$$

(similarly for (y, y') and (τ, ε)), where δ_x and $\delta_{x'}$ are random variables. In the simulations described in Refs. 31 and 58-61, the δ's are selected from Gaussian distributions; in the simulations described in Ref. 57, the δ-pairs are selected from samples distributed smoothly along ellipses. The variances of the Gaussians, or the dimensions of the ellipses, are chosen so that the equilibrium test-beam distribution, after many transformation cycles, agrees with the distribution expected from unperturbed single-beam storage ring theory, when beam-collision transformations are ignored and the μ's do not vary.

Beam-beam encounters: The jolt that a test particle receives when it encounters an opposing bunched beam head-on is modelled by a transformation of the form

$$\Delta x = \Delta y = \Delta\tau = \Delta\varepsilon = 0$$

$$\Delta x' = \partial_x F(x + (\varepsilon/E)\,\eta_x^*,\ y + (\varepsilon/E)\,\eta_y^*)$$

$$\Delta y' = \partial_y F(x + (\varepsilon/E)\,\eta_x^*,\ y + (\varepsilon/E)\,\eta_y^*)\ . \tag{22}$$

($\Delta\varepsilon$ can be nonzero when bunched beams cross at a nonzero angle. See Ref. 22.) The function F is defined in terms of the density, ρ, of charge in a bunch of the opposing beam per differential of area in the plane perpendicular to the path of the test particle, at an interaction point, by

$$F(x,y) = \text{constant} + \frac{2e}{E}\int d\bar{x}d\bar{y}\ \rho(\bar{x},\bar{y})\log[(x-\bar{x})^2+(y-\bar{y})^2]^{1/2} \tag{23}$$

or, equivalently,

$$\nabla^2 F = \frac{4\pi e}{E}\,\rho. \tag{24}$$

In using this form, one assumes that the interactions involved in the collision are instantaneous, ultrarelativistic and entirely Coulombic.* The coefficients η_x^* and η_y^* are the horizontal and vertical dispersions at the interaction point.

Note that the gradient of F at the origin is directly related to the space-charge parameters: $\partial_y F(0,0) = -4\pi\,\xi_y/\beta_y^*$, and similarly for x.

One sometimes changes the η^*'s from cycle to cycle, employing the first type of variation described in connection with the μ's, in order further to model the effects of storage ring irregularities. F can also be modulated in other, more situation-specific ways. One example will be described in Subsection 3b; another is discussed in Ref. 105.

We note that when ρ describes a Gaussian bunch of N particles, each carrying charge e',

$$\rho = \left(\frac{Ne'}{2\pi\sigma_x^*\sigma_y^*}\right)\exp -\frac{1}{2}\left(\frac{x^2}{\sigma_x^{*2}}+\frac{y^2}{\sigma_y^{*2}}\right), \tag{25}$$

then the integral (23) can be simplified considerably[7]

*Equation (23) follows from Eq. (5) of Ref. 3. Note that there is a misprint in Eq. (5) of Ref. 3: the right-hand-side should be divided by ρ^2 (not the same ρ as in (23) above).

$$F(x,y) = F(0,0) + \frac{ee'N}{E} \int_0^\infty dt\ (\sigma_x^{*2} + t)^{-1/2}\ (\sigma_y^{*2} + t)^{-1/2}$$

$$\cdot \left\{ -1 + \exp - \frac{1}{2} \left[\frac{x^2}{\sigma_x^{*2}+t} + \frac{y^2}{\sigma_y^{*2}+t} \right] \right\}. \qquad (26)$$

(For an elementary proof, see Ref. 83.)

When an ultrarelativistic test particle crosses a continuous opposing beam at an angle α, F depends only on y and is given by[7]

$$F(y) = \frac{2\pi e}{E \tan(\alpha/2)} \int d\bar{y}\ \rho(\bar{y})\ |y-\bar{y}| \qquad (27)$$

or

$$\frac{\partial^2 F}{\partial y^2} = \frac{4\pi e}{E \tan(\alpha/2)} \cdot \rho(y)\ , \qquad (28)$$

where the density $\rho(y)$ is defined in this context as follows: For small dy, $\rho(y)dy$ is the total current in the opposing beam, between vertical levels y and y + dy.

I am aware of three ways in which programmers have computed the potential F:

Weak-strong method: F is given by an approximation to the result of substituting a Gaussian density ρ into (23) or (27), and is held fixed for the entire length of the simulation. Physically, this is equivalent to assuming that all the test particles are contained in one beam whose current is so low ("weak") that it leaves the opposing ("strong") beam completely unaffected. The other two ("strong-strong") methods described below do not involve this assumption. They are accordingly more realistic, but also more costly, because they require that many different F's--not just one--be computed during the course of a simulation.

Although unrealistic, weak-strong simulations do help us sort out the physical phenomena that contribute to observed beam-beam phenomena. In particular, they indicate the extent to which observed phenomena can be understood in terms of the nonlinear dynamics of individual particles, because a weak-strong simulation makes no provision for cooperative effects--those that are associated with the simultaneous adjustment of two beams to one another.

Thus it is in principle significant that considerable beam blowup has been observed in weak-strong simulations of PETRA[58-61] and of SPEAR.[57] In Chapter V, I shall argue that the blowup reported in Ref. 57 may, to a large extent, be an artifact of an approximation. As far as I know, this particular objection does not apply to the blowup reported in Refs. 58-61.

Strong-strong method, according to Reference 31: One divides the test particles into two equally populated model beams. The charge distribution in each beam is thus, literally, the sum of a small number of delta-functions. The beam-beam kick received by a test particle in a transformation cycle is computed from an approximation to (23) or (27), where the density function ρ is obtained by fitting a Gaussian (or something nearly Gaussian) to the delta-function sum that describes the opposing beam after the preceding cycle. One fits the true distribution with a smooth function so that a model beam with a small number of test particles can mimic a real beam that contains many closely packed real particles.

Strong-strong method, according to Ref. 63: As before, one divides the test particles into two equally populated model beams. One also groups the transformation cycles into supercycles, each composed of some constant number, S, of ordinary cycles. (In Ref. 63, S = 300.) In a supercycle, the phase space coordinates of the test particles in one beam are held fixed; the test particles in the other beam are subject to S conventional transformation cycles, with the beam-beam kicks computed from an approximation to (23) or (27), where the density ρ is obtained by fitting a Gaussian to the delta-function sum corresponding to the positions in the fixed beam. In the next supercycle, the roles are reversed. One fixes the phase space coordinates of the particles that had just before been varying; their coordinates are fixed at the values they had at the end of the cycle just completed. One subjects the coordinates in the other beam (previously fixed) to S conventional cycles, with ρ computed by fitting a Gaussian to the delta-function distribution of the particles that are now fixed. And so on.

The advantage of the second strong-strong method is that it requires less computer time: A beam-beam kick $\vec{\nabla}F$ is fit 2S times less often than with the first strong-strong method. I do not know whether or not there are important physical effects that might be well simulated by the first method and missed by the second.

In either method, charge distributions are fit to Gaussians primarily for convenience. Although this should be a good approximation when currents are low, one should be aware that it may have a number of drawbacks. Here are two possibilities: First, as mentioned in Chapter II, real beam distributions are not Gaussian at high current in real colliders; however, the extent to which the Gaussian approximation distorts simulation results on this account depends on the extent to which details in ρ are washed out by the integral transform (23) or (27) that defines F. Second, the use of a Gaussian artificially imposes an inversion symmetry on the function F; if there are important effects related to asymmetries, this procedure may wash some of them out.[64]

2. Specific electron-positron simulations

In this section, we discuss four recent simulations of electron-positron storage rings. In order of decreasing complexity, they are: A strong-strong simulation of LEP,[31] a strong-strong

simulation of CESR,[63] a weak-strong simulation of PETRA,[58-61]* and a weak-strong simulation of SPEAR.[57]**

These all merit our attention because each showed behavior similar to that seen in real storage rings. In particular, all four exhibited substantial beam blowup at large current.

To my mind, the LEP, PETRA, and SPEAR simulations are especially notable for their authors' attempts to identify causes underlying their observations. The authors both of the LEP and of the PETRA simulations were able to identify storage ring irregularities as major contributors to beam blowup. The author of the SPEAR simulation discovered a resonant effect ("resonance streaming") that appeared to be entirely responsible for the blowup evident in his results. As mentioned in the preceding section, we shall argue in Chapter V that the blowup seen in this SPEAR study may have been, to a large extent, an artifact of one of the approximations used. Nevertheless, we will devote some time to this work because, despite its problems, resonance streaming represents one of the more sensibly conceived attempts to deduce an intuitive explanation of colliding beam phenomena from first principles.

a. LEP[31]

This computer study is unique in two major respects:

First, the program was used to simulate an uncommonly wide range of operating conditions. During the course of this study, beam energy, beam currents, numbers of bunches per beam, β^*'s, η^*'s, storage ring tunes, beam-center separations, and storage ring irregularities were all varied in small steps.

Second, the output (after all transformation cycles are completed) was analyzed in an uncommonly large number of different ways, and the analysis documented in an unusually thorough manner. The circulated report[31] contains many graphs showing luminosity, beam dimensions, tuneshifts, and beam distributions as functions of the parameters listed in the preceding paragraph.

The model beams in this study behave like real e^+e^- beams in some familiar ways. For example: The beams blow up vertically, but not horizontally; luminosity grows linearly with current at large current (Fig. 18), as is the case, for example, at CESR; beam-beam effects intensify as energy decreases (i.e., as damping time increases--Fig. 19); beam distributions (when smoothed, but not fit to standardized ρ's) are visibly not Gaussian in their tails.

As one expects from Eq. (8), the tuneshift saturates when L becomes proportional to I. The largest such saturation value observed in this study--and, accordingly, this study's prediction for ξ_{ymax} at LEP--is between .03 and .035.

These results were all obtained under a common set of assumptions regarding the total storage ring tunes, the dispersions

*A strong-strong simulation of PETRA (using Ref. 63's formalism) is mentioned in Ref. 60, but only very briefly.

**An earlier, less conclusive simulation is described in Ref. 65.

and beta functions at intersection regions, the unperturbed bunch length, the way that the μ's vary due to storage-ring irregularities, etc. We may form an impression of the extent to which the beam behavior is sensitive to changes in these assumptions by considering Figs. 20, 21, and 22.

Figure 20 shows a graph of simulation luminosity versus the total unperturbed vertical tune that the storage ring would have if there were no irregularities. In this simulation, irregularities contribute additional, randomly chosen increments to the total tune that are, per interaction, equal to .0125 in root-mean-square. The horizontal tune (defined in the same way), the energy per e^+ or e^-, the number of bunches per model beam, and the tuneshifts that either beam would generate in the absence of the other beam (a measure of beam currents and dimensions) are indicated at the top of the figure. Luminosity at fixed current is evidently a very erratic function of storage ring tune. This is reminiscent of what real storage ring operators report, as discussed briefly in the preamble to Chapter II, Section 2. Other e^+e^- simulations show similar tune-dependence when large enough tune irregularities and/or dispersions are assumed. (The (randomly generated) η^*'s used in obtaining Fig. 20 were equal, in root-mean-square, to 5 cm (horizontal) and 5 mm (vertical).) The data shown in Fig. 19 (resp. Fig. 18) was obtained by setting all parameters (resp. all parameters except current) equal to those corresponding to the luminosity maximum in Fig. 20.

Figure 21 shows a graph of simulation luminosity as a function of the rms irregularity-related variation in μ_x (all other parameters taking values corresponding to maximum luminosity in Fig. 20). Simulation luminosity apparently increases by about 60% when $(\Delta\mu_x)_{rms}$ is reduced to zero from the value--.0125--actually expected to characterize real LEP. (In this chapter, $\Delta\mu$ refers to time-dependent tune variation, not to tuneshift.) Simulation luminosity also increases by a comparable amount when $(\Delta\mu_y)_{rms}$, or the (η_x^*)'s, or the (η_y^*)'s are reduced to zero from their nominal values. This suggests that systematic reductions in dispersions and irregularities might enhance the performance of real storage rings as well. We shall consider more support for this possibility when we discuss the PETRA simulation, below.

Figure 22 shows graphs of luminosity versus the value taken by the vertical beta function at interaction points, for five different assumptions (1.2, 2.4, 3.6, 4.8, 6.0 cm) concerning the unperturbed length ($\mathfrak{L}$) of a model beam bunch. (As before, all other parameters are given values that correspond to maximum luminosity in Fig. 20.) In each case, luminosity rises as β_y^* falls, until $\beta_y^* \approx \mathfrak{L}$, when luminosity drops to zero (the test particles are rapidly dispersed).

One's naive expectation is that luminosity should rise indefinitely as β_y^* shrinks, since in single-beam theory,[4] smaller β_y^* means stronger focusing, i.e., denser beams. Indeed, this is the traditional rationale for using "low-beta insertions" to enhance luminosity in storage rings.[66] In view of the widespread use of low beta sections, it is important to understand the instability threshold seen in this simulation at $\beta_y^* \approx \mathfrak{L}$. Unfortunately, there is no good quantitative theory of this cutoff.

The most obvious way to estimate this threshold does very poorly. One argues as follows: For a typical particle, the longitudinal distance ℓ by which it leads its bunch center satisfies $\ell \simeq (\bar{\ell}/2) \cos(\Omega t + \text{phase})$ (at least for times so short that damping and noise can be neglected), where Ω is the angular frequency of synchrotron oscillations. Thus the variable $\Delta\ell$ used to compute the variation in μ_y is typically comparable, in order of magnitude, to $(\Omega/Cf)\cdot\bar{\ell}/2$. (In this context, $(Cf)^{-1}$ is the time between beam-beam interactions.) Thus one is erroneously led to expect luminosity problems only when β_y^* is as small as $1/2\pi\ (\Omega/Cf)\ \bar{\ell}/2 \ll \ell_0$.

If Eq. (17) were taken into account in this simulation--it is not--then one might argue as follows:[67] The vertical beta function seen by the typical particle at mid-collision is given roughly by $\beta_y^* + (\bar{\ell}/4)^2\ \beta_y^{*-1}$, which decreases as β_y^* decreases, until a minimum at $\beta_y^* = \bar{\ell}/4$, after which it increases. This would crudely explain why the curves in Fig. 22 have maxima where they do, but it would not explain why the luminosity falls to zero so precipitously just below the maxima.

In electron-positron storage rings, $\beta_y^* \simeq \bar{\ell}$ typically means $\beta_y^* \simeq$ a few centimeters; in bunched proton-proton and proton-antiproton colliders, it typically means $\beta_y^* \simeq$ a few meters.[68]

A modified version[106] of this program was the simulation to which the recent improvement in PEP luminosity was in part due.

b. CESR[63]

The computer program used in this work differs from that used in the LEP study primarily in the strong-strong methodology employed (as explained in the preceding section), and in the use of μ's that do not vary during the course of the simulation.

One should bear in mind that the authors of this simulation approximate the horizontal component of the collision transformation (22) by an expression that is independent of the vertical coordinate.[69] A similar approximation is made in the SPEAR simulation that we describe later. We shall argue in Chapter V that this approximation can greatly exaggerate the beam blowup exhibited by the SPEAR simulation. I do not know whether the same thing is true in the case at hand.

The only results shown explicitly in the published report[63] are reproduced here in Figs. 23 and 24. In either case, one bunch per beam is assumed; the beam energy is equivalent to 5.5 GeV per electron.[69]

Figure 23 shows a graph of simulation luminosity versus simulation current. This is superimposed on a scatter of data points accumulated in the course of actual CESR operation; a quadratic curve extrapolated from small current is also shown. The agreement between the real and simulated data seems to be reasonable. The threshold labelled "maximum current" is defined[69] as the current at which one test particle (out of an initial sample of 200) is lost during the course of the simulation (typically equivalent in real time to three damping periods). In this simulation, as in the LEP study, a test

particle was considered "lost" when either its vertical or horizontal coordinate exceeded a certain pre-determined "aperture" value. As far as I am aware, there was no serious attempt to calculate maximum currents from the results of the LEP simulation. As explained in Section 1, one particle in two hundred per three damping times is actually much more rapid than the loss rate that one usually associates with maximum current; therefore, either maximum current does not depend sensitively on maximum tolerable loss rate (above some threshold), or this particular point of agreement between simulation and experiment is somehow fortuitous.

Figure 24 shows a contour plot of simulation luminosity as a function of horizontal and vertical unperturbed storage ring tunes. The beam currents are such that the vertical tuneshift due to either beam, if unperturbed, would be $\xi_y = .08$. The luminosities displayed in this plot were, for convenience, computed with dispersions set equal to zero, and with the horizontal components of the collision transformations (22) also summarily set equal to zero. Presumably, this is why luminosity seems to be a much less erratic function of tune per interaction in this figure than it is in Fig. 20. (Tune per interaction is equal to tune divided by two in the present case, and divided by eight in the case of Fig. 20.)

c. PETRA [58-61]

Only one property of this simulation's output has been graphed or tabulated in publicly circulated reports: the rms height of the weak beam, i.e., the test particles' root-mean-square vertical displacement from the center of the strong beam. Published reports have so far not discussed beam distributions, horizontal beam widths, loss rates, etc.

Beam height in this simulation has been computed for a large number of closely-spaced storage ring tunes, as well as for small numbers of energies, strong-beam tuneshifts, numbers of bunches (B) per beam, and patterns of irregularities and dispersions. In an early phase of this project,[58] beam height was also computed as a function of time, for different values of various parameters, among them the decay time of the voltage initially applied to separate the beams.

Some features seen in the results of this simulation have been at least partially confirmed by real measurements made on PETRA itself.[60,62] We shall discuss these measurements shortly.

The graphs reproduced in Figs. 25, 26, and 27 are typical of the data generated by this simulation. From Figs. 25 and 26 one sees that simulation blowup increases with decreasing energy and with increasing (strong beam) current, just as real blowup does. The data shown in these figures were computed using nonzero $\Delta\mu$'s and η^*'s [0(.01) and 0(1 cm), respectively]. The importance of such irregularities and dispersions is strikingly demonstrated in Fig. 27, the left half of which shows beam heights calculated with $\Delta\mu$'s and η^*'s set equal to zero, while the right half shows beam heights computed with $\Delta\mu$'s and η^*'s that have orders of magnitude similar to

those just quoted. Notice the erratic nature of the curves shown in the right half of Fig. 27, and in Figs. 25 and 26. This is similar to the behavior seen in the graph of luminosity vs. tune generated by the LEP simulation.

It should be stressed that the blowup associated in Fig. 27 with the presence of nonzero dispersion is much larger than can be accounted for by the usual energy contribution, $(\eta_y^* \sigma_E/E)^2$, to the mean square unperturbed radiative beam height. In the case at hand, $(\eta_y^* \sigma_E/E)^2$ is about[7] $(1\text{cm}\times 10^{-3})^2 = 10^{-6}$ cm^2, while the unperturbed mean square beam height is about[7] 10^{-5} cm^2.

The right side of Fig. 27, corresponding to realistic irregularities, has this intriguing feature: There are "magic tunes" at which the beam does not blow up (at some it even shrinks). At present, it is not known how general this phenomenon might be. As far as I am aware, magic tunes have not been observed in the outputs of other simulations. It is conceivable that other effects not included in this PETRA model might eradicate magic tunes, or might shift them in an erratic and essentially unpredictable fashion.

In an attempt to reproduce in vivo the operating conditions that correspond most closely to this simulation, a number of real weak-strong experiments--i.e., experiments in which one (weak) beam has very low current--have been performed on PETRA.[62]* According to the published report, these experiments confirm that beam blowup decreases when dispersions are reduced. It is also claimed that the dependence of weak beam height on storage ring tune as measured in these experiments has a number of features in common with the tune dependence as computed by the simulation program.

Here are sample experimental results, obtained with two bunches per beam, a beam energy of 7 GeV per particle, and a strong-beam-induced tuneshift of $\xi_y = .015$: with the tunes held fixed at $\nu_x = 25.2$ and $\nu_y = 23.1$, weak beam vertical blowup was reduced from 5× to 2× when the mean horizontal η^* was reduced by special magnetic correction from 15 mm to 3 mm. With horizontal tune held fixed at 25.2, and the dispersions left unreduced, weak beam vertical blowup was reduced in a similar way when the vertical storage ring tune was shifted from 23.3 to 23.1.

It is claimed[62] that both these reductions were predicted by the simulation. Unfortunately, as far as I can tell, the documentation in Refs. 58-61 is not complete enough to permit one to evaluate the quantitative agreement between experiment and simulation for oneself. In view of the simulation data that is available, these measured effects seem rather large for such a relatively small tuneshift.

Similar results have been obtained in strong-strong measurements performed on PETRA. An example[60] of such data, in this case showing enhanced luminosity instead of decreased blowup, is reproduced in Fig. 28.

*An early effort to coordinate weak-strong simulation (of DORIS, in this case) with in vivo weak-strong measurement is discussed in Ref. 22.

d. SPEAR [57]

This is by far the simplest of the four e^+e^- simulations that we discuss: No provision is made for storage ring irregularities; and longitudinal coordinates--and therefore dispersions--are completely ignored.

In order further to simplify computation, this simulation employed an approximation in which both strong and weak beams were idealized as flat, and vertical focussing at interaction points was idealized as infinitely strong. I.e., all computation was done in the following limit:[70] $\sigma_y^* \to 0$, $\beta_y^* \to 0$, $y \to 0$, y/σ_y^* fixed, β_y^*/σ_y^* fixed. In real SPEAR,[7] $\sigma_y^*/\sigma_x^* \simeq 1/30$, and $\beta_y^*/\beta_x^* \simeq 1/10$.

In this limit, the last two equations in (22) become

$$\lim \Delta x' = \left(\frac{Ne^2}{E\sigma_x^*}\right) \underline{x} \int_0^\infty du\ (u+1)^{-3/2}\ u^{-1/2}\ \exp\left[\frac{-\underline{x}^2}{2(1+u)}\right] ,$$

$$\lim \Delta y' = \left(\frac{Ne^2}{E\sigma_x^*}\right) \underline{y} \left\{\int du (u+1)^{-3/2}\ \exp\left[\frac{-\underline{y}^2}{2(1+u)}\right]\right\} \exp(-x^2/2), \tag{29}$$

where we have used the Gaussian formula (26) for F (with $e = -e'$), neglected dispersion, and introduced the reduced variables $\underline{x} \equiv x/\sigma_x^*$ and $\underline{y} \equiv y/\sigma_y^*$. The two expressions on the right-hand side of (29) are especially convenient for numerical approximation because--unlike the full expression (26) for F--each can be factored into a product of single-variable functions. In particular, $\lim \Delta x'$ is a function of $\underline{x}$ alone, and $\lim \Delta y'$ is the product of $\exp(-\underline{x}^2/2)$ and a function of $\underline{y}$ alone.

Despite omissions and simplifications, the data generated by this simulation--like that generated by other simulations--shows substantial (weak) beam vertical blowup. This is illustrated in Fig. 29, which shows root-mean-square values of weak beam $\underline{y}$'s as computed by this simulation program for various model operating conditions.

In order to determine why the model beam was blown up, an effort was made to examine the shapes of phase-space paths that individual test particles follow during the course of this simulation. Because phase space in this case is four dimensional, and therefore not readily visualized, the decision was made to view only the projections of these paths on a plane (to be called the normalized amplitude plane) whose axes correspond to the variables a_x and a_y, where

$$a_y \equiv \frac{1}{\sigma_y^*} (y^2 + \beta_y^{*2} y'^2)^{1/2} = (\underline{y}^2 + y'^2(\beta_y^*/\sigma_y^*)^2)^{1/2}, \tag{30}$$

and similarly for a_x. (For reasons that will be explained in Chapter IV, Subsection 2a, a_x and a_y are often used as approximations to $\sqrt{2}\,x$ and $\sqrt{2}\,y$, respectively.) This is a rather novel step, because authors of e^+e^- simulations are usually concerned only with gross statistical properties of their model beams.

This kind of analysis has revealed an interesting pattern: During a run of this simulation, the normalized amplitudes of most test particles trace out erratic paths that, for the most part, remain within or close to a quarter-circle of radius $\sqrt{2}$, centered at the origin. (When there is no strong beam, all the test particles behave this way.) However, there is a small number of test particles, each of whose erratic motion is interrupted by a more ordered segment in which it travels a long distance in a short time. The nature of this ordered interlude is the same for all these exceptional particles: rapid streaming up a nearly vertical curve (the same curve for all such particles), on which is superposed a rapid and noisy oscillation transverse to the curve. The location of this curve, as well as the number of test particles that are attracted to it, depends on simulation parameters such as strong beam tuneshift, etc. Beam blowup in this simulation is due to these streaming particles--the large a_y's to which they stream dominate the average that defines the beam height.[70] The vertical orientation of the streaming segment seems to explain in a natural way why no horizontal blowup is observed.

An example of a trajectory with such a streaming segment is illustrated in Fig. 30, reproduced from Ref. 5. The strong beam tuneshifts in this case were $\xi_x = \xi_y = .06$, and the beam energy was 2.2 GeV per electron. The orbit shown represents about 75,000 revolutions of SPEAR (≈three damping times). One does not see any oscillations transverse to the streaming curve (close to the line $a_x \simeq 2.8$) because the oscillations have been averaged away.[70] (The path shown in this plot is interpolated between points that are obtained by averaging the normalized amplitudes that correspond to the results of 250 successive transformation cycles, i.e., 125 successive revolutions.) The time spent travelling along $a_x \simeq 2.8$ is about 8,000 revolutions.[70] 5% of the test particles in this run behaved as shown in this figure.

Tennyson observed[5] that the curve $a_x \simeq 2.8$ is close to the resonance $3\nu_x + \nu_y = 21$ (as we shall explain in Chapter IV, a resonance can be a set of points in amplitude space, as well as a relation among unperturbed storage ring tunes), and also provided a succinct mathematical explanation[71] of how such a resonance can channel the motion of test particles, under appropriate conditions. We shall present our own, somewhat more detailed form of this explanation in Chapter IV.

The possibility that resonance streaming might account for all blowup in all e^+e^- storage rings is an attractive one. However, we shall argue in Chapter V that the flat beam/infinite-focusing approximation can cause the distance covered by a particle streaming in a resonance to be exaggerated. We shall argue that when this approximation is corrected, the blowup due to streaming in the

simulation of Ref. 57 is substantially reduced, if not eliminated altogether.

3. Proton-antiproton and proton-proton simulations

In this section we discuss two weak-strong simulations of the Fermilab Tevatron $\bar{p}p$ collider,[72-74] and a weak-strong simulation of ISABELLE (operated as a bunched-beam collider).[75]

The common objective of these exploratory studies was information concerning strong instabilities to which pp and/or $\bar{p}p$ storage might be prone near their operating limits. In each case, special attention was paid to stochastic instabilities, which we already mentioned briefly near the end of Chapter II, and which currently constitute one of the main themes in the modern theory of strongly nonlinear dynamical systems.[76] Designers of proton/antiproton storage rings have traditionally been particularly fearful of stochastic effects,[15] because such effects are primarily long-time phenomena, and proton/antiproton storage rings lack obvious rapid relaxation mechanisms (such as the radiative processes that dominate e^+e^- machines) that might cut off such phenomena before they become fully developed. As indicated in Chapter II, Sec. 4, inconclusive attempts were made to interpret the results of nonstandard ISR studies in terms of stochastic behavior. The flexibility and precision of the digital computer make it a much more appropriate setting for pursuing such interpretations.

Note that neither storage ring irregularities, nor dispersions (nor, of course, radiative damping and noise) are included in any of the simulations to be discussed below. Realism has been traded for computing time, in order to maximize the number of storage ring revolutions simulated.

a. Tevatron [72-74]

The main difference between the two Tevatron simulations discussed here is this: One included no provision for longitudinal effects; the other provided for them in a limited way, by including in each μ a term that varies sinusoidally with time, at the synchrotron frequency. For simplicity's sake (as explained in Sec. 1) the amplitude and phase of this oscillation were assumed to be the same for all test particles. It was also assumed that a test particle encounters the strong beam only once per revolution (i.e., $C=1$).

(It should be mentioned that these two simulations are part of a more extensive series that also includes models in which the horizontal coordinate is neglected. (See for example Refs. 77 and 80.) The two models that we discuss here have generated the most interesting results.)

In each case, a round strong beam ($\sigma^*_x=\sigma^*_y\equiv\sigma^*$) and x-y symmetric beta-functions ($\beta^*_x=\beta^*_y\equiv\beta^*$) were used. (A simulation using an elliptical strong beam is in progress at this writing.[79]) This implies equal horizontal and vertical tuneshifts ($\xi_x=\xi_y\equiv\xi$), according

to Eq. (4). Note that for round beams, the derivative of (26) can be expressed directly in terms of elementary functions

$$\frac{\partial F}{\partial x} = - \frac{2Ne^2x}{E(x^2+y^2)} \cdot (1-e^{-(1/2)(x^2+y^2)/\sigma^{*2}}), \qquad (31)$$

and similarly for $\partial F/\partial y$ (we have set $e' = -e$, as is appropriate for $\bar{p}p$ collisions).

The variable-tune simulation was carried out for only one value (.01) of ξ, and only one pair of storage ring tune time-averages, but for several values of the amplitudes and phases of the oscillatory terms in μ_x and μ_y. In all cases, the oscillatory terms in μ_x and μ_y were equal in amplitude (ranging from .001 to .01), and differed in phase only by zero or π. The tuneshift was deliberately chosen high so that one might get a feel for the worst possible beam-beam effects that might be encountered in the Tevatron. All modulation amplitudes used--even the low ones--were also deliberately chosen to be larger than those actually expected for the real collider. The other simulation (no tune modulation) was carried out for several values (.005, .01, and .02) of ξ, and for many values of the storage ring tunes.

These simulations are especially noteworthy in two respects:

First, each test particle was subject to an unusually large number of transformation cycles--6×10^6 in the case of the modulated-tune simulation, and, on separate occasions, 10^5 and 6×10^7 in the case of the constant-tune simulation. This is equivalent, in real time, to two minutes, two seconds, and twenty minutes, respectively.

Second, the action of each transformation on the transverse phase-space coordinates of each test particle was calculated to extremely high accuracy--twenty-eight decimal places (double precision) in x and y, measured in millimeters, and in x' and y', measured in milliradians. (The natural scales, set by the strong beam, are $\sigma^* = .0816$ mm, and $\sigma^*/\beta^* = .0416$ mrad.) This was done to facilitate the identification of stochastic effects, which are typically characterized by orbital behavior that depends very sensitively on initial conditions.

(In hindsight, this level of accuracy may have been far more exacting than was actually necessary, because, as explained below, much of the observed stochastic behavior turned out to be apparent on a rather coarse scale. It has also been pointed out[80] that it is unphysical to pursue such classical calculations to an accuracy that exceeds five decimal places, because of quantum-mechanical uncertainty.)

Here are the main conclusions that the authors have drawn from the results of these simulations:

1. The emittances of the weak beam do not grow--or grow only very slowly--when tune modulation is absent. When tune modulation is included (at least for 180° phase difference between the oscillations

in μ_x and μ_y, and oscillation amplitudes greater than about .003) the emittance can grow very noticeably, doubling in as little real time as a fraction of a minute (a few hundred thousand transformation cycles).

Inasmuch as storage ring irregularity as modelled in e^+e^- simulations is just a very rapid tune modulation, it is tempting to conjecture that the ability of slow tune modulation to trigger time-dependent emittance growth in this $\bar{p}p$ simulation is related to the ability of storage ring irregularity to enhance the blowup seen in e^+e^- simulations. However, there are at present no concrete arguments to support such a contention.

2. When conditions are such that emittances grow noticeably, every test particle that, by the end of the simulation, reaches x,y, $x'\beta^*$ or $y'\beta^*$ substantially larger than σ^* does so by travelling along a "chaotic (or stochastic) trajectory."

The operational definition of "chaotic" employed by the authors of these simulations is as follows. Subject a quadruple of initial phase space coordinates to 100,000 transformation cycles, and then apply the inverses of these same 100,000 transformations, in reverse order. Evaluate the distance, in $(x, y, x'\beta^*, y'\beta^*)$-space, between the initial coordinates and the coordinates that result from the forward-backward process just described. This distance provides a measure of the degree to which the orbit beginning at the initial coordinates in question is sensitive to very small perturbations--the small perturbations are provided by computing inaccuracies. If the distance is greater than 10^{-10}mm, the trajectory that begins with the initial coordinates in question is defined to be chaotic.

The cutoff 10^{-10} is actually rather arbitrary. According to Refs. 72 and 74, for most initial conditions, the result of this "reversibility" test turned out, in the simulations at hand, to be either close to 10^{-20}, or close to unity.

It may be difficult for the reader to form a mental picture of stochasticity from this definition alone, especially since phase space in this case is four-dimensional. Examples of stochastic orbits that are much more easily visualized will be considered in the next subsection.

3. Properties of chaotic orbits are correlated with resonance conditions in several ways. For example:

-- Suitably defined "effective tunes" of divergent orbits (i.e., those orbits that reach large x,y, $x'\beta^*$ or $y'\beta^*$ when substantial emittance growth is observed) lie close to tunes that satisfy low order (≤ 6) resonance conditions. "Effective tunes" in this context are defined[72] as $(2\pi)^{-1}$ times the changes, per transformation cycle, in the angles $\tan^{-1}(x'\beta^*/x)$ and $\tan^{-1}(y'\beta^*/y)$, averaged over the first one thousand transformation cycles. (One thousand cycles was the period of synchrotron oscillation in this model.)

-- When tune modulation is omitted, a high percentage of weak-beam orbits are chaotic (none are divergent when modulation is omitted) only when the unperturbed storage ring tunes lie near tunes that satisfy two low-order resonance conditions simultaneously. Specifically, in the simulation of Ref. 72, more than ten percent of

the test particles followed chaotic orbits when the tunes lay near the intersections of fourth-and sixth-order resonances; one or more out of the one hundred test particles followed chaotic orbits when the tunes lay near intersections of resonances of order less than or equal to ten. No chaotic orbits were observed otherwise. These percentages were observed to apply to all three tuneshift values (ξ = .005, .01, .02) studied. According to modern nonlinear theory, such a correlation between stochasticity and the overlap of resonances is a common occurrence. For details, see Ref. 76.

b. ISABELLE [75]

This highly idealized model was formulated in order to simulate, in a schematic way, the operation of ISABELLE as a bunched-beam collider. As usual when protons are involved, the model does not include radiative noise and damping; nor does it include tune variation or dispersions; nor does it include the horizontal phase-space coordinates x and x'. This model has only one nontrivial feature: The function F that describes the kicks due to the strong beam is multiplied by a time-dependent factor of the general form $a+b\cos 2\Omega t$, where Ω is the angular frequency of synchrotron oscillations, and the coefficients a and b take the same pair of values for all test particles in the weak beam. The physical motivation for this particular mathematical structure is explained in Ref. 75. The advantage of bunching proton beams that cross at a nonzero angle (11.88 mrad in the case of ISABELLE[7]) is discussed in Ref. 81.

We introduce this model here in order to supplement the discussion of stochasticity presented in the preceding subsection. Inasmuch as its phase space is only two-dimensional, this simulation provides a much more accessible picture of chaotic orbits than one could have derived from the four-dimensional Tevatron studies.

Some representative output from this ISABELLE simulation is reproduced in Fig. 31. The different pictures correspond to different values, $\overline{\Delta\nu}_y$, of the average of the tuneshift over one modulation period (one modulation period here corresponds to 200 beam-beam encounters). Note that in each case the modulation is quite severe, because the parameters a and b have been set equal. The horizontal axis in each picture represents the vertical coordinate y; the vertical axis (labelled "V/ω_0") represents y' times β_y^*. Each picture is an overlay of from 1000 to 5000 "snapshots" of the phase-space locations of precisely ten test particles; the snapshots are taken once every modulation period, in order to ensure that all points in each picture have been recorded under identical conditions.

In the overlays corresponding to $\overline{\Delta\nu}_y$ = .0095 and .0102, particle orbits lie on simple closed curves. Curves that do not encircle the origin indicate the general outlines of the resonance regions to be discussed at length in the next chapter. When $\overline{\Delta\nu}_y$ = .0115, the overlay contains fuzzy structure that cannot be resolved into closed curves. This structure, which appears to fill a two-dimensional

piece of phase space, is precisely three chaotic orbits. When $\overline{\Delta\nu}_y$ = .0127, the space-filling structure comprises seven chaotic orbits; when $\overline{\Delta\nu}_y$ = .0318, only two orbits are nonchaotic for $|y|$ and $|y'\beta^*_y|$ less than 2. The sixth picture in Fig. 31 shows the $\overline{\Delta\nu}_y$ = .0318 phase plane on a larger scale, revealing that gross chaotic structure can extend to very large amplitude.

When all of phase space, with the exception of small isolated islands, is filled with chaotic structure, one says that a "stochastic transition" has occurred. According to the conventional lore,[76] a colliding beam model that has passed a stochastic transition is unstable in the following sense: The area, in phase space, that lies outside the small isolated islands is dense with points whose distances from the origin can be made arbitrarily large by the application of appropriate numbers of transformation cycles. I.e., in time, most of the beam strikes whatever aperture might be set up to contain it, even without external sources of noise.

For weak-strong beam-beam models that omit noise, damping, modulation of any kind, and horizontal phase space, the stochastic transition takes place at $\xi_y \lesssim .25$; when horizontal phase space is restored, the limit can be as low as $\lesssim .125$, depending on the beta functions, and on the transverse dimensions of the strong beam.[82] In either case, the limit is much higher than any tuneshift presently accessible at real storage rings, either e^+e^-, or $\bar{p}p$, or pp.

Figure 31 suggests that the stochasticity limit can be significantly reduced when some parameter is externally modulated. However, it is not clear that even this reduced limit has a meaningful impact on real storage ring behavior. According to Ref. 75, the threshold for a stochastic transition is significantly reduced only when modulation (generally proportional to the amplitude of synchrotron oscillation) is strong; and in a real beam, only a small fraction of the particles actually undergo synchrotron oscillation with a very large amplitude.

IV. TOPICS IN THEORY--BACKGROUND

1. Overview

Although more than a few published papers on beam-beam phenomena have been devoted to analytical theory (as opposed to computer simulation),* I have chosen--for reasons to be explained below--only two theoretical ideas for detailed discussion in the present report. These ideas will be worked out in the next chapter. In the present chapter--specifically, in Sections 2 and 3--we discuss some

*A listing of theoretical papers, pre-1980, that do not discuss stochasticity, can be found in Ref. 84. Papers--especially ones written in the Soviet Union--on theories of chaotic phenomena in storage rings are listed in Ref. 5. Theoretical studies can also be found among the articles contributed to Refs. I and II, and cited in Ref. 35. Some additional theoretical papers will be cited later in this section.

mathematical prerequisites. Before we proceed, however, let us briefly survey the theoretical literature as a whole.

Most published work on colliding beam theory falls under one of two headings.** They are: efforts to use elementary properties of resonant or stochastic instabilities in order to deduce the maximum currents storable in colliding beam rings; and efforts to use elementary properties of resonant and collective effects in order to establish that above some minimum current, colliding beams are not stable unless they are blown up beyond their unperturbed sizes. (The first category is by far the more popular.)

We describe these lines of research below; however, we shall not discuss them at great length, for the following two reasons: First, the goals of these kinds of work are very limited, and--at least in part--are removed from questions of immediate observational importance. Efforts of the first kind might in principle tell us about the currents beyond which colliding beams cannot be stored, but they have not yet shown how to predict the behavior of colliding beams at normal operating currents, below the maxima. Efforts of the second kind might in principle tell us if beams necessarily blow up, but they have not yet shown how quantitatively to evaluate the size of the blowup. Second, even within these limitations, neither direction has produced a clear success. Research of the first kind (which is always conducted entirely in the weak/strong approximation), where it has yielded specific results at all, has generally overestimated maximum colliding currents; research of the second kind has so far not yielded concrete predictions.

Here are capsule sketches of these theoretical mainstreams:

Maximum current theories: Authors who attempt to attribute storage limitations to resonant instability calculate maximum current by substituting tuneshifts plus unperturbed storage ring tunes into resonant conditions of the form (13). As mentioned in Chapter II, in the course of our discussion of rules of thumb for ξ_{ymax}, this kind of procedure in its simplest form was introduced--and discredited--about twenty years ago. Later variations, involving coherent oscillation,[84] have also been unsuccessful. (It has recently been suggested, however, that such calculations be reconsidered.[85])

Authors who attempt to attribute storage limitations to stochastic transitions calculate maximum currents by applying a "resonance overlap" criterion[76] that we shall describe briefly in the next section. As we have already mentioned, stochasticity calculations can be made to yield maximum tuneshifts comparable to those observed in real storage rings, but only at the expense of artificial assumptions regarding the strength of longitudinal oscillations.

**There are exceptions, of course. The theory of overlap-knockout resonance[47] is an example.

This general approach to colliding-beam physics suggests a number of possibilities that, as yet, have either not been considered, or have been considered only superficially:

It is possible that even though these ideas incorrectly predict the precise current at which beam is lost too rapidly for practical storage, some such mechanism might be able to account for the rate at which such rapid loss takes place. (Reference 86 proposes a calculation of the loss rate due to a resonant instability; however, in this calculation the function F of Eq. (22) is approximated by a low-order polynomial in x and y, which, as we shall see in the next section, can badly distort particle behavior at large x and y, where losses actually take place.)

It is also possible that the maximum tuneshifts observed in real storage rings have more to do with the approach to an instability, than with the instability itself. Stochastic beam loss below the stochastic transition, known as "Arnol'd diffusion,"[76] has been investigated in the abstract by various authors, but as yet there have been no phenomenological calculations that can be meaningfully compared with observation.

Finally, it is also possible that radiative noise somehow lowers the threshold for the stochastic transition. However, the interplay between chaotic and radiative effects has received very little attention. (Kheifets[35,87] and Ruggiero[88] have proposed similar approximate mathematical formalisms in which these two kinds of effects are both incorporated naturally; however, neither of these formalisms has predictive power, since each involves an unknown parameter h whose dependence on energy and current can only be determined from experimental data. An earlier version of such a formalism, proposed by Hereward and by LeDuff, is described in Section IV of Ref. 16. For a related attempt (unsuccessful) to model an SPS beam-beam effect as the result of an effectively random nonlinear process, see Ref. 37. In this connection, see also Ref. 113.)

Blow-up onset theories: Several authors[89,90] have attempted to explain the tendency of colliding beams to blow up as a cooperative phenomenon involving simultaneous unstable oscillations of the charge distributions of both beams. (This type of analysis employs mathematical techniques from Plasma Physics.) In view of computer simulations in which weak beams blow up considerably even when strong beams do not vary at all, it is likely that cooperative mechanisms can at best account only for a fraction of the blowup observed in real storage rings.

It has been suggested[91] that the nonlinear resonances $2\mu_x - 2\mu_y$ = integer play an important role in the onset of vertical beam growth (at least in weak/strong systems) because they can facilitate transfer of particles from large horizontal amplitude to large vertical amplitude. This is supported, in the absence of dispersions and irregularities, by data from the PETRA simulation of Refs. 58-61, as one can easily see in the left half of Fig. 27; however, when dispersions and irregularities are present, as in the right half of Fig. 27, the situation is not so clear. The role of this resonance

may be exaggerated by the analysis of Ref. 91, which employs the same polynomial approximation to F as is employed in the loss rate calculations of Ref. 86, mentioned above.

The calculations that will be highlighted in Chapter V lie outside these mainstreams, but in my view represent the most complete attempts so far to confront issues of immediate phenomenological significance in colliding beam physics on the basis of first principles. They are: a calculation of collision-related beam loss during routine operation of the ISR; and a calculationn of beam blowup at SPEAR, by application of a theoretical correction to the results of the computer simulation described in Sec. 2d of the preceding chapter.

Let us now proceed to the mathematical preliminaries.

2. Resonant behavior in weak-strong systems--basics

Each of the two calculations in Chapter V will refer to weak-beam test particles that circulate through a storage ring having the following idealized characteristics: The distance, ℓ_0/C, between adjacent interaction regions is also a repeat-period of the storage ring in the absence of the strong beam; longitudinal effects, including dispersion, are not present; the only nonlinearities are those in the strong-beam kicks, idealized as in Eq. (22). In one of the calculations (the ISR model) the storage ring will be idealized further--radiative damping and noise will be absent. In the present section we derive some basic properties of such an idealized noiseless, undamped, weak/strong system. In the next section we shall explore some ways in which these properties are modified when radiative damping is taken into account.

a. Action and angle variables

In what follows, we shall find it more convenient to represent the motion of a test particle by the action (or amplitude) and angle variables $\{I_x, I_y, \theta_x, \theta_y\}$ of unperturbed betatron oscillations, than by its Cartesian phase-space coordinates $\{x,y, x', y'\}$. The action (or amplitude) and angle variables are defined in terms of the Cartesian variables by

$$x \equiv \sqrt{2I_x\beta_x(s)}\ \cos(\Phi_x(s) + \theta_x)$$

$$\beta_x(s)x' - (1/2)\beta_x'(s)\, x \equiv -\sqrt{2I_x\beta_x(s)}\ \sin(\Phi_x(s) + \theta_x), \tag{32}$$

and similarly for y. The azimuthal parameter s is, as usual, equal to the product of time and the speed of light. The beta functions β_x and β_y are periodic in s, with period ℓ_0/C. The periodic functions Φ_x are defined by

$$\Phi_x(s) \equiv \int_0^s \frac{ds'}{\beta_x(s')} - \frac{2\pi C\mu_x}{\ell_0} \cdot s \ , \tag{33}$$

and similarly for y. (For simplicity of normalization, we assume in what follows that s = 0 is an interaction point.)* Between beam-beam kicks, the I's and Φ's satisfy

$$\frac{dI_x}{ds} = 0$$

$$\frac{d\theta_x}{ds} = \frac{C}{\ell_0}(2\pi\mu_x) \equiv \omega_x \ , \tag{34}$$

and similarly for y. This is well-known.[92]

In certain contexts, the quantities x and $\sqrt{2I_x\beta_x}$ (or y and $\sqrt{2I_y\beta_y}$), or x and $\sqrt{I_x\beta_x}$ (or y and $\sqrt{I_y\beta_y}$), are often used interchangeably, even though, according to (32), they are not identically equal. The reason is that the cosine in (32) generally oscillates rapidly, so that x quickly swings between $\pm\sqrt{2I_x\beta_x}$, and the average (at a fixed (mod ℓ_0/C) value of s) of x^2 over times short compared to the scale on which I_x varies, is $I_x\beta_x$. Thus, for example, if the half-width of a beam pipe is W, it is common to conclude that a test particle is effectively lost as soon as $\sqrt{2I_x\beta_x} \geq W$; similarly, it is common to equate the average of $\beta_y I_y$ over the test particles in the weak beam with the beam's mean-square half-height. We shall adopt these conventions in what follows, because it is often easier to derive the behavior of the I's (by averaging--when one can--over rapidly varying angles) than it is to derive the behavior of x and y directly.

b. The colliding-beam Hamiltonian; resonant amplitudes

Equations (22) and (32) imply that when a test particle encounters the k^{th} interaction region ($1 \leq k \leq C$), its actions and angles change according to

$$\begin{aligned} \Delta I_x &= -\partial F^k/\partial\theta_x + O((F^k)^2) \\ \Delta\theta_x &= +\partial F^k/\partial I_x + O((F^k)^2), \end{aligned} \tag{35}$$

The normalized amplitudes a_x and a_y defined in Eq. (30) are related to I_x and I_y as follows: $a_x = 2I_x\beta^*_x/\sigma^{*2}_x$, $a_y = 2I_y\beta^*_y/\sigma^{*2}_y$. The σ's here are the dimensions of the strong beam.

and similarly for y. $F^k(I_x, I_y, \theta_x, \theta_y)$ is obtained by solving Eq. (32) for x and y, and substituting the results into $F^k(x,y)$ (F^k is the integral defined in (23), evaluated using the distribution of the strong beam as seen at region k). For convenience, we shall--as is customary--neglect the terms of quadratic and higher order in the right-hand-side of Eq. (35).* Note, however, that we shall not make additional truncations of this kind in cumulative effects of many beam-beam encounters. More precisely: In what follows, we employ approximate dynamical equations that omit terms of $O(F^2)$ and higher, but we do not systematically omit higher-order terms in solutions to these approximate equations.

This approximation permits us to combine the two parts, (34) and (35), of the equations of motion of the I's and θ's in the following simple way

$$\frac{dI_x}{ds} = -\frac{\partial H}{\partial \theta_x}$$

$$\frac{d\theta_x}{ds} = +\frac{\partial H}{\partial I_x} \qquad (36)$$

(and similarly for y), where the Hamiltonian H is defined by

$$H = \omega_x I_x + \omega_y I_y + \sum_{k=1}^{C} F^k(I_x, I_y, \theta_x, \theta_y) \sum_{n=-\infty}^{+\infty} \delta(s - n\ell_0 - k(\ell_0/C)) \; . \qquad (37)$$

(By allowing H to depend explicitly on s (beyond the dependence in the δ-functions), one can use the same type of system of equations to describe models in which external parameters (such as tunes) vary with time.)

Because the θ's are periodic variables, it is appropriate to expand H in a Fourier series

$$H(\vec{I},\vec{\theta}) = \vec{\omega}\cdot\vec{I} + \frac{1}{\ell_0} \sum_{n_x, n_y, n=-\infty}^{+\infty} F_{\vec{n},n}(\vec{I}) \cos(\vec{n}\cdot\vec{\theta} + n(\tfrac{2\pi}{\ell_0})s + \delta_{\vec{n},n}(I)). \qquad (38)$$

*Let us estimate, for example, the remainder $\Delta I_y + \partial F^k/\partial\theta_y$. To begin, we quote estimates, given, either explicitly or implicitly, in Problem 4 of Ref. 3: For a typical test particle at an interaction point (where $\beta_y' = 0$), $\Delta y'/y \sim -4\pi\xi_y^k/\beta_y^*$, where $\Delta y'$ is defined by (22) and ξ_y^k is the vertical tuneshift at region k (similarly for x). Thus, $\Delta I_y/I_y = (1/\sqrt{2})\,[\sqrt{1+(1-4\pi\xi_y^k)^2} - \sqrt{2}]$, to be compared with the $O(F^k)$ expression, $-2\pi\xi_y^k$. For $\xi_y^k = .06$ (SPEAR's maximum), the latter exceeds the former by about 25%.

The vectors $\vec{I}$, $\vec{\theta}$, $\vec{\omega}$, and $\vec{n}$ represent the pairs (I_x, I_y), (θ_x, θ_y), (ω_x, ω_y), and (n_x, n_y). The numbers n_x, n_y and n are all integers. The coefficients $F^x_{\vec{n},n}{}^y$ and phases $\delta^x_{\vec{n},n}{}^y$ are defined by

$$F_{\vec{n},n}(\vec{I}) \equiv \left| \sum_{k=1}^{C} e^{-2\pi ikn/C} \oint \frac{d\theta_x}{2\pi} \oint \frac{d\theta_y}{2\pi} F^k(\vec{I},\vec{\theta})\, e^{-i\vec{n}\cdot\vec{\theta}} \right|$$

$$\delta_{\vec{n},n}(\vec{I}) \equiv \mathrm{Arg}\left\{ \sum_{k=1}^{C} e^{-2\pi ikn/C} \oint \frac{d\theta_x}{2\pi} \oint \frac{d\theta_y}{2\pi} F^k(\vec{I},\vec{\theta})\, e^{-\vec{n}\cdot\vec{\theta}} \right\}. \tag{39}$$

Definition (39) implies that, in general, the amplitudes $F_{\vec{n},n}$ and phases $\delta_{\vec{n},n}$ have the properties

$$\begin{aligned} F_{-\vec{n},-n}(\vec{I}) &= F_{\vec{n},n}(\vec{I}) \\ \delta_{-\vec{n},-n}(\vec{I}) &= -\delta_{\vec{n},n}(\vec{I}), \end{aligned} \tag{40}$$

because the functions F^k are real. The integer n and the multiplier $1/\ell_0$ in Eq. (38) arise from Fourier decomposition of the sum of delta functions in (37):

$$\sum_{n=-\infty}^{+\infty} \delta(s-n\ell_0) = \frac{1}{\ell_0} \sum_{n=-\infty}^{+\infty} e^{in(2\pi/\ell_0)s} . \tag{41}$$

Note that when all the $F^k(x,y)$ are even functions of x and y, then $F_{\vec{n},n}$ is zero unless both n_x and n_y are even. Note also that if all interaction regions are identical, i.e., if all F^k are equal to the same function F, then $F_{\vec{n},n}$ is zero unless n is an integral multiple of C, in which case

$$F_{\vec{n},n}(\vec{I}) e^{i\delta_{\vec{n},n}(\vec{I})} = C \oint \frac{d\theta_x}{2\pi} \oint \frac{d\theta_y}{2\pi} F(\vec{I},\vec{\theta}) e^{-i\vec{n}\cdot\vec{\theta}} . \tag{42}$$

Also, when the F^k are all even, and all equal, $F_{\vec{n},n} e^{i\delta_{\vec{n},n}}$ is real for all $\vec{n}$ and n (i.e., $\delta_{\vec{n},n}$ is an integral multiple of π, when $F_{\vec{n},n}$ is not zero).

Define

$$H_0 \equiv \vec{\omega}\cdot\vec{I} + \frac{1}{\ell_0} F_{0,0}(\vec{I}) , \tag{43}$$

and $H_1 \equiv H - H_0$. It is conventional[5,84] to replace H by H_0, in Eq. (36), in order to obtain a simple first approximation to the trajectory of a test particle.* With this replacement, $\vec{I}$ is independent of time, and $\vec{\theta}$ varies at a constant angular rate

$$\begin{aligned} \dot{\vec{I}} &= 0 \\ \dot{\vec{\theta}} &= \vec{\nabla} H_0(\vec{I}) . \end{aligned} \tag{44}$$

The effective tunes of such a trajectory are the components of $(\ell_0/2\pi)\,\vec{\nabla}H_0(\vec{I})$. (This is the leading effect to which we referred in the discussion surrounding Eq. (12). The function $\delta\mu_y$ appearing in Eq. (12) is equal to $(\ell_0/2\pi C) \times (-\omega_y + \partial H_0/\partial I_y)$ and similarly for $\delta\mu_x$.)

An $\vec{I}_0(s)$ and $\vec{\theta}_0(s)$ that satisfy (44) approximate a solution of (36) as closely as can be expected, provided the cosines that constitute $H_1(\vec{I}_0(s), \vec{\theta}_0(s))$ oscillate rapidly enough. (The quantitative meaning of "rapidly enough," and of "far enough" and "near enough," below, will be made explicit in the next subsection.) In this case--as a perturbative calculation in powers of H_1 indicates--the corrections to $\vec{I}_0$ and $\vec{\theta}_0$ also oscillate, and are $O(F)$ in amplitude, where F is any parameter (for example, vertical tuneshift) describing the overall scale of the F^K.

If $\vec{I}_0$ and $\vec{\theta}_0$ satisfy (44), the angular rate at which a cosine in H_1 oscillates is

$$\frac{d}{ds}\left(\vec{n}\cdot\vec{\theta}_0 + \frac{2\pi n s}{\ell_0}\right) = \vec{n}\cdot\vec{\nabla} H_0(\vec{I}_0) + \frac{2\pi n}{\ell_0} , \tag{45}$$

for some integral $\vec{n}$ and n. H_1 oscillates "rapidly enough" when the expression above is far enough away from zero for all integers n_x, n_y, and n.

An $\vec{I}_0$ for which the right-hand-side of (44) vanishes is called "resonant" (see footnote next page); the set of all $\vec{I}_0$ for which the right-hand-side of (44) vanishes is called "the $(\vec{n},n)$ resonance curve in the action plane." Near a resonant action, perturbative

*It is not obvious that $H = H_0 + H_1$ is the only systematic decomposition of H that permits an analysis similar to that developed in the remainder of this chapter. It is the only one that I have encountered.

calculations in powers of H_1 are ill-defined because they yield expressions that contain factors of the form (45) in denominators. A more careful calculation, to be carried out in the next subsection, shows that when $\vec{I}_0$ is near enough to resonant(and $\vec{\theta}_0$ is appropriate), the corrections to $\vec{I}_0$ and $\vec{\theta}_0$ can oscillate with amplitudes as large as $O((F)^0)$.

Note that when the strong beam is symmetric about the design orbit--i.e., when $F^k(x,y)$ is an even function of x and y for all k, so that $F_{n,n}^{\rightarrow} = 0$ for odd n_x or odd n_y--there are large resonant oscillations only for even n_x and n_y. This is the origin of the expectations, expressed in Chapter II, Section 3, concerning the constraints that beam symmetries impose on observable effects in circular colliders.

Note also that when the storage ring, <u>including</u> the strong beam, has repeat period ℓ_0/C--i.e., when all the F^k are equal, so that $F_{n,n}^{\rightarrow} = 0$ when n is not an integral multiple of C--there are large resonant oscillations only when n/C is integral. This is the origin of the expectations, also expressed in Chapter II, Section 3, concerning the constraints that storage ring periodicity imposes on observable effect in circular colliders.

Resonant actions play central roles in each of the calculations to be discussed in the next chapter. Accordingly, the remainder of the present chapter is devoted to the theory of oscillations about resonant actions.

c. <u>Nearly-resonant motion; resonance overlap; frequency and width</u>

In this subsection we derive and analyze a simple approximation to the behavior of nearly resonant solutions to Eq. (36). Most of the discussion in this subsection will refer explicitly to models in which provision is made for both horizontal and vertical degrees of freedom. Near the end of this subsection, we shall point out some simplifications that arise for models--such as the ISR model discussed in the next chapter--in which the horizontal degree of freedom is omitted.

Thus, let $\vec{I}$ and $\vec{\theta}$ represent actions and angles that solve (36); let $\vec{I}_r$ be a constant vector such that

$$\vec{n}\cdot\vec{\nabla}H_0(\vec{I}_r) + \frac{2\pi}{\ell_0}\, n = 0, \tag{46}$$

for some integers n_x, n_y, and n ; and let $\vec{i} \equiv \vec{I}-\vec{I}_r$.

For small $\vec{i}$, it is conventional[5,84] to approximate Eq. (36) by modifying $H = H_0 + H_1$ as follows

*Note that this differs from the standard accelerator-physics usage[93] of "resonant," which most often refers to conditions of the form (13), to be satisfied by storage ring tunes, rather than by phase-space coordinates. (This footnote refers to preceding page.)

$$H_0 \to H_0(\vec{I}_r) + \vec{i}\cdot\vec{\nabla} H_0(\vec{I}_r) + 1/2\,(\vec{i}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)\ , \tag{47}$$

$$H_1 \to 2\ell_0^{-1} F_{\vec{n},n}(\vec{I}_r) \cos\left[\vec{n}\cdot\vec{\theta} + \frac{2\pi n}{\ell_0} s + \delta_{\vec{n},n}(\vec{I}_r)\right]\ , \tag{48}$$

where, among all triples of integers (associated with nonzero Fourier coefficient F) in the same ratios, the triple $(\vec{n},n)$ is one of the two (the other is $(-\vec{n},n)$) with the smallest common divisor. This is usually sufficient to guarantee that, in (48), one approximates H_1 by the largest of the Fourier terms that would be slowly varying if $\vec{\theta}$ satisfied (43) with $\vec{I} = \vec{I}_r$. The factor of two is present because the terms corresponding to $(\vec{n},n)$ and $(-\vec{n},-n)$ are identically equal. (48) contains no $\vec{i}$-dependence because $\vec{i}$ is assumed small. Note that the gradients in (47) do not act on $\vec{i}$.

The rationale for approximating H_0 by the first three terms in its expansion in powers of $\vec{i}$ is this: If one retains only the first term, then substitution into (36) (using (48)) gives $\vec{\theta}' \equiv 0$, which is far from the starting premise $\vec{\theta}' \sim \vec{\nabla}H_0(\vec{I}_r)$ (a prime indicates differentiation with respect to s). If one retains only the first two terms, then substitution into (36) (using (48)) gives an $\vec{I}(s)$ that grows linearly with s, i.e., $\vec{I}$ grows without limit; however, this is not reasonable, because if $\vec{I}$ deviates considerably from $\vec{I}_r$, then $\vec{n}\cdot\vec{\nabla}H_0(\vec{I}_r) + n(2\pi/\ell_0)$ deviates considerably from zero, and so the resonant substitution (48) becomes inappropriate. Three is the smallest number of power-series terms needed to avoid these difficulties (see footnote next page).

With modifications (47) and (48), Eq. (36) becomes

$$\vec{i}' = -\left(\frac{2}{\ell_0}\right)\vec{n}\, F_{\vec{n},n}(\vec{I}_r) \sin\left[\vec{n}\cdot\vec{\theta} + \frac{2\pi n}{\ell_0} s + \delta_{\vec{n},n}(\vec{I}_r)\right],$$

$$\vec{\theta}' = \vec{\nabla} H_0(\vec{I}_r) + (\vec{i}\cdot\vec{\nabla})\,\vec{\nabla} H_0(\vec{I}_r). \tag{49}$$

To simplify this, we recombine the variables: Let $\vec{c}$ be a fixed vector tangent to the $(\vec{n},n)$ resonance curve at $\vec{I}_r$ (in particular,

$$(\vec{c}\cdot\vec{\nabla})\,(\vec{n}\cdot\vec{\nabla})\, H_0(\vec{I}_r) = 0\ ), \tag{50}$$

and define i_a, i_t, θ_a, and θ_t by

$$\vec{i} \equiv i_t\,\vec{n} + i_a\,\vec{c}\ ,$$

$$\vec{\theta} \equiv s\vec{\nabla}H_0(\vec{I}_r) + \left[\frac{\theta_t - \delta_{\vec{n},n}(\vec{I}_r)}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)}\right] (\vec{n}\cdot\vec{\nabla})\vec{\nabla}H_0(\vec{I}_r) \tag{51}$$

$$+ \theta_a (\vec{c}\cdot\vec{\nabla})\vec{\nabla}H_0(\vec{I}_r).$$

The subscripts "a" and "t" stand for "along" and "transverse:" By definition of $\vec{c}$, $i_a\vec{c}$ is a small displacement along the resonance curve, and--as long as $\vec{n}$ and $\vec{c}$ are linearly independent (i.e., as long as $(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r) \neq 0$)--$i_t\vec{n}$ is a small displacement transverse to the resonance curve; similarly, the vectors $(\vec{c}\cdot\vec{\nabla})\vec{\nabla}H_0(\vec{I}_r)$ and $(\vec{n}\cdot\vec{\nabla})^2$.

*Some authors[86,91] model resonant effects by making the replacements

$$H_0 \to H_0(\vec{I}=0) + \vec{I}\cdot\vec{\nabla} H_0(\vec{I}=0) + 1/2\, I_k I_l \partial_k \partial_l H_0(\vec{I}=0), \tag{47a}$$

$$H_1 \to 2\ell_0^{-1} A_{\vec{n},n} I_x^{|n_x|/2} I_y^{|n_y|/2} \cos[\vec{n}\cdot\vec{\theta} + \frac{2\pi n}{\ell_0} s + \delta_{\vec{n},n}(\vec{I}=0)]. \tag{48a}$$

The expression on the right-hand-side of (47a) is the sum of the first three terms in the Taylor expansion of H_0 in powers of $\vec{I}$; the expression on the right-hand-side of (48a) is the same Fourier term as in (48), but with $F_{\vec{n},n} e^{i\delta_{\vec{n},n}}$ replaced not by its value at $\vec{I}=\vec{I}_r$, but by its leading behavior as $\vec{I}\to 0$. These are the power-series approximations mentioned in Section 1 of this chapter. Let us briefly sketch the main weaknesses in these approximations.

The quadratic approximation (47a) simplifies the procedure of solving Eq. (46) for $\vec{I}_r$. However, the $\vec{I}_r$'s so obtained are likely to be much smaller than the corresponding exact solutions, for the following reason: The approximation to $\vec{\nabla} H_0$ derived from (47a) grows linearly with $\vec{I}$ for large $\vec{I}$; however, in reality, $\vec{\nabla} H_0$ approaches the constant $\vec{\omega}$ as $\vec{I}\to\infty$, because (roughly speaking) for large $\vec{I}$ a weak beam particle is--on the average--far from the strong beam, and therefore can only suffer small perturbations because of the strong beam. Thus, (47a) overestimates $\vec{\nabla}H_0$ for most $\vec{I}$, and therefore underestimates solutions to $\vec{n}\cdot\vec{\nabla}H_0 + (2\pi n/\ell_0) = 0$ for most $(\vec{n},n)$. I.e., if one uses approximation (47a), one is likely to conclude that a given resonant effect is important for smaller actions--and therefore for more particles--than is really the case.

For $|n_x| + |n_y| > 4$, and for large $\vec{I}$, (48a) overestimates H much more badly than (47a) overestimates H_0, but for essentially the same reason. Thus, if one uses (48a), one is likely to conclude that resonant effects extend to larger actions--and therefore transport weak beam particles to larger distances from the beam center--than is really the case. (This footnote refers to preceding page.)

$\vec{\nabla}H_0(\vec{I}_r)$ indicate the direction in which $\vec{\theta}$ would change if Eq. (44) were in force, and $\vec{I}_r$ suffered a small displacement along the resonance curve and across the resonance curve, respectively.

In terms of i_a, i_t, θ_a, and θ_t, Eq. (49) is

$$i_t' = -\frac{2}{\ell_0} F_{n,n}(\vec{I}_r) \sin\theta_t \ , \tag{52a}$$

$$\theta_t' = i_t (\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r) \ , \tag{52b}$$

$$i_a' = 0 \ , \tag{52c}$$

$$\theta_a' = i_a \ . \tag{52d}$$

One obtains an equation involving only θ_t by differentiating both sides of (52b) with respect to s, and then using (52a). The result is

$$\theta_t'' = -\omega_r^2 \sin\theta_t \ , \tag{53}$$

where

$$\omega_r^2 \equiv \frac{2}{\ell_0} F_{n,n}(\vec{I}_r) [(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)]. \tag{54}$$

(In what follows, we assume, for convenience, that $\omega_r^2>0$. No generality is sacrificed in this way, because the two cases $\omega_r^2>0$ and $\omega_r^2<0$ can be transformed into one another by the shift $\theta_t \to \theta_t + \pi$.)

Equation (53) is familiar from elementary mechanics[94]--it describes the angular coordinate of a simple pendulum. The conserved "energy" is

$$\varepsilon \equiv \frac{1}{2}(\theta_t')^2 - \omega_r^2 \cos\theta_t = \frac{1}{2}[(\vec{n}\cdot\nabla)^2 H_0(\vec{I}_r)]^2 i_t^2 - \omega_r^2 \cos\theta_t \ . \tag{55}$$

For $\varepsilon<\omega_r^2$, θ_t oscillates regularly about $2\pi m$, for some integer m. For $\varepsilon>\omega_r^2$, θ_t rotates, i.e., θ_t (and therefore also $[\theta_t(s) - s\langle\theta_t'\rangle]$) oscillates, but with nonzero mean value. $|\langle\theta_t'\rangle|$ is a monotonic function of ε, approaching $\sqrt{2\varepsilon}$ for large ε. The function $\theta_t(s)$ is completely determined by $\langle\theta_t'\rangle$ and $\langle\theta_t - \langle\theta_t'\rangle s\rangle$.

Substituting Eqs. (52b-d) into (51), and bearing in mind the foregoing discussion, we see that $\vec{I}$ and $\vec{\theta}$ oscillate periodically about $\vec{I}_c$ and $\vec{\theta}_c$ (the subscript "c" stands for "center"), given by

$$\vec{I}_c = i_a \vec{c} \ ,$$

$$\vec{\theta}_c = s\vec{\nabla}H_0(\vec{I}_r) + i_a\,(s-s_a)\,(\vec{c}\cdot\vec{\nabla})\,\vec{\nabla}H_0(\vec{I}_r)$$

$$+ \left[\frac{2\pi m - \delta_{\vec{n},n}(\vec{I}_r)}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)}\right](\vec{n}\cdot\vec{\nabla})\,\vec{\nabla}\,H_0(\vec{I}_r) \qquad (56)$$

$$= s\vec{\nabla}H_0(\vec{I}_r+\vec{i}_c) - i_a s_a\,(\vec{c}\cdot\vec{\nabla})\,\vec{\nabla}H_0(\vec{I}_r)$$

$$+ \left[\frac{2\pi m - \delta_{\vec{n},n}(\vec{I}_r)}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)}\right](\vec{n}\cdot\vec{\nabla})\,\vec{\nabla}H_0\,(\vec{I}_r) + O(i_c^2),$$

for $\varepsilon<\omega_r^2$, and by

$$\vec{I}_c = \langle i_t\rangle\,\vec{n} + i_a\vec{c} \ ,$$

$$\vec{\theta}_c = s\,\vec{\nabla}\,H_0(\vec{I}_r) + \langle i_t\rangle\,(s-s_t)\,(\vec{n}\cdot\vec{\nabla})\,\vec{\nabla}H_0\,(\vec{I}_r) +$$

$$+ i_a\,(s-s_a)\,(\vec{c}\cdot\vec{\nabla})\,\vec{\nabla}\,H_0(\vec{I}_r)$$

$$= s\,\vec{\nabla}H_0\,(\vec{I}_r+\vec{i}_c) + \vec{\theta}_c(s=0) + O(i_c^2), \qquad (57)$$

for $\varepsilon>\omega_r^2$. The coefficients i_a, $\langle i_t\rangle$, s_a, and s_t are independent constants of integration. The oscillatory remainders $\vec{I}-\vec{I}_c$ and $\vec{\theta}-\vec{\theta}_c$ are determined entirely by i_a, $\langle i_t\rangle$, s_a, s_t, and by a solution, $\theta_t(s)$, of Eq. (53). For $\varepsilon>\omega_r^2$, θ_t is, in turn, entirely determined by $\langle i_t\rangle$ and s_t.

Expressions (56) and (57) are both of the same general form--solutions of (43) (to $O(i_c)$), with $\vec{I} = \vec{I}_r + \vec{i}_c$. However, they differ in that $\vec{i}_c$ and $\vec{\theta}_c(s = 0)$ are unrestricted in (57), while in (56) $\vec{i}_c$ must lie along $\vec{c}$, and the component of $\vec{\theta}_c(s = 0)$ along $(\vec{n}\cdot\vec{\nabla})\vec{\nabla}H_0(\vec{I}_r)$ must take a special value. Trajectories of the special form (56) will be called (stable) "resonant trajectories." A region of phase space that, at some fixed time, is centered on the locations of resonant trajectories that have in common the same index m and resonance numbers $(\vec{n},n)$, and satisfies $\varepsilon<\omega_r^2$, will be called a "resonant region." As time evolves, resonant regions follow resonant

trajectories through phase space. The boundary of a resonant region, where $\epsilon = \omega_r^2$, will be called a "separatrix."

The main conceptual difference between solutions of (36) inside and outside resonant regions is this: For $\epsilon > \omega_r^2$, there is a one-to-one correspondence between solutions of Eq. (36) (modified according to (47) and (48)) and the solutions to Eq. (44) about which they oscillate; i.e., in this case, solutions to (36) are simply deformations of solutions to (44). For $\epsilon < \omega_r^2$, by contrast, there are more solutions to Eq. (36) than there are solutions to Eq. (44) about which they oscillate; a two-dimensional set (labelled by the initial conditions $i_t(s = 0)$ and $\theta_t(s = 0)$) of solutions to Eq. (36) oscillates about each resonant trajectory.

Thus, resonant trajectories are stable configurations (i.e., centers of stable oscillation) in phase space; the domains of stability (i.e., the sets that oscillate about the stable centers) are the resonant regions. This is why resonance figures prominently in each model that we discuss in the next chapter--a resonant region can compete with the neighborhood of the design orbit for a share of the beam population. If a very stable resonant trajectory is located far enough from the beam center, the observational consequence can be expansion of the beam size, or enhancement of the rate at which particles strike the containing pipe walls, or both.

When, in the preceding section, I referred to phase space points that are "near enough" to resonant, or "far enough" from resonant (equivalently, for which the cosines in H_1 oscillate "rapidly enough") I was trying loosely to characterize phase space points that are, respectively, inside and outside resonant regions.

It should be noted that resonant regions also figure prominently in the theory of stochastic behavior, through the "resonance overlap"[76] criterion: Subsets of phase space in which two resonant regions overlap are especially likely to contain many points that lie on chaotic orbits. Since neither calculation in Chapter V will require an application of this criterion, we shall not consider it further here. A very detailed treatment can be found in Ref. 76.

In the next chapter we shall need to refer to results obtained using formulae for the frequency of small oscillations in a resonant region, and for the width of such a region's projection onto $\vec{I}$-space. Let us derive these formulae and comment briefly on some of their features.

Frequency: The behavior of small oscillations is determined by Eq. (53), linearized about $\theta_t = 0$. The angular frequency of the corresponding harmonic vibrations is clearly ω_r.

The reader may find it instructive to see an order-of-magnitude estimate of ω_r. For this purpose, we rewrite the definition (54) of ω_r entirely in terms of ℓ_0 and the F^k, using the definition (42) of H_0:

$$\omega_r = \frac{1}{\ell_0}\sqrt{2F_{n,n}^{\rightarrow}(\vec{I}_r)[(\vec{n}\cdot\vec{\nabla})^2 F_{o,o}^{\rightarrow}(\vec{I}_r)]}. \qquad (58)$$

We may estimate (58) as follows: Let σ be a rough common approximation to both σ_x^* and σ_y^*, the strong beam (and unperturbed weak beam) Gaussian half widths, at the interaction points; and let β be a rough common approximation to both β_x^* and β_y^*. According to the definitions (23) and (39) of the F^k and of $\vec{F}_{n,n}$, the following simple estimates follow from dimensional considerations, for $\vec{I}_r \sim O(\sigma^2/\beta)$, an amplitude typical of most particles in the beam:

$$\vec{F}_{n,n}(\vec{I}_r) \sim O\left(\frac{e^2N}{E}\right) \sim O\left(\frac{eI}{Ef}\right),$$
$$(\vec{n}\cdot\vec{\nabla})^2 \vec{F}_{o,o}(\vec{I}_r) \sim O\left[\left(\frac{\beta}{\sigma^2}\right)^2\left(\frac{e^2N}{E}\right)\right] \sim O\left[\left(\frac{\beta}{\sigma^2}\right)^2\left(\frac{eI}{Ef}\right)\right]. \tag{59}$$

As in Chapter II, the variables N, I, f, E refer, respectively, to the number of particles of charge e in the strong beam, the current in the strong beam, the storage ring revolution frequency, and the energy per beam particle. In formulating (59) we have neglected--among many other things--any $\vec{n}$-dependence. When (59) is substituted into (58), we obtain

$$\omega_r \sim O\left[\frac{1}{\ell_0}\left(\frac{eI\beta}{Ef\sigma^2}\right)\right] \sim O(2\pi\xi/\ell_0). \tag{60}$$

The motivation for the second approximate equality in (60) is provided by the form of the equation, (4), that defines the tuneshifts; ξ is a rough common approximation to the tuneshifts ξ_x and ξ_y.

(A similar estimate, $\omega_r \sim O(\sqrt{2\pi\xi}/\ell_0)$, applies to storage rings--e.g., the ISR--in which continuous beams cross at a non-zero angle. In such cases, Eqs. (27) and (4a), rather than (23) and (4), must be used in defining and estimating $\vec{F}_{n,n}$, $(\vec{n}\cdot\vec{\nabla})^2\vec{F}_{o,o}$, and ξ.)

Let us compare this order of magnitude to other rates characteristic of storage rings, for the case of e^+e^- colliders: The rate of betatron oscillations is typically $\lesssim O(10)$ vibrations per beam-beam crossing; thus $(1/2\pi)\,\omega_\beta \sim O(10/\ell_0)$, where ω_β is a rough measure of the angular frequency (per azimuthal distance s, not per time) of either vertical or horizontal betatron oscillations. The transverse damping time γ^{-1} is typically $\sim O(10^3\text{-}10^4)$ beam-beam crossings; thus $\gamma \sim O(10^{-4}\text{-}10^{-3}/\ell_0)$ (per azimuthal distance s). I.e., for $\xi \sim O(10^{-2})$,

$$\frac{1}{2\pi}\omega_\beta : \frac{1}{2\pi}\omega_r : \gamma \sim 10^5 : 10^2 : 1\text{-}10 \ . \tag{61}$$

Width: The projection of a resonant region onto the $\vec{I}$-plane is a tube that surrounds the corresponding resonance curve. According to

(51), the width of this tube at $\vec{I}_r$, along the direction $\vec{n}$, is equal to the product of $|\vec{n}|$ $(\equiv\sqrt{\vec{n}\cdot\vec{n}})$ and the length of the interval defined by all values of i_t for which $\varepsilon\leq\omega_r^2$ can be satisfied for some real value of θ_t. According to the definition (55) of ε, the inequality $\varepsilon\leq\omega_r^2$ is equivalent to

$$|i_t| \leq \frac{2\omega_r|\cos\theta_t|}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)} \leq \frac{2\omega_r}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)} \quad . \tag{62}$$

Thus, the resonance width W, along $\vec{n}$, is equal to

$$W_r = \frac{4|\vec{n}|\omega_r}{(\vec{n}\cdot\vec{\nabla})^2 H_0(\vec{I}_r)} = 4|\vec{n}|\sqrt{\frac{2F_{\vec{n},n}(\vec{I}_r)}{(\vec{n}\cdot\vec{\nabla})^2 F_{o,o}(\vec{I}_r)}} \quad . \tag{63}$$

(The true geometric (perpendicular) width of the tube is $W_r[1-(\vec{n}\cdot\vec{c})^2/|\vec{n}|^2|\vec{c}|^2]^{1/2}$.)

An order-of-magnitude estimate similar to that described above would give $W_r \sim O(\sigma^2/\beta) \sim O(\vec{I}_r)$. If taken literally, such an estimate would cast doubt on much of the picture that we have just derived, since that picture is based on the assumption that $\vec{I}$ is small. However, the rough comparison $W_r \sim O(\vec{I}_r)$ is an overestimate, since[75] in fact the Fourier coefficients $F_{\vec{n},n}$ fall rapidly with increasing $|\vec{n}|$, for $\vec{I}_r \sim O(\sigma^2/\beta)$.

Let us note some simplifications that arise when the analysis described in this subsection is applied to models--such as the ISR model discussed in the next chapter--in which no provision is made for the horizontal variables x and x'.

In such cases, the vectors $\vec{I}$, $\vec{I}_r$, $\vec{i}$, $\vec{\theta}$, $\vec{n}$, etc. have only y-components, and $\vec{c}$ is irrelevant. Thus, the general resonance condition (46) becomes

$$n_y\left(\frac{dH_0}{dI_y}\right)(I_y = I_{yr}) = -n\frac{2\pi}{\ell_0} \quad , \tag{64}$$

and the general expression (56) for a resonant trajectory becomes

$$i_{yc} = 0 \ ,$$

$$\theta_{yc} = s\frac{dH_0}{dI_y}(I_{yr}) + \frac{1}{n_y}[2\pi m - \delta_{n_y,n}(I_{yr})] \tag{65}$$

$$= -s\ \frac{2\pi}{\ell_0}\ \frac{n}{n_y} + \frac{1}{n_y}\,[2\pi m - \delta_{n_y,n}(I_{yr})].$$

It follows from (65) that resonant trajectories with m's that differ by n_y are identical, because the angular variable θ_{yc} has period 2π; thus, there are precisely n_y resonant trajectories associated with n_y and n at I_{yr}. For the same reason, each of these trajectories is periodic in s with period at most* $(n_y/d)\ell_0$, where d is the greatest common divisor of the integers n_y and n. Neither of these statements is true in general when the storage ring model involves both vertical and horizontal degrees of freedom.

To make the discussion in this subsection more concrete, a schematic representation of the two-dimensional phase space of such an x-independent model is shown in Fig. 32. In this figure one sees the locations (at one value of s) of four resonant trajectories (corresponding to $n_y = 4$), and the separatrices of the associated resonant regions.[y] A movie of such a phase plane would show that as time (=s/c) passes, the points within each resonant region swirl about the resonant trajectory at the region's center, and the whole "island chain" of resonant regions revolves--with some periodic deformation--about the origin. The revolution and deformation are such that the pattern of resonant trajectories and separatrices repeats in a time interval $\Delta t = \ell_0/c$--the island that occupies position number p in the chain at time t evolves smoothly into the island that occupies position number p + n at time $t + \ell_0/c$.

3. Resonant behavior in weak-strong systems-damping included.

In this section we derive some basic properties of nearly resonant behavior in damped systems. For convenience, we shall assume, in what follows, that the rates of vertical and horizontal damping are exactly equal. In real storage rings, they are very close, but not identical.[4]

We shall largely ignore noise processes, although, in passing, we shall indicate the way in which they modify the conclusions derived here. The theory of noise effects in the present context is not well developed. In the next chapter, when we shall need to take radiative noise into account quantitatively, we shall have to do so semiphenomenologically.

*One might have naively expected that the period is always $(n_y/n)\ell_0$. However, in obtaining the physical orbital coordinates y and y' from I_y and θ_y, one must use Eq. (32), which depends explicitly on s with period $\Delta s = \ell_0/C$. Thus--unless, for some special reason, the period of $\beta_y(s)$ and $\Phi_y(s)$ is actually less than ℓ_0/C--the true period of a resonant trajectory is the smallest common integral multiple of ℓ_0/C and $(n_y/n)\ell_0$, i.e., $(n_y/d)\ell_0$.

We shall see that the damped case is similar to the undamped case in that there is a subset of phase space within which the vector separation, parallel to the resonance direction $\vec{n}$, between $\vec{I}$ and some given resonance curve oscillates about zero. However, the damped and undamped cases also differ in several respects. The most important such difference is this: Without damping, resonant-region oscillation centers, when projected onto the action plane, are (according to Eq. (52c)) stationary points on the resonance curve. With damping, the oscillation centers, when projected onto the action plane, also lie on the resonance curve, but they are no longer stationary--instead, they drift steadily toward resonant actions $\vec{I}_r$ that satisfy

$$\frac{I_{rx}}{I_{ry}} \simeq \frac{n_x}{n_y} \quad . \tag{66}$$

The precise point at which a particle's orbit stops drifting in this way is determined by noise. If noise were absent, such a drift would persist indefinitely, gradually slowing as the oscillation center approaches (66).

This drift--"resonance streaming"--is the mathematical phenomenon introduced by Tennyson[5] in an attempt to interpret the ordered, nearly vertical normalized-amplitude motion associated with the beam blowup observed in the SPEAR simulation discussed in the preceding chapter. Tennyson was led to this interpretation by comparing such ordered trajectories with maps that show a number of resonance curves in the a_x-a_y plane*, calculated from the flat-beam equations of motion used in his simulation. In all cases, the ordered parts of trajectories were observed to follow resonance curves. The resonance map for 2.2 GeV per particle and $\xi_x = \xi_y = .06$ is reproduced here (from Ref. 5) as Figure 33, for direct comparison with the test-particle trajectory shown in Fig. 30. One sees clearly that the long vertical segment in Fig. 30 lies along the resonance $(n_x, n_y, n) = (3, 1, -21)$.

The work of this section is organized as follows. In subsection 3a, we shall show how one modifies the $\vec{I}$-$\vec{\theta}$ equations of motion (36) in order to include the effects of radiative damping. In subsection 3b, we shall use the approximate techniques introduced in subsection 2c in order to analyze the modified equations of motion obtained in subsection 3a. In particular, we shall derive the existence of the streaming effect.** We shall also derive--among other things--the following approximate formula for the rate at which streaming proceeds (as long as noise can be neglected): Let $\vec{I}_r(s)$, lying on a

*Recall that when translated into the notation used in this chapter, $a_x = (1/\sigma_x)\sqrt{2I_x\beta_x^*}$, $a_y = (1/\sigma_y)\sqrt{2I_y\beta_y^*}$.

**The original explanation of streaming, due to Tennyson,[5,71] involved a geometrical argument that we shall not reproduce here. The analytical treatment presented in Subsection 3b enables one to be more quantitative.

resonance curve, be the time-dependent action of the center of some resonant-region oscillation; let $\vec{p}$ be a fixed vector orthogonal to the resonance vector $\vec{n}$; let γ be the common value of the horizontal and vertical radiative damping rates (per azimuthal distance); then

$$\vec{p}\cdot\vec{I}_r(s) \simeq (\text{constant})\cdot e^{-2\gamma s} \ . \tag{67}$$

This formula will play an important part in the SPEAR calculation to be discussed in the next chapter.

Note that Eq. (66) is a consequence of Eq. (67). According to (67), $\vec{p}\cdot\vec{I}_r$ approaches zero as s becomes infinite; but $\vec{p}\cdot\vec{I}_r = 0$ is equivalent to (66), because $\vec{n}\cdot\vec{p} = 0$, and because--for our purposes--action space is two-dimensional.

a. $\vec{I}$-$\vec{\theta}$ equations of motion in the presence of damping.

In terms of Cartesian phase space coordinates, the equations that describe the damped linear motion of test particles between strong-beam encounters are

$$\frac{dz}{ds} = z' \ ,$$

$$\frac{dz'}{ds} = K_z(s)z - \Gamma_z(s)z', \tag{68}$$

where z represents x or y. The function $K_z(s)$ measures the focussing strength of quadrupole magnets; it is related to $\beta_z(s)$ as follows[4]

$$\beta_z'' - \frac{\beta_z'^2}{2\beta_z} - 2K_z\beta_z - \frac{2}{\beta_z} = 0. \tag{69}$$

The damping coefficient $\Gamma_z(s)$ is in general s-dependent--typically, it takes its largest values in RF cavities.[4] The average, $\langle\Gamma_z\rangle$, of Γ_z over one repeat period (ℓ_0) of the model storage ring is equal[4] to twice the transverse damping rate γ.

To write (68) in terms of $\vec{I}$ and $\vec{\theta}$, one substitutes the definition (32) of the actions and angles into (68). After some algebra (that we shall not reproduce here), in the course of which Eq. (69) must be used, one obtains the result

$$I_z' = -2I_zS_z\left(S_z - \frac{1}{2}\beta_z'C_z\right)\Gamma_z \ ,$$

$$\theta_z' = \omega_z - C_z\left(S_z - \frac{1}{2}\beta_z'C_z\right)\Gamma_z \ , \tag{70}$$

where C_z and S_z represent the cosine and sine, respectively, of $\Phi_z + \theta_z$.

The $\vec{I}$ - $\vec{\theta}$ equations of damped motion in the presence of a counter-rotating strong beam are obtained by combining (70) and (35), just as we combined (34) and (35) in deriving the undamped equations (36). The result is

$$I_z' = -\frac{\partial H}{\partial \theta_z} - 2 I_z S_z (S_z - \frac{1}{2}\beta'_z C_z)\Gamma_z ,$$

$$\theta_z' = +\frac{\partial H}{\partial I_z} - C_z (S_z - \frac{1}{2}\beta'_z C_z)\Gamma_z , \qquad (71)$$

where H is defined in Eq. (37).

As in Subsection 2b, it is appropriate to expand not only H, but also the damping terms in (71), as Fourier series in powers of the periodic variables $\exp(2\pi i s/\ell_0)$, $\exp(i\theta_x)$, and $\exp(i\theta_y)$. For simplicity, we shall, in what follows, retain only terms of zeroth order in the Fourier expansions of the damping corrections. Thus, the damping term in the I_z' equation will be replaced by $-\langle\Gamma_z\rangle I_z = -2\gamma I_z$; and the damping term in the θ_z' equation will be replaced by the s-independent number $\delta\omega_z \equiv 1/4 \langle \Gamma_z \beta'_z \rangle$.

With these reductions, Eq. (71) becomes

$$I_z' = -\frac{\partial \tilde{H}}{\partial \theta_z} - 2\gamma I_z ,$$

$$\theta_z' = +\frac{\partial \tilde{H}}{\partial I_z} , \qquad (72)$$

where $\tilde{H}$ is defined by

$$\tilde{H} \equiv H + \vec{I}\cdot\delta\vec{\omega} .$$

In what follows, a resonant action $\vec{I}_r$ will always be defined as a solution of Eq. (46), with $\tilde{H}$ substituted for H. (However, since the Γ's are generally small, this substitution should shift the $\vec{I}_r$'s only slightly.) Note that the zeroth Fourier components, H_0 and $\tilde{H}_0$, of H and $\tilde{H}$ satisfy

$$\tilde{H}_0 - H_0 = \tilde{H} - H.$$

Thus the remainders H_1 and $\tilde{H}_1$ are equal.

b. Streaming, etc..

Let us now apply the resonance approximations (47) (with $\tilde{H}_0$ replacing H_0) and (48) to Eqs. (72). In terms of the variables defined in (51) (with H_0 again replacing H_0), Eqs. (72), in this approximation, are equivalent to

$$i_t' = -\frac{2}{\ell_0} F_{\vec{n},n}(\vec{I}_r) \sin\theta_t - 2\gamma i_t - 2\gamma I_t \ , \tag{73a}$$

$$\theta_t' = i_t \, (\vec{n}\cdot\vec{\nabla})^2 \, \tilde{H}_0 \, (\vec{I}_r) \ , \tag{73b}$$

$$i_a' = -2\gamma i_a - 2\gamma \, I_a \ , \tag{73c}$$

$$\theta_a' = i_a \ . \tag{73d}$$

The parameters I_a and I_t are obtained from $\vec{I}_r$ in the same way that i_a and i_t are obtained from $\vec{i}$, i.e.,

$$\vec{I}_r \equiv \vec{n} \, I_t + \vec{c} \, I_a \ . \tag{74}$$

As in Subsection 2c, one can obtain an equation involving only θ_t: Differentiate both sides of (73b), and then insert (73a), as well as (73b) itself. The result is

$$\theta_t'' = -\omega_r^2 \sin\theta_t - 2\gamma \, [(\vec{n}\cdot\vec{\nabla})^2 \tilde{H}_0(\vec{I}_r)] \, I_t - 2\gamma\theta_t' . \tag{75}$$

If the initial values of $\vec{I}$ and $\vec{\theta}$ lie in a resonance region--appropriately deformed because of the damping terms in (75)--then θ_t and (because of (73b)) i_t oscillate; and, as s becomes infinite, θ_t approaches one of the static values

$$\tilde{\theta}_m \equiv 2\pi m - \text{Arcsin} \, \{\frac{2\gamma}{\omega_r^2} \, [(\vec{n}\cdot\vec{\nabla})^2 \tilde{H}_0(\vec{I}_r)] \, I_t\}, \tag{76}$$

and i_t approaches zero. (We shall not derive an analytical expression--to replace $\epsilon = \omega_r^2$--for the boundary of a resonance region in a damped system, because we shall not need such an expression in the calculations to be discussed in Chapter V.)

Within a resonance region, the center of oscillation in the $\vec{I}$ plane is $\vec{I}_r + \vec{c} i_a$. Thus, the oscillation center streams along the resonance curve because, according to (73c), $i_a \neq 0$, unless $i_a = \pm -I_a$, where streaming stops. Note that $i_a = -I_a$ means that $\vec{I}_r + \vec{c} i_a = \vec{n} \, I_t$, i.e., that Eq. (66) is satisfied.

Unless I_a is small, however, $i_a = -I_a$ violates the small-$\vec{i}$ assumption that underlies the resonance approximations (47) and (48). When analyzing long-range streaming, one should modify this assumption and these approximations. Here is an appropriate set of such modifications: Write

$$\vec{I} = \vec{I}_r(s) + \vec{i}\ , \tag{77}$$

where $\vec{I}_r$ is located on a resonance curve, as before, but is now a time-dependent dynamical variable; and $\vec{i}$, as before, is small, but constrained--$\vec{i}$ must now lie along $\vec{n}(\vec{i} \equiv \vec{n} i_t)$. (Since $\vec{I}_r$ is now variable, a term $\vec{i}_a c$ would be redundant.) As for (47) and (48), leave them unchanged in form (except for the replacement $H_0 \to \tilde{H}_0$), but reinterpret the variables on which they depend, according to the ansätze symbolized by the decomposition (77).

These modified approximations have this in common with the original ones: In either case, $\vec{I}' + 2\gamma\vec{I}$ is proportional to $\vec{n}$. This means that

$$\frac{d}{ds}(\vec{p}\cdot\vec{I}) + 2\gamma(\vec{p}\cdot\vec{I}) \approx 0, \tag{78}$$

for $\vec{p}\cdot\vec{n} = 0$. Upon integration, Eq. (78) yields Formula (67).

V. TOPICS IN THEORY-CALCULATIONS

In this chapter, we discuss two calculations of quantitatively measurable phenomena in colliding-beam storage rings: collision-enhanced beam loss at the ISR,[95-97] and beam blowup at SPEAR. In each case, we shall choose input parameters so that the results can be compared directly with published experimental data.

These calculations are similar in that, in each case, the phenomenon in question is attributed to the transfer of some fraction of the beam from the beam center to a single distant stable resonant region.* However, they differ in the mechanism by which this transfer takes place.

In Section 1, below, we introduce and contrast these two transfer mechanisms. The details of the calculations will be presented in Sections 2 and 3.

1. Comparative discussion of transfer mechanisms

In the SPEAR model,[5] a particle reaches large amplitudes in the following way: Radiative noise transports the particle to the low-amplitude end of a resonance curve, and then (if the particle enters the corresponding resonant region) radiative damping causes it to stream towards the high-amplitude end.

There is no such pathway available to ISR particles, for two reasons. First, there is no damping process to drive streaming. Second, as explained in Chapter II (Section 3a), the horizontal coordinates x and x' do not couple through the beam-beam interactions, so that action space may be regarded as effectively one-dimensional; in particular, a "resonance curve" is nothing more than an isolated point--it has no "near end" or "far end."

In the ISR, according to the calculation that we present in the next section, noise drives particles to large amplitude in a less direct way: Because of noise (due primarily to intrabeam scattering[98,99]), a particle's energy diffuses. Because a particle's unperturbed storage-ring tune is a function of its energy ($Ed\nu/dE \approx 2.8$ in the ISR at 26 GeV[93]) the tune also diffuses. As the tune diffuses, the locations of resonant regions in phase space diffuse as well, because H_0--which, through (46), determines the values of resonant actions--itself depends explicitly on the unperturbed tune, according to definition (43). A proton then reaches large amplitude by first becoming trapped in a resonant region when the tune is such that the region is close to the beam center; and then by being swept away as the center of the resonant regions diffuses outward. We shall explain "trapping and sweeping" more precisely in the next

*One can probably generalize either of these calculations in order simultaneously to take into account the effects of more than one well-separated resonant region. I do not know how one would proceed if the main resonant regions overlapped.

section.*

It is hard to see how a similar process could be responsible for important effects in an electron-positron storage ring (although this was suggested in Ref. 97), because phase stability (due to the RF system) prevents the electrons' and positrons' energies--and therefore also their tunes--from wandering very far from the energies and tunes of the e^- and e^+ synchronous orbits.

2. ISR beam loss[95-97]

In this section, we estimate beam loss at the ISR, operated near the fifth-order resonance $5\nu_y = 43$, following Refs. 95-97. As indicated in the preceding section, the entire effect will be attributed to the five resonant regions associated with the (time-dependent) resonant action I_{yr} defined by $5(dH_0(I_{yr})/dI_{yr}) - 43(2\pi/\ell_0) = 0$.

Similar calculations have also been done for operation near other, higher order resonances. We shall confine ourselves to the fifth-order case in order to avoid certain complications that would otherwise be present.

Note that the calculation will apply to times soon (a few tens of minutes) after injection, because, as yet, there are no adequate techniques for predicting the long-time behavior of beams--such as those in proton storage rings--that are not acted on by a fast relaxation process, such as radiative damping.

We shall compare the result of our calculation with published ISR data[93] taken at 26 GeV per proton, and about 8-9A of current per beam. Under those conditions, a loss rate of about 20-60 parts per million has been observed in at least one of the beams. This is roughly ten times the loss rate that is registered when only one beam circulates.[44]

The work of this section will be presented in three parts. In Subsection 2a, below, we explain the basic starting assumptions. In Subsection 2b, we explain in detail how moving isolated resonant regions can sweep particles out of a storage ring beam. Finally, in Subsection 2c, we combine the material covered in 2a and 2b, in order to complete the calculation.

a. Basic premises

We shall carry out this calculation as if the ISR were a weak-strong system, even though the experimental situation involves beams of roughly equal currents. This is permissible because the tuneshifts are very small (about $\xi_y = -.0004$ per interaction, for the operating conditions in question here) so that the changes that beam number one induces in beam number two make a negligible contribution to the changes that beam number two induces in beam number one, and vice versa.

* Intrabeam noise can also act directly on a particle's transverse coordinates, without a resonance as intermediary, but much less rapidly, according to Ref. 96.

In the course of this calculation, we shall require an ansatz for the distribution of an ISR beam, both in transverse phase space and in energy, immediately following injection. For the transverse distribution, we shall assume the Gaussian expression

$$\bar{\rho}(y,y',s)\, dy\, dy' \equiv (e^{-I_y/I_y^*}) \frac{dI_y}{I_y^*} \frac{d\theta_y}{2\pi} \tag{79}$$

for the normalized probability that a test particle can be found in the differential area dy dy' about (y,y'). The constant I_y^* is related to the beta-function and beam-half-height at an interaction point, by $I_y^* \equiv (\sigma_y^*)^2/\beta_y^*$. At the ISR,[93] under the operating conditions with which we shall be concerned, σ_y^* is 1 mm and β_y^* is 14 m. Also, the largest vertical amplitude that a proton can have before it strikes the beam pipe is*

$$I_{ymax} \simeq 10\, I_y^* \tag{80}$$

(i.e., the largest possible vertical displacement from the beam center is $\sqrt{2\beta_y^* I_{ymax}} \simeq 4.5$ times the rms beam half-height).

For the probability in energy space, immediately after injection, we shall assume a uniform distribution over an interval of finite size. For a 26 GeV ISR beam of 8-9 A, I estimate the width, Δ_E, of this interval to be given by**

$$\frac{\Delta_E}{E} \simeq .01\ . \tag{81}$$

Correspondingly, the unperturbed storage ring tunes are distributed uniformly in an interval whose width Δ_ν is given by

$$\Delta_\nu \simeq E \frac{d\nu}{dE} \frac{\Delta_E}{E} \simeq (2.8) \frac{\Delta_E}{E} \simeq .03. \tag{82}$$

As indicated in the preceding section, we shall assume that $\tilde{\rho}$, the probability distribution in tune space, evolves in time, after injection, according to a diffusion equation

---

*I have taken this estimate from Ref. 93. Reference 96 uses the comparable estimate $\sqrt{I_{ymax}/I_y^*} = 3$. According to Ref. 97, however, the maximum amplitude is given by $\sqrt{I_{ymax}/I_y^*} = 7$. I do not know the origin of this disagreement. I expect Ref. 93, a CERN report, to be more reliable on this point.

**This has been obtained by extrapolation from Guignard's discussion[93] of a 30 Amp beam. According to our description (Chapter II) of the manner in which the ISR is filled, Δ_E should roughly scale with the first power of the beam current.

$$D_\nu \frac{\partial^2 \tilde{\rho}}{\partial \nu^2} = \frac{\partial \tilde{\rho}}{\partial t} \tag{83}$$

(i.e., the individual values of ν execute random walks with very small steps[100]), where the diffusion constant D_ν is given, approximately, by*

$$D_\nu \simeq 5 \times 10^{-10} \text{ sec}^{-1}. \tag{84}$$

We assume that immediately following injection ($t = 0$), the distributions $\tilde{\rho}$ and $\bar{\rho}$ are independent.

Finally, let us discuss some assumptions that we shall make concerning the Fourier coefficient, $F_{5,-43}$, that sets the scale for the resonant effects to be analyzed below.

We begin by noting that $F_{5,-43}$ is nonzero because the ISR beams collide slightly off-center--a typical vertical offset is[93] $\pm$ .2 mm--and because the offset is not the same at all interaction points. If the functions F^k describing beam-beam encounters were all symmetric under $y \to -y$, or if all interaction regions were identical, then $F_{5,-43}$ would vanish identically because five is not even, and forty-three is not divisible by eight (the number of ISR interaction points).

In order to model the dependence of $F_{5,-43}$ on the beam offsets δy_k, we assume that all the strong-beam F^k are related to a single "master" function F (symmetric under $y \to -y$) via

$$F^k(y) \equiv F(y - \delta y_k) \simeq F(y) - F'(y)\delta y_k , \tag{85}$$

where the prime in (85) indicates differentiation with respect to y. We shall take F to be given by Eq. (27), for $\rho(\bar{y})$--following the foregoing discussion--a Gaussian, centered at $\bar{y} = 0$, with half-width σ_y, and with total integrated weight equal to the beam current 8-9A.

*I have taken this number from Ref. 93. In Refs. 96 and 97, numbers an order of magnitude smaller are quoted. I can account for the shortfall in Ref. 97. In Ref. 97 the figure quoted for D_ν was really D_E, the diffusion constant (with $(26 \text{ GeV})^2$ factored out) for the distribution in energy; this is clear because the results in Ref. 97 do not follow from the input unless the factor $D_\nu/D_E = (E\, d\nu/dE)^2 \simeq 8$ is inserted in the appropriate places. The number quoted in Ref. 93, extracted from Ref. 98, corresponds to a beam intensity of 30A. For simplicity, we shall ignore the variation of D_ν with current (a proper analysis of the relevant formulae in Refs. 98 or 99 is beyond the scope of the present review). One should be aware that the derivations in Ref. 98 do not actually refer explicitly to diffusion. As far as I can tell, diffusion is an appealing, but not rigorous interpretation of the theory of intrabeam scattering.

It follows from (85) and (39) that $F_{5,-43}$ is given by

$$F_{5,-43} = \left\{\left|\sum_{k=1}^{8} \delta y_k \, e^{2\pi i k(+43/8)}\right|\right\} \left\{\left|\frac{d\theta_y}{2\pi} F'(\sqrt{2I_y}\beta_y^{*-}\cos\theta_y)e^{-5i\theta_y}\right|\right\}. \tag{86}$$

For future reference, we note that if all δy_k have the same magnitude δy, but arbitrary signs, the largest value that the first factor (to be called h) in (86) can take is

$$\max h \equiv \max \left|\sum_{k=1}^{8} \delta y_k \, e^{2\pi i k(43/8)}\right| = [8(2+\sqrt{2})]^{1/2}\delta y = (5.2)\delta y. \tag{87}$$

(This can easily be derived by directly enumerating the possibilities.) Under the same constraint, the average of h^2, over all independent choices of the signs of the δy_k, is easily seen to be

$$\langle h^2 \rangle = 8\,(\delta y)^2 \,. \tag{88}$$

In what follows, we shall neglect the dependence of H_0 on the δy_k.

b. Trapping and sweeping

In this subsection, we discuss in detail how ISR protons are swept to large amplitude by (5, -43)-resonant regions that move outward because the protons' unperturbed tunes wander.

For most of this discussion, we shall focus on (the vanishingly small population of) protons associated with only a single wandering tune $\nu_y(t)$. The loss rate for the full beam will be obtained by inserting the appropriate results of this analysis into an integral over the initial distribution of tunes.

Before we proceed, we shall need to make explicit some qualitative features of the relation between the strong-beam-independent unperturbed storage ring tune ν_y and the resonant action, I_{yr}, defined by

$$0 = 5\,H_0'(I_{yr}) - 43(2\pi/\ell_0) = 5(2\pi/\ell_0) + (5/\ell_0)F'_{0,0}(I_{yr}) - 43(2\pi/\ell_0). \tag{89}$$

(The primes in (89) indicate differentiation with respect to I_{yr}.) For this purpose, we show, in Fig. 34, a graph (reproduced from Ref. 97) of $(8\xi_y)^{-1}\,(2\pi)^{-1}\,F_{0,0}(I_{yr})$, as a function of $\sqrt{I_y/I_y^*}$. It follows from Fig. 34 that $(\ell_0/2\pi)H_0$ always lies between $\nu_y \mp 8\xi_y$,

corresponding to $I_y = 0$ (recall that ξ_y is the tuneshift per interaction, and that there are eight interaction regions), and $\nu_y + 4\xi_y$, corresponding to $I_y = I_{ymax} = 10I_y^*$. Thus, there is no accessible resonant action unless ν_y lies between (43/5) - $4\xi_y$ = 8.6 - $4\xi_y$, and 8.6 - $8\xi_y$. I_{yr} increases from zero to I_{ymax} as ν_y decreases (recall that $\xi_y < 0$) from 8.6 - $8\xi_y$ to 8.6 - $4\xi_y$.

Let us now begin by examining the time-dependent changes that the mobile (5, -43) resonant regions make in the initial transverse distribution (79), under the special assumptions that $\nu_y(t)$ is greater than 8.6 - $8\xi_y$ when t = 0, and that $\nu_y(t)$ decreases monotonically thereafter. We make this assumption only for simplicity, so that the main conclusions will not be obscured by technical complications. We shall indicate later how the conclusions are modified when these assumptions are removed.

As indicated above, there are no (5, -43) resonant regions in transverse phase space until $\nu_y(t)$ reaches 8.6 - $8\xi_y$. As soon as $\nu_y(t)$ passes 8.6 - $8\xi_y$, an "island chain" of five resonant regions grows out of the origin $I_y = 0$. Any proton caught in such a region swirls about the region's center, while the center itself revolves about the origin.

Because of this swirling, many protons originally on the low-amplitude sides of these regions are moved to the high-amplitude sides, and many protons originally on the high-amplitude sides are moved to the low-amplitude sides. Since the initial transverse distribution (79) becomes more rarified as I_y increases, this high-low interchange produces a net movement of protons outward, to larger amplitude.*

When ν_y is still close to 8.6-$8\xi_y$, this net movement is very small, because the resonant regions are close to $I_y = 0$, where the initial Gaussian distribution (79) is nearly uniform--i.e., where the number of protons that are available to be displaced from higher to lower action is nearly equal to the number of protons that are available to be displaced from lower to higher action.

This net movement begins to be appreciable when ν_y is such that $I_{yr} \simeq (1/2)\, I_y^*$, because that is where the initial transverse distribution (expressed in cartesian coordinates) has its steepest gradient. Note that, according to Fig. 34, this corresponds to $\nu_y \simeq$ 8.6 - $7\xi_y$.

As ν_y moves beyond 8.6 - $7\xi_y$, the resonant regions move beyond $I_{yr} = (1/2)I_y^*$, into a domain of phase space in which the initial distribution is sparse. Thus, because of the continuing interchange of high and low amplitude particles within the moving resonant

---

*Note that this interchange can take place only if the resonant regions move outward very slowly. Otherwise, the time spent by a proton inside a resonant region cannot be long enough for swirling to have a significant effect. A discussion (in terms of the frequency, ω_r, of resonant oscillation) of the maximum value of dI_{yr}/dt beyond which the distribution of protons is largely unaffected by the passing resonances has been given by Chao and Month in Ref. 95, with some refinement in Ref. 97. We shall not consider this point further here.

regions, as in a game of leapfrog, the moving regions carry with them, to large amplitude, a substantial fraction of the protons that contributed to the original net outward movement at $I_{yr} \simeq (1/2)I_y^*$, without picking up many new protons. When ν_y reaches $8.6 - 4\xi_y$, corresponding to $I_{yr} = I_{ymax}$, this population strikes the beam pipe and is lost.

This transport mechanism is what we have referred to as "trapping and sweeping." For direct evidence that this phenomenon can actually take place, the reader is referred to the computer studies described in Ref. 95.

In accordance with the picture* we have just described, Chao and Month[95] propose the following recipe for obtaining a crude estimate of the number of protons that are swept out of the beam in this way: Multiply the total (infinitesimal) number of protons associated with $\nu_y(t)$ by the total integrated weight assigned by the initial probability distribution (79) to the five (5,-43) resonant regions corresponding to $\nu_y \simeq 8.6 - 7\xi_y$. We shall employ this estimate in what follows.

We may easily determine how this integrated weight depends on the beam offsets: As we saw in the preceding subsection, the resonant Fourier amplitude $F_{5,-43}$ depends on the offsets through the multiplicative factor h. Thus, according to Eq. (63), the width of the (5, -43) resonance in action space is proportional to $\sqrt{h}$. When the offsets are small--so that h, and therefore the action-space width, is small--the radial widths of the resonant regions in Cartesian (y,y') coordinates are proportional to the action-space width, and are therefore also proportional to $\sqrt{h}$. (The angular widths are always $2\pi/5$.) When the Cartesian widths are small, the integrated probability is approximately proportional to the Cartesian area occupied by the resonant regions; and since, for small width, the area scales with the width, one finally concludes that the integrated probability is approximately proportional to $\sqrt{h}$, for small offsets. According to Ref. 97, the proportionality constant** is about $(.06)/\sqrt{\sigma_y^*}$.

For typical values of h (substituting $\delta y = .2$ mm into Eq. (88)) this probability is quite small, about .05, although (using (87)) it could be as high as about .08. We shall use .05 in the rough estimates that follow.

A nearly identical informal analysis (which we need not reproduce here) can be applied when the tune trajectory $\nu_y(t)$ does

*Actually, this picture, and the recipe that follows, are somewhat oversimplified. However, the subtleties that have been omitted here are only relevant to the higher-order resonances with which we are not concerned here. For more details, see Ref. 97.

**This number is extracted from Ref. 97's Fig. 6, which shows a graph of the integrated resonant-region probability, normalized to $(h/\sigma_y^*\sqrt{8})^{1/2}$ as a function of $\sqrt{I_{yr}/I_y^*}$. Guignard[93] quotes a number larger than this by an order of magnitude; this can be attributed to an incorrect replacement of $(d^2F_{0,0}/dI_y^2)$ and $F_{5,-43}$ by the leading terms in their expansions in powers of I_y.

not satisfy the assumptions imposed at the beginning of this subsection. The general conclusion is this: Whenever $\nu_y(t)$ reaches $8.6 - 4\xi_y$ after having earlier passed $8.6 - 7\xi_y$, with no other encounter with $8.6-4\xi_y$ in the interim, then approximately 5% of the protons associated with $\nu_y(t)$ are lost from the beam.

c. Total loss rate

We now estimate the number $R(t)dt$ of random-walking tunes that (a) have encountered the point $8.6 - 7\xi_y$ at least once between time zero and time t, and (b) have not reached $8.6 - 4\xi_y$ between time t and the last encounter with $8.6 - 7\xi_y$, and (c) reach $8.6 - 4\xi_y$ between times t and t + dt (for small dt). Our estimate for the total rate at which protons are lost will then be five percent of our estimate of R.

For this purpose, let us divide the initial tune interval into two subintervals A and B, defined by

$$8.6 - 7\xi_y + \Delta_A \geq \nu_y \geq 8.6 - 7\xi_y \ , \tag{89a}$$

$$8.6 - 7\xi_y > \nu_y \geq 8.6 - 7\xi_y - \Delta_B \ , \tag{89b}$$

where the positive numbers Δ_A and Δ_B satisfy

$$\Delta_A + \Delta_B = \Delta_\nu \ . \tag{90}$$

The initial tune probability distribution $\tilde{\rho}$ takes the value $1/\Delta_\nu$ inside the union of these intervals, and zero outside.

Let us also decompose R into the infinite series

$$R = R_1^A + R_1^B + R_2^A + R_2^B + \ldots \ , \tag{91}$$

where $R_n^A(t)dt$ ($R_n^B(t)dt$) is the number of random-walking tunes, initially in the interval A(B), that (a) have encountered $8.6 - 7\xi_y$ precisely n times between times zero and t, and (b) have not reached $8.6 - 4\xi_y$ between time t and the last (n'th) such encounter with $8.6-7\xi_y$, and (c) reach $8.6- 4\xi_y$ between times t and t + dt. We shall see that it is easier to evaluate the constituent rates R_n^A and R_n^B than it is to evaluate their sum. In any case, we shall also see that the sum is dominated by its first two terms, at least for $t \lesssim$ a few tens of minutes.

Let us also define

$$P(\nu_o, \nu, t) \equiv \frac{1}{\sqrt{4\pi D_\nu t}} \exp\left[\frac{-(\nu-\nu_0)^2}{4D_\nu t}\right] . \tag{92}$$

As a function of ν, P solves the diffusion Eq. (83), with boundary condition

$$P(\nu_0,\nu,t = 0) = \delta(\nu-\nu_0). \tag{93}$$

Thus, $Pd\nu$ is the probability that a tune ν_y, located initially at $\nu_y = \nu_0$, will, because of random walk, lie in an interval of width $d\nu$ about $\nu_y = \nu$ at time t.

We now proceed to determine R.

The most easily constructed contribution to R is the first term, R_1^A, in the right-hand-side of (91). Because every tune initially in interval A must pass $8.6 - 7\xi_y$ before it encounters $8.6 - 4\xi_y$ even once, R_1^A is simply the rate at which tunes, initially in interval A, reach $8.6 - 4\xi_y$ for the first time, without ever having reached it before.

It is easy to convince oneself that this is equal to the rate at which tunes, initially in A, would be lost altogether if a fictitious, perfectly absorbing wall were placed at $\nu_y = 8.6 - 4\xi_y$. This is a productive way to rephrase the definition of R_1^A because the problem of random walk in the presence of an absorbing barrier is a standard one in the theory of probability.[100] Following the treatment of absorbers in Ref. 100, R_1^A is given by

$$R_1^A = \int_{8.6-7\xi_y}^{8.6-7\xi_y+\Delta_A} d\nu_0 [N\tilde{\rho}(\nu_0)] \left[2D_\nu \frac{\partial}{\partial\nu} P(\nu_0,\ \nu=8.6-4\xi_y,t)\right]$$

$$= \frac{-2ND_\nu}{\Delta\nu} \int_{8\cdot6-7\xi_y}^{8.6-7\xi_y+\Delta_A} d\nu_0 \left[\frac{\partial}{\partial\nu_0} P(\nu_0,\ 8.6-4\xi_y,t)\right]$$

$$= \frac{2ND_\nu}{\Delta_\nu} [P(8.6-7\xi_y,\ 8.6-4\xi_y,t) - P(8.6-7\xi_y+\Delta_A,\ 8.6-4\xi_y,t)]$$

$$= \frac{2ND_\nu}{\Delta_\nu} [P(-3\xi_y,0,t) - P(\Delta_A-3\xi_y,0,t)], \tag{94}$$

where N is the total number of particles in the beam.

We now proceed to the construction of R_1^B. For this purpose, note that every random walk that contributes to R_1^B at time t consists of these two parts: an initial trajectory that starts in B at time zero, and ends with its first arrival at $8.6 - 7\xi_y$ at some intermediate time; and a final trajectory that starts with $8.6 - 7\xi_y$ at the intermediate time and ends with its first arrival at $8.6 - 4\xi_y$ at time t. Note also that the same decomposition, characterized by the same length scales, applies to the paths that contribute to the rate at which diffusing tunes, initially in the fictitious interval

$$8.6 - 7\xi_y + \Delta_B \geq \nu_y > 8.6 - 7\xi_y \ . \tag{95}$$

encounter $8.6 - 4\xi_y$ for the first time, without having encountered it before. Thus, R_1^B and the latter rate are equal. Since the latter rate corresponds to the same kind of process to which R_1^A corresponds, we may evaluate it in the same way. The result is

$$R_1^B = \frac{2ND_\nu}{\Delta_\nu} [P(-3\xi_y,0,t) - P(\Delta_B - 3\xi_y,0,t)] \ . \tag{96}$$

In a similar fashion, one obtains

$$R_n^A = \frac{2ND_\nu}{\Delta_\nu} [P(-3\xi_y(2n-1),0,t) - P(\Delta_A - 3\xi_y(2n-1),0,t)] \ , \tag{97}$$

$$R_n^B = \frac{2ND_\nu}{\Delta_\nu} [P(-3\xi_y(2n-1),0,t) - P(\Delta_B - 3\xi_y(2n-1),0,t)] \ .$$

In order to estimate R numerically, let us assume that

$$(3\xi_y)^2/4\ D_\nu t \geq O(1) \ , \tag{98}$$

i.e., (using the data provided in subsection 2a) t ≤ a few tens of minutes. This guarantees that the exponents obtained when (92) is inserted into (97) are much larger for $n > 1$ than they are for $n = 1$, so that we may approximate R by $R_1^A + R_1^B$. If we also assume that both Δ_A and Δ_B, like Δ_ν, are much larger than $3\xi_y$, then in addition we may neglect the second terms in each of the formulae (94) and (96) for R_1^A and R_1^B. As a result of these assumptions, we have

$$R \sim R_1^A + R_1^B \sim \frac{2ND_\nu}{\Delta_\nu\sqrt{\pi D_\nu t}} \exp\ (-9\xi_y^2/4\ D_\nu t) \ . \tag{99}$$

Finally, following Month,[96,97] we estimate R by the maximum value of the approximation (99)

$$R \lesssim \frac{2}{3}\sqrt{\frac{2}{\pi \exp(1)}}\ \frac{ND_\nu}{\xi_y \Delta_\nu}\ \bar{N}(1.4\times 10^{-5})\sec^{-1}. \qquad (100)$$

(This maximum is attained when $(3\xi_y)^2/4\ D_\nu t = 1/2$, i.e., when $t \simeq 24$ minutes.) Thus, our estimate for the fractional beam loss rate is

$$-\frac{\dot{N}}{N} \lesssim (.05)\left(\frac{R}{N}\right) \lesssim 7 \times 10^{-6}\ \sec^{-1}$$

$$\lesssim 40 \text{ parts per million per minute} . \qquad (101)$$

This appears to be in rough agreement with the observed rates of 20-60 ppm per minute quoted in the prologue to this section.

It should be noted that although the bound (101) refers explicitly only to times comparable to half an hour, the observed loss rates quoted above can persist for longer than ten hours (see Fig. 35 for an example of relevant time-dependent experimental data). Presumably, this is due to the $R_n^{A,B}$ with $n>1$, which are not counted in (101).

3. SPEAR beam blowup

In this section, we shall attempt to estimate beam height at SPEAR by applying a (semiphenomenological) theoretical correction to results of the computer simulation[57] described in Chapter III, Section 2d. Our aim is to obtain more information about the role of resonance streaming in storage ring behavior. Accordingly, the correction, to be derived from Eq. (67), will be based on the assumption that all beam enlargement observed in the simulation is due to streaming up the resonance curve $(n_x,n_y,n) = (3,1,-21)$ that appears to determine the ordered structure seen in Fig. 30. As will be explained later in more detail, a correction of some sort is necessary in order to compensate for large distortions that are inherent in the particular flat-beam limit--$\sigma_y^* \to 0$, $\beta_y^* \to 0$, y/σ_y^* and β_y^*/σ_y^* fixed--in which the simulation was carried out.* Our beam height estimate relies on extrapolation from the results of a computer simulation because, as yet, there is no adequately self-contained theory of resonance streaming in the presence of noise.

The work of this section will be presented in two parts: In Subsection 3a, below, we discuss in detail how the results of our correction are to be compared with published experimental data from

*In this section, we continue to use the symbols σ_x^* and σ_y^* for the half-widths of the Gaussian strong beam at interaction points.

SPEAR. In Subsection 3b, we explain the flat-beam distortion and calculate the correction itself.

a. Scheme of comparison with experiment

The correction to be calculated in the next subsection will be applied only to flat-beam simulation data corresponding to weak/strong operation at 2.2 GeV per particle, and with horizontal and vertical strong-beam tuneshifts both equal to .06. These are operating conditions for which the simulation output made available[5,70] to me has been the most detailed. The corrected results will describe weak/strong operation at the same energy, with the same tuneshifts, and also with

$$r \equiv \frac{\sigma_y^*}{\sigma_x^*} = \frac{\beta_y^*}{\beta_x^*}\left(\frac{\beta_x^*\sigma_y^*}{\beta_y^*\sigma_x^*}\right) = \frac{\beta_y^*}{\beta_x^*}\frac{\xi_x}{\xi_y} = \frac{\beta_y^*}{\beta_x^*} \neq 0 \ . \qquad (102)$$

(In the third equality above, we have used Eq. (4) and its horizontal counterpart, together with the inequality $\sigma_y^* << \sigma_x^*$.)

The parameter r should be determined by the experimental data with which we want to compare the results of our calculation. However, both the choice of data for comparison, and the assignment of a specific value to r, are problematical, for the following reasons.

(i) Strictly speaking, there is no value of r for which the conditions enumerated above reproduce conditions under which published SPEAR data has been obtained, because all published SPEAR measurements have been made during strong/strong operation, with equal currents in the two beams. If the experimental observations had revealed that (at least for E = 2.2 GeV) one of the two strong beams was not significantly enlarged, then we might have been able to interpret the situation as effectively weak/strong. But, as explained in Chapter II, most SPEAR measurements are taken only after the RF phases have been adjusted so that both beams are blown up equally.

(ii) Moreover, even if weak/strong experimental data were available, a single value of r could not consistently be assigned because the dimensions of unperturbed beams at SPEAR satisfy $\sigma_y^*/\sigma_x^* \simeq 1/30$, while the beta functions satisfy $\beta_y^*/\beta_x^* \simeq 1/12 \simeq (2.5)\sigma_y^*/\sigma_x^*$, contradicting (102).

Our strategy for dealing with these problems is as follows:

(i) There is some indication[17] that if the RF phases are readjusted so that the enlargement of one of two strong SPEAR beams is eliminated, then the vertical dimension of the other SPEAR beam is between one and two times as great as it is when (at the same energy and current) the two beam sizes are matched (see footnote next page). Accordingly, we shall guess that at 2.2 GeV, with a strong-beam tuneshift of $\xi_y = .06$, the height of a weak beam at SPEAR lies

somewhere between these two values: On the low side, the height that either one of two matched 2.2 GeV beams has when the current is such that the vertical tuneshift due to either beam would be .06 if it were not blown up; and on the high side, twice this low value.

Let us determine these values from the SPEAR data shown in Fig. 1. (Although the energy--1.94 GeV per particle--at which this data was taken is not exactly equal to the energy--2.2 GeV--with which we are immediately concerned, it should be close enough for the rough estimates that we formulate here.) We begin by observing that Fig. 1 has been marked to indicate that the tuneshift .023 corresponds to a current--2.2 mA--at which luminosity follows the low-current I^2 law, and therefore at which neither matched beam is blown up. Since (following (4)) tuneshift would be proportional to the first power of current if the beam dimensions were not to blow up, the current that a beam would have if it gave rise to a vertical tuneshift of .06 without being blown up is (2.2 mA)·(.06/.023) 5.7 mA. From the ratio of unperturbed quadratic luminosity to the actual luminosity shown in Fig. 1 (i.e., according to Eq. (2), from the ratio of actual area to unperturbed area, recalling that horizontal blowup is negligible) one finds that the matched-beam vertical blowup at 5.7 mA is about 1.6. Thus, we shall guess that at $\xi_y = .06$ and $E = 2.2$ GeV, the vertical dimension of a weak beam at SPEAR is between about 1.6 and 3.2 times as great as it is when the strong beam is absent (see first footnote next page).

(ii) As for r: In the next subsection, we shall present the results of our correction for all values of r between the actual value, 1/30, of σ_y^*/σ_x^*, and the actual value, 1/12, of β_y^*/β_x^*. We will be guessing (for lack of a better procedure) that if we could properly extrapolate the simulation output to $\beta_y^*/\beta_x^* = (2.5)\sigma_y^*/\sigma_x^*$, the resulting blowup would lie somewhere between these two extremes.

b. Distortion and correction

In order to understand how the flat beam limit can distort beam enlargement associated with resonance streaming, let us first rewrite the fundamental streaming equation (67) in terms of the normalized amplitudes a_x and a_y, because, unlike the I's, they remain finite in the limit $\sigma_y^* \to 0$, $\beta_y^* \to 0$, y/σ_y^* fixed, β_y^*/σ_y^* fixed. The result is

$$P_x\left(\frac{a_{xr}^2\sigma_x^{*2}}{2\beta_x^*}\right) + P_y\left(\frac{a_{yr}^2\sigma_y^{*2}}{2\beta_y^{*2}}\right) = \text{constant} \times e^{-2\gamma s}\ , \qquad (103)$$

*One reaches a stronger conclusion from data generated by the LEP simulation discussed in Chapter III. Figure 36 (reproduced from Ref. 31) shows how the height (normalized to the unperturbed height) of a simulation beam varies as a function of the beam's current, when the current in the opposing beam is held fixed. One sees that when one beam has very low current, its height is about twice the height of either beam when the two currents are matched. (This footnote refers to preceding page.)

where $\vec{p}$ is orthogonal to the vector $\vec{n}$ that, together with the number n, defines the resonance curve on which $\vec{a}_r$ is constrained to lie. In the flat-beam limit, the second term on the left-hand side of (103) vanishes, so that this equation may be simplified further, becoming

$$a_{xr}(s) = a_{xr}(s=0) \cdot e^{-\gamma s} . \qquad (104)$$

We can now see that the flat beam limit distorts blowup when there are resonance curves--such as the one associated with $(n_x,n_y,n) = (3,1,-21)$ in Fig. 33--that are nearly vertical. In such cases, when large changes in a_{yr} accompany small changes in a_{xr}, a streaming equation that, like (104), makes no explicit reference to a_{yr} describes test particles that move very rapidly to very large vertical amplitudes, dominating the averages that determine beam heights.

In order to predict, for the particular case of interest here, how this distortion is reduced when σ_y^* and β_y^* become non-zero, we shall have to assume that a number of things (in the model used in the simulation) which at present can only be determined by computer, change very little when r departs very slightly from zero, with E,ξ_x and ξ_y fixed at 2.2 GeV, .06, and .06 respectively. They are

-- The location of the (3,1,-21) resonance curve in the $\vec{a}$ plane. When r = 0, this is, approximately, the vertical line $a_{xr} = 2.8$, according to Fig. 33.

-- The percentage of all test particles that are streaming up the (3,1,-21) resonance curve at any moment. When r = 0, this is[70] about 5%.

-- The time that a typical streaming particle actually spends streaming up the (3,1,-21) resonance curve in any one stretch. When r = 0, this is[70] about one third of the transverse damping time.

Let us now use these assumptions to estimate weak beam height as a function of r.

We begin by computing the largest value of a_{yr} typically reached by a particle that streams up the (3,1,-21) resonance curve. To do this, we argue as follows: Let $\vec{a}^i$ and $\vec{a}^f$ be the points at which the particle begins and ends, respectively, an episode of streaming. According to the assumptions above, $a_x^i \simeq a_x^f \simeq 2.8$. We shall set $(a_y^i)^2 \sim 2$** because we expect that when a particle first gets caught by the resonance, its vertical amplitude is typical of the

*Probably closer to 3.2 than to 1.6, if one can generalize from the LEP simulation. (This footnote refers to preceding page.)

**In Fig. 30, a_y^i is approximately equal to 4, not to $\sqrt{2}$. However, in view of unpublished data[70] that we shall not reproduce here, this appears to be an exception. In any case, our final upper bound (following inequality (107)) for SPEAR blowup will not be very sensitive to a_y^i, even for a_y^i as large as 4.

unperturbed dynamics (there is no reason for it to be otherwise, when one assumes, as we do, that the main effect of the strong beam is localized in the $\vec{p} = (3,1,-21)$ resonance curve). To determine a_y^f, we use Eq. (103) (with $\vec{p} = (1,-3)$) and the assumption that the particle streams only for a time interval Δt satisfying $\gamma c \Delta t \simeq 1/3$:

$$\frac{(2.8)^2\sigma_x^{*2}}{2\beta_x^*} - \frac{3(a_y^f)^2\sigma_y^{*2}}{2\beta_y^*}$$

$$= (e^{-2/3})\left[\frac{(2.8)^2\sigma_x^{*2}}{2\beta_x^*} - \frac{3\cdot 2\cdot\sigma_y^{*2}}{2\beta_y^*}\right] . \qquad (105)$$

I.e., using $e^{-2/3} \simeq 1/2$,

$$(a_y^f)^2 \simeq 1 + (1.2)\left(\frac{\sigma_x^*}{\sigma_y^*}\right)^2\left(\frac{\beta_y^*}{\beta_x^*}\right)$$

$$= 1 + (1.2)/r. \qquad (106)$$

(It should be pointed out that the apparent singularity at $r = 0$ is an artificial consequence of our having assumed that the resonance curve is exactly vertical. A more careful analysis[107] reveals that the $1/r$ singularity washes out when r becomes comparable to the square of the typical true slope ($\simeq 1/20$ from Fig. 33) of the resonance curve, i.e., $r \simeq 1/400$.) According to (106), $(a_y^f)^2$ lies between about 15 and 37 when r lies between about 1/30 and 1/12.

Finally, to obtain the contribution of resonance streaming to the rms weak beam blowup itself, we argue as follows: According to Eq. (30), the mean square blowup (i.e., the mean square weak beam height, normalized to $(\sigma_y^*)^2$) is equal to one-half the mean square value of a_y. According to the assumptions, the mean value of a_y^2 should be close to its unperturbed value (=2) for 95% of the test particles, because, at any time, 95% of all particles are not participating in streaming. The mean value of a_y^2 for the remaining 5% should be somewhere between two and $(a_y^f)^2$. Thus we have

$$\langle y^2\rangle^{1/2}/\sigma_y^* \lesssim [(95\%)\cdot 1 + (5\%)\cdot 1/2\cdot(a_y^f)^2]^{1/2} . \qquad (107)$$

When r lies between 1/30 and 1/12, the right-hand side of (107) lies between about 1.2 and 1.4 (a significant reduction from the

blowup of about 4.5, observed in the simulation[70] when r = 0.) Note that this interval does not overlap the interval [1.6, 3.2], towards the high end of which we guessed that the experimental blowup would lie. This suggests that resonance streaming does not make the dominant contribution to the beam blowup seen in real electron-positron colliders.

Acknowledgements

I am grateful to my wife Jonina for her patience and encouragement.

I wish to thank J. Bjorken, L. Durand, D. Edwards, M. Month, and A. Ruggiero for reviewing the manuscript. I also wish to thank Sue Grommes for typing the manuscript with accuracy and speed.

Part of this work was carried out at the Aspen Center for Physics.

REFERENCES

I. M. Month and J. C. Herrera, eds., Nonlinear Dynamics and the Beam-Beam Interaction, AIP Conference Proceedings no. 57.

II. Proceedings of the Beam-Beam Interaction Seminar, Stanford Linear Accelerator Center, May 22-23, 1980, SLAC-PUB-2624.

III. Proc. 11th Int. Conf. High Energy Accelerators, Geneva, 1980 (Birkhauser, Basel, 1980).

IV. Proc. Int. Conf. High Energy Accelerators, SLAC, 1974.

V. R. A. Carrigan, Jr., F. R. Huson, and M. Month, eds., Physics of High Energy Particle Accelerators (Fermilab Summer School, 1981), AIP Conference Proceedings no. 87.

VI. M. H. Blewett, ed., Proc. Int. School of Particle Accelerators, Erice, 1976, CERN 77-13.

1. CERN Courier, May 1981, pp. 143-144.

2. R. Donaldson, R. Gustafson, and F. Paige, eds., Proc. 1982 DPF Summer Study on Elementary Particle Physics and Future Detectors, Snowmass, Colorado, June 28-July 16, 1982. Spec., contributions of P.M. Tuts et al., pp. 62-64; F. Pauss and P. M. Tuts, pp. 411-416; and S. Herb, pp. 417-419.

3. J. F. Schonfeld, in V, pp. 314-344.

4. M. Sands, in Physics with Intersecting Storage Rings, B. Touschek, ed. (Academic Press, New York, 1971) pp. 257-411.

5. J. L. Tennyson, in V, pp. 345-394.

6. J. C. Herrera and M. Month, in I, pp. 1-28.

7. J. C. Herrera, in I, pp. 29-41.

8. P. Morton, private communication.

9. Proceedings of the Workshop on Polarized Electron Acceleration and Storage March 22-27, 1982, Hamburg, DESY publication M-82/09.

10. H. Wiedemann, in II, pp. 33-39.

11. J. Seeman, Cornell Laboratory of Nuclear Studies preprint 82/531.

12. J. Rees, IEEE Trans. Nucl. Sci., NS-28, 1989 (1981).

13. A. Piwinski, DESY preprint 81-066.

14. J. D. Bjorken and S. D. Drell, Relativistic Quantum Mechanics (McGraw-Hill, New York, 1964).

15. E. D. Courant et. al., D. C. Pellegrini, ed., Proceedings of the 1979 Workshop on Beam Current Limitations in Storage Rings, Brookhaven National Laboratory, July 16-27, 1979, pp. 16-20.

16. H. Wiedemann, in I, pp. 84-98.

17. M.H.R. Donald and J. M. Paterson, IEEE Trans. Nucl. Sci. NS-26, 3580 (1979).

18. S. Tazzari, in I, pp. 128-135.

19. M. Cornacchia, in I, pp. 84-98.

20. PETRA Storage Ring Group, in III, pp. 26-27.

21. H. Wiedemann, private communication.

22. A. Piwinski, in I, pp. 115-127.

23. H. Zyngier, in I, pp. 136-142.

24. F. Amman and D. Ritson, in Proc. Int. Conf. High Energy Accelerators, Brookhaven, 1961, pp. 471-475.

25. E. D. Courant, IEEE Trans. Nucl. Sci. NS-12, 550 (1965).

26. B. Richter, in Proc. Int. Symp. Electron and Positron Storage Rings, Saclay, 1966, pp. (I-1-1) - (I-1-24).

27. CESR Design Report, Cornell Lab. for Nuclear Studies, Ithaca (1977).

28. PEP Conceptual Design Report, LBL-4288, SLAC-189 (T/E/A) UC-28 (1976).

29. Design Study of a 22 to 130 GeV e^+e^- Colliding Beam Machine (LEP), CERN/ISR-LEP/79-33 (1979).

30. U. Amaldi, in R. A. Carrigan, Jr., F. R. Huson and M. Month, eds., The State of Particle Accelerators and High Energy Physics (Fermilab Summer School, 1981), AIP Conference Proceedings no. 92, pp. 79-100.

31. S. Myers, CERN preprint CERN-ISR/RF/82-06, LEP NOTE 362 (1982).

32. J.-E. Augustin et. al., in Proc. Workshop on Possibilities and Limitations of Accelerators and Detectors, Fermilab, 1979, pp. 87-105.

33. J. Rees, private communication.

34. H. Wiedemann, unpublished.

35. S. Kheifets, in C. W. Horton, Jr., L. E. Reichl, and V. G. Szebehely, eds., Long-Time Prediction in Dynamics (Wiley, New York, 1983), pp. 397-426.

36. E. Keil and R. Talman, CERN-ISR-TH/81-33 (Rev.) (1982).

37. L. Evans and J. Gereyte, CERN SPS/82-8 (DI-MST).

38. A. Chao, in I, pp. 42-68.

39. J. LeDuff, in VI, pp. 377-410.

40. E. Keil, Nucl. Instr. and Meth. 188, 9 (1981).

41. J.-Y. Hemery et. al., IEEE Trans. Nucl. Sci., NS-28, 2497 (1981).

42. J.-P. Koutchouk, ISR Performance Report, Runs 1217 and 1221 (19 March, 1982).

43. J. LeDuff, in II, pp. 80-98.

44. K. Johnsen, in IV, pp. 32-36.

45. G. Guignard, in I, pp. 69-83.

46. J. P. Gourber, H. C. Hereward and S. Myers, IEEE Trans. Nucl. Sci. NS-24, 1405 (1977).

47. J.-P. Koutchouk, private communication.

48. K. R. Symon and A. M. Sessler, in E. Regenstreif, ed., Proc. CERN Symp. on High Energy Acceleration and Pion Physics, CERN, 1956, pp. 44-58.

49. L. Evans and J. Gareyte, CERN report SPS/82-8 (DI-MSI) (1982).

50. K. Hübner in IV, pp. 63-65.

51. A. Hofman, L. Vos and B. Zotter, in III, pp. 713-717.

52. B. Zotter, in Proc. Int. Conf. High Energy Accelerators, Protvino (USSR), 1977, pp. 23-28.

53. E. Keil and G. Leroy, IEEE Trans. Nucl. Sci. NS-22, 1370 (1975).

54. J. P. Delahaye et. al., SPS Improvement Report No. 180 (1980), as reported by D. Boussard et. al., in III, pp. 627-632.

55. E. Keil, in III, p. 770.

56. A. Hutton, private communication.

57. J. L. Tennyson, in II, pp. 1-20.

58. A. Piwinski, in III, pp. 751-753.

59. A. Piwinski, DESY preprint 80/131.

60. A. Piwinski, IEEE Trans. Nucl. Sci. NS-28, 2440 (1981).

61. A. Piwinski, DESY Internal Report M-81/31.

62. D. Degèle et. al., DESY Internal Report M-81/03.

63. S. Peggs and R. Talman, Phys. Rev. D24, 2379 (1981).

64. K. Brown, private communication.

65. E. Close, in I, pp. 210-221.

66. K. W. Robinson and G. A. Voss, Cambridge Electron Accelerator report CEAL-1029 (1966).

67. D. Neuffer, private communication.

68. Design Report, Tevatron Phase 1 Project, Femilab, Batavia (1980).

69. R. Talman, private communication.

70. J. Tennyson, private communication.

71. J. Tennyson, Physica 5D, 123 (1982).

72. D. Neuffer, A. Riddiford, and A. Ruggiero, Fermilab Report FN-359 (1982).

73. D. Neuffer, A. Riddiford, and A. Ruggiero, IEEE Trans. Nucl. Sci. NS-28, 2494 (1981).

74. D. Neuffer, A. Riddiford, and A. Ruggiero, Fermilab Report FN-363 (1982).

75. J. Tennyson, in I, pp. 158-193.

76. B. Chirikov, Phys. Repts. 52, 263 (1979).

77. D. Neuffer, A. Riddiford, and A. Ruggiero, Fermilab Report TM-1007 (1980).

78. D. Neuffer, A. Riddiford, and A. Ruggiero, Fermilab Report FN-357 (1982).

79. A. Riddiford, private communication.

80. A. Ruggiero, private communication.

81. M. Marx, in Proc 1977 ISABELLE Summer Study, Brookhaven report BNL 50721, pp. 287-290.

82. A. Ruggiero, in IV, pp. 419-423, and references therein.

83. K. Takayama, Lett. Nuovo Cimento 34, 190 (1982).

84. A. Chao, in I, pp. 42-68.

85. E. Keil, in III, pp. 759-762.

86. G. Guignard, IEEE Trans Nucl. Sci. NS-28, 2506 (1981).

87. S. Kheifets, IEEE Trans. Nucl. Sci. NS-26, 3615 (1979).

88. A. Ruggiero, unpublished.

89. H. S. Uhm and C. S. Liu, Phys. Rev. Lett. 43, 914 (1979); and in II, pp. 114-143.

90. N. S. Dikansky and D. V. Pestrikov, Particle Accelerators 27, 12 (1982).

91. B. W. Montague, in III, pp. 731-735.

92. E. D. Courant and H. S. Snyder, Ann. Phys. 3, 1 (1958).

93. G. Guignard, CERN 78-11 (1978).

94. H. Goldstein, Classical Mechanics (Addison-Wesley, Reading, Mass., 1950).

95. A. W. Chao and M. Month, Nucl. Instr. and Meth. 121, 129 (1974).

96. M. Month, in IV, pp. 402-404.

97. M. Month, IEEE Trans. Nucl. Sci. NS-22, 1376 (1975).

98. A. Piwinski, in IV, pp. 405-409.

99. J. D. Bjorken and S. Mtingwa, FERMILAB-Pub-82/47-THY.

100. S. Chandrasekhar, Rev. Mod. Phys. 15, 1 (1943).

101. F. Amman, IEEE Trans. Nucl. Sci. NS-20, 858 (1973).

102. P. Marin, in III, pp. 742-743.

103. S. Tazzari, IEEE Trans. Nucl. Sci. NS-28, 2420 (1981).

104. E. Keil, in V, pp. 805-862.

105. F. M. Izraelev, Physica 1D, 243 (1980).

106. A. Hutton, PEP-NOTE-375 (1982).

107. J. F. Schonfeld, unpublished.

108. R. Helm et. al., SLAC-PUB-3070, contributed to the Particle Accelerator Conference, Santa Fe, March 21-23, 1983.

109. J. M. Paterson, private communication.

110. L. Evans and J. Gareyte, private communication.

111. J. Buon, in IV, pp. 83-86.

112. F. Amman, in Proc. Int. Conf. High Energy Accelerators, CERN, 1971, pp. 63-71.

113. L. C. Teng, in II, pp. 99-113.

FIGURE CAPTIONS

Fig. 1. Luminosity vs. current per beam at SPEAR, from Ref. 10. The letter ξ refers to ξ_y.

Fig. 2. Luminosity vs. current per beam at CESR for four values of β_y^*, from Ref. 11.

Fig. 3. Luminosity vs. current per bunch at PEP, from Ref. 12. The symbols β_x and β_y refer to β_x^* and β_y^*.

Fig. 4. Specific luminosity vs. current per bunch at PETRA, for several energies, from Ref. 13.

Fig. 5. Television photographs of beam cross-sections at SPEAR, from Ref. 16.

Fig. 6a. Horizontal (R, for "radial") and vertical (V) beam density profiles at ADONE, for beam current below the blowup threshold, from Ref. 18.

Fig. 6b. Same as in 6a, but for beam current above the blowup threshold.

Fig. 7. Maximum vertical beam-beam tuneshift vs. energy at SPEAR, from Ref. 10.

Fig. 8. Vertical beam-beam tuneshift vs. current per bunch at PETRA, for several energies, from Ref. 13. The subscript z corresponds to y in the present review.

Fig. 9. Phenomenological fit (solid line) to experimental data from various e^+e^- storage rings. Redrawn from unpublished plot due to H. Wiedemann (1980).

Fig. 10. Schematic drawing of the DCI system, from Ref. 43. The crosses indicate collision points.

Fig. 11. Results of several experiments performed at DCI, from Ref. 43. The parameter $\bar{\nu}$ refers to the (equal, in this case) fractional parts of the unperturbed storage ring tunes: $\nu_x = 3.73$, $\nu_y = 1.73$.

Fig. 12. Schematic representation of an ISR beam (a) as viewed head-on, and (b) as viewed at a nonzero crossing angle.

Fig. 13. Results of an ISR overlap-knockout experiment, from Ref. 46. I_{beam} refers to the current in a coasting beam. $(\Delta p/p)_b$ refers to the fractional difference between the momentum of a counter-rotating bunched beam and some reference

momentum. $(\Delta p/p)_b$ decreases with time; the absolute value of its time derivative is indicated in the figure. $(\Delta p/p)_p$ is the fractional momentum spread in the coasting beam. V_{RF} is the peak voltage in the bunching RF cavities. Q is the unperturbed vertical storage ring tune. The energy is not specified.

Fig. 14. Antiproton beam lifetime vs. unperturbed horizontal tune at the SPS, from Ref. 49.

Fig. 15. ISR current loss rate and beam lifetime vs. vertical beam-beam tuneshift, as measured during a high-β_y^* experiment, from Ref. 52. The subscript z is as in Fig. 8.

Fig. 16. ISR current loss rate vs. vertical beam-beam tuneshift, as measured during a nonlinear-lens experiment, from Ref. 53. The subscript v is as in Figs. 6.

Fig. 17. Schematic representation of a head-on collision of two identical bunches. The points c_1 and c_2 are the centers of the two colliding bunches; c_0 is midway between the two bunches; p lies in the horizontal midplane of bunch no. 1, and is located a distance ℓ in advance of the bunch center c_1. (a) shows the bunches before contact. (b) shows that the point p reaches the center of bunch no. 2 when it is a distance $\ell/2$ (or a time $\ell/2c$) beyond the collision center c_0.

Fig. 18. Luminosity vs. current per beam at LEP, according to numerical simulation, from Ref. 31.

Fig. 19. Luminosity and vertical beam-beam tuneshift vs. vertical damping time at LEP, according to numerical simulation, from Ref. 31. The subscript z is as in Fig. 8.

Fig. 20. Luminosity vs. unperturbed vertical tune at LEP, according to numerical simulation, from Ref. 31. Q_{xo} is the unperturbed horizontal tune. K_B is the number of bunches per beam. ξ_{xo} and ξ_{zo} are the horizontal and vertical tuneshifts that either beam would generate if the opposing beam were absent. The subscript z is as in Fig. 8.

Fig. 21. LEP luminosity vs. root-mean-square irregularity in horizontal tune per interacton region, according to numerical simulation, from Ref. 31. The symbol μ_x used in this figure is equivalent to 2π times the μ_x used in the present review.

Fig. 22. LEP luminosity vs. vertical interaction-point beta function (measured in meters), for different bunch lengths, according to numerical simulation, from Ref. 31. "Bunch lengthening factor = n" means that the bunch length is equal to (1.2)n cm. The expected LEP bunch lengthening factor is four.[31] The subscript z is as in Fig. 8.

Fig. 23. Luminosity vs. current per beam at CESR according to numerical simulation, from Ref. 63.

Fig. 24. Contour plot of luminosity vs. the fractional parts of unperturbed tunes at CESR, according to numerical simulation, from Ref. 63. The numerals along the contours indicate luminosity values, in units of 10^{30} $cm^{-2}sec^{-1}$. Other markings in this figure are explained in Ref. 63.

Fig. 25. Weak beam height (normalized to unperturbed beam height) vs. vertical unperturbed storage ring tune at PETRA, according to numerical simulation of weak-strong operation, for various energies, from Ref. 60. "4IP" means 4 interaction points (C = 4). The subscript z is as in Fig. 8.

Fig. 26. Weak beam height (normalized to unperturbed beam height) vs. vertical unperturbed storage tune at PETRA, according to numerical simulation of weak-strong operation, for various tuneshifts, from Ref. 60. "4IP" is as in Fig. 25. The subscript z is as in Fig. 8. Q_s refers to the tune of synchrotron oscillations.

Fig. 27. Weak beam height (normalized to unperturbed beam height) vs. unperturbed storage ring tunes at PETRA, according to numerical simulation of weak-strong operation, with (right) and without (left) dispersions and tune irregularities, from Ref. 61. The subscript z is as in Fig. 8. Q_s is as in Fig. 26. The energy is 17.9 GeV per particle, the strong-beam tuneshift is .04, and C = 4.

Fig. 28. Luminosity vs. current per bunch as measured at PETRA (strong-strong) under three different operating conditions, from Ref. 60. Curve (a) corresponds to an unperturbed vertical tune of 23.3. Curve (b) corresponds to an unperturbed vertical tune of 23.1. Curve (c) corresponds to $\nu_y = 23.1$ and, in addition, to a special reduction of η_x^*'s. In all cases, $\nu_x = 25.2$.

Fig. 29. Weak beam height (normalized to unperturbed beam height) vs. vertical strong beam tuneshift at SPEAR, according to numerical simulation of weak-strong operation, from Ref. 57. The parameters D_x and D_y are explained in Ref. 57.

Fig. 30. A streaming trajectory from the SPEAR simulation of Ref. 57. Reproduced from Ref. 5.

Fig. 31. Orbits in an ISABELLE simulation, from Ref. 75.

Fig. 32. Schematic representation of fourth-order stable resonant trajectories (large dots) and separatrices (football-shaped enclosures) at a single time, in a system with

one degree of freedom. The long arrow indicates that the trajectories and separatrices revolve about the origin as time progresses; the short arrow indicates that phase space flows around the resonant trajectory within each resonant region. Revolution of resonant trajectories is clockwise (counterclockwise) if dH_0/dI_{yr} is positive (negative); flow in resonant regions is clockwise (counterclockwise) if d^2H_0/dI_{yr}^2 is positive (negative).

Fig. 33. Resonance curves from the SPEAR model used in the simulation of Ref. 57, from Ref. 5. (Note: Based on a comparison between this figure and Fig. 6 of Ref. 57, it seems to me that the labels "$2\nu_y + 6\nu_x = 2$" and "$10\nu_y = 2$" should be replaced by "$2\nu_y + 6\nu_x = 5$" and "$10\nu_y = 6$," respectively. The variables ν_x and ν_y used in this Figure are what we would call $\mu_x - 2$ and $\mu_y - 2$ in the present report. In terms of our usage of the symbols ν_x and ν_y ($\nu_x = 2\mu_x$ and $\nu_y = 2\mu_y$ at SPEAR), these replacements are equivalent to $\nu_y + 3\nu_x = 21$ and $5\nu_y = 26$, respectively. These replacements are assumed in the text.)

Fig. 34. $(16\pi\xi_y)^{-1}\, dF_{0,0}/dI_y$ ($D(\sigma)$ in the figure) vs. $\sqrt{I_y/I_y^*}$ (σ in the figure), for the ISR model discussed in Chapter V, from Ref. 97.

Ref. 35. Current loss rates (and beam height) vs. time near a fifth-order resonance in the ISR, from Ref. 45.

Fig. 36. Heights (normalized to unperturbed height) of two colliding LEP beams vs. the current of one of them (A), the current of the other (B) being fixed, according to numerical simulation, from Fig. 31. All fixed parameters--except horizontal and longitudinal damping times--assumed in this graph are identical to those that correspond to the luminosity maximum in Fig. 20. The horizontal damping time is half that used in generating Fig. 20; the longitudinal damping time is 25% greater than that used in generating Fig. 20.

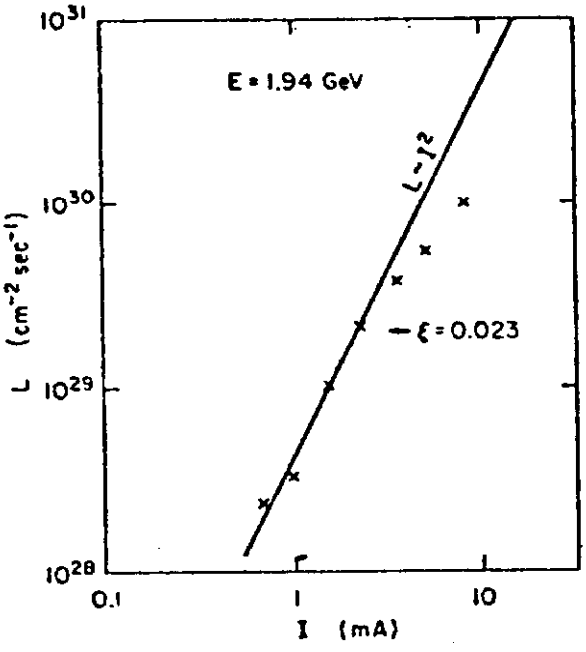

10^{31}
E = 1.94 GeV
$L \sim I^2$
10^{30}
L $(cm^{-2} sec^{-1})$
$\xi = 0.023$
10^{29}
10^{28}
0.1
1
10
I (mA)

Fig. 1

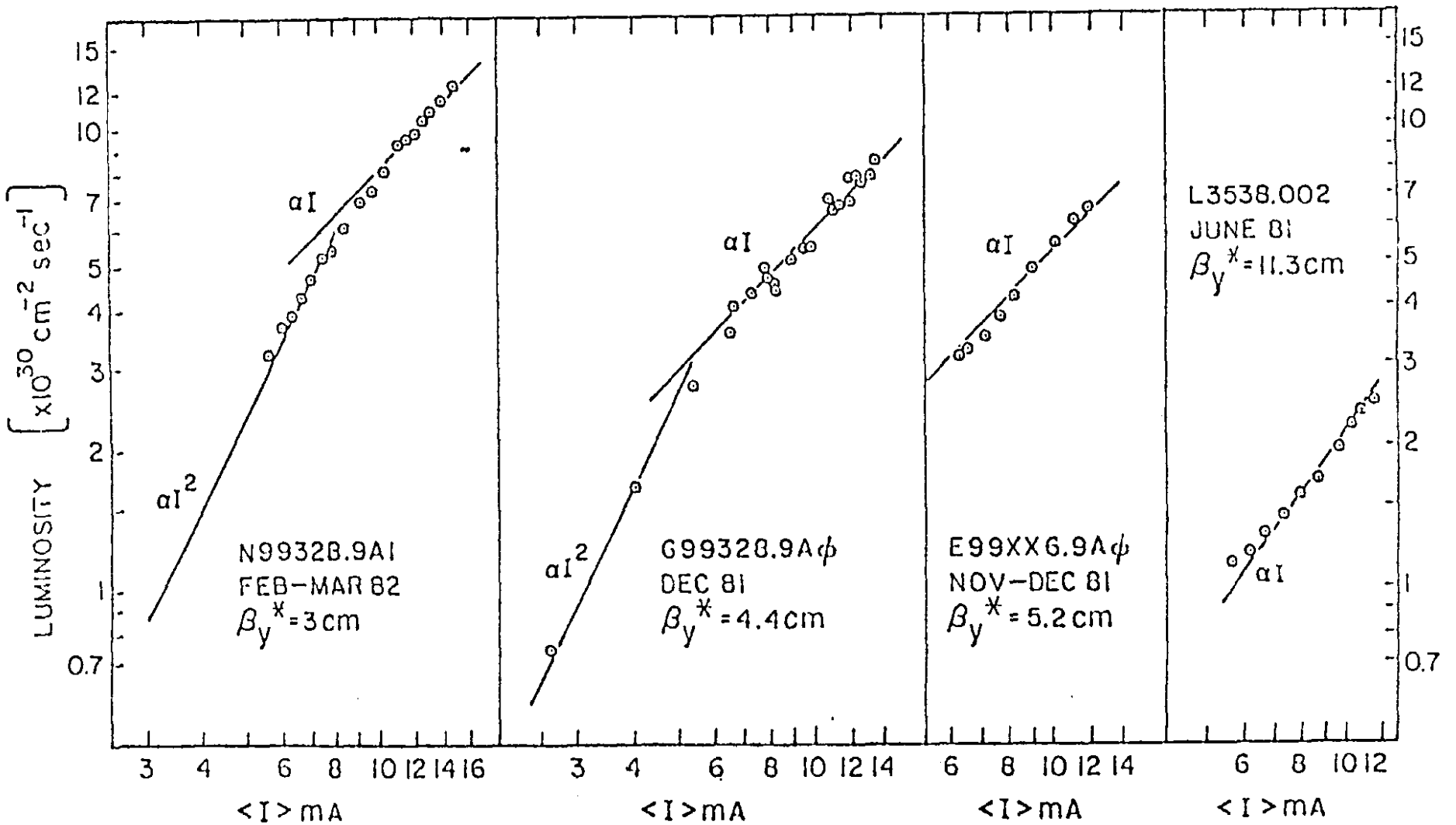

LUMINOSITY [x10^30 cm^-2 sec^-1]
aI
aI^2
N9932B.9AI
FEB–MAR 82
β_y* = 3cm
G9932B.9Aφ
DEC 81
β_y* = 4.4cm
E99XX6.9Aφ
NOV–DEC 81
β_y* = 5.2cm
L3538.002
JUNE 81
β_y* = 11.3cm
<I> mA

Fig. 2

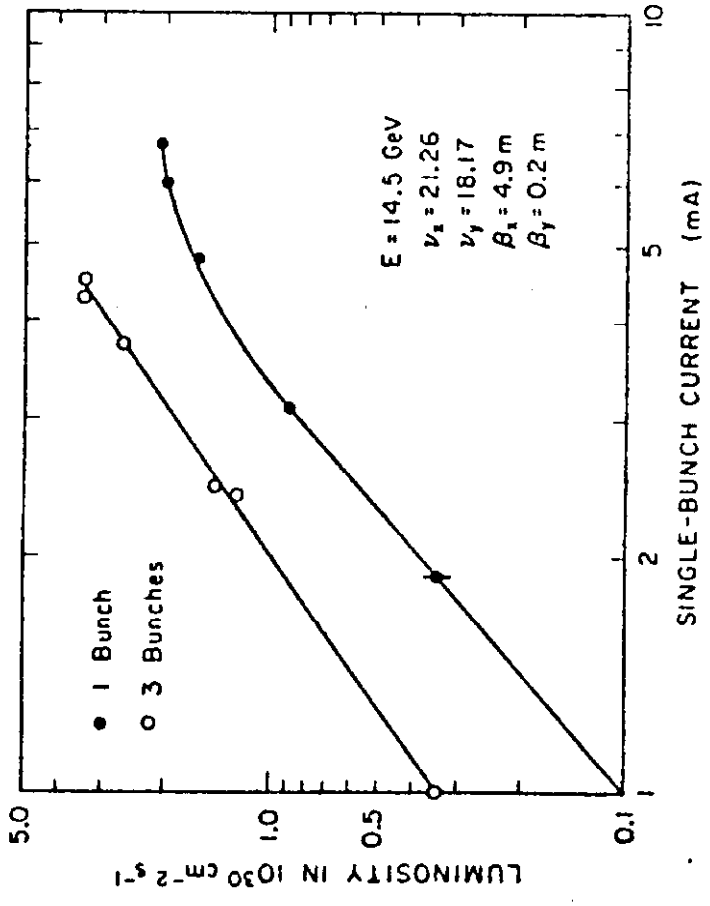

LUMINOSITY IN 10^{30} cm^{-2} s^{-1}
5.0
1.0
0.5
0.1
1
2
5
10
SINGLE-BUNCH CURRENT (mA)
E = 14.5 GeV
ν_x = 21.26
ν_y = 18.17
β_x = 4.9 m
β_y = 0.2 m
● 1 Bunch
○ 3 Bunches

Fig. 3

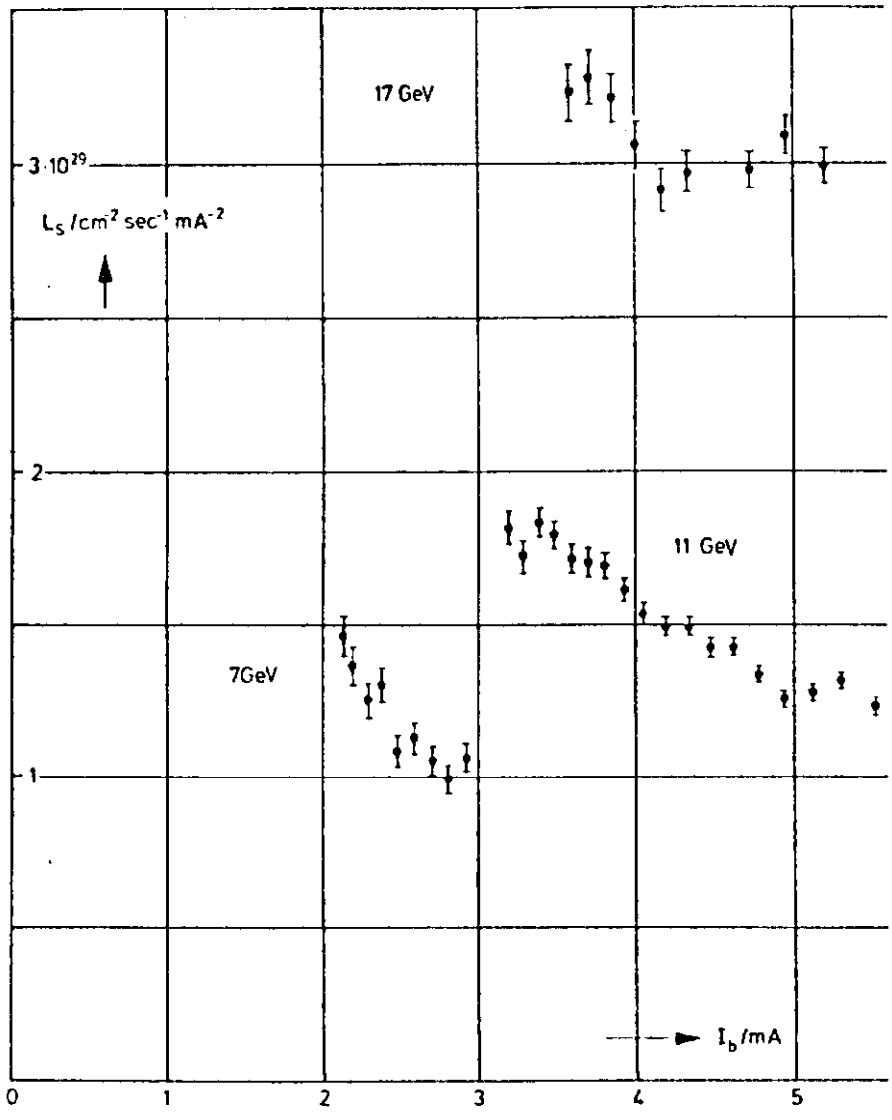
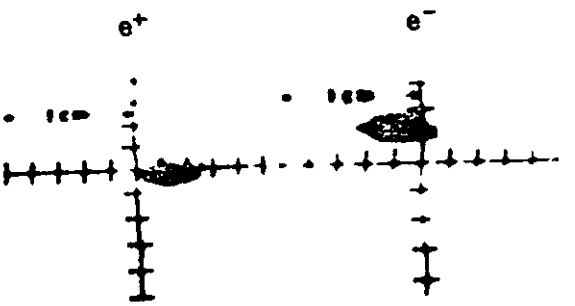
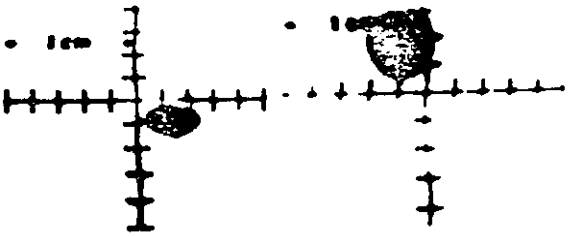


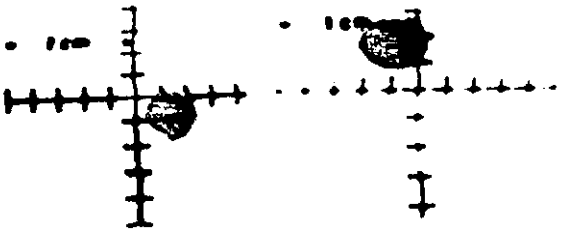
Fig. 4



(a) Separated Beams



(b) Colliding Beams With Flip Flop Effect



(c) Colliding Beams Flip Flop Balanced

Fig. 5

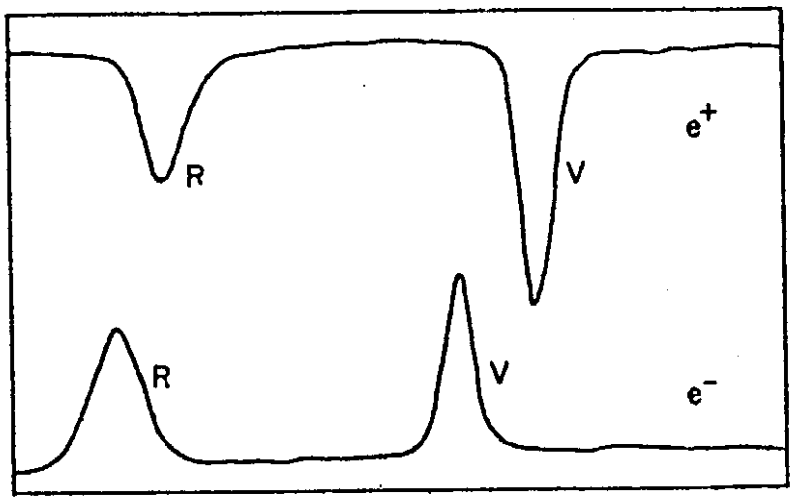


Fig. 6a

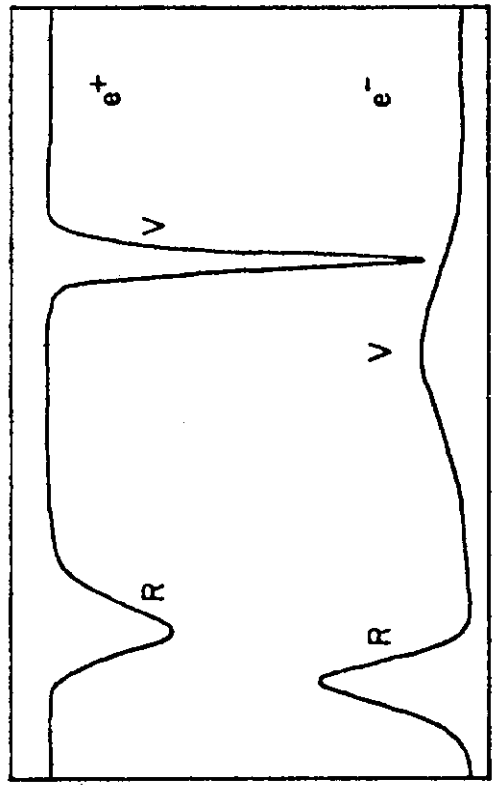


Fig. 6b

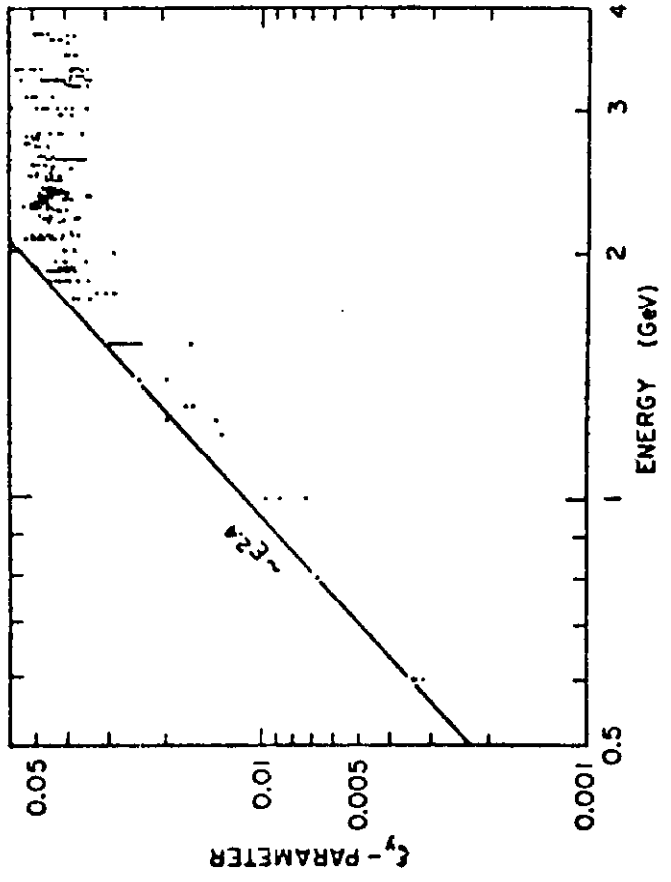

0.05
0.01
0.005
0.001
0.5
1
2
3
4
ENERGY (GeV)
ϵ_γ - PARAMETER
$E^{-2.3}$

Fig. 7

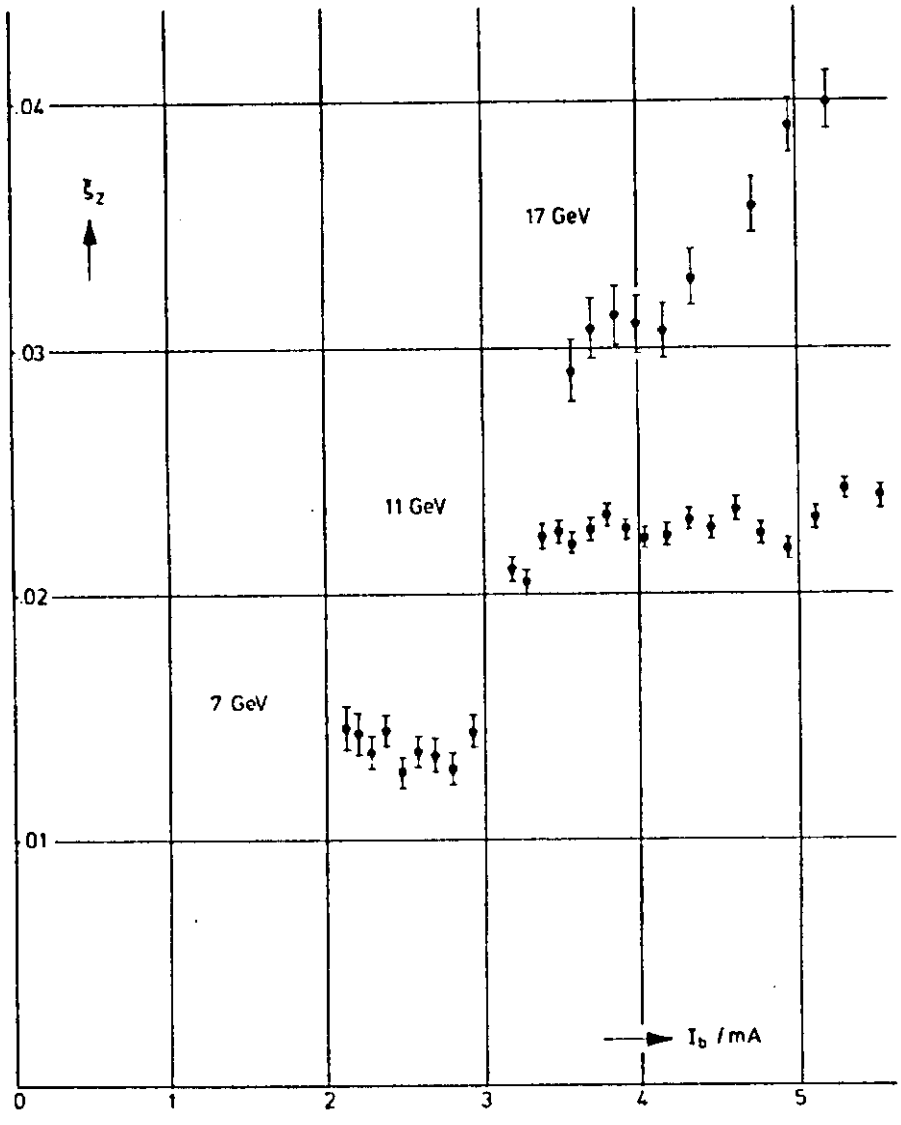


Fig. 8

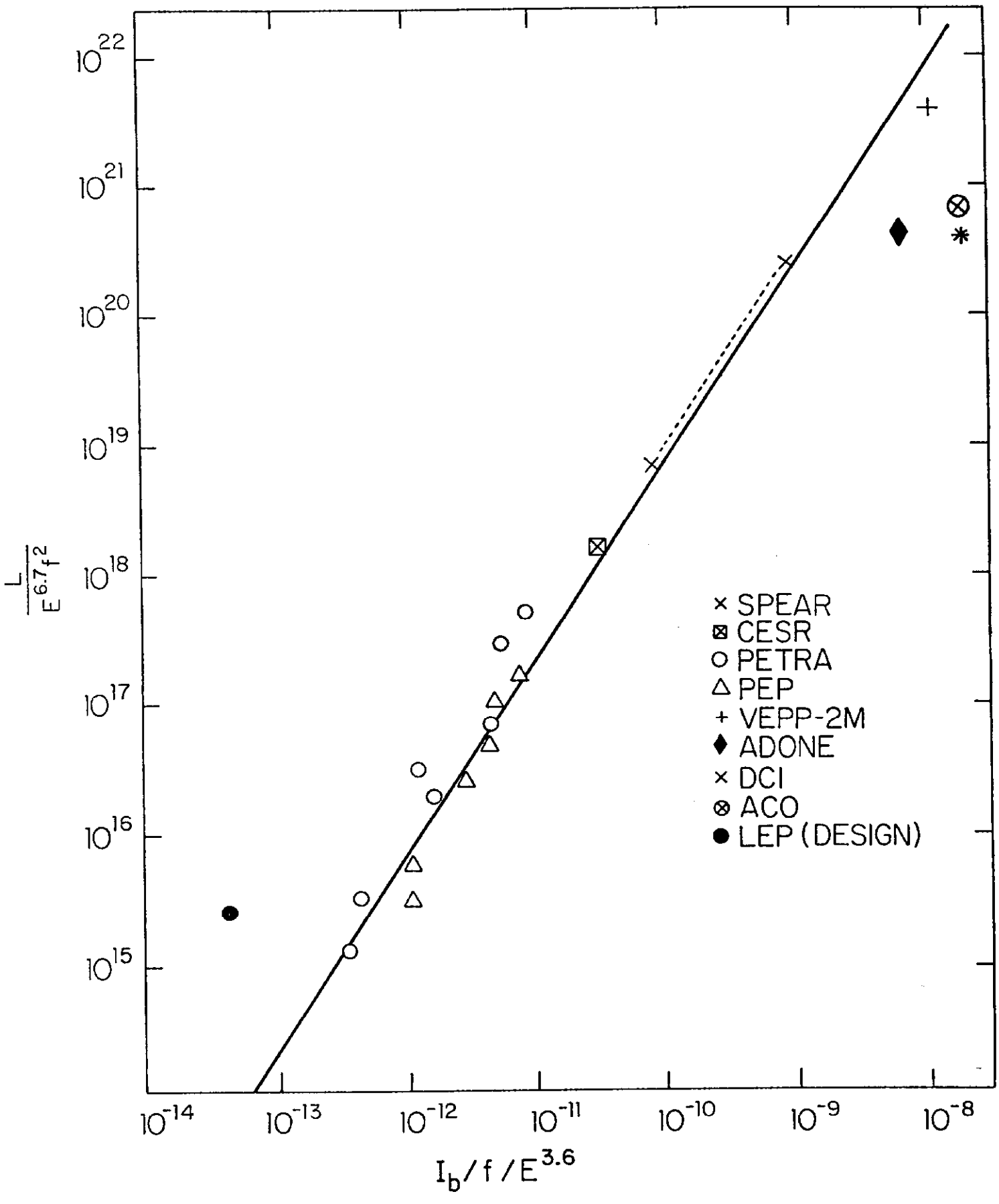

10^{22}
10^{21}
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
10^{15}
$\frac{L}{E^{6.7} f^2}$
10^{-14}
10^{-13}
10^{-12}
10^{-11}
10^{-10}
10^{-9}
10^{-8}
$I_b / f / E^{3.6}$
× SPEAR
⊠ CESR
○ PETRA
△ PEP
+ VEPP-2M
♦ ADONE
× DCI
⊗ ACO
● LEP (DESIGN)

Fig. 9

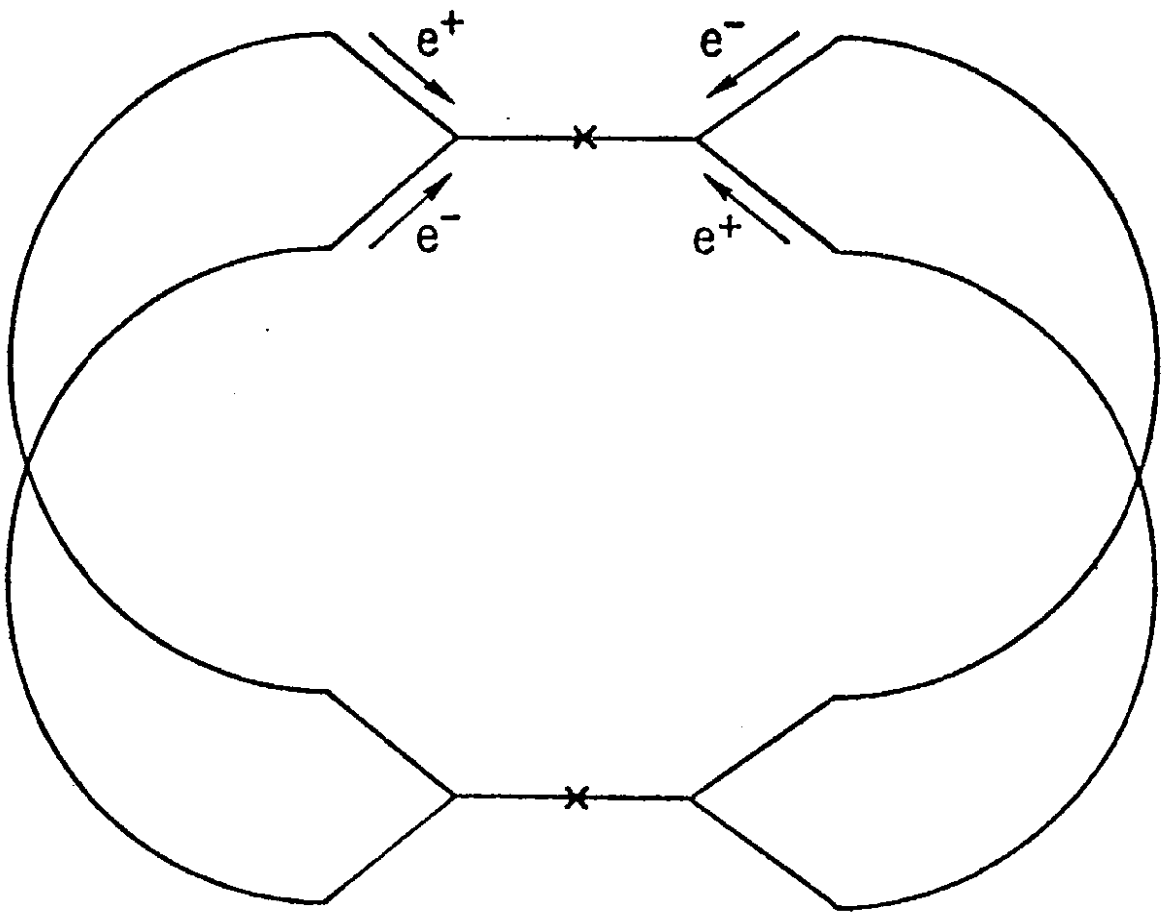


Fig. 10

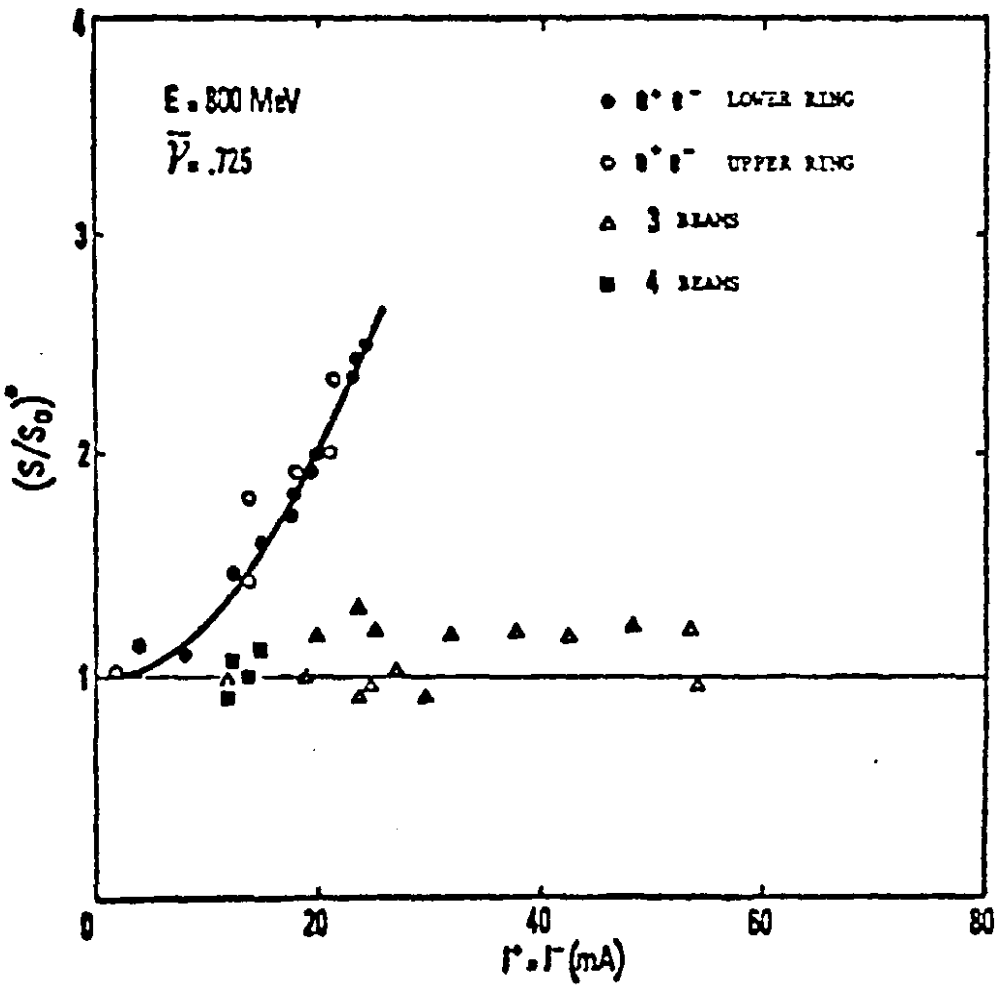

E = 800 MeV
$\bar{\nu}$ = .725
e⁺ e⁻ LOWER RING
e⁺ e⁻ UPPER RING
3 BEAMS
4 BEAMS
$(S/S_0)^*$
$I^+ = I^-$ (mA)
0
20
40
60
80
1
2
3
4

Fig. 11

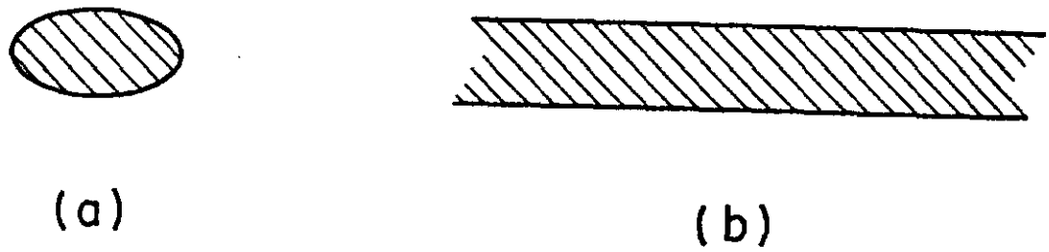


Fig. 12

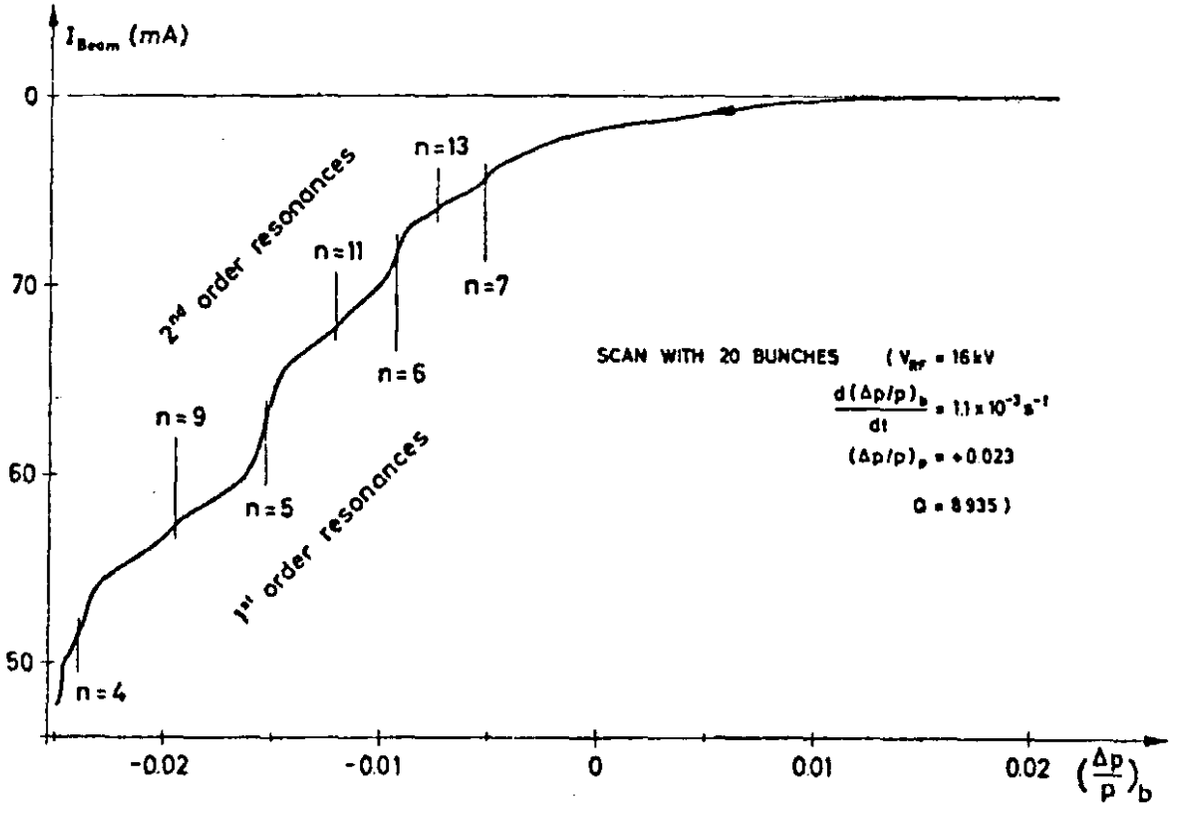

I_{Beam} (mA)
0
70
60
50
n=13
n=11
n=7
n=6
n=9
n=5
n=4
2nd order resonances
1st order resonances
SCAN WITH 20 BUNCHES (V_{RF} = 16 kV
$\frac{d(\Delta p/p)_b}{dt} = 1.1 \times 10^{-3} s^{-1}$
$(\Delta p/p)_s = +0.023$
Q = 8.935)
-0.02
-0.01
0
0.01
0.02
$(\frac{\Delta p}{p})_b$

Fig. 13

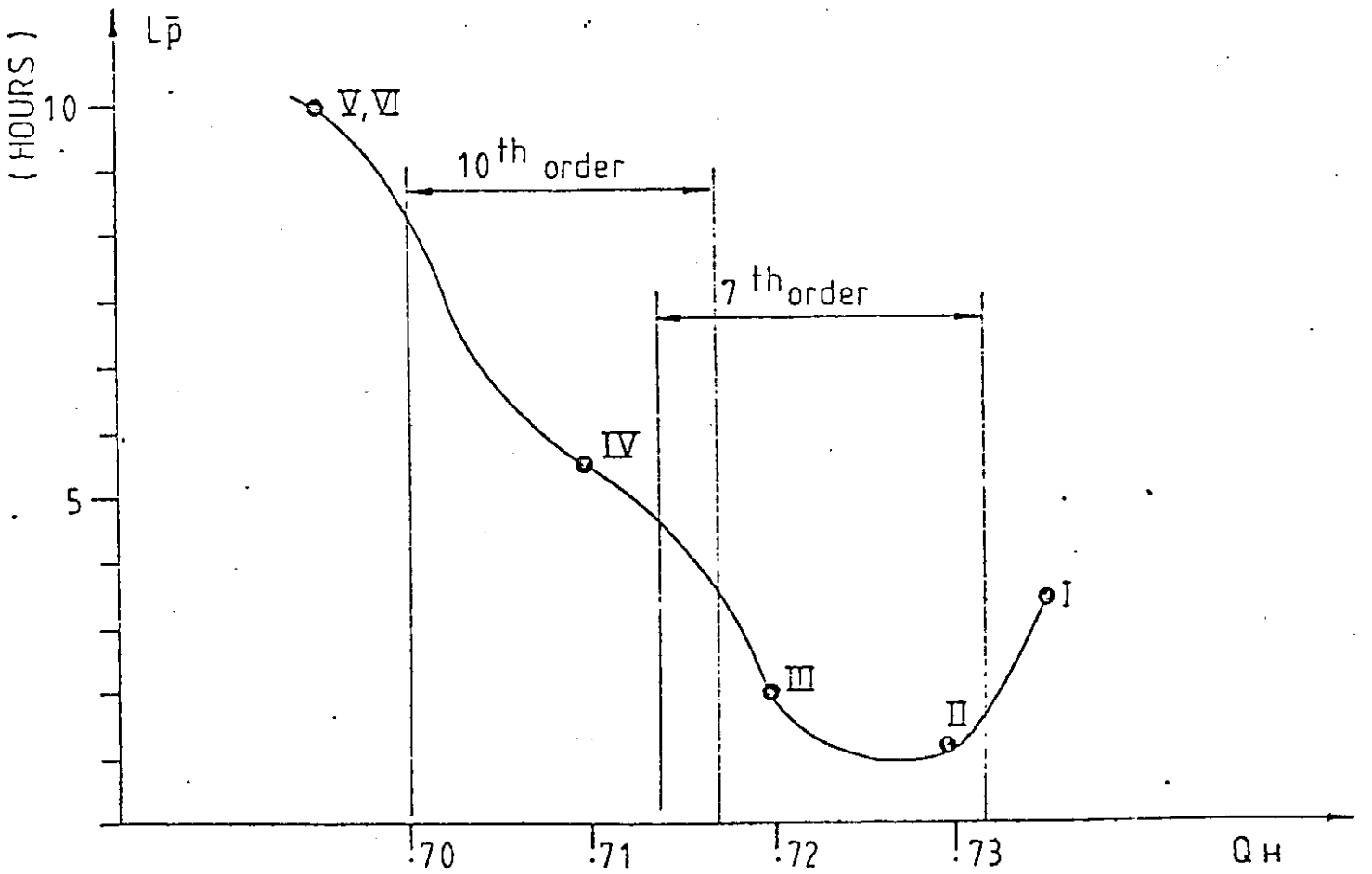


Fig. 14

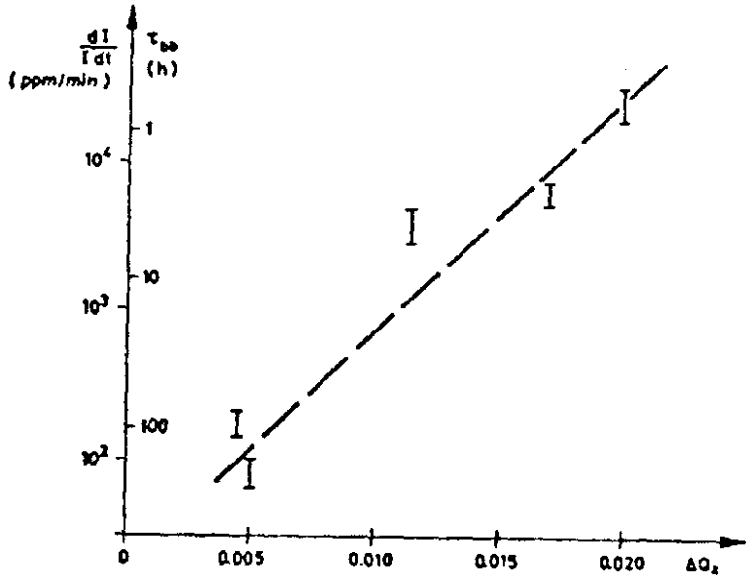

$\frac{dI}{I\,dt}$
(ppm/min)
τ_{ba}
(h)
1
10^{4}
10
10^{3}
100
10^{2}
0
0.005
0.010
0.015
0.020
ΔQ_z

Fig. 15

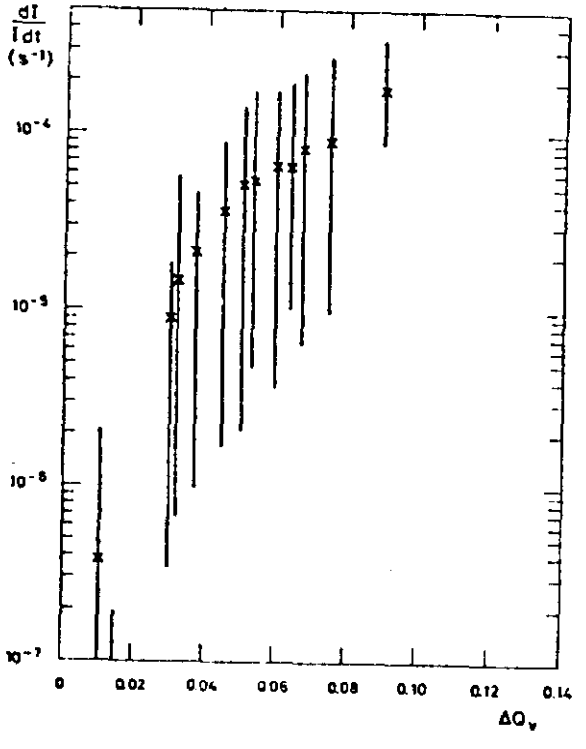
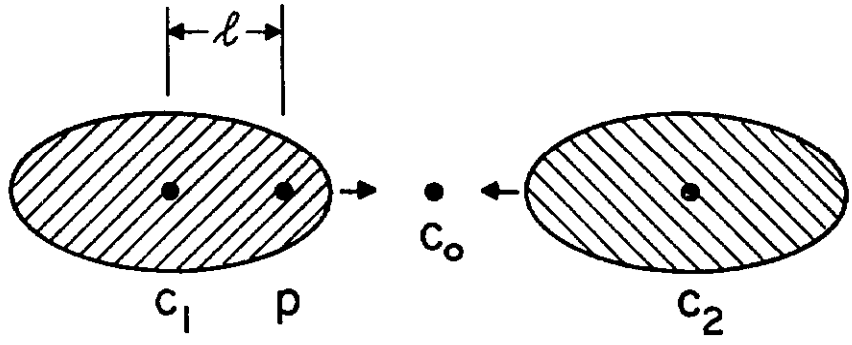
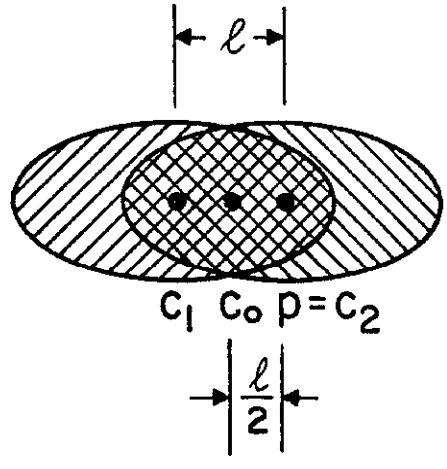

$\frac{dI}{I\,dt}$ (s^{-1})
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
0.02
0.04
0.06
0.08
0.10
0.12
0.14
ΔQ_y

Fig. 16



(a)



(b)

Fig. 17

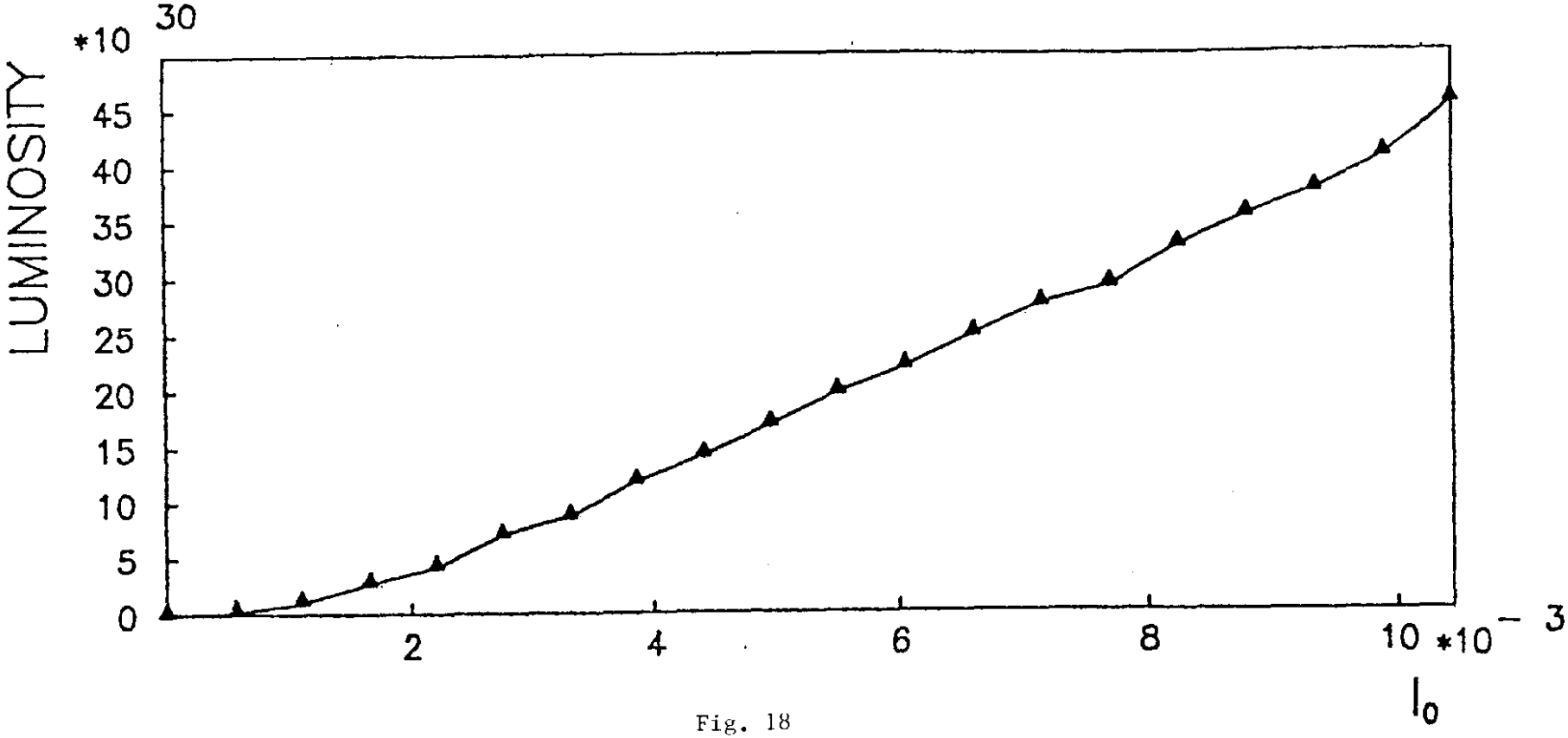


Fig. 18

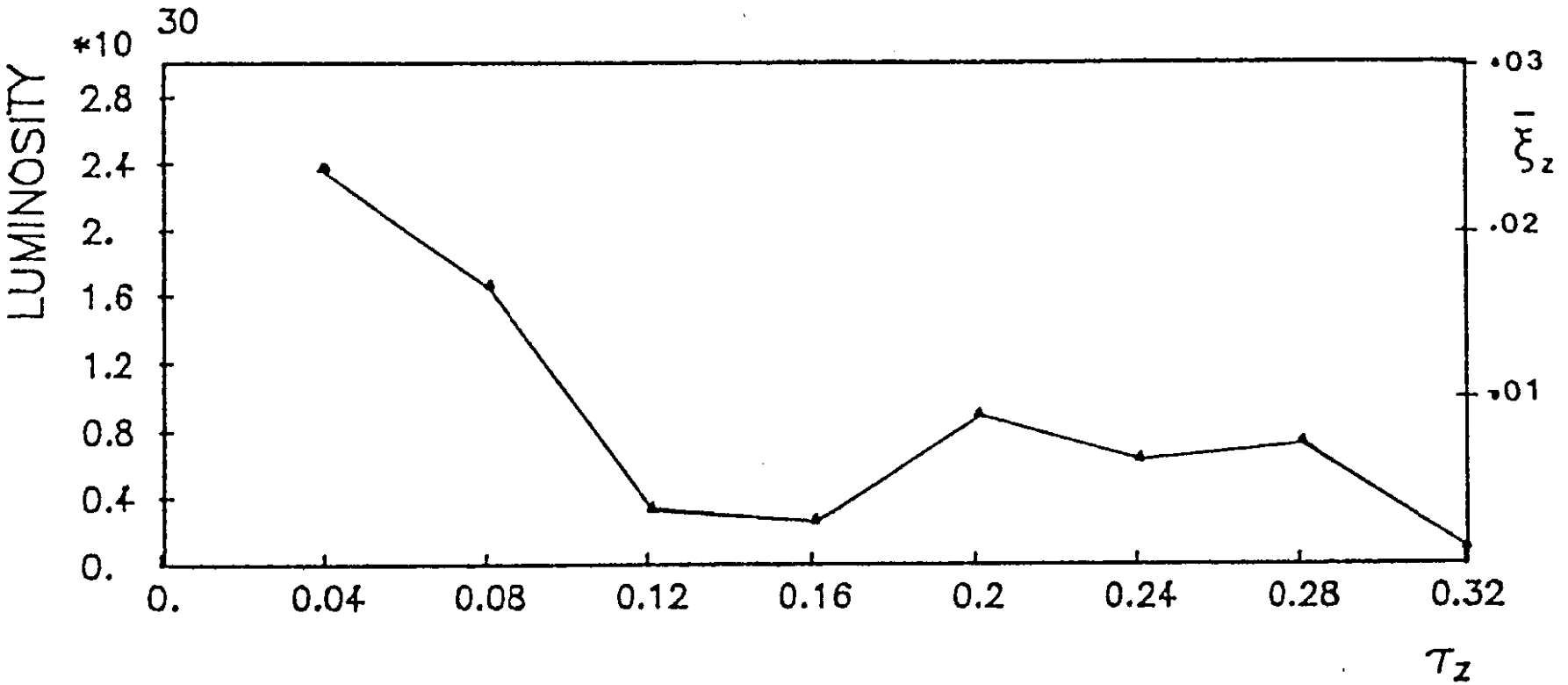

LUMINOSITY
*10 30
2.8
2.4
2.
1.6
1.2
0.8
0.4
0.
0.
0.04
0.08
0.12
0.16
0.2
0.24
0.28
0.32
τ_z
.03
$\bar{\xi}_z$
.02
.01

Fig. 19

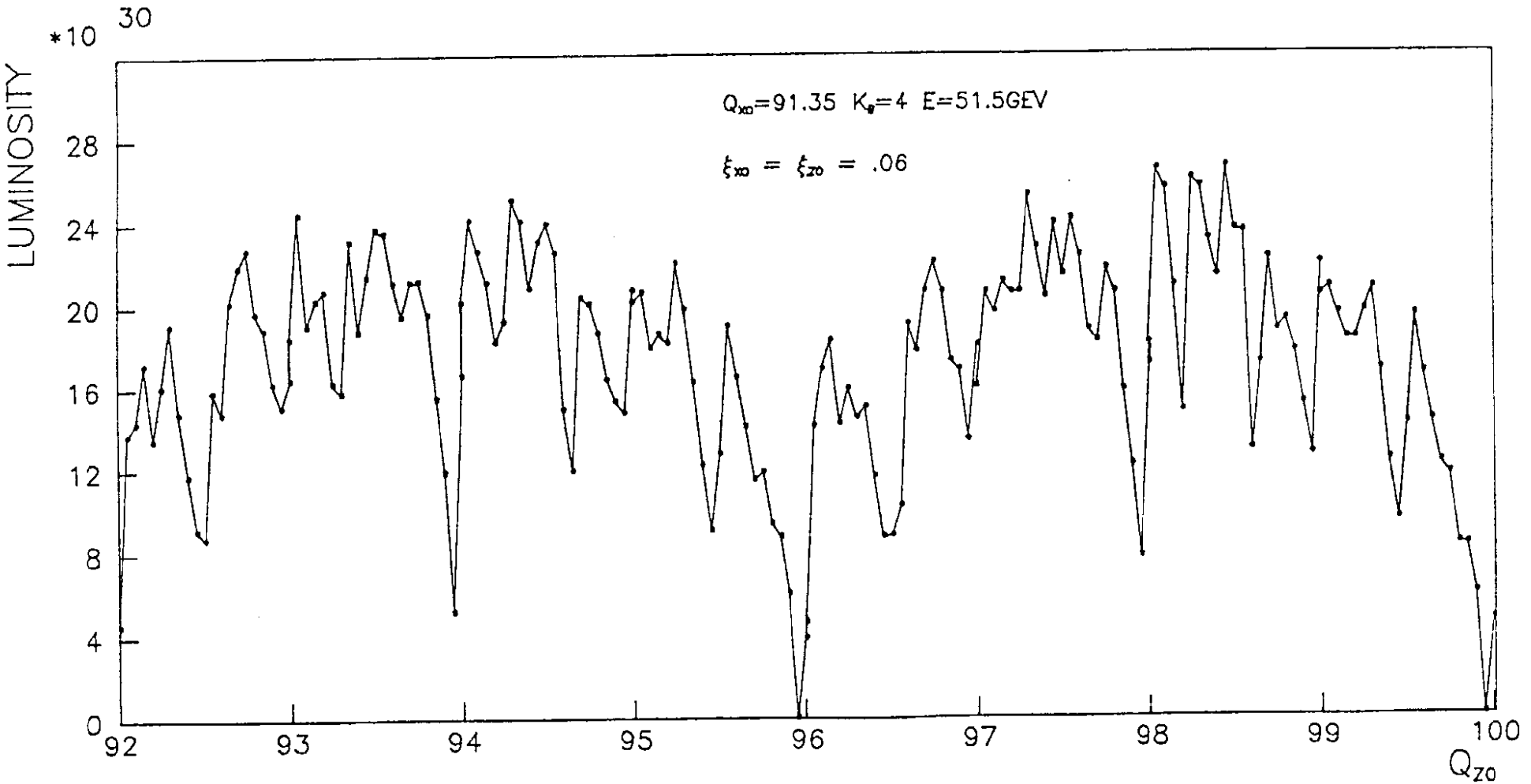

LUMINOSITY
*10 30
Q_{x0}=91.35 K_β=4 E=51.5GEV
$\xi_{x0} = \xi_{z0} = .06$
0
4
8
12
16
20
24
28
92
93
94
95
96
97
98
99
100
Q_{z0}

Fig. 20

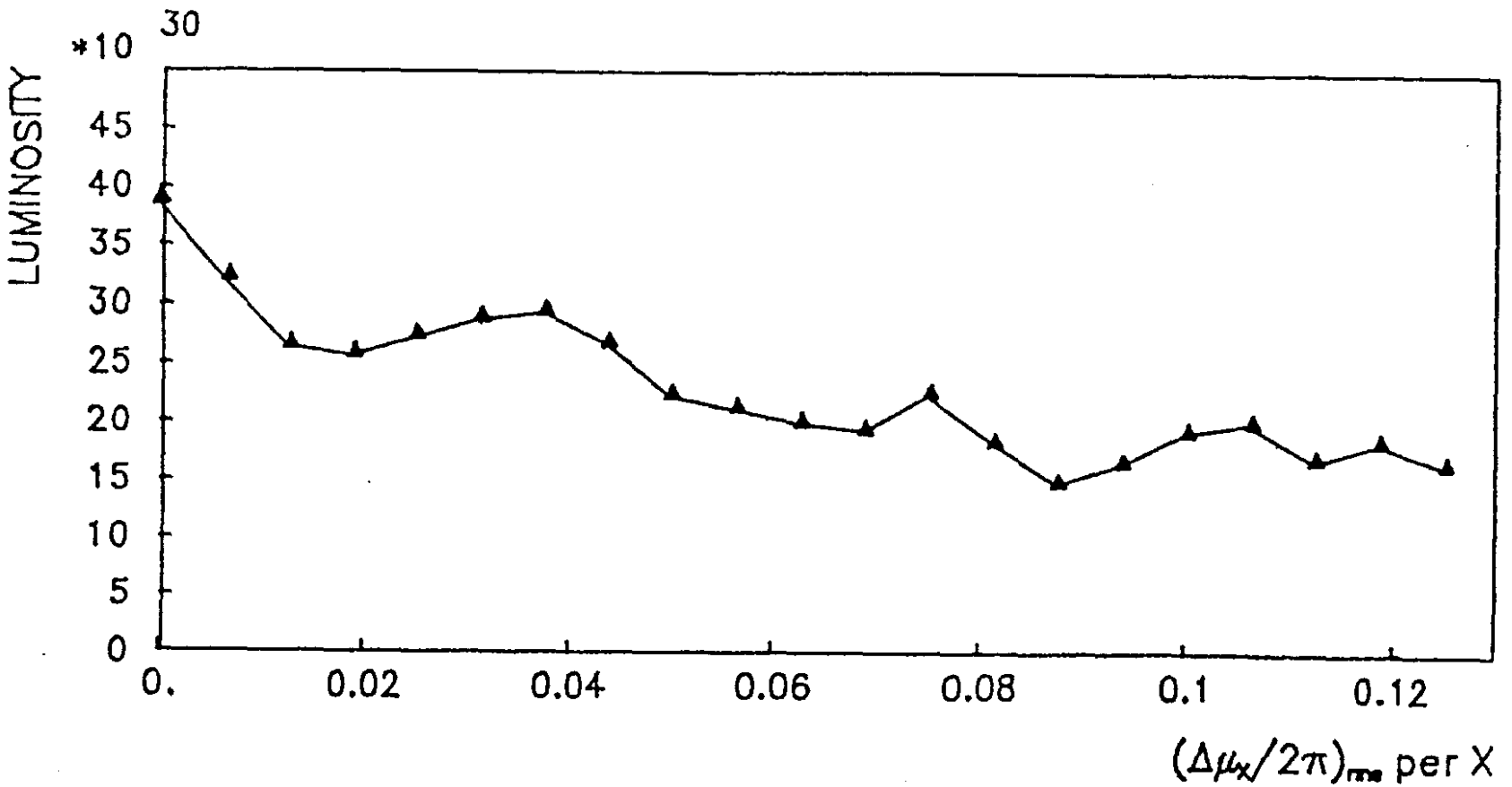

LUMINOSITY
*10
30
45
40
35
30
25
20
15
10
5
0
0.
0.02
0.04
0.06
0.08
0.1
0.12
$(\Delta\mu_x/2\pi)_{rms}$ per X

Fig. 21

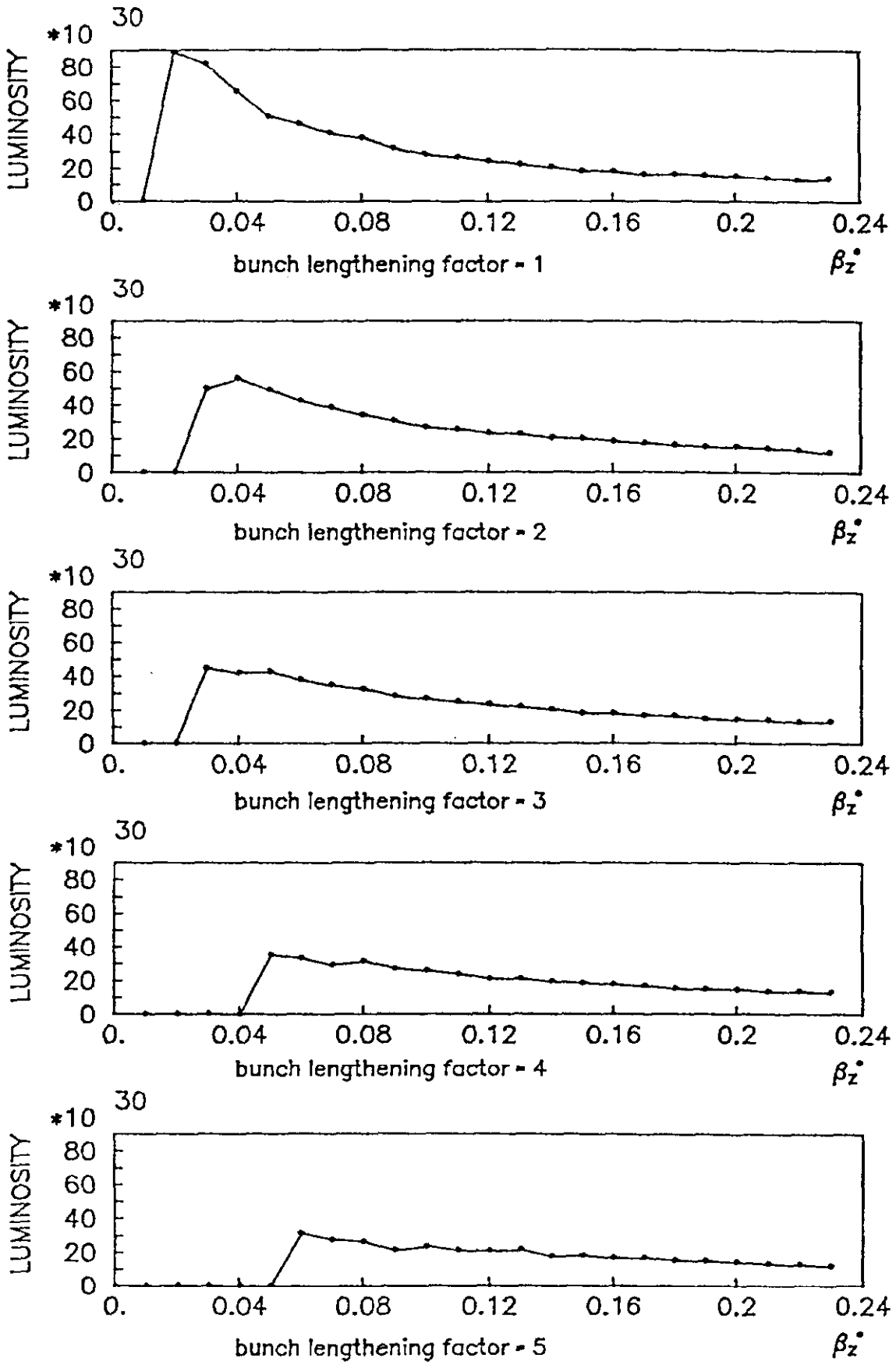

LUMINOSITY
*10 30
80
60
40
20
0
0.
0.04
0.08
0.12
0.16
0.2
0.24
β_z^*
bunch lengthening factor = 1
bunch lengthening factor = 2
bunch lengthening factor = 3
bunch lengthening factor = 4
bunch lengthening factor = 5

Fig. 22

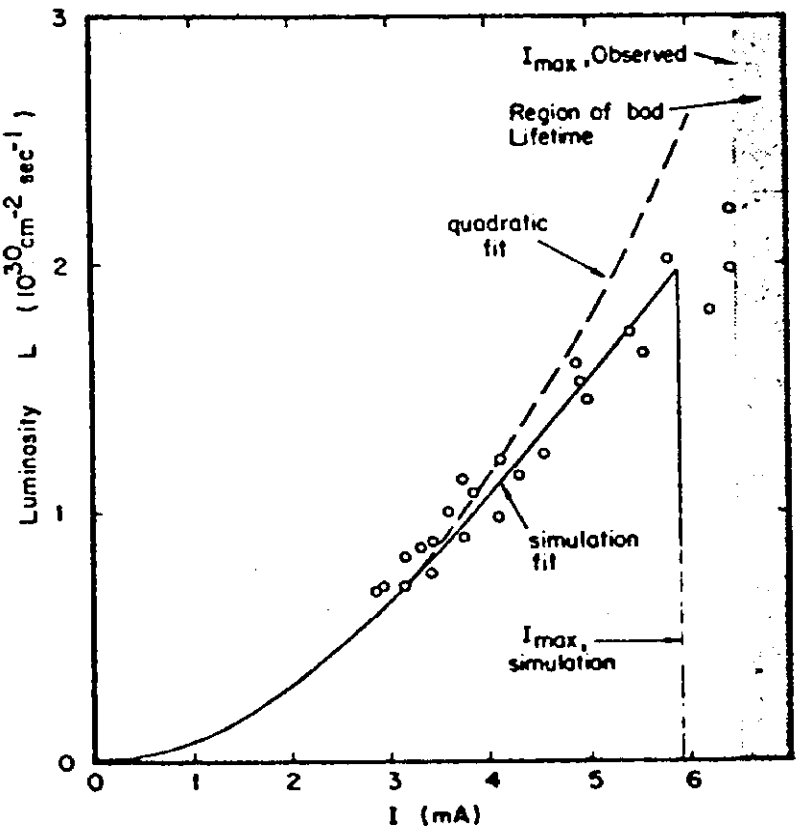

Luminosity L (10^{30}cm^{-2} sec^{-1})
I (mA)
I_{max}, Observed
Region of bad Lifetime
quadratic fit
simulation fit
I_{max}, simulation
0
1
2
3
4
5
6

Fig. 23

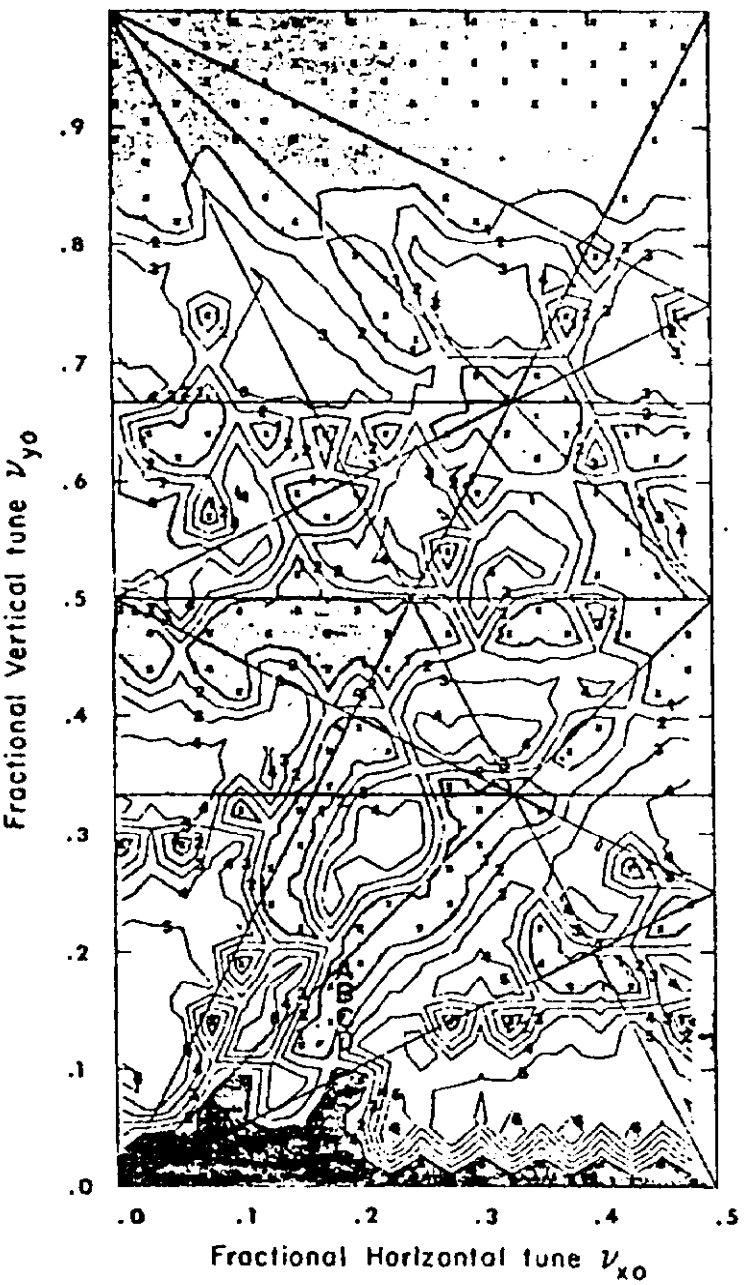

.9
.8
.7
.6
.5
.4
.3
.2
.1
.0
Fractional Vertical tune ν_{yo}
A
B
C
.0
.1
.2
.3
.4
.5
Fractional Horizontal tune ν_{xo}

Fig. 24

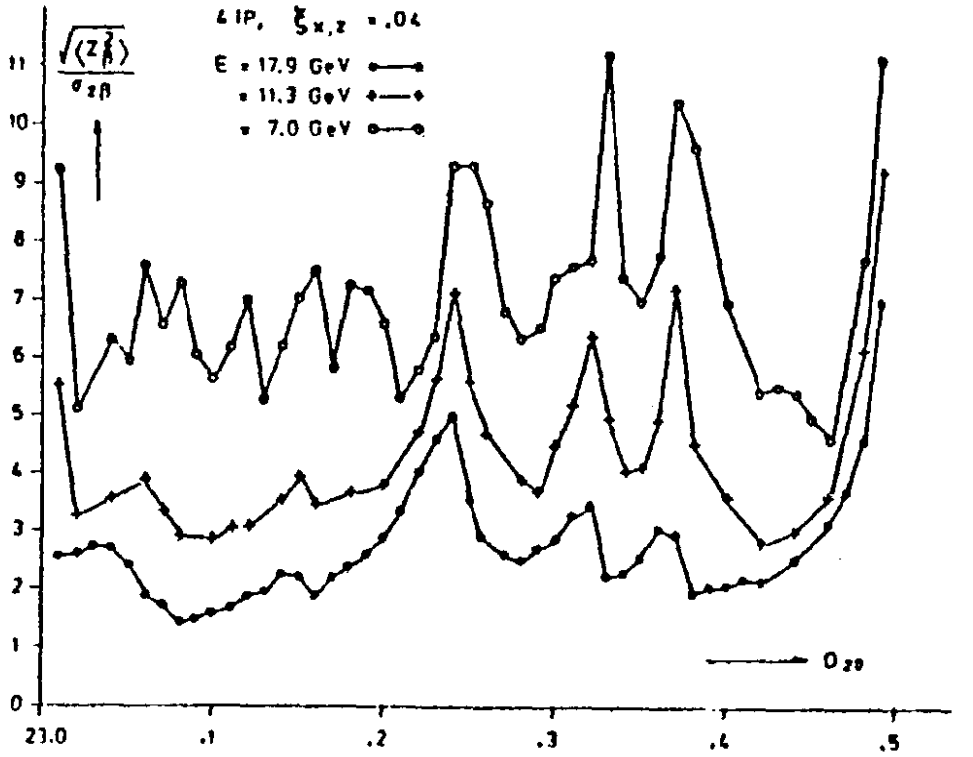

4 IP, $\xi_{x,z}$ = .04
E = 17.9 GeV
= 11.3 GeV
= 7.0 GeV
$\frac{\sqrt{\langle z_\beta^2 \rangle}}{\sigma_{z\beta}}$
$\sigma_{z\beta}$
0
1
2
3
4
5
6
7
8
9
10
11
23.0
.1
.2
.3
.4
.5

Fig. 25

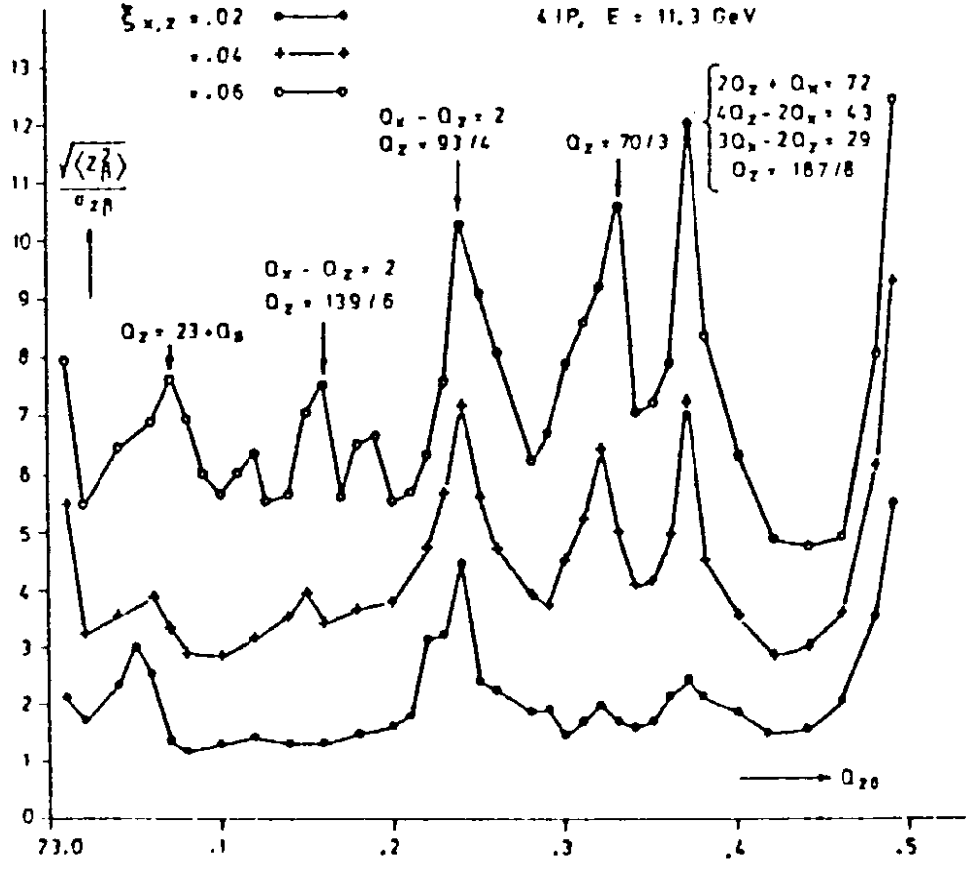

ξx,z = .02
.04
.06
4 IP, E = 11.3 GeV
$\frac{\sqrt{\langle z_\beta^2 \rangle}}{\sigma_{z\beta}}$
$Q_x - Q_z = 2$
$Q_z = 93/4$
$Q_z = 70/3$
$2Q_z + Q_x = 72$
$4Q_z - 2Q_x = 43$
$3Q_x - 2Q_z = 29$
$Q_z = 187/8$
$Q_z = 23 + Q_\beta$
$Q_x - Q_z = 2$
$Q_z = 139/6$
Q_{z0}
0
1
2
3
4
5
6
7
8
9
10
11
12
13
23.0
.1
.2
.3
.4
.5

Fig. 26

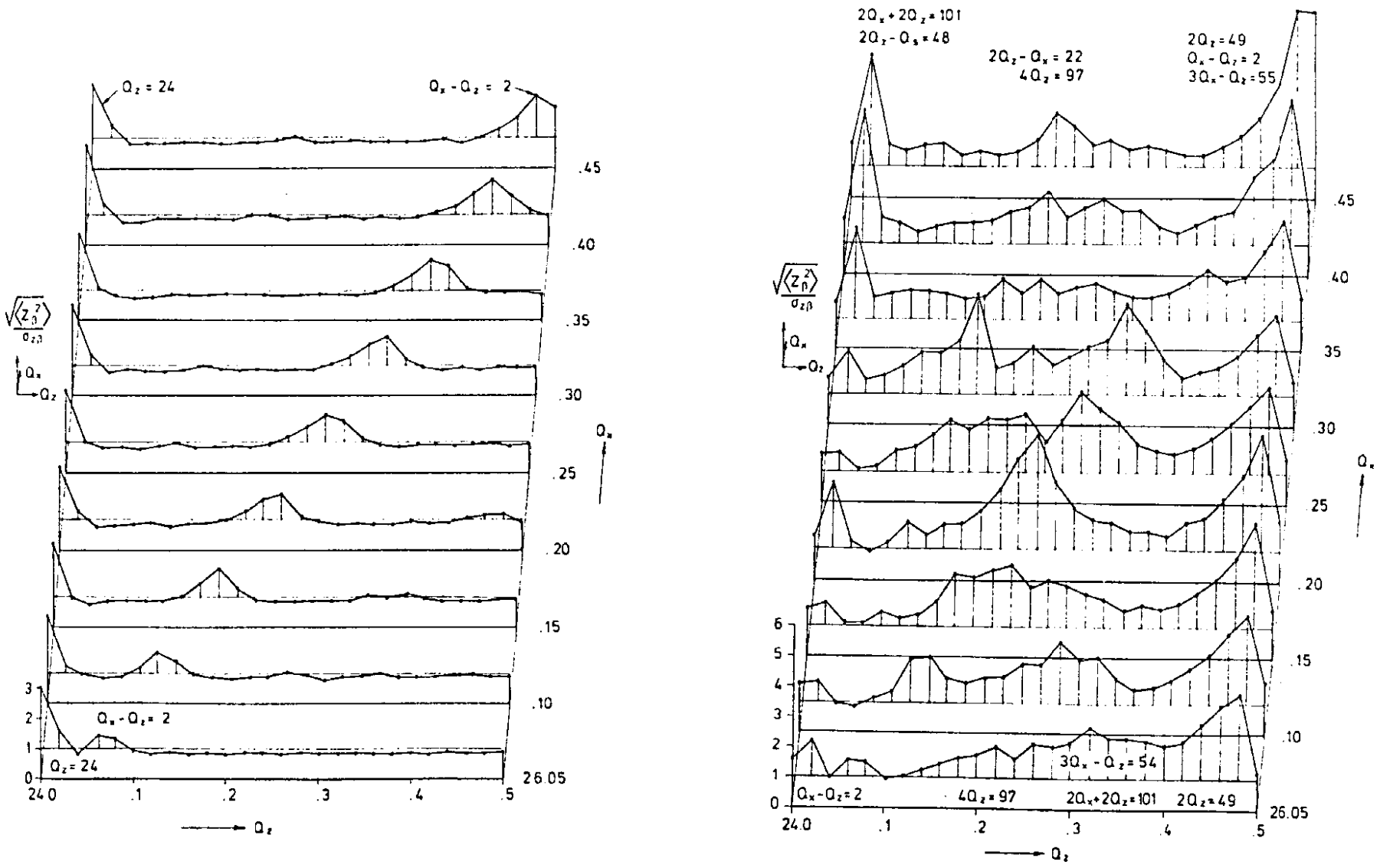

$Q_z = 24$
$Q_x - Q_z = 2$
.45
.40
.35
.30
.25
.20
.15
.10
26.05
$\sqrt{\langle z_\beta^2 \rangle}$
$\sigma_{z\beta}$
Q_x
Q_z
$Q_x - Q_z = 2$
$Q_z = 24$
24.0
.1
.2
.3
.4
.5
Q_z
$2Q_x + 2Q_z = 101$
$2Q_z - Q_x = 48$
$2Q_z - Q_x = 22$
$4Q_z = 97$
$2Q_z = 49$
$Q_x - Q_z = 2$
$3Q_x - Q_z = 55$
$3Q_x - Q_z = 54$
$Q_x - Q_z = 2$
$4Q_z = 97$
$2Q_x + 2Q_z = 101$
$2Q_z = 49$

Fig. 27

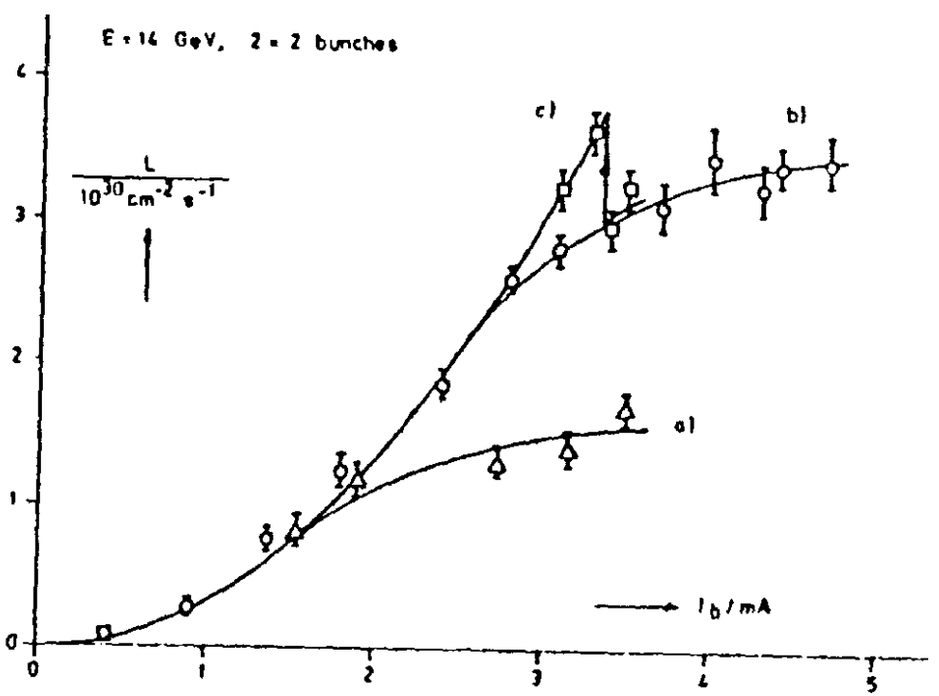


Fig. 28

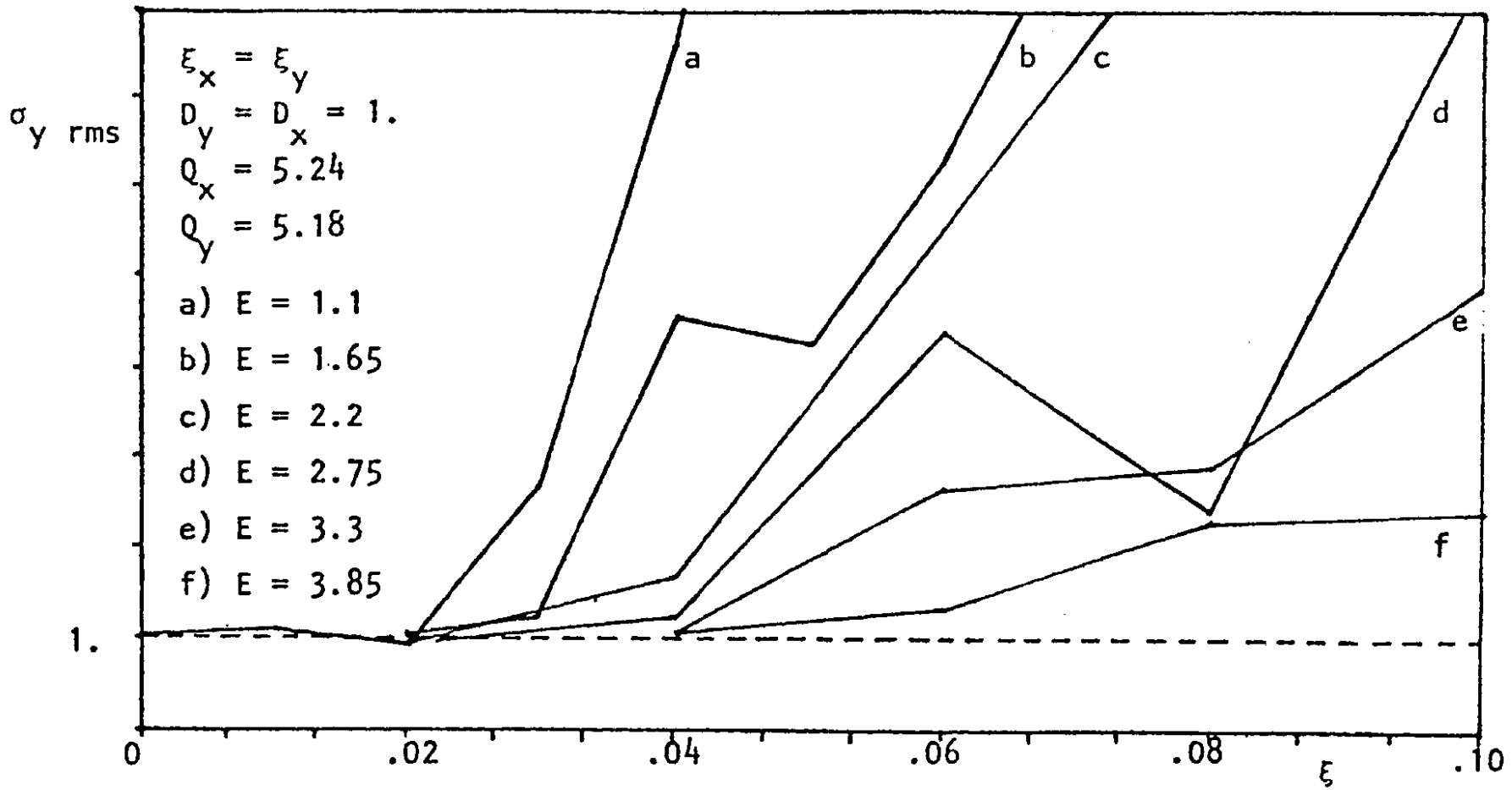

σ_y rms
$\xi_x = \xi_y$
$D_y = D_x = 1.$
$Q_x = 5.24$
$Q_y = 5.18$
a) E = 1.1
b) E = 1.65
c) E = 2.2
d) E = 2.75
e) E = 3.3
f) E = 3.85
a
b
c
d
e
f
1.
0
.02
.04
.06
.08
.10
ξ

Fig. 29

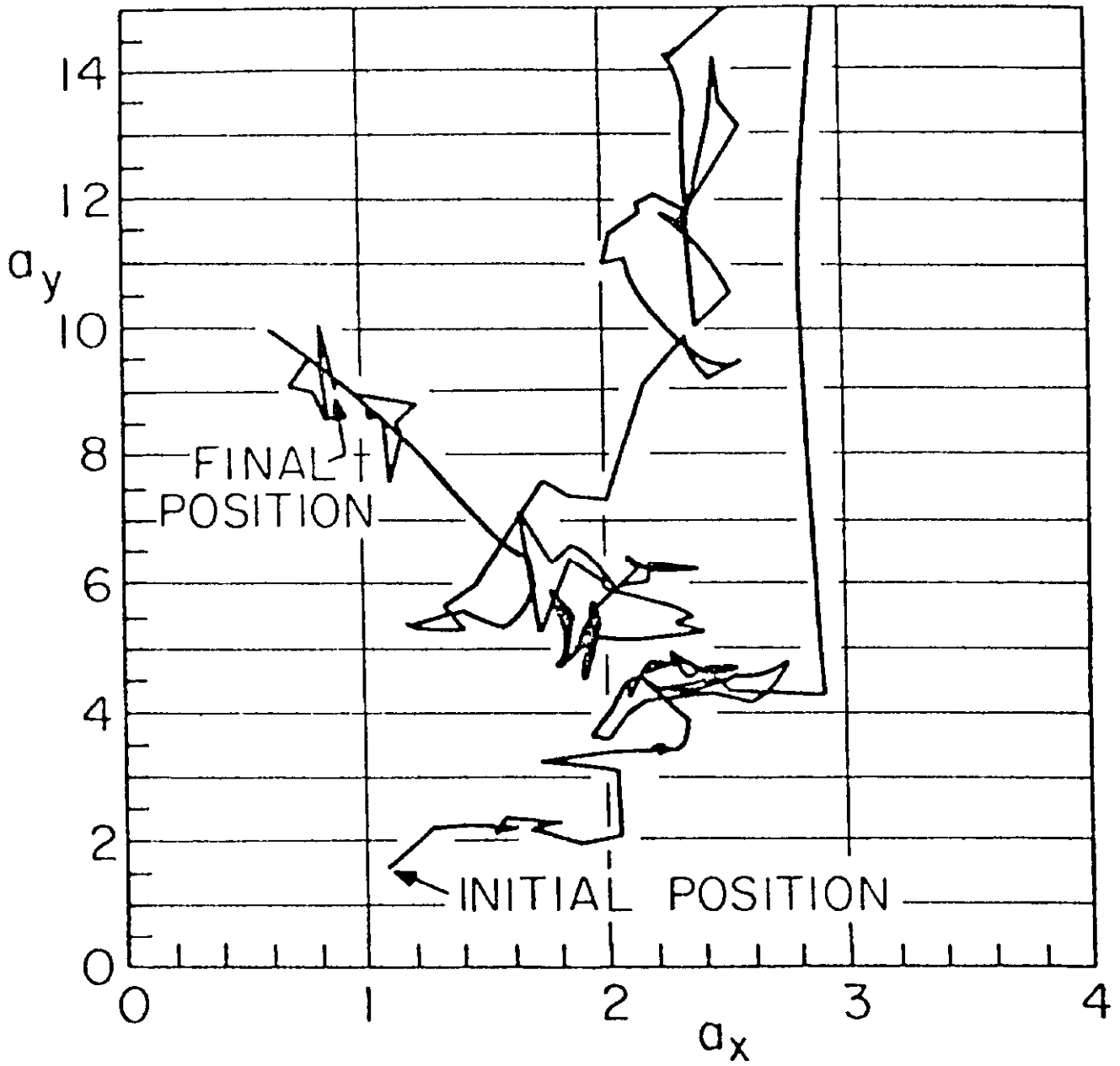


Fig. 30

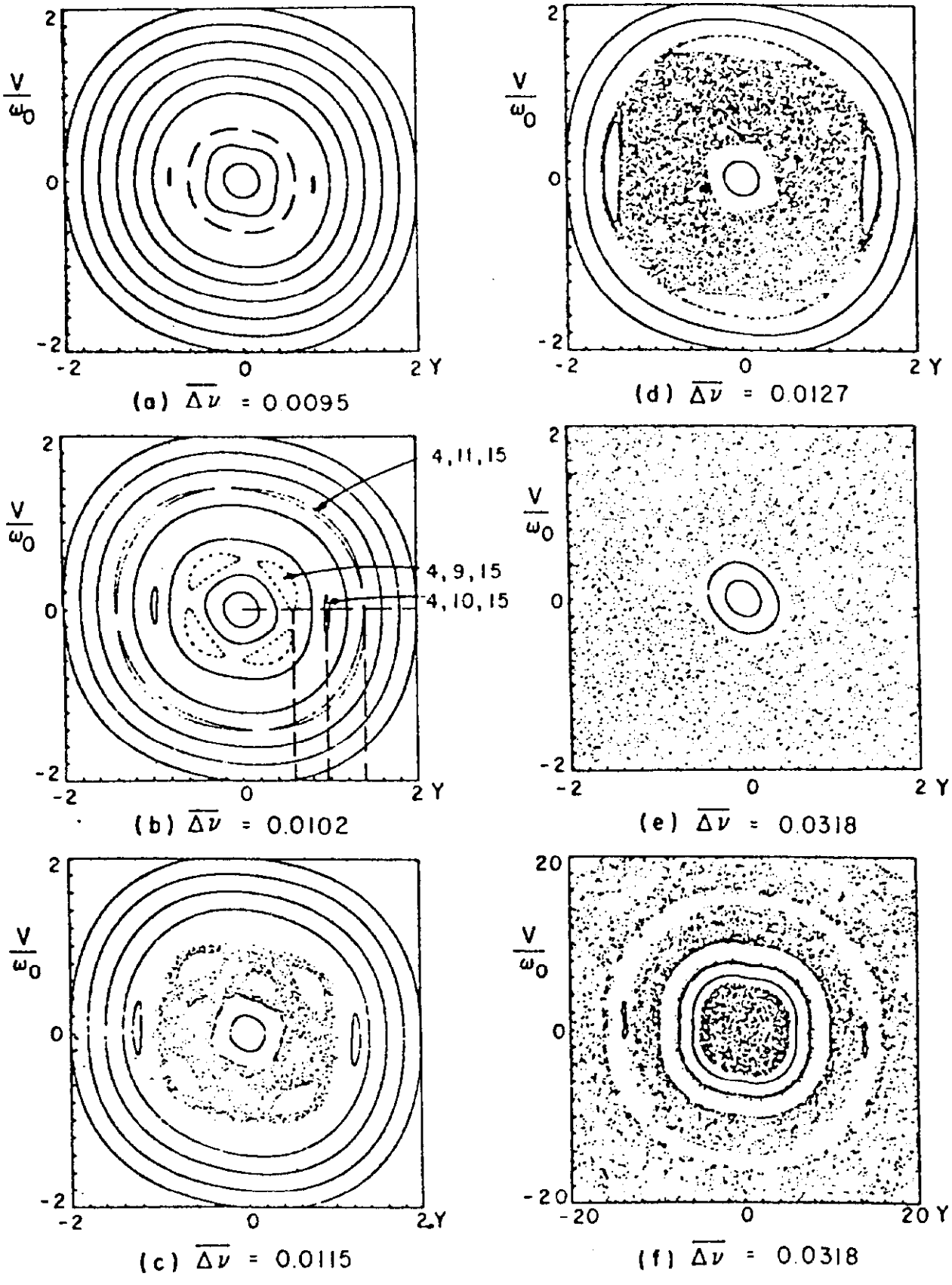

V/ω0
2
0
-2
-2
0
2 Y
(a) Δν = 0.0095
(d) Δν = 0.0127
4,11,15
4,9,15
4,10,15
(b) Δν = 0.0102
(e) Δν = 0.0318
(c) Δν = 0.0115
20
-20
(f) Δν = 0.0318

Fig. 31

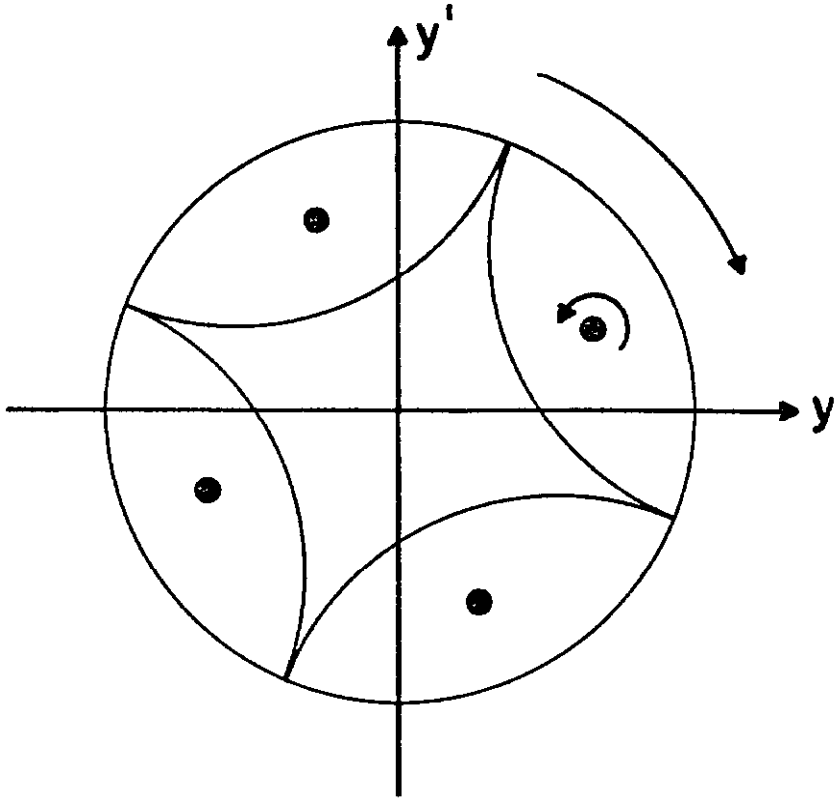


Fig. 32

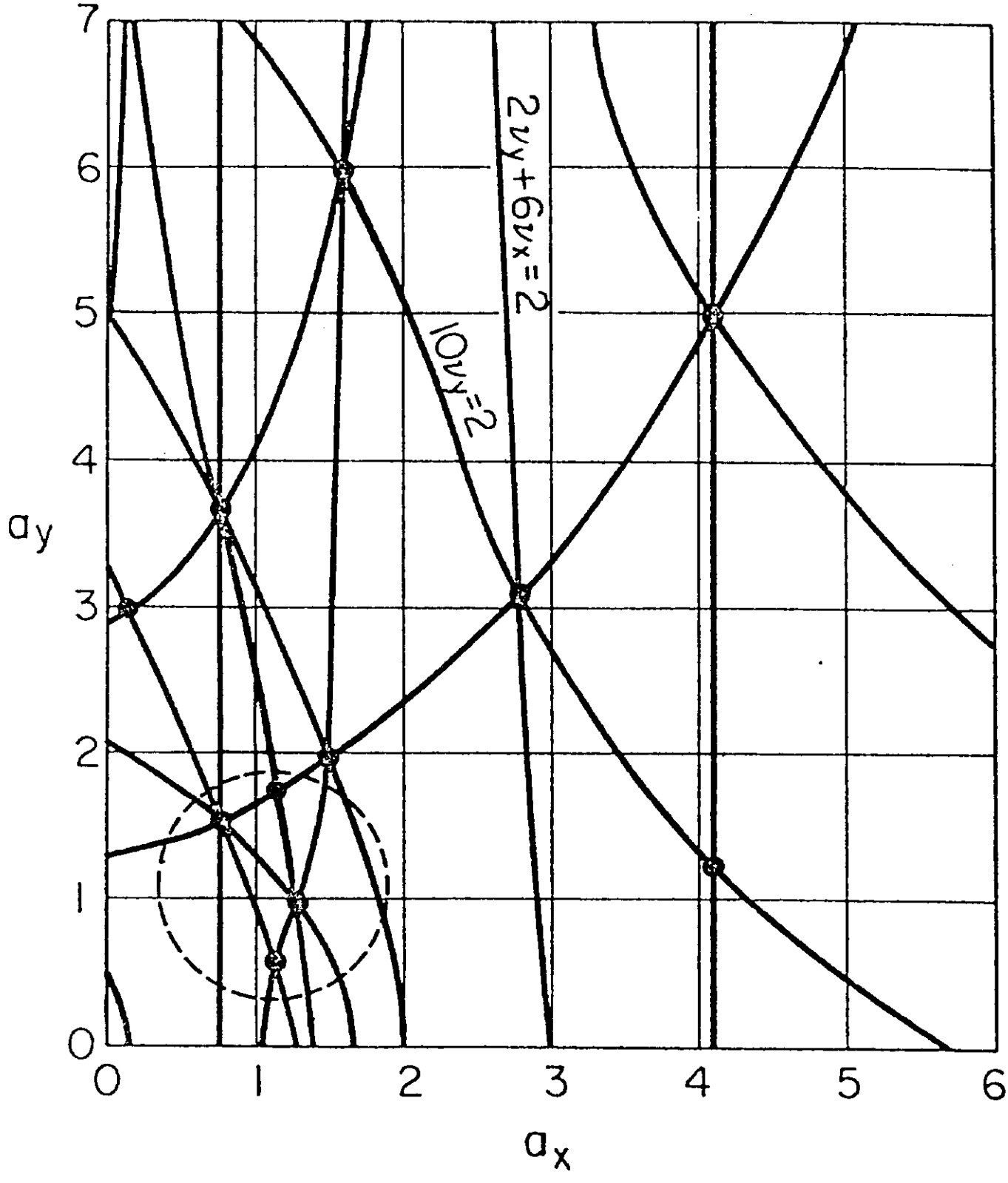


Fig. 33

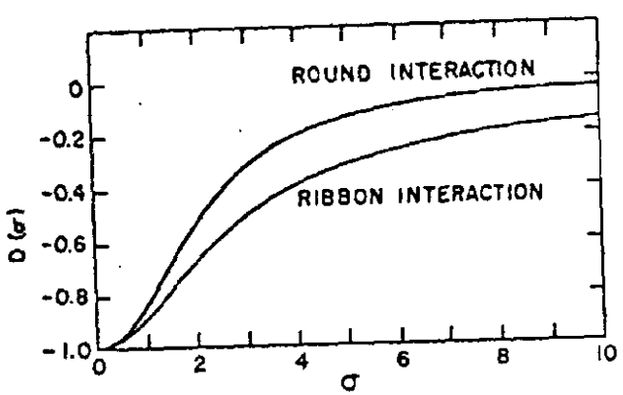

ROUND INTERACTION
RIBBON INTERACTION
0
-0.2
-0.4
-0.6
-0.8
-1.0
D (σ)
0
2
4
6
8
10
σ

Fig. 34

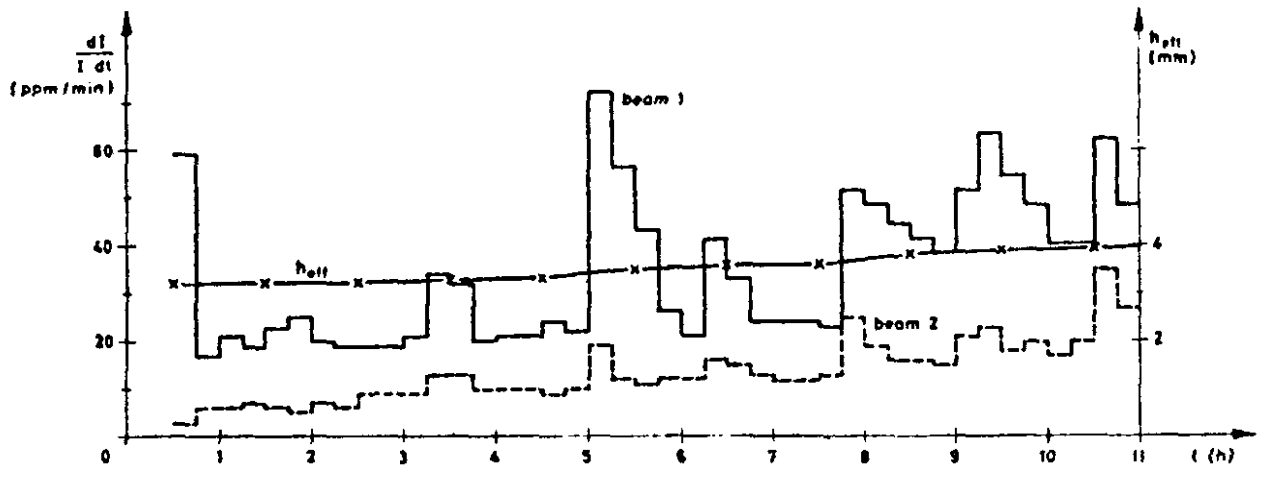

$\frac{dI}{I\,dt}$
(ppm/min)
beam 1
beam 2
h_{eff}
h_{eff} (mm)
80
40
20
0
1
2
3
4
5
6
7
8
9
10
11
t (h)
4
2

Fig. 35

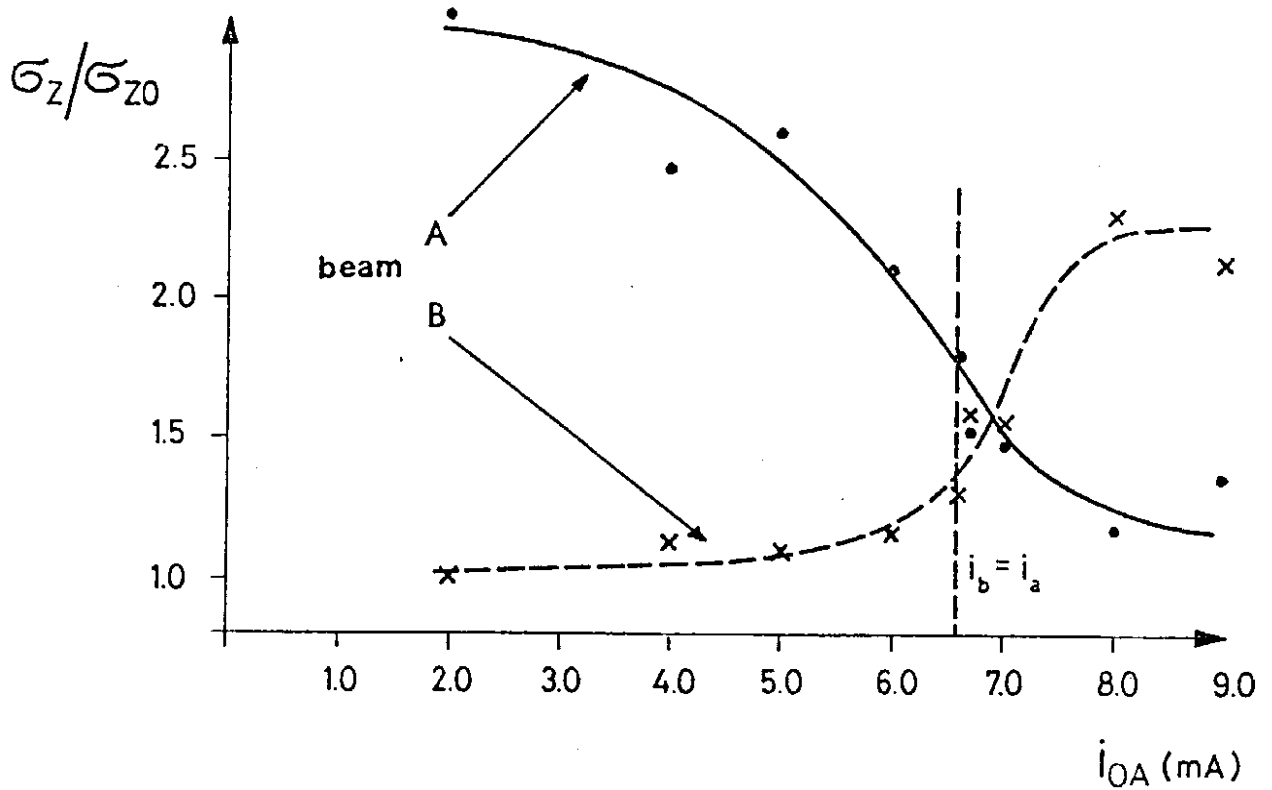

σ_z/σ_{z0}
2.5
2.0
1.5
1.0
beam
A
B
$i_b = i_a$
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
i_{OA} (mA)

Fig. 36