Fermi National Accelerator Laboratory

FERMILAB-CONF-81-56-T
OCTOBER 1, 1981

LUMINOSITY IN ELECTRON-POSITRON COLLIDING-BEAM STORAGE RINGS: INTRODUCTION TO THE PROBLEM

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

Invited talk presented at the 1981 Summer School on High Energy Particle Accelerators, Fermilab, July 13-24, 1981.

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

The goal of the colliding-beam storage ring designer and operator is optimization of luminosity, which is defined as the number of elementary-particle events of unit cross-section that take place at a single beam-beam encounter point, per unit time.

At the newest electron-positron storage rings - CESR, PEP, and PETRA - the luminosities observed have so far been at best about an order of magnitude below the design values[1] of $\sim 10^{32}$ cm^{-2} sec^{-1}. This means that high energy physics experiments, so dependent on good statistics for analyzing rare processes, must run much longer than had been originally expected for acceptable results.

In the first two sections of this pedogogical report, I shall discuss two assumptions on which these design expectations were based, and specific ways in which subsequent operating experience has shown these assumptions to be naive. One assumption led to an overestimate of luminosity at a given current, while the other led to an overestimate of the largest current that could be stored. (I shall consider in detail only phenomena related directly to beam-beam collisions; single-beam effects will be mentioned only in passing.) In the third and final section I shall describe some recent theoretical attempts to go beyond these assumptions.

I have chosen to concentrate here exclusively on e^+e^- machines, because there is published data on high-energy e^+e^- storage from at least seven separate laboratories, collected over a period of, by now, about twenty years. By contrast, there is as yet no significant data on $\bar{p}p$ colliders; and all the pp data comes

from only one facility - the ISR - whose operating conditions differ from those of electron-positron rings.

I. HOW CONVENTIONAL DESIGN PROCEDURE OVERESTIMATES LUMINOSITY AT A GIVEN CURRENT

Mathematically, luminosity is defined by the formula

$$L = \left(\frac{I_1}{A_1}\right)\left(\frac{I_2}{A_2}\right)\frac{A}{e^2 Bf} \quad , \tag{1}$$

where f is the revolution frequency of a stored particle (i.e., speed of light divided by circumference of ring); e is the positron charge; B is the number of bunches per beam; I_1 and I_2 are the two beam currents; A_1 and A_2 are the two beam areas, transverse to the collision axis; and A is the transverse area in which the two beams overlap upon intersection. (When the beam particles are not distributed uniformally in the transverse plane, these A's are effective areas, defined by appropriately weighted averages. See ref. 35.)

It has been standard practice, when computing anticipated luminosity, to ignore the effect that one beam has on the size and shape of the other.[2] This is the non-perturbation assumption. It means, in particular, that A_1, A_2 and A are all equal and independent of the number of particles in the storage ring. For equal currents ($I_1=I_2=I$) we may then write

$$L = I^2/e^2 BfA \quad . \tag{2}$$

In particular, luminosity in this configuration should increase

with current as (constant) $\times I^2$, where the proportionality constant is computable from first principles.

It turns out that this is not consistent with experiment: Direct measurements of luminosity show that L always grows much more slowly than (constant) xI^2 at large current. Direct observations of beam size show that the cross-sectional area of at least one of two colliding beams always increases significantly with I at large current. (When both beams blow up equally, equation (2) is still valid, but one may no longer take A to be independent of I.) These trends are illustrated in figures 1-12, to which we now turn.

(It should be noted that by unperturbed A I mean the single-beam cross-sectional area as computed from the beta-functions and emittances of the real storage ring. Because of single-beam instabilities,[33] (among other things), these β's may have to be held at larger values than anticipated by the designers. Thus one should really distinguish two principal reasons for low luminosity, relative to design: colliding-beam effects, which lower the luminosity vs. current curve relative to some quadratic; and single-beam beta-function limitations, which lower the normalization of the reference quadratic itself.)

Figures 1-5 show data on luminosity as a function of current, from several laboratories. Figure 1 represents data taken at SPEAR.[3] The data point with the largest current (about 8mA) falls short of the quadratic curve (extrapolated from small current) by a factor of about three. Figure 2 represents data taken at CESR.[4]

For β_v^* (what we have been calling β_y^*) = 29 cm, the data point with the largest current falls short of the corresponding quadratic extrapolation by a factor of about 1.6. Figure 3 shows data taken at PETRA.[5] For these graphs, PETRA was operated with two bunches per beam. (The SPEAR and CESR data[6] described above were taken with one bunch per beam.) In each graph, the abscissa corresponds to the single-beam current per bunch, and the ordinate corresponds to the quotient of luminosity and the square of the single-bunch current. If L were proportional to I^2, then the curves in Fig. 3 would be straight horizontal lines. Instead, the 15.3 GeV curve is lower at 5 mA, by a factor of about two, than it is at very small current. The effect is clearly much worse at the lower energy of 11 GeV. Figures 4 and 5 represent data taken at PEP, with two different sets of magnetic lattice parameters. The two curves in each figure correspond to operation with one and three bunches per beam. In Fig. 4, the one-bunch curve is quadratic at small current, but falls below the extrapolation of the quadratic curve by a factor of about two at the largest current shown; in the same figure, the curve interpolated between the three-bunch data points grows more slowly than quadratic (more like $I^{1.6}$) even at small current. At the highest current for which data with both bunch numbers is shown, the three-bunch curve exceeds the two-bunch curve by a factor of two, not the naively expected three. (Three would arise in formula (2) from combining a factor of nine, corresponding to the ratio of I^2 in the two cases, with a factor of one-third, corresponding to the ratio of the two values of B^{-1}.) In Fig. 5,

the single-bunch curve at the largest current falls short of the quadratic extrapolation from small current by a factor of about 5.4; the ratio of three-bunch to single-bunch luminosity falls to 2.2 at the largest current for which the comparison can be made.

(A number of workers have attempted to fit the high-current luminosity of different high-energy e^+e^- storage rings to a single phenomenological scaling law. See references 3 and 34 for details.)

Figures 6-12 show data on the increase of transverse beam size with current. Let me first make a few comments before discussing this data in detail.

When the currents in two colliding beams are not equal ("weak-strong" case), the more tenuous ("weak") beam is always the one to be more enlarged. When the two currents are equal ("strong-strong" case) there is no general rule for predicting which beam (if either) will expand more than the other, and a perturbation (for example momentary excitation at a betatron frequency) can turn the thicker beam thinner and vice versa. At SPEAR[7] this kind of interchange can be induced in a gradual and controlled manner by slowly varying the phases between RF cavities. In particular, for any energy there is a setting of the relative phases for which the two beams are blown up equally, and this setting turns out to maximize luminosity. (There are indications that the same thing might be possible at PEP.[6]) Unless otherwise noted, all the SPEAR data discussed in this report were obtained with the beams deliberately matched in this way. (Data from other machines may

also have been taken under conditions of equal blowup, but only at SPEAR is this situation claimed to be a result of conscious manipulation.)

This suggests the first of a number of exercises that will be proposed in the course of this paper.

Problem 1

How should one place two identical RF cavities in a given storage ring so that any variation of their relative phase will not affect the location of the point at which the electron and positron synchronous design orbits cross?

While you're at it, a second exercise:

Problem 2

Even without colliding-beam effects, the shape of an electron bunch in a storage ring is not an entirely trivial matter. For example: A common practice is the deliberate rotation of a number of quadrupole magnets in order to couple horizontal and vertical betatron motion and thus enhance the vertical extent of the beam. (Even with such enhancement, electron and positron beams are much wider than they are high.) This linear xy coupling also has the effect of tilting transverse profiles of the bunches. Look up the theory of equilibrium beam distributions in linearly focused storage rings (Ref. 8), and then use the figures in Table I to estimate the magnitude of the transverse tilt (it should be small)

at an interaction point in PEP. Show that counter-rotating electron and positron bunches tilt in the same sense (i.e., both inclined toward the same general area above or below the center of the ring), when they pass the same azimuth. If they did not tilt in the same sense, then the two beams would not overlap totally at the interaction point even at low currents, and the incompleteness of this overlap would be one reason for low luminosity.

We now return to experimental data.

Figures 6-11 are obtained from observations of the synchrotron light radiated by beam particles. The radiation due to an ultra-relativistic charged particle is (to a good approximation) emitted tangent to the particle's path, so that (when the detector is sensitive enough) a picture of this light can be interpreted as a graph of the transverse beam distribution, with light intensity serving as a measure of particle density. Such radiation profiles provide the clearest indications (essentially self-evident) of the extent and shape of beam enlargement.

Figure 6 shows six television photographs of electron and positron synchrotron images taken under various conditions at SPEAR.[9] (Current and energy are not specified.) The indication "with flip-flop effect" means that the RF relative phases have not been adjusted so as to deliberately match the e^+ and e^- profiles. "Flip-flop balanced" means that such an adjustment has been made. These photos are not fine enough to reveal details of light intensity, for which one is referred to Fig. 7, depicting the

results of photometric measurements[9] of bunch dimensions at one and four mA of current per colliding beam. The number under each peak in Fig. 7 is its full width at half-maximum, in millimeters.

The T.V. photographs shown in Fig. 8 were taken at PETRA[11] (the beam energy is not indicated, but is probably around 7 GeV), as was the data represented in the graph of bunch heights vs. time[5] shown here as Fig. 9. Figure 10, similar to Fig. 7, shows synchrotron light profiles measured at ADONE.[12] The energy is not indicated, but is in the vicinity of 1 GeV. ("r" and "v" stand for "radial" and "vertical", equivalent to "x" and "y" in Fig. 7.) Figure 11 shows strong and weak beam profiles for various values of vertical tune, as observed by a television monitor at DCI.[13] (Energy is once again around 1 GeV; current not specified.) Similar blow-up effects have been described in reports from CESR[4] and from PEP.[6]

The results of some less direct SPEAR measurements of the growth of beam height with current are shown in Fig. 12. (As is clear upon inspection of figures 6-11, beam width varies very little, if at all, with current.) I introduce this figure more to illustrate the range of experimental and interpretative techniques that have been applied to this phenomenon, than to add to the evidence of beam blow-up already indicated in the preceeding six figures.

The two curves in Fig. 12 were obtained in two entirely different types of measurements, so there is nothing a priori problematical in their disagreement.

The black dots in Fig. 12 represent a reassembly of data that we've already examined, rather than a direct and independent measurement of beam size. They were obtained by inserting luminosity data (like that shown in Fig. 1) and the unperturbed value of σ_x^* (the rms bunch width) into the right-hand side of the following formula for σ_y^* (the rms beam height)

$$\sigma_y^* = I^2/L\sigma_x^* e^2 Bf \quad . \tag{3}$$

Equation (3) is derived from (2) by writing the effective area A as $4\pi\sigma_x^*\sigma_y^*$, where the assumption of a gaussian distribution accounts for the exact number 4π as a correction for the presence of particles beyond the rms distances.

The white dots in Fig. 12 represent calculations based on beam-lifetime measurements, which are independent of luminosity measurements, and are more directly related to beam size, being immediately sensitive to the flux of particles at large y. The calculations were based on the following formula (derived again assuming a gaussian bunch distribution)[2]

$$\tau = \tau_\beta \left(\frac{\sigma_y}{h}\right)^2 \exp - \frac{1}{2} \left(\frac{h}{\sigma_y}\right)^2 \quad . \tag{4}$$

Equation (4) expresses the beam lifetime τ in terms of: the vertical betatron damping time τ_β; the half-height, h, of the beam pipe at the location of a scraper (assuming a pipe much wider than high); and the rms beam height, σ_y, at that point. The data for τ

at different beam currents is substituted into the left-hand-side of (4), the known values of τ_β and h are substituted into the right-hand side, and the resulting equation is solved for σ_y to produce the open dots in Fig. 12.

(Note that the values of $(\sigma_y)^2$ shown in Fig. 6 are normalized to β_y. According to the theory of quantum (photon) noise fluctuations, σ_y and σ_x are proportional, as functions of storage-ring azimuth, to $\sqrt{\beta_y}$ and $\sqrt{\beta_x}$ (at least for vanishing dispersion). Thus, although the open and closed dots correspond to data taken at two different points (the interaction point and the scraper location), one may interpret the figure as if all the data was extracted at a single location.)

One should be aware that both curves are intrinsically inaccurate, in that both presuppose gaussian charge distributions, while other, independent experiments have shown nongaussian behavior at large distances from the bunch center.[9,10]

II. HOW CONVENTIONAL DESIGN PROCEDURE OVERESTIMATES MAXIMUM ACHIEVABLE CURRENT

To explain the method by which maximum storable colliding currents have conventionally been predicted, it is first necessary to introduce the concept of "beam-beam linear tuneshifts," which are measures of the effect that the whole of one beam has on a single particle of the other beam. (Thus, as there are two beams, there are, strictly speaking, two sets of beam-beam tuneshifts.)

The conventional prediction of maximum current is based on a picture of instability that is most immediately concerned with maximum achievable tuneshifts. This point of view is becoming obsolete today, but has in the past enjoyed considerable influence, despite a very poor predictive record.

The tuneshift is a notion most naturally associated with beams whose distributions have reached a steady state, and whose bunches are not significantly tilted relative to the usual comoving axes. Because of the steady state, a single particle in one beam sees the perturbation due to the whole of the opposing beam as periodic in the usual path length parameter s. (The period is the path length between interaction points, which is the same as the circumference of the ring divided by twice the number (B) of identical bunches in a beam.) Because the tilt angles are negligible, the linear part of this perturbation does not couple x and y; so if xy coupling in the ring magnets is also neglected, then the motion of a test particle around the magnetic lattice and through the perturbation is described in linear approximation by two uncoupled linear 2nd order ordinary differential equations with periodic coefficients. Such a system can be described, as usual, in terms of (horizontal and vertical) tunes per period. The differences, ξ_x and ξ_y, between these tunes and the unperturbed ones are the horizontal and vertical linear beam-beam tune-shifts (per bunch per collision - the total tuneshifts would be these numbers times 2B) due to the beam in question.

Twenty years ago, it was expected that colliding-beam performance would be limited primarily by linear instability,[14] so

that the maximum achievable tuneshift (x or y) would be the smallest number that would give an integer or half-integer when added to the corresponding unperturbed tune. Storage rings before SPEAR encountered instabilities at tuneshifts much smaller than those predicted by the linear theory.[15] The obvious next guess - that nonlinear resonances are responsible for upper limits on the ξ's - never developed into a body of unambiguous criteria. Instead, the idea that the maximum ξ's could be easily computed for each machine from first principles was supplanted by the idea that - for whatever reason - all e^+e^- storage rings are characterized by the same maximum achievable ξ_y. (As will be explained later, rules of thumb for ξ_x are not commonly employed in estimates of the maximum luminosity of e^+e^- machines.) Accordingly, the designers of SPEAR based their maximum current predictions on the maximum vertical tuneshift - .025 - reached at any of the earlier e^+e^- colliders.[16] SPEAR II subsequently performed up to $\xi_y \sim .06$ (although not at all energies - see Fig. 13) so the designers of CESR,[17] PEP,[18] and PETRA[5] based their optimum performance predictions on the supposition that .06 would be the largest ξ_y achievable. At present, none of these machines have exceeded $\xi_y \sim .03$.

In view of this record, how one should predict (short of large-scale computer simulation) the limiting parameters of the next generation of storage rings is not clear.

We now derive expressions for tuneshifts in terms of beam current and beam dimensions, so that we can see explicitly how an overestimate of maximum vertical tuneshift contributes to an

overestimate of maximum current and luminosity. For convenience, we shall assume that the beams are Gaussian. The first two steps in the derivation are proposed as problem 3.

Problem 3

a) Consider a collision between a positron and an electron, initially very far apart, with initial velocities $\vec{v}$ and $-\vec{v}$, and initial impact parameter $\vec{\rho}$ (perpendicular to $\vec{v}$, pointing from positron to electron). Show that in the limit of very high (relativistic) energy (i.e. $v \sim c$), the total change (accumulated from very large negative to very large positive times) in electron momentum due to this collision is

$$\Delta p \sim -\vec{\rho} \left(\frac{2e^2}{c\rho}\right) \tag{5}$$

($\Delta\vec{\rho} = 0$ in the same approximation.)

b) Now replace the positron by a bunch of N positrons whose density function in the spatial plane perpendicular to $\vec{v}$ is a Gaussian of rms widths σ_x and σ_y. Show that when (5) is integrated over the entire positron distribution (beware of singular integrals), the result is, in linear approximation,

$$\Delta p_x \sim \frac{-2e^2 xN}{c\sigma_x(\sigma_x+\sigma_y)} \tag{6}$$

$$\Delta p_y \sim \frac{-2e^2 yN}{c\sigma_y(\sigma_x+\sigma_y)},$$

where x and y are the horizontal and vertical transverse displacements of the electron relative to the positron bunch center.

Formula (6) is applicable to the typical collision in a storage ring when (as is usually the case) the bunch lengths (along $\vec{v}$) are small compared to the scale on which β_x and β_y vary in the interaction region. In such cases, one commonly regards the collisions between single particles and whole bunches as instantaneous; so that one may relate the momentum changes to the more conventional variables of beam dynamics according to $\Delta p_x = mc\gamma\Delta x'$ (and similarly for y), where m is the electron mass and $mc^2\gamma$ its energy; and one may replace the width and height, σ_x and σ_y, by σ_x^* and σ_y^*, their values at the point where the e^+ and e^- synchronous design orbits cross. In toto, for the linear approximation

$$\Delta x' = -\left(\frac{2r_o N}{\gamma\sigma_x^*(\sigma_x^* + \sigma_y^*)}\right) x \qquad (7)$$

$$\Delta y' = -\left(\frac{2r_o N}{\gamma\sigma_y^*(\sigma_x^* + \sigma_y^*)}\right) y,$$

where $r_o = e^2/mc^2 = 2.82 \times 10^{-13}$ cm. Of course the words "electron" and "positron" in the foregoing may be freely interchanged.

Before we continue, you might want, as proposed below in problems 4 and 5, to estimate a couple of orders of magnitude that might give you a feel for the scale of the beam-beam interaction.

Problem 4

How does the typical change in x' or y' due to a beam-beam encounter in PEP compare with the components of transverse velocity of the typical stored electron? To estimate the typical changes in x' and y', use formula (7): γ and N can be computed from Table I and equation (12); x and σ_x^* may be estimated from the table by setting $x \sim \sigma_x^* \sim$ unperturbed σ_x^*, and similarly for y and σ_y^*. To estimate the typical velocity components, use

$$\sqrt{\langle (y')^2 \rangle^*} \sim \sqrt{\langle y^2 \rangle^*}/\beta_y^* \sim \sigma_y^* \Big|_{\text{unperturbed}} /\beta_y^* \qquad (8)$$

(and similarly for x), suggested (at least when β' can be neglected) by the form

$$W_y \equiv \frac{y^2}{\beta_y} + \beta_y \left(y' - \frac{\beta_y'}{2\beta_y} y\right)^2, \qquad (9)$$

of the Courant-Snyder invariant.[19]

Problem 5

How does the electromagnetic energy radiated as a result of the abrupt kick (7) compare with the energy lost to synchrotron radiation during a single turn through the storage ring? To estimate the energy radiated because of the collision, use, for

example, formulae in Ref. 20, assuming that the time elapsed during the kick is of the order of the unperturbed bunch length divided by the speed of light. (For consequences of this kind of radiation—known as "beam-strahlung"—see reference 36.) To estimate the energy radiated during one machine revolution, use[2] the approximation (energy radiated per turn) $\sim E/\tau_e f$, where τ_e is the energy damping time.

We continue toward the tuneshifts:

Following (7), transit of an electron (in linear approximation) through one machine period can be represented by the transformation

$$\begin{pmatrix} y \\ y' \end{pmatrix} \to \begin{pmatrix} \cos 2\pi\mu_y & \beta_y^* \sin 2\pi\mu_y \\ -\frac{1}{\beta_y^*} \sin 2\pi\mu_y & \cos 2\pi\mu_y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{-2r_o N}{\gamma\sigma_y^*(\sigma_y^* + \sigma_x^*)} & 1 \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} \quad (10)$$

(and similarly for x). The first matrix factor corresponds to linear transport between interaction points (μ_y is the unperturbed tune per period; the total unperturbed tune is $\nu_y = 2B\mu_y$), and the second factor corresponds, according to (7), to transport through an interaction region. The last step in the calculation of the ξ's proceeds from (10), and is proposed as problem 6:

Problem 6

Show that the eigenvalues of the transformation in (10) are $\exp 2\pi i (\mu_y + \xi_y)$, where

$$\cos 2\pi(\mu_y + \xi_y) = \cos 2\pi\mu_y - \left(\frac{r_o N\beta_y^*}{\gamma\sigma_y^*(\sigma_y^*+\sigma_x^*)}\right) \sin 2\pi\mu_y.$$

For small current (i.e., small N), this reduces to

$$\xi_y \sim \frac{r_o N\beta_y^*}{2\pi\gamma\sigma_y^*(\sigma_y^* + \sigma_x^*)} \tag{12}$$

and similarly for x.

Equation (12) (together with $N = I/eBf$ and $\sigma_y^* << \sigma_x^*$, which implies that $\sigma_y^*(\sigma_y^* + \sigma_x^*) \sim \sigma_y^*\sigma_x^* = A/4\pi$) leads directly to

$$I \simeq \left(\frac{\gamma eBf}{2r_o\beta_y^*}\right) \xi_y A, \tag{13}$$

which is the basis for the standard prediction of maximum current (and, through (2), maximum luminosity). The maximum current for each of CESR, PEP, and PETRA was predicted by substituting $\xi_{y\ max} = .06$, and the unperturbed A, into the right-hand-side of (13).

How accurate are these predictions? As an example, let us compare prediction with outcome for three-bunch operation of PEP, according to the data shown in Fig. 4. Prediction gives $I_{max} \sim 29$ mA, when $\xi_{y\ max} = .06$ and $A_{unpert} \sim .0029$ cm^2 are substituted into (13) for B = 3 (A_{unpert} is determined by comparing the quadratic part of the one-bunch curve with formula (2)), while the highest three-bunch current shown in Fig. 4 is 3 x 5 = 15 mA, i.e. a shortfall of about 50% (actually 50% over and above single-beam effects, as explained earlier).

In accordance with (13) (despite misgivings one might have about our initial gaussian approximation) we can identify the one-half with a product of one-third from the ratio of $\xi_{y\ max}$ observed at PEP[6] to that predicted, and three-halves from the ratio of A observed (at maximum current) to that predicted. (At least for a rough calculation like this, the observed area can be determined from the three-bunch curve in Fig. 4 by applying formula (2), since for this data the two beams in PEP turned out to be about equally enlarged.[33]) According to (2), a maximum current 50% of that predicted, and a cross-sectional area 150% of that predicted, together mean an 83% shortfall in luminosity (not counting single-beam effects).

At this point a comment on experimental methodology is in order: Colliding beam facilities do not all employ the same criterion in determining the current beyond which the machine should not be operated. At SPEAR,[21] PEP,[21] and PETRA,[5] one considers the threshold to be crossed when background is excessive in high-energy-physics experimental detectors. The rise in background near maximum current is sharp enough that appreciable changes in the cutoff on background would result in only small changes in the measured maximum beam current.[21] A number of reports[11,12,13] from other labs have mentioned beam lifetime as the quantity used to operationally define the threshold.

The skeptical reader may wonder if there is in fact much difference between the maximum currents achievable in single- and colliding-beam modes. In this connection we note that at PETRA,[5]

as of 1980, 20 mA of electrons could be stored in single-beam operation at about 7 GeV, but only up to about 5 mA per beam could be maintained at the same energy in colliding mode.

III. THEORETICAL TRENDS

This concludes our discussion of the experimental situation. I now describe some of the main themes in current theoretical work. This is not meant to be a comprehensive review (see References I, II, and 34). I want simply to give you a feeling for what seem to me to be the more ambitious recent efforts.

The focus will be on three general (and overlapping) areas: computer simulation, models with reduced numbers of degrees of freedom, and instabilities of single-particle motion.

COMPUTER SIMULATIONS

Three computer analyses, numeral simulations of CESR, PETRA, and SPEAR, have attracted particular attention in the last year.

The CESR analysis simulated strong-strong operation. Motion was fully three-dimensional. Radiation damping and quantum noise were taken into account. The published report[22] includes a figure showing a computer-generated curve of luminosity vs. current (at an unspecified tune and energy), superimposed on a scatter of data points from real machine operation. This graph is reproduced here as Fig. 14; the agreement between simulation and experiment is

encouraging. The simulation also produced a value for maximum storable colliding current that comes within about 10% of an experimentally observed value.

The PETRA analysis simulated both strong-strong and weak-strong operation. As in the Cornell study, three degrees of freedom, radiation damping, and quantum noise were included. In addition, some effects of machine imperfection and finite bunch length were taken into account. In the published reports,[23] most of the simulation results are presented in graphs of beam heights vs. time (t=0 is the onset of collisions), and of large-time steady-state beam heights vs. betatron tunes, for different values of energy, tuneshifts, machine imperfection parameters, etc. The accompanying texts contain the suggestion that machine imperfections are the major cause of beam blow-up. These reports claim good agreement between some of the computer results and measurements made directly on PETRA. Unfortunately, very little real storage ring data is actually shown, so a reader cannot evaluate the agreement for himself.

The computer analysis of SPEAR[24] simulated weak-strong operation, and was simplified in that longitudinal oscillations and energy oscillations were neglected. The novel aspect of the work was the use of the x and y damping and fluctuation strengths as variable parameters. The results appear to indicate that vertical beam blow-up is driven primarily by horizontal damping and quantum noise, and is largely insensitive to vertical noise. Spurred by this observation, Tennyson[25] has attempted to formulate an

intuitive picture of beam blow-up based on the notion of "resonance streaming." (No comparable heuristic ideas have yet developed from the other storage ring computer simulations.) Since resonance streaming is to be discussed in Tennyson's contribution to this summer school, it will not be considered further here.

When personally assessing the importance of computer simulation - as compared with analytical work - the reader ought to bear in mind first that those who actually program and interpret simulations do not all seem to agree on what is "at the bottom" of colliding-beam effects in real machines. Second, successes of electron-positron computer simulations do not ipso facto ensure comparable successes in the proton-antiproton or proton-proton cases. The amount of computer time required for a simulation grows with the number of storage-ring turns to be simulated; a meaningful study models at least the number of revolutions in a radiation damping time. For electrons in CESR, this number is about 10^3. For protons in existing or foreseeable high-energy machines, this number is larger by several orders of magnitude.

Problem 7

Compute the number of machine revolutions in the radiation damping time of a proton in the CERN SPS; in the Fermilab Main Ring; in the Tevatron, presently under construction at Fermilab. (See Ref. 32 for relevant specifications.)

MODELS WITH FEW DEGREES OF FREEDOM

Many analytical investigations of colliding beam effects have concentrated on simplified model systems having fewer than three degrees of freedom per particle.

The one-dimensional model involves motion only in the vertical coordinate. Such a simplification is justified to the extent that one can: neglect finite bunch lengths; neglect xy coupling in the magnetic lattice; and take σ_x^*, for either beam, to be so much larger than σ_y^* that in the integral of the impulse (5) over the beam distribution, one need retain only the leading order in $1/\sigma_x^*$, i.e.

$$\Delta y' \simeq -\frac{2r_o N}{\sigma_x^* \gamma} \int_0^{y/\sigma_y^*} e^{-u^2/2}\, du, \Delta y \simeq 0 \qquad (14a)$$

$$\Delta x' \simeq 0 \simeq \Delta x, \qquad (14b)$$

which clearly does not couple x and y. In the one-dimensional model, the development of y and y' upon transit of an electron through one machine period is described by the functional composition of (14a) with the linear transformation corresponding to the first of the two matrix factors in (10). The importance of ξ_y, compared to ξ_x, in the conventional approach to estimates of anticipated maximum luminosity is based on the approximations underlying the one-dimensional model.

The one-and-one-half dimensional models are refinements of the one-dimensional model. They were introduced in order to study the

effect of synchrotron oscillations on the stability of motion in the vertical coordinate. They differ from the one-dimensional case in the appearance of explicit time-dependence in the transformation that describes the development of y and y' upon transit through a machine period. The explicit time dependence is typically periodic, at the synchrotron frequency. (The time parameter must of course be limited to those discrete values at which the beam-beam collisions actually take place; i.e. to $\{ nt_o + \delta \}$, where n is an unrestricted integer, t_o is the time required for a beam particle to travel between interaction points, and δ is an arbitrary but fixed offset.)

Problem 8

In positing the kind of explicit, steady time-dependence that is customary in such models, one tacitly assumes that all particles in a given bunch undergo uninterrupted harmonic synchrotron oscillations, all with the same amplitude, all in phase. This can be, at best, a tentative working hypothesis because there are too many particles in a bunch to be correlated in this way, because the synchrotron oscillations can have an anharmonic component, and because the synchrotron oscillation of even a single particle is subject to frequent random disturbance by photon emission. Estimate, for a typical particle in single-beam operation of PEP, the times required for random photon emission to substantially change the amplitude and phase of linear synchrotron oscillation; do the same for horizontal and vertical betatron oscillation. (Phase and amplitude decorrelations due to nonlinearities are harder to estimate, although they may be more important.[27]) Presumably, any one-and-one-half-dimensional

instability that develops slowly compared with such a decorrelation time is of a priori questionable importance in a real storage ring.

Three forms of time-dependence are most frequently discussed;

1) Energy oscillations of the transverse coordinates are modeled by making the replacement

$$y \to y + a\sin\Omega t \tag{15}$$

in the right-hand-side of (14a). Ω is the synchrotron frequency; a is proportional to the vertical dispersion (η_y^*) at the interaction point. The relevence of this substitution is debatable, as reports often claim $\eta_y \equiv 0$.

2) Energy oscillations of the tune are modeled by making the replacement

$$\mu_y \to \bar{\mu}_y + m_y \sin\Omega t \tag{16}$$

in the factor that corresponds to magnet transport in the right-hand-side of (10). The amplitude m_y is proportional to the chromaticity.

3) Finite bunch lengths are modeled by making the decompositions

$$\begin{aligned} \beta_y^* &= \bar{\beta}_y^* + b_y \sin\Omega t \\ \beta_x^* &= \bar{\beta}_x^* + b_x \sin\Omega t \end{aligned} \tag{17a}$$

(and therefore

$$\left.\begin{aligned} \sigma_y^* &\propto \sqrt{\bar{\beta}_y^* + b_y \sin\Omega t} \\ \sigma_x^* &\propto \sqrt{\beta_x^* + b_x \sin\Omega t} \end{aligned}\right), \tag{17b}$$

in the right-hand-side of (14a), and in the first matrix factor in (10). The amplitudes b are determined by the extent to which the β-functions near the interaction points vary over a scale set by the bunch length. At SPEAR, such variations can reportedly be appreciable.[27]

Problem 9

As is well-known,[2] the beta functions in an interaction region obey

$$\beta_{x,y} = \beta^*_{x,y} \left(1 + (s/\beta^*_{x,y})^2\right), \tag{18}$$

where s is distance, measured along the beam pipe, from the intersection point of the e^+ and e^- synchronous design orbits. Using (18) and the parameter values in Table I, compare β_x and β_y at s = 0 with β_x and β_y at s = one bunch length, for PEP.

Problem 10

For certain parameter values, a one-and-one-half dimensional model can be reduced to a one dimensional model by a change of

variables. Since the typical one dimensional model is thought to be more stable than the typical one-and-one-half-dimensional model (see below), it may be desirable to build machines that can be operated in such exceptional configurations.[28] Consider, for example, the type-3 one-and-one-half-dimensional model of an imaginary storage ring whose horizontal and vertical β-functions are identically equal. Show that the transformation describing transit of an electron through one period of this machine, in this model, loses its explicit time dependence when expressed as a transformation on $\tilde{y} \equiv y/\sqrt{\beta_y^*}$ and $\tilde{\tilde{y}} \equiv y'\sqrt{\beta_y^*}$, rather than on y and y'.

INSTABILITY STUDIES

A number of authors[26,29,30,31] have performed numerical and analytical studies of the long-time stability of single-particle motion in one- and one-half-dimensional models. These analyses all neglect damping and noise. In mathematical terms, they are concerned only with maps of (y,y')-space formed by iterating, over and over, the single-machine-period transformations. For a number of years, efforts have been underway to determine what relations, if any, exist between the results of such stability analyses and the observed behavior of real e^+e^- storage rings.

Studies of the one-dimensional model have revealed that for $\xi_y \left[\equiv (r_o \beta_y^* N)/(2\pi\gamma q_x^* \sigma_y^*)\right]$ less than a critical value near .25, all orbits in vertical phase space are bounded. An orbit is a

sequence of points in phase space, each related to the one before it by an application of the single-period transformation. (In the one-and-one-half-dimensional case, the time parameter in each application of the transformation exceeds that in the preceeding application by the time required for a beam particle to travel between interaction points.) For ξ_y significantly less than the critical value, the typical orbit appears to the eye to lie on a closed curve, approximated well by an ellipse

$$\frac{y^2}{\beta_y^*} + (y')^2 \beta_y^* = \text{constant.} \qquad (19)$$

(The left-hand-side of (19) is just the Courant-Snyder invariant for the linear system obtained by setting N=0.) For ξ_y just below the critical value, there are orbits that appear to the eye to densely fill two-dimensional regions of the phase plane. The dynamics in such a region are highly unstable, in the sense that no two orbits that momentarily approach one another there can stay close indefinitely. For ξ_y greater than the critical value, one of these regions contains both the origin and points with arbitrarily large y or y' - so that most orbits that begin near the origin are unbounded.

One would have liked to identify this critical ξ_y with the maximum tuneshifts observed in electron-positron colliders, although as yet there is no solid theoretical justification for doing so; in any case, .25 is too large. In this regard, the one-and-one-half-dimensional models are more appealing. There are one-and-one-half-dimensional models exhibiting the same type of

stability-to-instability transition, but at values of $< \xi_y >$ (i.e. ξ_y averaged over a modulation period) considerably smaller than the one-dimensional critical value.

OUTLOOK

At present, it is hard to designate any one theoretical program as the most promising. The criterion should be agreement with experiment. So far, only the largest numerical simulations have produced results that can be meaningfully compared with experimental observations. Analytical theory remains primitive.

ACKOWLEDGEMENTS

I am grateful to J. D. Bjorken, M. Month, and C. Quigg for encouraging my interest in this problem, and to J. Rees and J. Tennyson for helpful conversations.

REFERENCES

I. M. Month and J. C. Herrera, eds., Nonlinear Dynamics and the Beam-Beam Interaction, AIP Conferrence Proceedings no. 57 (1979).

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. Proceedings of the Workshop on Long-Time Prediction in Nonlinear Conservative Dynamical Systems, University of Texas, Austin, Match 16-19, 1981 (to be published).

1. CERN Courier, May 1981, pp. 143-144.
2. M. Sands, in Physics With Intersecting Storage Rings, B. Touschek, ed (Academic Press, New York, 1971) pp. 257-411.
3. H. Wiedemann, in II, pp. 33-39.
4. CESR Operations Group, in III, pp. 26-37.
5. PETRA Storage Ring Group, in III, pp. 16-25.
6. J. Rees, invited paper presented at 1981 Particle Accelerator Conference, Washington, D.C., March 11-13, 1981, SLAC-PUB-2684, PEP-NOTE-347.
7. M. H. R. Donald and J. M. Paterson, IEEE Trans. Nucl. Sci. NS-26, 3580 (1979).
8. A. W. Chao and M. J. Lee, J. Appl. Phys., 47, 4453 (1976); IEEE Trans. Nucl. Sci., NS-24, 1203 (1977).

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

10. M. Cornacchia, in I, pp. 99-114.

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

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

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

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

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

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

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

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

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

20. J. D. Jackson, Classical Electrodynamics (Wiley, New York, 1962).

21. H. Wiedemann, private communication.

22. S. Peggs and R. Talman, in II, pp. 21-32.

23. A. Piwinski, in III, pp. 751-75; DESY preprint 80/131; DESY internal report M 81 G.

24. J. Tennyson, in II, pp. 1-20.

25. J. Tennyson, in IV.

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

27. J. Tennyson, private communication.

28. J. B. Vasserman, F. M. Izraelev, G. M. Tumaikin, preprint I.

Ya. F. 79-74 (in Russian), Institute for Nuclear Physics, Novosibirsk.

29. B. V. Chirikov, Phys. Repts. 52, 263 (1979).

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

31. A. Ruggiero, in Proc. 9th Int. Conf. High Energy Accelerators, Stanford, 1974, pp. 419-423.

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

33. J. Rees, private communication.

34. S. Kheifets, in IV (SLAC-PUB-2700, PEP-NOTE-346).

35. A. Ruggiero, Fermilab internal report FN271 (1974).

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

TABLE I

Selected PEP Design Parameters

(for use in doing the exercises,

from Ref. 18)

f	= 136 kHz
E	= $mc^2\gamma$ = 15 GeV
B	= 1 or 3
β_x^*	= 3.7 m
β_y^*	= 0.2 m
σ_x^*	= 1.1 mm
σ_y^*	= .06 mm
ν_x	= 18.77
ν_y	= 19.26
τ_β	= 8.2 msec
τ_ε	= 4.1 msec
bunch length	= 4 cm
Ω	= .4 radian/revolution
$\xi_{y\text{-max}}$	= .06 (but use more realistic .02 in the exercises)
κ	= xy linear coupling = .27
α	= momentum compaction factor = .004

FIGURE CAPTIONS

Fig. 1: Luminosity vs. current per beam at SPEAR, from Ref. 3.

Fig. 2: Luminosity vs. current per beam at CESR, from Ref. 4.

Fig. 3: Specific luminosity vs. single-bunch current at PETRA, from Ref. 5.

Fig. 4: Luminosity vs. single-bunch current at PEP, from Ref. 6.

Fig. 5: Luminosity vs. single-bunch current at PEP, from Ref. 6.

Fig. 6: Television photographs of beam cross-sections at SPEAR, from Ref. 9.

Fig. 7: Transverse beam distributions at SPEAR, from Ref. 9.

Fig. 8: Television photographs of beam cross-sections at PETRA, from Ref. 11.

Fig. 9: Beam heights vs. time at PETRA, from Ref. 5.

Fig. 10: Transverse beam distributions at ADONE, from Ref. 12.

Fig. 11: Transverse beam shapes vs. fractional part of vertical tune at DCI, from Fig. 13.

Fig. 12: Square of beam height, normalized to β_y, vs. current per beam at SPEAR. Solid points derived from luminosity data; open points derived from lifetime measurements. From Ref. 3.

Fig. 13: Maximum vertical tuneshift vs. energy at SPEAR, from Ref. 3.

Fig. 14: Luminosity vs. current at CESR, according to numerical

simulation (solid curve) and experimental observation (open dots), from Ref. 22.

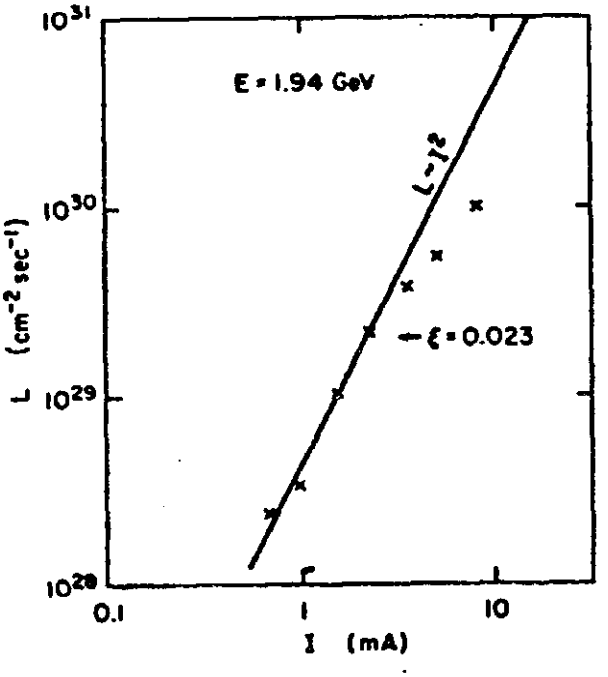


Fig. 1

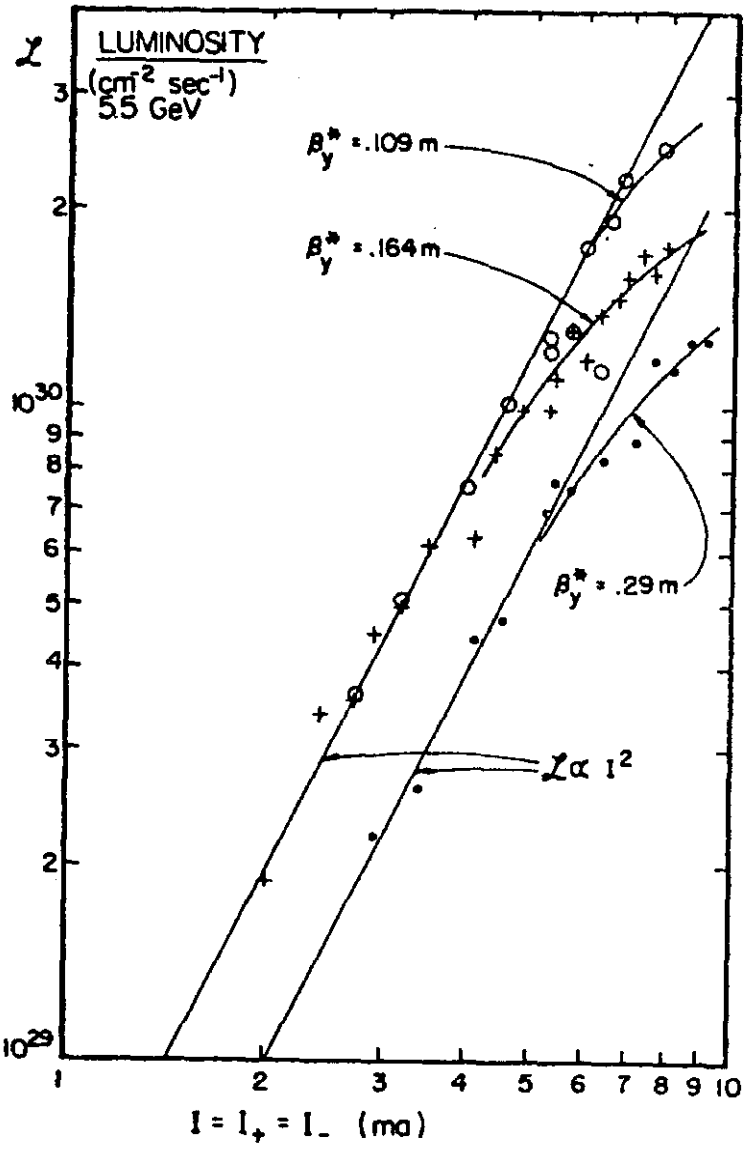

LUMINOSITY
(cm⁻² sec⁻¹)
5.5 GeV
$\mathcal{L}$
β_y^* = .109 m
β_y^* = .164 m
β_y^* = .29 m
$\mathcal{L} \propto I^2$
10^{30}
10^{29}
I = I_+ = I_- (ma)

Fig. 2

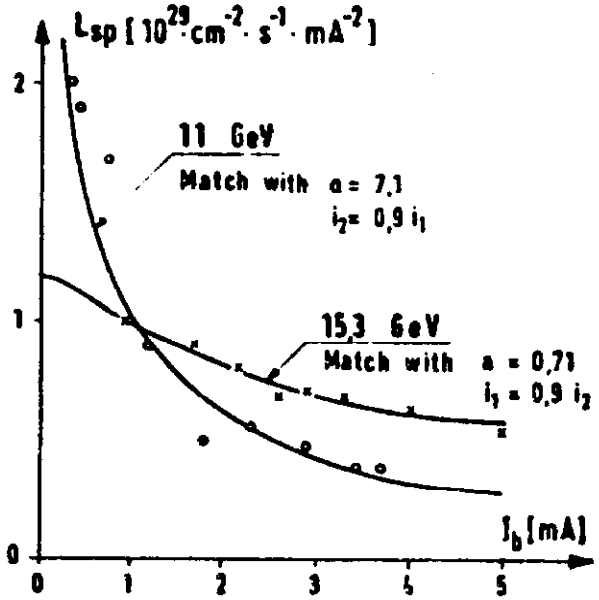

L_{sp} [10^{29} $cm^{-2} \cdot s^{-1} \cdot mA^{-2}$]
2
1
0
11 GeV
Match with a = 7,1
i_2 = 0,9 i_1
15,3 GeV
Match with a = 0,71
i_1 = 0,9 i_2
I_b [mA]
0
1
2
3
4
5

Fig. 3

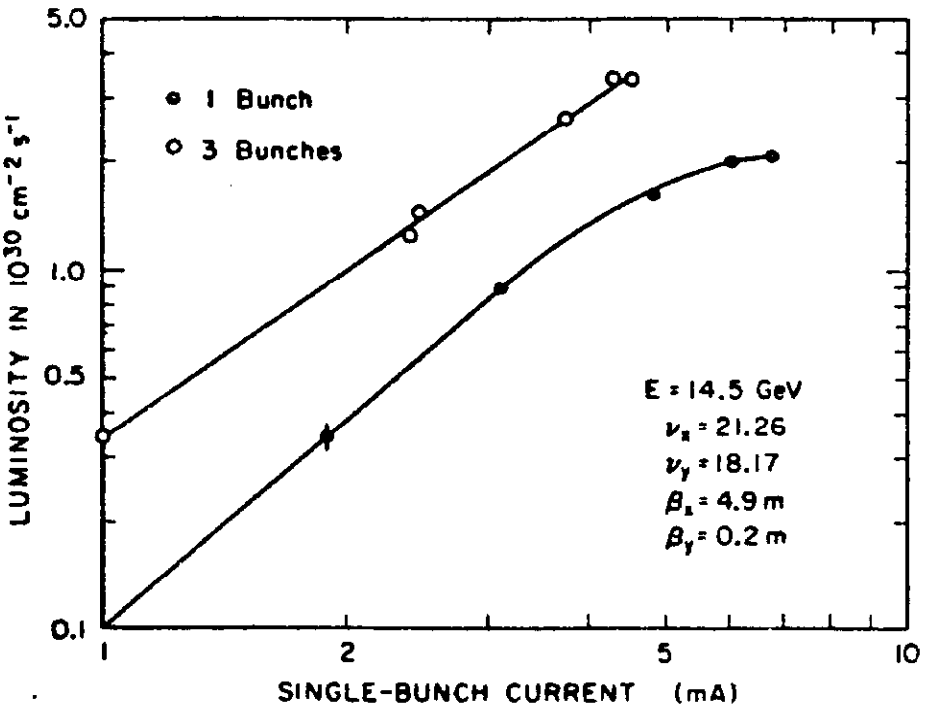

5.0
1.0
0.5
0.1
LUMINOSITY IN 10^{30} cm^{-2} s^{-1}
1 Bunch
3 Bunches
E = 14.5 GeV
ν_x = 21.26
ν_y = 18.17
β_x = 4.9 m
β_y = 0.2 m
1
2
5
10
SINGLE-BUNCH CURRENT (mA)

Fig. 4

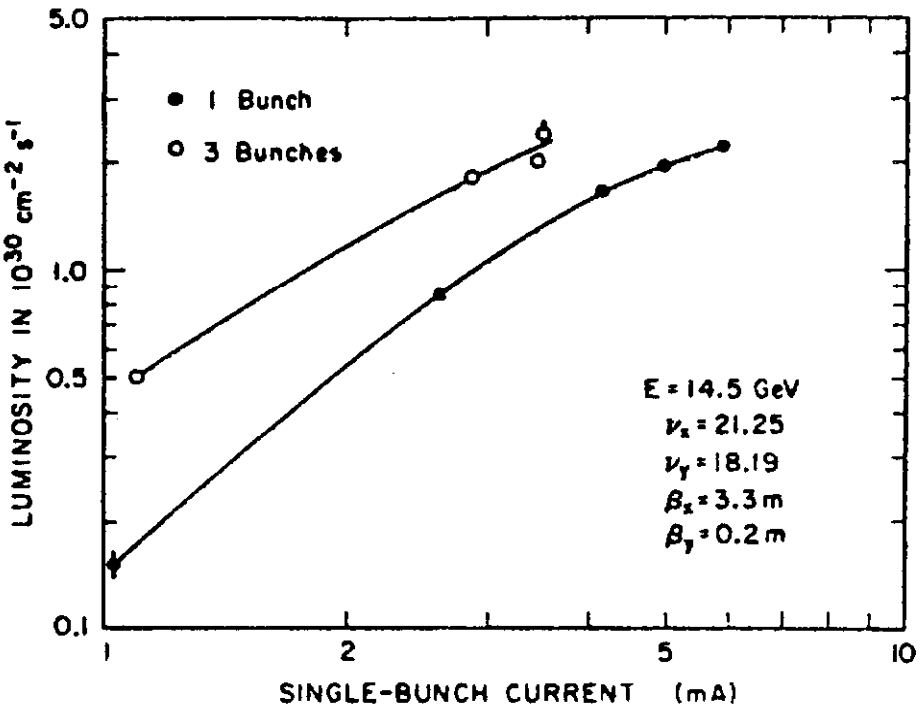
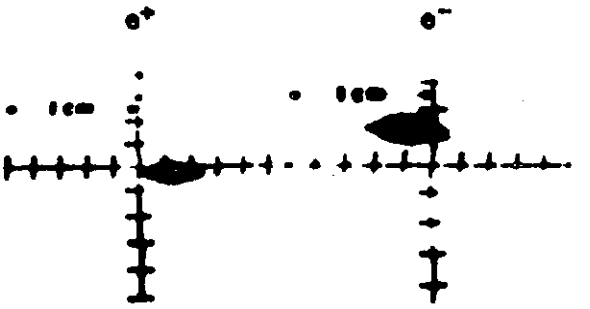
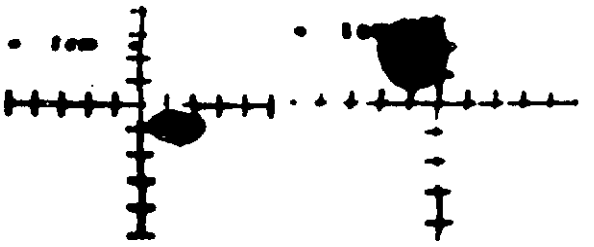

5.0
1.0
0.5
0.1
1
2
5
10
LUMINOSITY IN 10^{30} cm^{-2} s^{-1}
SINGLE-BUNCH CURRENT (mA)
● 1 Bunch
○ 3 Bunches
E = 14.5 GeV
ν_x = 21.25
ν_y = 18.19
β_x = 3.3 m
β_y = 0.2 m

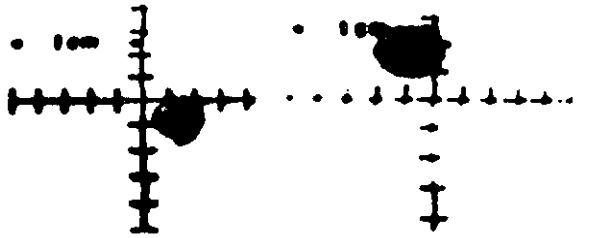
Fig. 5



(a) Separated Beams



(b) Colliding Beams With Flip Flop Effect



(c) Colliding Beams Flip Flop Balanced

Fig. 6

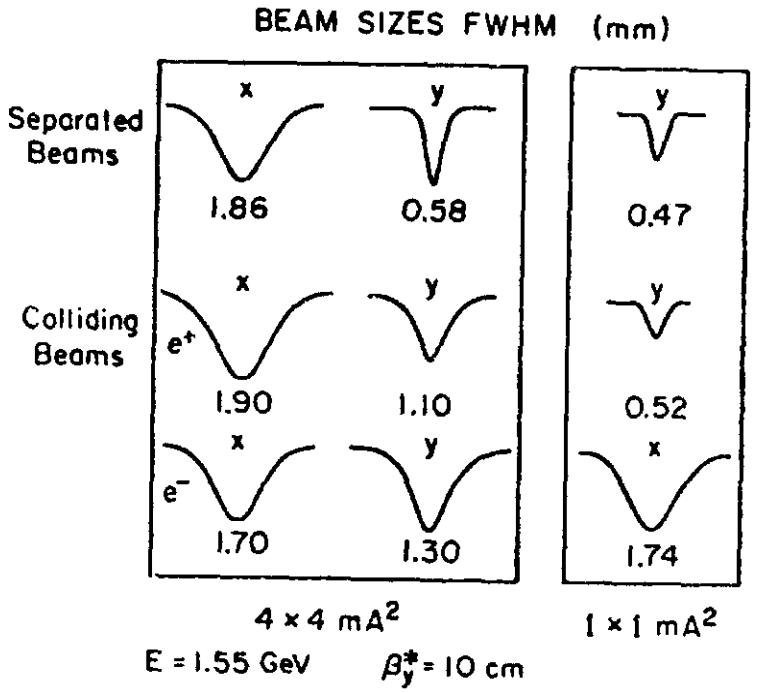
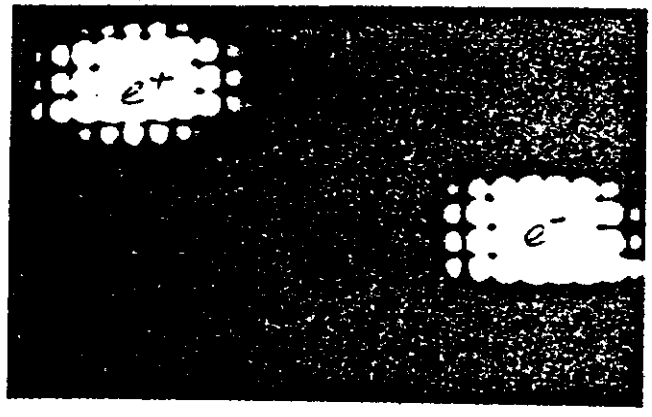
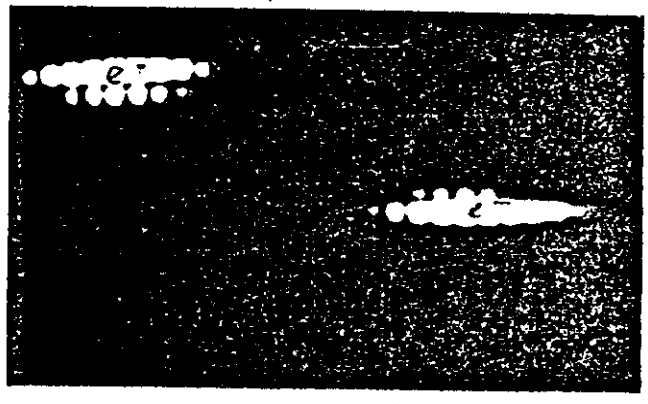

BEAM SIZES FWHM (mm)
Separated Beams
Colliding Beams
x
y
1.86
0.58
0.47
e+
1.90
1.10
0.52
e−
1.70
1.30
1.74
4 × 4 mA2
1 × 1 mA2
E = 1.55 GeV
β*y = 10 cm

Fig. 7



Beam dimensions at $I^+ = I^- = 0.3$ mA.



Beam dimensions at $I^+ = I^- = 0.15$ mA.

Fig. 8

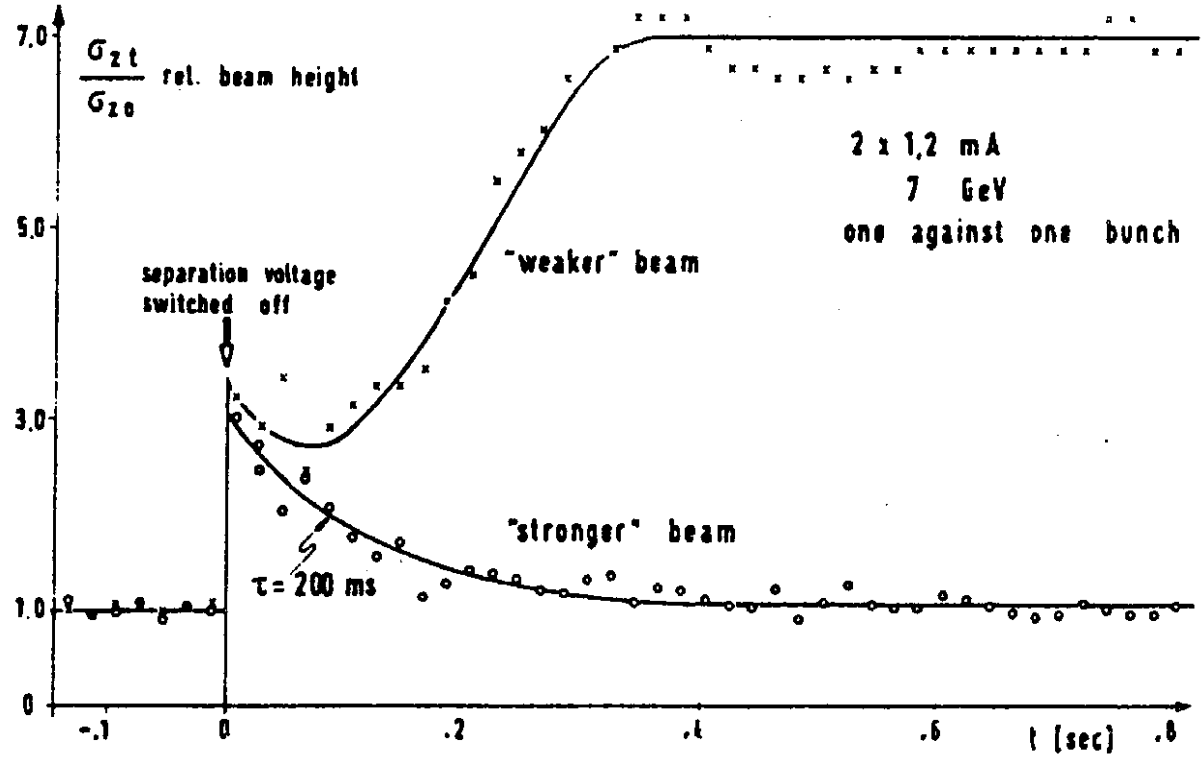
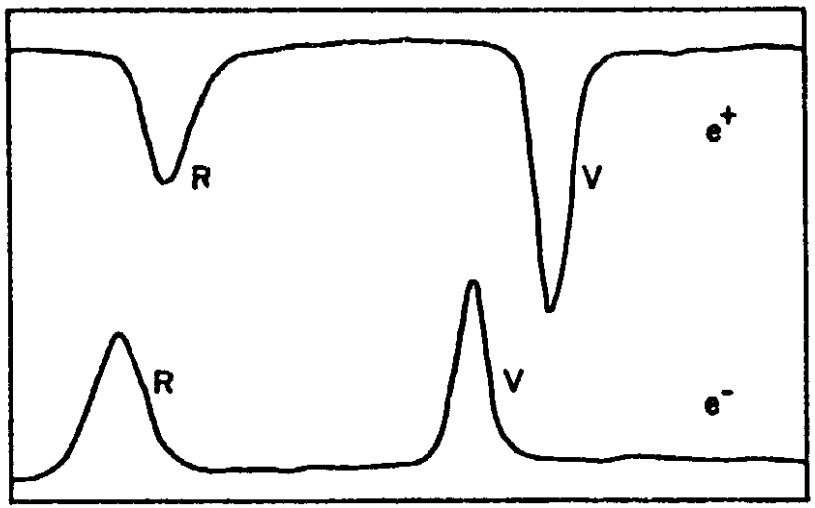

7.0
5.0
3.0
1.0
0
$\frac{\sigma_{z\,t}}{\sigma_{z\,0}}$ rel. beam height
2 x 1,2 mA
7 GeV
one against one bunch
separation voltage
switched off
"weaker" beam
"stronger" beam
τ = 200 ms
-.1
0
.2
.4
.6
.8
t [sec]

Fig. 9



Interacting strong beams at or near the space charge limit.

Weak beam is "flipped" due to excessive current in the strong beam. The strong beam has its unperturbed vertical dimension.

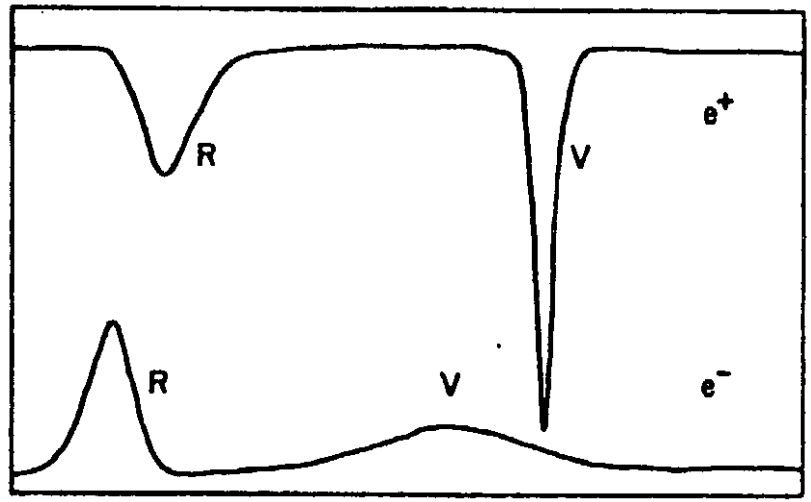


Fig. 10

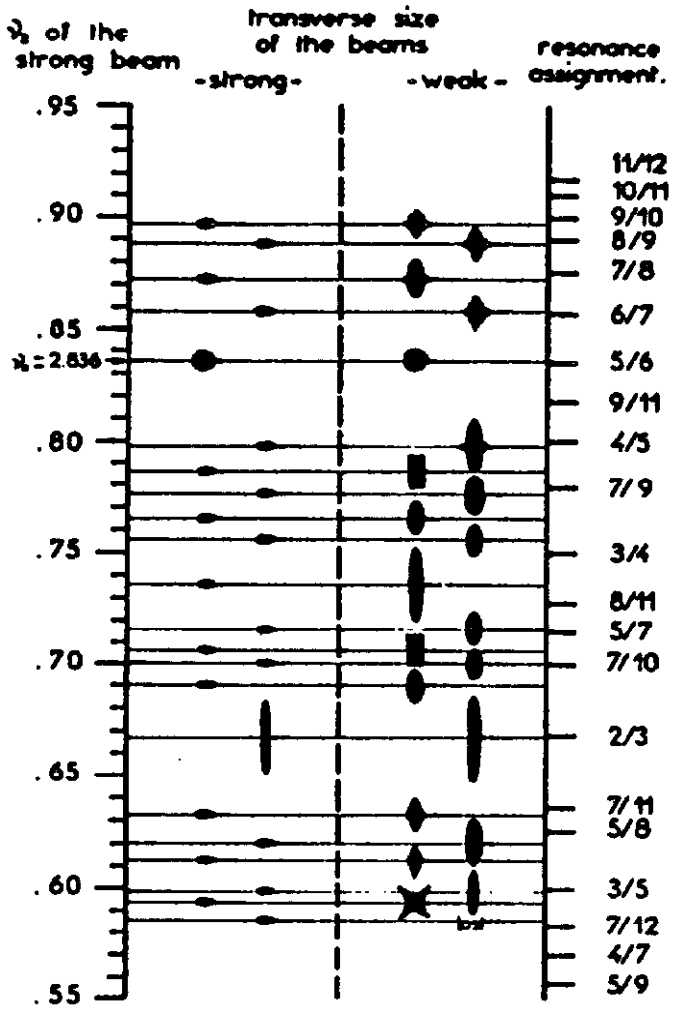

ν_x of the strong beam
transverse size of the beams
-strong-
-weak-
resonance assignment.
.95
.90
.85
$\nu_x = 2.836$
.80
.75
.70
.65
.60
.55
11/12
10/11
9/10
8/9
7/8
6/7
5/6
9/11
4/5
7/9
3/4
8/11
5/7
7/10
2/3
7/11
5/8
3/5
7/12
4/7
5/9

Fig. 11

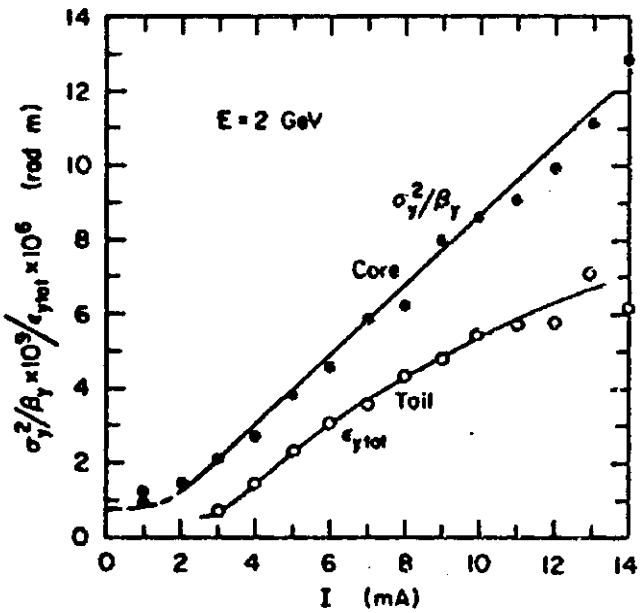

E = 2 GeV
σ_y^2/β_y
Core
Toil
$\epsilon_{y\,tot}$
$\sigma_y^2/\beta_y \times 10^9/\epsilon_{y\,tot} \times 10^6$ (rad m)
I (mA)
0
2
4
6
8
10
12
14

Fig. 12

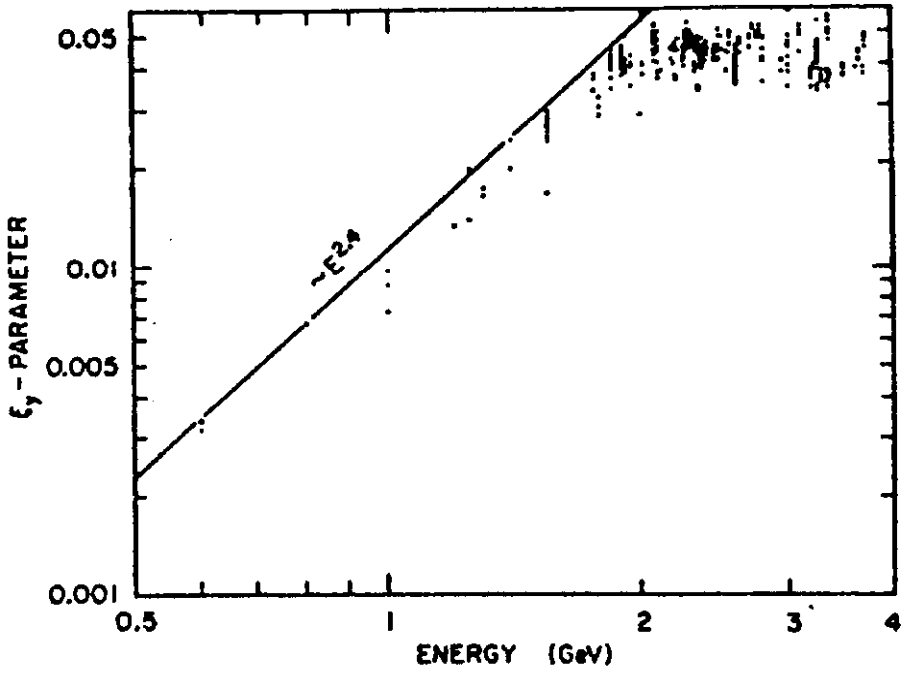


Fig. 13

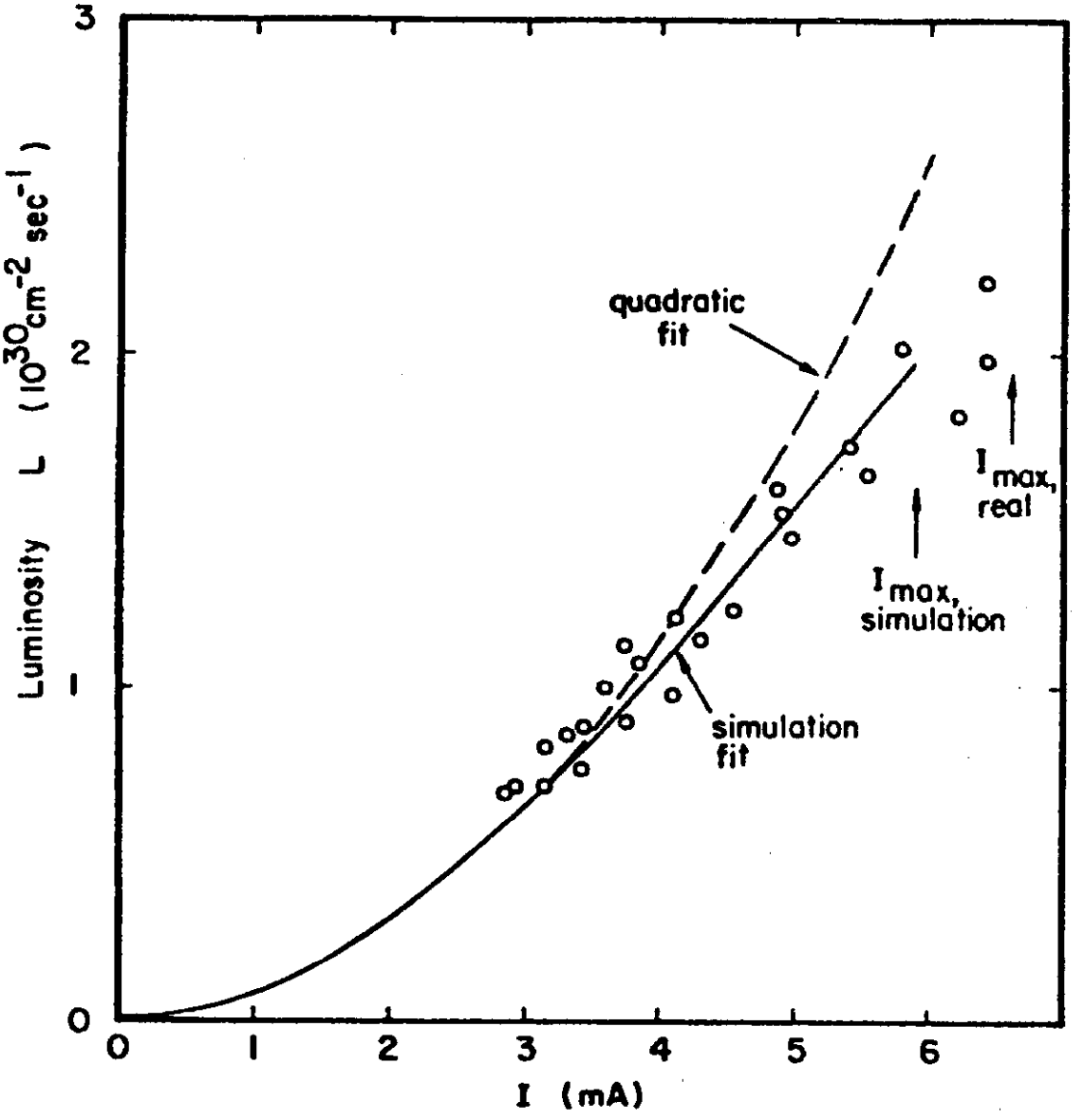

Luminosity L (10^{30}cm^{-2} sec^{-1})
I (mA)
quadratic fit
simulation fit
$I_{max, real}$
$I_{max, simulation}$
0
1
2
3
4
5
6

Fig. 14