Particle Accelerators, 1996, Vol. 52, pp. 65–93
Reprints available directly from the publisher
Photocopying permitted by license only

BEAM-BEAM SIMULATIONS AT LARGE AMPLITUDES AND LIFETIME DETERMINATION

DMITRY SHATILOV

Budker Institute of Nuclear Physics, 630090, Novosibirsk, Russia

(Received 31 March 1995; in final form 31 July 1995)

A special tracking technique for beam-beam simulation in circular e^+e^- colliders is discussed. This technique emerged from a concept proposed earlier by J. Irwin.[1] It is mainly intended for determination of beam halo and lifetime and allows us to reduce the required CPU time by several orders of magnitude. Equilibrium distribution within the amplitude space is built step by step beginning from the core region. During a step, only those particles are tracked which fall into "external" region of the amplitude space. The border of this region is moved to large amplitudes step by step and special border conditions are used to "sew" together distributions for "internal" and "external" regions. These border conditions are taken from the previous step. A special method of building such borders is proposed to avoid a possible loss of accuracy. The technique was tested and applied to several working colliders and projects.

Keywords: Beam-beam effects, simulation, tail distribution, lifetime

1 INTRODUCTION

We study incoherent beam-beam effects. Strong nonlinearity of the interaction leads to resonances and stochasticity, so that the behavior of the particles in the phase space becomes very complicated for analytical estimations. This is the reason for a wide employment of computer simulation for studying the beam-beam effects.

The perturbations of the equilibrium distribution can be conventionally divided into two parts: the perturbations at small amplitudes (core region) and the ones at large amplitudes (tails). The former cause an emittance growth, the latter result in the lifetime decrease. Both effects restrict the beam current and, therefore, the luminosity, but the connection between them is rather weak. This means that one can have a strong beam size growth with a very good lifetime and vice versa, a bad lifetime with almost an unperturbed core region. The lifetime determination is the most difficult problem here, because a huge CPU time is required to correctly determine the distribution in the tails. Besides, almost all the CPU time is spent

for improving core distribution. Such situation motivates the search for possibility of "CPU redistribution" for large amplitudes. Naturally, the question of accuracy arises since the tracking algorithm is changed artificially. Maintaining accuracy is the main problem we shall discuss below.

The base idea of the technique described below has been proposed in 1989 by J. Irwin.[1] It was developed later by T. Chen, J. Irwin and R. Siemann.[9] In INP this concept was developed independently by the author. Actually, this paper is a simple translation of the preprint INP 92–79 (in Russian[7]). It should be noticed, that justification and realization of Chen *et al.*'s and the author's methods are rather different, although recently carried out comparison indicated good agreement between both of them and the brute force technique.

Section 2 contains the detailed description of the technique for the two-dimensional case. The first part of this section can be considered as the rendering of J. Irwin's paper,[1] but this is important for the future discussion and cannot be omitted. In Section 3 the technique is extended to include the scattering on the residual gas. Section 4 discusses the question of how to extend the technique to the "strong-strong" case. In Section 5 the tracking for three-dimensional case is considered. Section 6 presents some results of the technique application.

2 DESCRIPTION OF THE METHOD

There are different ways to simulate the equilibrium distribution of the beam particles. The first one, the ensemble of N independent particles is considered. If the simulation time is much greater than the damping time, the initial positions of these particles do not matter. Otherwise, they will influence the results and one has to care about correct initialization. Another way is watching for a single particle only. If the simulation time here is chosen N times greater than in the first case, both of them become equivalent, but there is no the initialization problem in the last one. For future discussion it does not matter which way is used, but for definiteness we consider a single particle for a long time (hundreds of damping times).

Besides coordinates and momenta, we use normalized amplitudes and phases as well:

$$x_\beta = \sigma_{\beta x} A_x \cos\varphi_x \qquad x'_\beta = (\sigma_{\beta x}/\beta_x) A_x (\sin\varphi_x + \alpha_x \cos\varphi_x) \tag{1}$$

$$z_\beta = \sigma_{\beta z} A_z \cos\varphi_z \qquad z'_\beta = (\sigma_{\beta z}/\beta_z) A_z (\sin\varphi_z + \alpha_z \cos\varphi_z) \tag{2}$$

$$s = \sigma_l A_s \cos\varphi_s \qquad \frac{\Delta E}{E_0} = (\sigma_E/E_0) A_s \sin\varphi_s \tag{3}$$

Here all the parameters (σ, β, α) are related to an unperturbed beam. During simulation, the particle moves along a very complicated trajectory within a 6-dimensional phase space. The equilibrium distribution and the density of the flows in this space are the most complete information one can extract, but not all of it is of equal importance. The distribution in the space of amplitudes is much more important because the amplitudes are "slow" variables. The amplitudes define the nonlinear tune shifts and, therefore, the falling within a resonance. The distribution in the space of phases is not so informative since the phases are "quick" variables, but this distribution is also important for the motion within resonances. As far as the flows are concerned, the most important one is the flow to the aperture (which is defined as the limit values of the amplitudes). This flow, not the distribution in the tails, defines the lifetime, though there is a connection between these two.

For the sake of simplicity, let us consider the two-dimensional case (four-dimensional phase space). The three-dimensional case will be discussed separately in Section 5. We observe the test particle and locate its position in the plane of amplitudes at certain moments each turn (at the Interaction Point), so its motion looks like "jumps" from point to point. These jumps are produced by damping, noise and kicks from the opposite beam. The lattice nonlinearity, if any, also makes a contribution.

Now we draw a special line, or boundary, to split the amplitude plane into two parts: *internal* (I) and *external* (II) (see Figure 1). During the first step of the new algorithm we record all particle's coordinates and momenta each time it leaves the internal region for the external one (at once after crossing the boundary). These points (we call them *outflights*) form an original halo above the boundary (see Figure 2). After some time, we get enough statistics of such outflights to proceed to the next step. Let us imagine that the internal region becomes a "black box" (or *hidden* region), so that we cannot watch a particle within it. In this case, the particle's trajectory looks like a large number of completely independent pieces, which starts from the points of outflights and ends when the particle falls into the hidden region. During the second step, we put the test particle to one of the outflight points immediately when it leaves for the hidden region. Such a procedure is called *restart*. The points for the restarts are chosen from the complete outflights statistics by using a random number generator.

It should be noted that this algorithm violates the course of events with time. This means that the observer is able to distinguish a real trajectory from our simulation. The reason is as follows: when the particle falls into the hidden region, it spends some time near the boundary and there is a high probability of crossing the boundary many times in close proximity to this place. In other words, the outflights are usually made by local groups, before the particle goes away from the boundary. This means that there is a correlation between the location of outflights and the time, when they

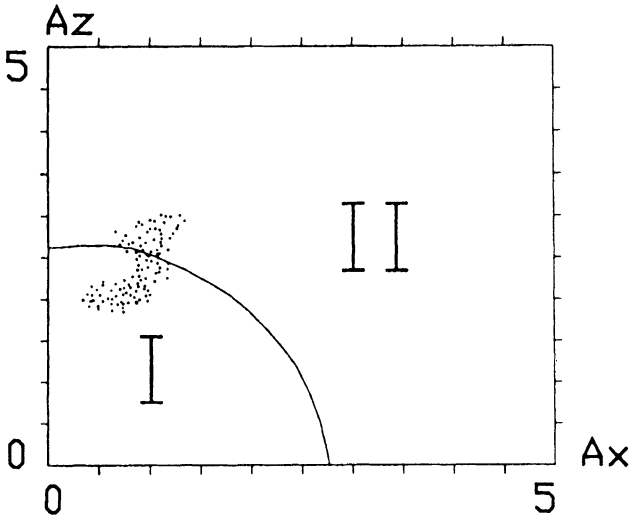

FIGURE 1: A plane of amplitudes is split into internal (I) and external (II) regions. A small fraction of the particle's trajectory is shown by points.

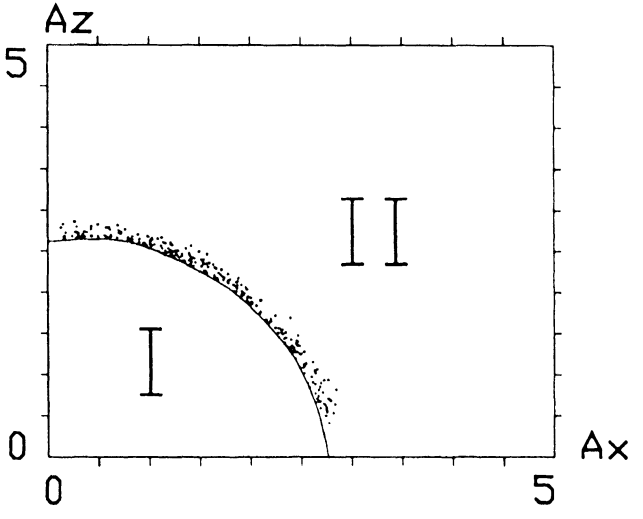

FIGURE 2: The outflights are shown by points, which form a specific halo above a boundary. In future these points will be used as the positions for restarts.

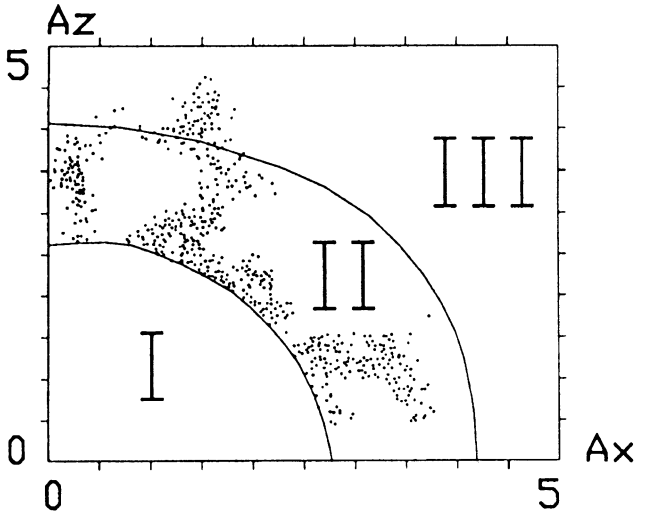

FIGURE 3: The particle's trajectory during the second step, consisting of a large number independent pieces. I — hidden, II — internal, III — external regions.

take place. But this correlation vanishes since we use a random number generator to choose the points for restarts. Luckily, such rearrangement of events in time has no influence on the motion characteristics: we obtain quite correct distribution and flows in the phase space during the second step. The reason is that the process under consideration is a Markovian process (without memory). The particle's behavior does not depend on its history, and since we reproduce the probability of falling into a certain cell of the phase space when the particle leaves the hidden region, we must obtain the correct density and flows within this space.

Now it is easy to understand how to proceed to large amplitudes. After completion of the first step the regions change their meaning: the internal region becomes hidden, and the next boundary is drawn to split the "old" external region into internal and external ones (see Figure 3). During the second step the position of the particle is checked each turn. As soon as it leaves for the hidden (I) region, a restart is produced. In case it moves from the internal region (II) to the external one (III), all its coordinates and momenta are registered to accumulate the statistics of outflights across the second boundary. Later on, we call the boundary between the hidden and internal regions as R-boundary (boundary for restarts); the boundary between the internal and external regions we call C-boundary (boundary for crossings). For the third step, the regions I and II are joined to form the new hidden region (so the C-boundary becomes the R-boundary), and the third boundary (it is the C-boundary

for the third step) is drawn to split region III into the internal and external regions (their numbers are III and IV, respectively), and so on. The reduction of CPU time can be calculated as a ratio between the number of restarts and the number of corresponding outflights. Usually this value is about 5–10 for each step.

Approximately in this form this technique has been already suggested by J. Irwin.[1] He used simple circular arcs as boundaries and the radius of the arc was chosen so that particles spent 90% of the CPU time in the internal region and 10% in the external one. This solution is very attractive due to its simplicity and can give good result, but sometimes loss of accuracy is possible in the case where the lifetime is calculated. Maintaining accuracy is the main problem here. The difficulties result from our wish to develop a universal method which must work correctly for systems with very different (practically arbitrary) "phase space portraits".

Rigorously speaking, the only source of errors is the incorrect (insufficient) statistics of outflights from the hidden region. However, the accuracy of the results depends on many factors: simulation time, damping times, shape and location of boundaries within the amplitude plane. It is easy to guess that there is direct connection between the accuracy and the decrease in the CPU time this method provides. The accuracy can be increased at the expense of efficiency, for example, by increasing the simulation time for each step and decreasing the distance between the boundaries. The aim is to achieve an optimal compromise here: to get high efficiency with good enough guaranteed accuracy.

The amplitude of noise measured in units of normalized amplitude is defined by the damping time according to the formula:

$$\delta_A = 2\sqrt{\alpha}. \tag{4}$$

Here α is the damping decrement and δ_A is the r.m.s. of amplitude change due to noise after a single turn. Without loss of generality, we can assume a Gaussian distribution for kicks caused by the noise. The damping time (in units of revolution time) is defined as $1/\alpha$ and this is a natural time scale of the system. The damping times are different for different degrees of freedom, but all of them are of the same order of magnitude. The largest one is denoted by τ and the simulation time for each step, measured in units of τ, is denoted by T (we assume it to be approximately the same from step to step).

Now let us consider the example shown in Figure 4 to understand the influence of the shape and location of the boundaries. Here we can see two strong resonances disturbing substantially the equilibrium distribution. The first one is located at small amplitudes (it is shown as the disturbance of lines of equal density). The second one (shown as resonance line and resonance vectors) plays a main role in the process of

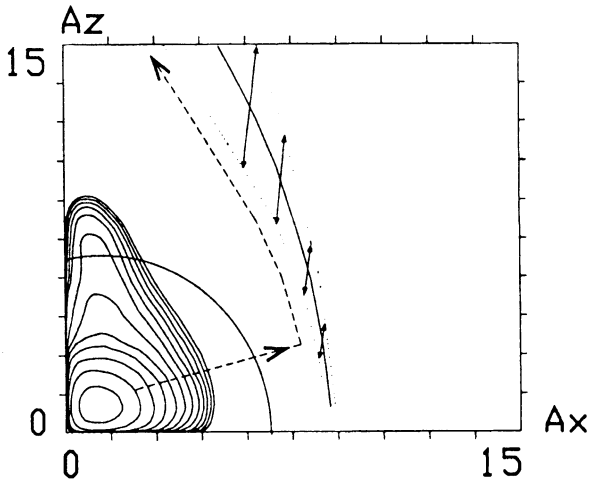

FIGURE 4: An example of location of the boundary which leads to a loss of accuracy. The most probable way of a particle going to the aperture is shown by dashed line.

losing particles at the aperture. The dashed line shows the Most Probable Way of a particle going to the Aperture (MPWA): first of all it moves to the second resonance (due to the noise only), is captured in it, and then moves to the high amplitudes along the resonance by streaming.[2] The first resonance leads to a very strong change in the density (several orders of magnitude) along the boundary which is chosen as a circular arc. As a result, since the simulation time is finite, all the outflights across this boundary are concentrated at the place where the boundary crosses the first resonance. This means that there are no outflights in the region where the MPWA is located. So, we lose this path for all following steps and the error in the lifetime determination can reach several orders of magnitude! By increasing the simulation time, we can get few outflights in the direction we need, but the insufficient statistics of such events leads to incorrect probability for this process and the accuracy of lifetime determination will not be good enough.

What is the optimal shape of the boundary? On the face of it, the main condition is the equal linear density of the outflights along the boundary. Indeed, a large number of outflights under this condition prevents from the above situations. Nevertheless, this choice has some essential disadvantages. First of all, it does not perfectly correspond to our goal of moving to large amplitudes because there is a possibility of forming spacious areas at small amplitudes with very weak flows out from the core. As a result, the line of equal linear density of the outflights can have such a form that these areas will be located in the external region for several steps in spite

of the high density of equilibrium distribution within them. Such a situation leads to the essential decrease in the efficiency of the method. After that it is not so easy to build such a boundary since each particle's jump within the amplitude plane can be an outflight or not, depending on the location of the boundary which we do not know *a priori*.

The other approach turns out to be more successful. Instead of the line of equal linear density of the outflights, we use the line of equal distribution density (or level line) as a boundary. To get this distribution, we should divide the plane of amplitudes into small rectangular cells (it is convenient to use cells of size δ_A). During the simulation we will account (individually for each step) how many times a test particle falls within each cell. After completion of the step, we will have an array of such numbers (N_p), which describe the equilibrium distribution outside the hidden region. The dispersion of N_p along the C-boundary must be sufficiently small since the boundary was chosen as a level line. The mean value of N_p on the C-boundary is marked by a tilde. It can be affirmed that the main condition, we must satisfy to get good statistics and solve our problems, is

$$\widetilde{N}_p \geq \overline{N} \gg 1, \tag{5}$$

where $\overline{N}$ is the constant (to be defined below), while $\widetilde{N}_p$ can vary from step to step. Indeed, on the one hand, this condition guarantees the representation of all the details of the "phase space portrait" since these details cannot be greater than δ_A. In other words, this is a criterion of statistical reliability of the obtained level lines. On the other hand, for the next step we will get a high value of N_p along all the border of the hidden region (or R-boundary). This means that we reproduce correctly the probability of any particle's "journey" which begins inside the hidden region and crosses the area with the high "density" N_p. It is worth of noting two important details: $\widetilde{N}_p$ keeps approximately the same value from step to step and the simulation time T, necessary to satisfy condition $\widetilde{N}_p \geq \overline{N}$, does not depend on the damping time τ for the two-dimensional case.

Let us consider now Figure 5, in which the situation similar to Figure 4 is shown. Here the dotted lines show the flows of phase convection, which result in almost all the outflights located at a certain place in spite of the boundary (bold) is a level line. The MPWA is connected with capture in the second resonance and streaming along it, as well as in Figure 4. As is seen, the short way from the core to this resonance is improbable. Nevertheless, even if there is a "channel" crossing the boundary at the place where no outflights have been obtained, during the next step we reproduce the probability of getting to this channel correctly since the value of N_p around the hidden region is sufficiently high (at least several times greater than $\overline{N}$).

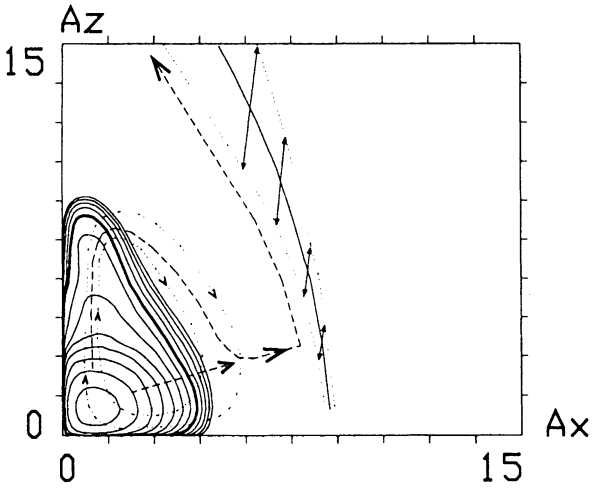

FIGURE 5: The boundary (full) coincides with a level line. Nevertheless, almost all the outflights are concentrated at a certain place due to the convective flows (dotted). The short dashed line shows a possible (but improbable) channel of going away. Since this channel crosses the area with high value of N_p, we correctly reproduce the probability of getting to it.

It is convenient to measure the distance between the boundaries through the logarithm of density change. Indeed, any boundary can be defined by a single value, i.e., the number of the corresponding level. Let us introduce for each step an individual reference system of levels Q_i in which they are read from the border of hidden region. So, the level $Q_i = 0$ corresponds to the R-boundary for the i-th step, the level $Q_i = k$ corresponds to the density $\exp(k)$ times less than the density at the level $Q_i = 0$, and so on. In these terms, the distance between the neighboring boundaries j and $(j + 1)$ (i.e. R- and C-boundaries for the $(j + 1)$-th step) is just a value of Q_j at the boundary $(j + 1)$. In future these distances are assumed to be equal to the same constant q for all the steps. Now we have to understand what the values q and the simulation time for each step T must be. We should keep in mind the following:

- The condition $\widetilde{N}_p \geq \overline{N}$ must be satisfied, otherwise we will get a wrong distribution. This means that with the distance between boundaries q, the value of T must also increase exponentially.
- The value of N_p obtained in time T at the level q depends on the shape of equilibrium distribution, so we cannot know it *a priori*.

- We need a lot of the outflights. If the boundaries are chosen correctly, as it was described above, the accuracy of the lifetime is determined by the statistics of the outflights. We can estimate it as $\sqrt{C_g/C}$, where C is the number of outflights and C_g is the mean number of outflights within a group (as a rule the outflights occur by local groups, which we can consider as the particular events).
- Deviations of the boundary from the true level line, if they are not so large (the true density data disagreement by a factor of 2–3 along the boundary), have practically no influence on the accuracy since we keep high value of N_p along the boundary. To calculate the lifetime accuracy, we need to take into account only the statistics of the outflights.

To estimate the optimal distance between the boundaries, the same time $T = T_0$ is assumed to be necessary for each step to achieve the number of particles per cell $N_p = \overline{N}$ at the level $Q_i = 0$ (just above the border of the hidden region). In this case, the full simulation time to move to m levels is as follows:

$$t = \frac{m}{q} T_0 \exp(q). \tag{6}$$

Here m/q is the number of steps and $T_0 \exp(q) = T$ is the simulation time for each step to achieve $N_p = \overline{N}$ at the level $Q_i = q$. We have to set $q = 1$ to minimize the time t. This means that the density falls down by a factor of e between the boundaries. Pay attention to that in this case we optimize q without taking into consideration the accuracy of the lifetime, which is defined through the statistics of the outflights. Fixing the final accuracy of the lifetime, we get the other expression for q. Indeed, on each step we have a statistical error of the outflights $\sqrt{C_g/C} \ll 1$. Since we follow the rules of building boundaries, the accuracy of the lifetime can be estimated as $(m/q)\sqrt{C_g/C}$. Here we assume (for estimation only) the equal number of the outflights across each boundary if $\widetilde{N}_p$ keeps the same value from step to step. So, the simulation time T is proportional to the square of the number of steps and for the full time we get

$$t \sim \frac{\exp(q)}{q^3}. \tag{7}$$

In this case, the minimum is achieved at $q = 3$.

To estimate the value of T_0, let us consider an unperturbed beam. The equilibrium distribution density (including the phase volume) is represented by the equation

$$\rho(A_x, A_z) = A_x A_z \exp\left(-\frac{A_x^2 + A_z^2}{2}\right). \tag{8}$$

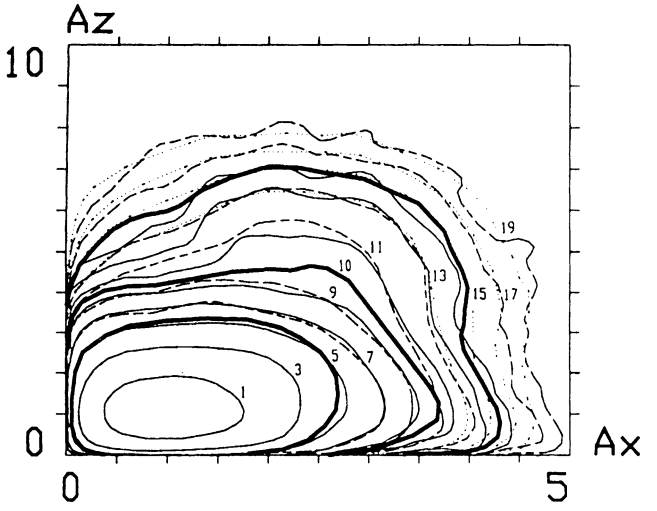

FIGURE 6: The level lines obtained during three sequential steps (1 – solid, 2 – dashed, 3 – dotted) are shown. Increasing of the line number by two corresponds to the equilibrium density decreasing by a factor of e. Boundaries (5, 10, 15) are shown in bold.

The maximum density, which is equal to $1/e$, is achieved at the point $(A_x, A_z) = (1, 1)$. The number of particle's falling within a cell of size δ_A at the point (1,1) is $N_p = 4T_0/e$ (we assume here the same damping decrements for both directions). Defining the value of $\overline{N} = 100$, we get $T_0 \approx 70$.

Practically, the distance between the boundaries is chosen as $q = 2.5$. Under this condition, the simulation time must be $T \sim 1000$ damping times for each step. The determination of the lifetime of several hours takes usually 6–8 steps and even if the accuracy of the outflight statistics is 5% for each step, we obtain a final lifetime accuracy of 50% or better, which seems to be good enough.

An important advantage of the method is the possibility of controlling the accuracy during the simulation: we can vary the simulation time T from step to step in such a way that the necessary value of $\widetilde{N}_p$ and number of outflights are achieved. Moreover, we can get the knowledge of the accuracy of the method directly from the results (see Figure 6). Here we can see the level lines obtained from several sequential steps and three boundaries. The first step actually consists of two parts which can be called "zero" step and the first step itself. The zero step is necessary to obtain the initial distribution at small amplitudes and to draw the first boundary. During the first step we get the outflights across this boundary and improve the statistics of the distribution, so the first boundary will not exactly correspond to the

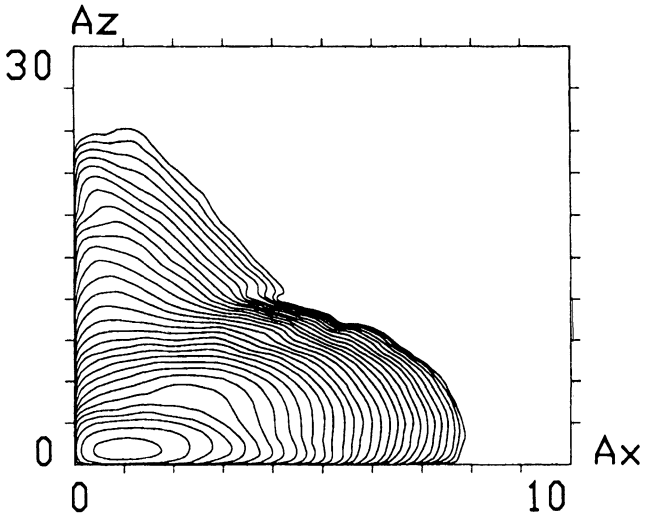

FIGURE 7: The final level lines (see also Figure 6) for the working point: $\{Q_x\} = 0.545$, $\{Q_z\} = 0.575$, $Q_s = 0.02$, $\xi_x = 0.005$, $\xi_z = 0.06$, $\sigma_x/\sigma_z = 1000$, $\sigma_l/\beta_z = 0.5$. The equilibrium density changes by a factor of e from line to line. The results of eleven sequential steps are used. The benefit in the CPU time was $\sim 10^8$.

final level line $Q_1 = q$. The second boundary is built after completion of the first step and goes along the level line $Q_1 = 2q$. After completion of the second step, we get the new lines (dashed, see Figure 6) which have the statistical reliability much better than the same lines on the first step. Nevertheless, close to the first boundary they are in good agreement, which means that both sets are reliable here. When going away from the first boundary to large amplitudes, the differences between the two sets of lines grow. This is a consequence of the insufficient statistics at large amplitudes after completion of the first step. As we can see, the second boundary has been determined well enough since it is close to the level line $Q_2 = q$ obtained after completion of the second step, as well as the third boundary ($Q_2 = 2q$) which was built after the second step is close to the level line $Q_3 = q$ obtained after completion of the third step (dotted lines in Figure 6).

The final level lines for this working point are shown in Figure 7. In the core region the results of the first step are used. Between the first and the second boundaries the results of the second step are used. Between the second and the third boundaries we use the results of the third step, and so on. As is seen, the advance to large amplitudes takes place in the regions, where the equilibrium density of distribution was defined well enough during the previous step. This condition ensures good "sewing" of all the steps.

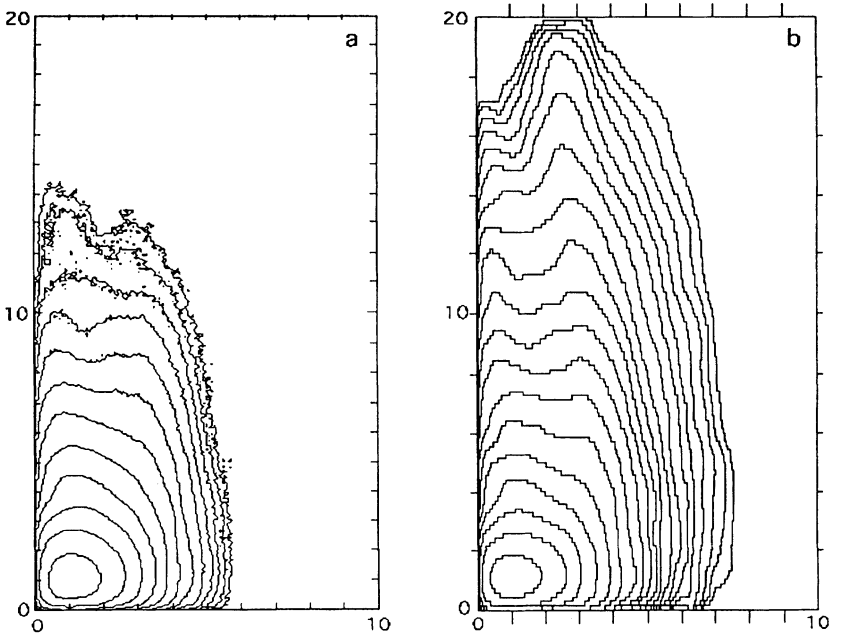

FIGURE 8: The two-dimensional amplitude distribution for PEP-II B-factory working point.[11] (a): TRS, brute force technique; (b): LIFETRAC, proposed technique.

Obviously, the decisive test of the technique validity is a comparison against the brute force technique. Recently such a comparison has been carried out for three beam-beam codes[11]:

- TRS[10] is a multiparticle strong-strong code, which does brute force tracking.
- LIFETRAC[7] is a weak-strong code, which uses the described above tracking technique.
- LFM[9] is a weak-strong code, which uses tracking algorithm similar (but not identical) to LIFETRAC.

The two-dimensional amplitude distributions for PEP-II B-factory working point[11] are shown in Figure 8. The results from TRS correspond to $5 \cdot 10^9$ particle-turns and took 818 CPU minutes to run on Cray-2S/8128. The results from LIFETRAC correspond to an effective number of $4 \cdot 10^{11}$ particle-turns and took 100 CPU minutes on VAX-6610 computer. The results from LFM are similar to the ones from LIFETRAC (see Tech. Note[11] for more information). The agreement among the codes is quite good. We hope that the next comparisons will be performed and will stimulate the progress in the tracking techniques.

In conclusion we discuss what the maximum amplitudes we can "climb" using this method are. Up to now we assumed the uniform density within a cell of size δ_A. However, at high amplitudes the level lines join together so close that this condition is violated. This means that we need the other criteria for the necessary value of $\widetilde{N}_p$ and the final lifetime accuracy (maybe, the simulation time shall increase with the increase in the step number). But we do not go deep into this problem, because it is not important for practical purposes. Indeed, the typical value of the damping time for e^+e^- colliders is about $10^3 - 10^4$ turns, which corresponds to the amplitude of noise $\delta_A \leq 1/30$. So, even for the unperturbed distribution the distance between the neighboring levels remains greater than δ_A up to the amplitudes $A \sim 30$, which is usually greater than the real aperture limit.

3 SCATTERING ON THE RESIDUAL GAS

Scattering has an essential influence on the equilibrium distribution, especially at large amplitudes. Just scattering defines the lifetime without an opposite beam. If both the beam-beam interaction and scattering occur, the interference between these two perturbations can have an effect. For example, if a strong resonance is located on large amplitudes, the scattering can essentially increase the probability of particle achieving it, providing particle's falling within the resonance directly from the core. So, including the scattering into simulation seems to be useful and important, but this is not possible directly in the new method. To clarify the problem, let us consider the elastic scattering without beam-beam effects. The scattering angle is inversely proportional to the impact parameter, so the probability of getting the amplitude jump ΔA is inversely proportional to the magnitude of this jump. Without scattering, the equilibrium density decreases as $\exp(-A^2/2)$. Hence, it can be seen that, beginning from a certain value of the amplitude, the distribution is defined only by a single scattering from the core region. In our method this core turns out to be inside the first boundary, and scattering to large amplitudes can be considered as a particular case of the outflight. Although this process plays a main role for the distribution at high amplitudes, its probability is very small and we may get no such specific outflights in time T during the first step. This means that we lose the way of going to large amplitudes for all the next steps, as in the case shown in Figure 4.

Nevertheless, slight modification of the technique allows us to successfully include the scattering. The idea is to simulate the scattering from the hidden region independently of the outflights statistics. We can do it correctly because we know the distribution inside this region.

Before discussing this technique, we would like to consider in detail the ordinary scattering outside the hidden region. We take into account only the elastic scattering

because it has a greater influence on the equilibrium distribution (inelastic scattering is very important for lifetime definition, but actually it does not disturb the equilibrium distribution since almost all the particles are lost immediately after such scattering). The cross-section for $\gamma \gg 1$ and $\theta \ll 1$ is classical:

$$d\sigma = 4Z^2 \frac{r_e^2}{\gamma^2} \cdot \frac{\mathrm{d}o}{\theta^4}, \tag{9}$$

whence

$$\theta = 2Z \frac{r_e}{\gamma \rho}. \tag{10}$$

Here Z is the nucleus charge, θ is the scattering angle, ρ is the impact parameter, $\mathrm{d}o$ is the solid angle. Each act of the scattering causes the change of momenta of the electron which can be written in normalized variables as follows:

$$\Delta p_x = \frac{\beta_x}{\sigma_x} \cdot \theta \cos\varphi, \tag{11}$$

$$\Delta p_z = \frac{\beta_z}{\sigma_z} \cdot \theta \sin\varphi. \tag{12}$$

Here φ is the angle between the horizontal and scattering planes. It follows that the elastic scattering, as well as the beam-beam effects, mainly disturbs the vertical distribution (for flat beams) due to the relation $\beta_z/\sigma_z \gg \beta_x/\sigma_x$. The scattering at small angles, when repeated many times, results in small normal (Gaussian) noise which can be combined with quantum fluctuations of synchrotron radiation into certain common noise defining the beam sizes. A special approach must be applied to scattering at amplitudes comparable with δ_A or greater, on which strong perturbations of noise distribution arise (we can consider scattering as a specific noise with very long tails).

Thus, we simulate only the scattering at the angles greater than a certain angle. As a border value, we take such an angle θ_0, that $\Delta p_z = \delta_A/10$, and estimate the probability of scattering at the angle $\theta \geq \theta_0$ during a single turn. For example, we use VEPP-4 parameters: $\gamma = 10^4$, $\tau \sim 3000$ turns, perimeter P=366 m, $\langle \beta_x/\sigma_x \rangle \sim 10^4$, $\langle \beta_z/\sigma_z \rangle \sim 10^5$, $Z = 7.5$, and a pressure of the residual gas of 10^{-8} torr. As a result, we get the probability $W_s \sim 10^{-2}$. This means the following:

- Such scatterings have no contribution to the r.m.s. of noise δ_A.
- We guarantee correct reproduction of noise distribution because we take into account the tails of noise beginning from small amplitudes $\Delta p \ll \delta_A$.

- We need not to consider the multiple scattering on such angles during a single turn due to very low probability of these events.
- The increase in the CPU time due to simulation of the above scatterings is insignificant.

The simulation with account of scattering (outside the hidden region) is now as follows: first of all, the probability of scattering at the angles $\theta \geq \theta_0$ is calculated using parameters of the ring. During the simulation, a random number generator is used at each turn to decide whether the scattering occurs. If yes, the particular angle is defined according to the formula

$$\theta = \theta_0/\sqrt{R}, \qquad 0 < R < 1, \tag{13}$$

where R is the other random number. The third random number is used then to define the scattering plane inclination angle φ and the fourth random number to define the azimuth of the collider on which the scattering takes place. To calculate the jumps in the normalized momenta, we have to multiply the obtained values of θ_x and θ_z with the corresponding relations β_x/σ_x, β_z/σ_z taken at the scattering point. In practice, the mean values of these relations are used, although a real lattice and distribution of ions along the ring (it can be quite different for different places) can be readily taken into consideration.

Now we come back to the scattering from the hidden region. When a particle falls inside, it has exactly three possibilities at each turn, one of them is realized without fail:

(1) The particle leaves the hidden region without scattering (ordinary outflight).
(2) The particle is scattered at the angle $\theta \geq \theta_0$. As a result, it can leave the hidden region (a particular case of the outflight), and can remain inside as well.
(3) The particle is not scattered and remains inside the hidden region.

The latter case is the most probable, and the essence of the above-described method is the consideration of only such events, where the particle leaves the hidden region without spending the CPU time for the third case. Now we only should take into account one more possibility of leaving the hidden region due to scattering. To reproduce this process correctly, the probability of the first case is necessary.

The algorithm of simulation is now as follows: during the first step (see Figure 1), in time T_1, we get a number of the outflights C_1 which can be divided into two parts: the outflights with scattering (C_{1s}) and the ones without it (C_{1r}). So, we have $C_1 = C_{1s} + C_{1r}$ and only the last events are saved as positions for future restarts. Besides, the fraction of time the particle has spent in the external region $V_1 < 1$ is accounted. On the second step (see Figure 3), the probability of leaving the hidden region without scattering during a single turn is

$$W_1 = \frac{C_{1r}}{(1 - V_1) \cdot T_1 \cdot \tau}. \tag{14}$$

Here the denominator is the number of turns the particle has spent within the internal region on the first step. The process of leaving is simulated in the following way: using a random number generator, we decide, according to the probabilities W_1 and W_s, what possibility (1 or 2) is realized at the moment. If it is the 1st possibility, a conventional restart is produced according to the statistics of the outflights without scattering. Otherwise, the point from which the particle is scattered is chosen randomly, according to the equilibrium distribution inside the hidden region (this has been already known), then the particular angle of scattering is defined, and so on (see the scattering outside the hidden region). After all, we get new amplitudes of the particle and check whether it leaves the hidden region or not. In the last case, the whole algorithm is repeated beginning with the choice from the 1st and 2nd possibilities. As a result, the "process of restarting" becomes longer due to a possibility of producing a few idle scatterings inside the hidden region without leaving it. But this time has no effect on the simulation time T_2 which takes into account only the motion outside the hidden region.

During the second step, the particle makes R_2 restarts and $C_2 = C_{2s} + C_{2r}$ outflights. At the same time it spends the fraction of time $V_2 < 1$ in the external region (now it is region III). The time which is necessary to achieve the same statistics, when a conventional tracking technique is used, is called *equivalent time*. For the 2nd step it is as follows:

$$T_{2eq} = T_1 \cdot \frac{R_2}{C_1}. \tag{15}$$

The benefit for the CPU time due to application of the new method can be calculated as T_{2eq}/T_2. On the third step, the probability of the ordinary outflight (1st possibility) is

$$W_2 = \frac{C_{2r}}{(1 - V_1 \cdot V_2) \cdot T_{2eq} \cdot \tau}. \tag{16}$$

The restarts are produced like on the second step, but using W_2 instead of W_1. In this case, the equivalent time and benefit are:

$$T_{3eq} = T_{2eq} \cdot \frac{R_3}{C_2}, \tag{17}$$

$$\frac{T_{3eq}}{T_3} = \frac{T_1 \cdot R_2 \cdot R_3}{T_3 \cdot C_1 \cdot C_2}. \tag{18}$$

Thus, we have the recurrent formulae to define the lifetime and the probability of the outflight without scattering W_k, which allows us to simulate the restart correctly.

Here we would like to make a remark concerning the value of W_k. At the first sight it looks strange because only near the border of the hidden region the particle is able to leave it without scattering. In spite of this, W_k was defined by the full time which the particle spends within the hidden region, and the most part of this time is spent in the core, far from the border. To clarify this situation, it is necessary to remember the main condition that should be satisfied to obtain the correct distribution outside the hidden region. Namely, we must reproduce the relations between the probabilities to find the particle in a certain place of the phase space at once after it has left the hidden region. We are not interested in how and where from the particle jumps, but what time correlation between such events is. We need only the relations between the probabilities and the value of W_k was defined quite correctly for this purpose.

Nevertheless, one more step must be done to successfully include the scattering into simulation. The reason is a decrease in the efficiency of the method when moving to large amplitudes. Indeed, the probability W_k falls down exponentially from step to step while the probability of scattering at the angle $\theta \geq \theta_0$, W_s, keeps the same value for all the steps. As a result, we obtain $W_s \gg W_k$ for $k \gg 1$. On the other hand, to leave the hidden region from the core, larger scattering angles are necessary with the distance of the border of the hidden region from the core. So, a lot of the CPU time will be spent for a huge number of idle scatterings within the hidden region without leaving it. To solve this problem, we use a division of the hidden region produced by sequential boundaries. It is quite easy to find the minimum scattering angles θ_{km} which can force the particle out from the region m to the region k (i.e. outside the hidden region) on the k-th step ($k > m$). Besides, we can calculate the relations between the probabilities to find the particle within these regions since we know the absolute probability for each m-th region:

$$\rho_m = V_0 \cdot V_1 \cdots V_{m-1} \cdot (1 - V_m), \qquad V_0 = 1. \tag{19}$$

This allows us to define the probability W_{km} of scattering from the region m at the angle $\theta \geq \theta_{km}$ during a single turn for the k-th step. The whole probability of scattering at the angles greater than the limit is a sum of W_{km}, and this value must be used instead of W_s:

$$\widetilde{W}_s = \sum W_{km}. \tag{20}$$

Now we get $\widetilde{W}_s$ and W_k approximately of the same order of magnitude for all the steps. Now the algorithm of the restart must be changed a little. If the random number generator decides to produce the scattering, first of all we have to choose,

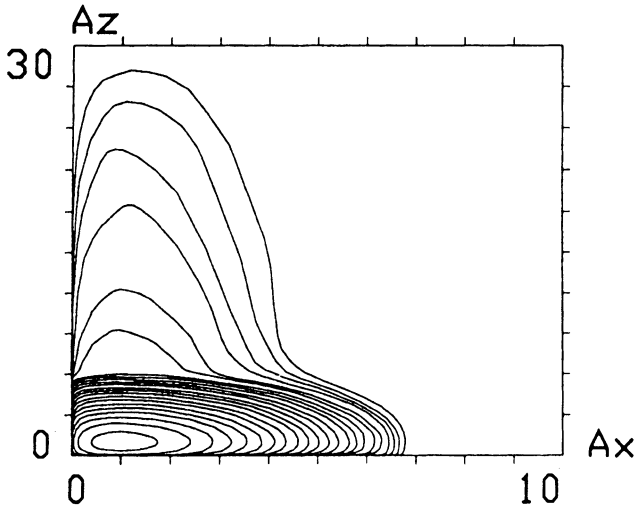

FIGURE 9: Equilibrium distribution for VEPP-4 without beam-beam effects. The density of the residual gas is 10^{-8} torr.

according to probabilities W_{km}, where from (considering the region number m) this scattering must be done. Then we choose an arbitrary (according to the distribution that we already know) point within this region and define the particular angle of scattering as

$$\theta = \theta_{km}/\sqrt{R}, \qquad 0 < R < 1, \tag{21}$$

where R is the other random number. The inclination angle of a scattering plane and the collider azimuth are defined as usual. Finally, we obtain the new particle amplitudes. After all, we have to check whether the particle leaves the hidden region or not (the scattering on the angle $\theta \geq \theta_{km}$ does not guarantee the leaving).

For example, the equilibrium distribution without beam-beam effects is shown for VEPP-4 in Figure 9. The lifetime was $4 \cdot 10^{10}$ turns for a vertical aperture of $30\sigma_z$ while the probability of a single scattering from the core to the aperture was $3.5 \cdot 10^{-11}$. The increase in the CPU time because of a more complicated restart algorithm was about 1%. The benefit in the CPU time due to this method was $\sim 10^5$. Figure 10 shows the result of interference between the beam-beam interaction and the scattering (to be compared with Figure 7, where scattering is turned off).

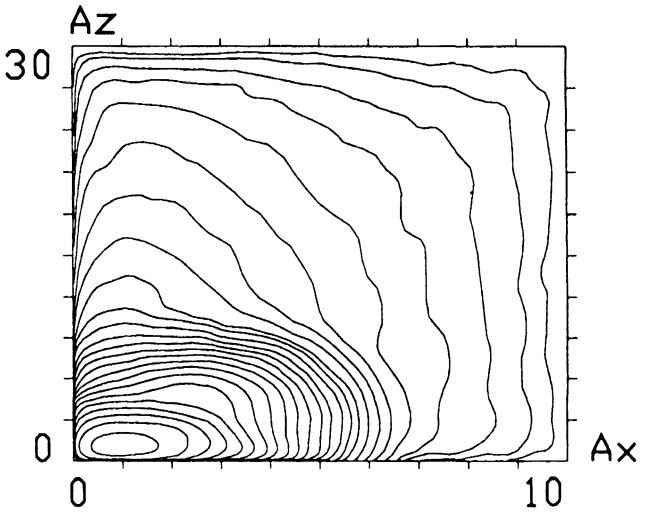

FIGURE 10: The same working point as in Figure 7 with account of the elastic scattering. The density of the residual gas is 10^{-8} torr. The roughness of level lines at large amplitudes caused from statistical errors, but anyway these are smaller than the distance between neighboring lines.

4 "STRONG-STRONG" CASE

Although a "weak-strong" model seems to be good enough for simulation of the beam halo and lifetime, sometimes it is important to take into consideration the perturbations of the "strong" beam; thus we come to a "strong-strong" model. We will not consider here the coherent beam-beam effects, collective instabilities, etc. The only thing we would like to take into account is the distortion of equilibrium distribution of the "strong" beam, that gives the change in the value of a kick which a test particle experiences due to the beam-beam interaction.

To obtain the equilibrium distribution for both beams, we have to simultaneously simulate a large number of particles for every beam. As a rule, the number of such particles is $P \sim 10^3$. There are different ways of how to calculate the value of the kick which the particle suffers from the opposite beam, but anyway we have to consider the beam as consisting of P "macroparticles", so that the charge of each macroparticle is equal to $1/P$ part of the full beam charge. In this case, the value of the kick is a function of coordinates of all these macroparticles. We will not consider in detail the question of the form of this function and of how to "smooth" it to avoid singularities at the locations of the macroparticles. Our goal is the development of only the technique.

During the first step, as usually (see Figure 1), we do not interrupt the particle motion and save all the outflights for all P particles. At the same time, P points (amplitudes and phases) within the external region are chosen randomly and saved. These are used as the initial positions for the next step. Besides, the fraction of time the particles have spent in the external region $V_1 < 1$ is accounted. In the general case, when there is the asymmetry between two beams (for example, they may have different energies, currents, lengths and so on), the boundaries, statistics of the outflights, etc. are defined separately for each beam.

Since in the first step the particles are mainly concentrated in the core, we can neglect the influence of the tails on the kick from the opposite beam. Therefore, after completion of the first step the distribution in region I (see Figures 1,3) is known well enough, although the distribution outside the core (which has an effect on the distribution inside the core) is still unknown. On the second step, the part of the beam charge in the hidden region is $1 - V_1$, which can be different for each beam. Now we can define the contribution of the hidden region to the value of the kick which the particle suffers from the opposite beam. Probably, the best solution in this case is to build a two-dimensional grid of forces, which allows us to calculate quickly this part of the kick as a function of the particle's coordinates. Besides, we have to take into account the particles outside the hidden region and their contribution to the kick value. With this purpose, we continue watching for P particles during the second step, the charge of each particle being V_1/P part of the full beam charge. As soon as any particle falls inside the hidden region, the restart is produced, so the number of particles outside the hidden region is always P. As well as on the first step, the outflights across the C-boundary and the points within the external region (initial positions for the next step) are saved, also a fraction of time the particles have spent within the external region V_2 is accounted.

The simulation algorithm for the third and all the next steps is similar to the second step, except two points:

- For the i-th step we have to rebuild our grid to calculate the hidden region's part of the kick value. Only a contribution of the (i-1)-th region (it is an internal region on the (i-1)-th step, which intends to be joined with the hidden region on the i-th step) can be added to the old grid.
- The charge of each particle outside the hidden region is equal now to a $V_1 \cdot V_2 \cdots V_{i-1}/P$ part of the full beam charge, and their contribution to the kick must be proportional to this value.

The above-described technique has one more application: this is the scattering on the neighboring particles of the same bunch (Touschek effect). Since the probability of such scattering depends on the distribution density at the scattering point, which is unknown *a priori*, we should simulate a large number of particles even for

the "weak-strong" model. In this case, the tracking technique[a] looks like in the "strong-strong" model with scattering (see Section 3).

5 THREE-DIMENSIONAL CASE

Up to now we considered the two-dimensional amplitude space. Evidently, all the results can be generalized for the three-dimensional case as well. This means that we have to build a distribution within the three-dimensional amplitude space and use two-dimensional boundaries (i.e., surfaces instead of lines). The only problem arises from the fact that probability to fall within a three-dimensional cell of size δ_A is much less than in the case of a two-dimensional cell. This leads to that the simulation time for each step T must be much greater in the three-dimensional case to satisfy the condition $\widetilde{N}_p \geq \overline{N}$. Moreover, the simulation time T depends now on the damping time τ. Besides, there are some technical difficulties which do not allow us to use the three-dimensional version of the method at the present time. As a compromise, we use a "three-dimensional tracking with two-dimensional distribution". This approach has a limited range of application, but anyway, even outside this range, one can get the essential information.

The algorithm we use can be presented as consisting of three independent parts: the simulation itself (outside the hidden region), building the level lines in the amplitude space, and the simulation of a particle leaving the hidden region (i.e. restarts). The first part represents the physical nature of the phenomenon under investigation and must be three-dimensional because of the important role of the longitudinal motion. Concerning the second part, we can integrate the distribution with respect to the third amplitude and build the boundary similar to that in the two-dimensional case. This means that the boundary is a cylinder within the three-dimensional amplitude space. Nevertheless, each outflight includes six values: all the amplitudes and phases (or coordinates and momenta). What is the effect of such simplification of the boundary? The finite statistics of the outflights can lead to losing some ways (see Figure 4). For example, reproducing the probability of being captured in the three-dimensional resonance, whose location with respect to the third axis is far from the core, can be incorrect. Besides, the lifetime can be determined by an aperture for the third dimension, but we keep low third amplitude for all the steps.

Nevertheless, there are areas of parameters where the system becomes really two-dimensional (or "almost two-dimensional"). Particularly, the distribution is exactly symmetrical with respect to arbitrary rotations in the plane (A_x, A_z) for

[a] At the present time, the "strong-strong" model, as well as the Touschek effect, is not included in the code. It will probably be done in the future.

round beams. In this case, we build the distribution in the plane (A_r, A_s), where $A_r = \sqrt{A_x^2 + A_z^2}$. Besides, the system becomes "almost two-dimensional" in the case, where the dispersion at the IP is so high, that the synchrotron transverse beam size is much greater than the betatron one (so-called *monochromatization*[3]). In this case, the horizontal displacement of the particle is mainly defined by the longitudinal motion, and we have to build the distribution in the plane (A_s, A_z). The probability of the particle going to the aperture through large amplitude A_x vanishes in this case, although the resonances are really three-dimensional.

The significance of the longitudinal motion remains essential even without dispersion at the IP due to modulation of the betatron phases. However, the betatron motion seems to be more important and the distribution must be produced in the plane (A_x, A_z). The case, where the synchrotron and betatron beam sizes at the IP becomes comparable, is most difficult. Here all three dimensions are of the same importance and the two-dimensional distribution can result in serious errors in the lifetime determination. Nevertheless, even under such conditions the new method can be helpful. The thing is that all the errors caused by two-dimensional simplification of the distribution give only the increase in the lifetime. Therefore, we can use the method to quickly search for "bad" regions of the parameters (i.e. the regions, where the lifetime is short), while in "good" regions we should use a conventional tracking technique to correctly define the lifetime. The topology of such "bad" regions in the space of various parameters can be an important source of information for future investigations.

6 SOME RESULTS OF TECHNIQUE APPLICATION

The study of beam-beam effects with monochromatization for the Novosibirsk B-Factory Project has given impetus to the development of the proposed technique. As it was shown (see Gerasimov[3]), for flat beams the width of the resonance $l \cdot Q_x + m \cdot Q_z + n \cdot Q_s = k$ depends on the monochromatization parameter λ (it is defined as the ratio of synchrotron and betatron beam sizes at the IP) as follows:

$$\Delta A_z \sim \lambda^{-l/2}. \tag{22}$$

This formula mathematically represents the fact that significance of the horizontal betatron motion decreases (i.e. the system becomes "almost two-dimensional") in case of increasing the synchrotron beam size at the IP. However, there were some doubts concerning the dispersion at the IP, since a series of strong synchro-betatron resonances arose. The simulation results obtained in 1990–1991 by using a new technique essentially clarified the situation.

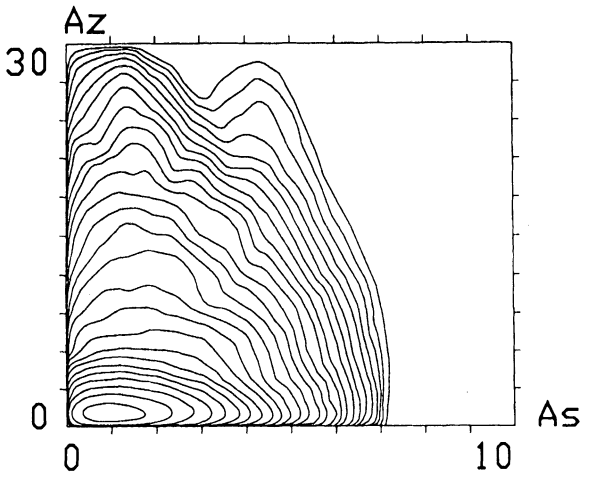

FIGURE 11: Equilibrium distribution for VEPP-4M working point: $\{Q_x\} = 0.53$, $\{Q_z\} = 0.57$, $Q_s = 0.02$, $\xi_x = 0.005$, $\xi_z = 0.06$, $\sigma_x/\sigma_z = 80$, $\lambda = 2$.

It is likely that VEPP-4M is the first collider with big dispersion at the IP. This dispersion arises constructively because of a magnet spectrometer for scattered electrons[4] and was considered earlier as a disadvantage. However, we expect now that the permissible tune shift parameter ξ will be sufficiently high. The simulation results for VEPP-4 are shown in Figures 11, 12. The only distinction between these figures is the monochromatization parameter $\lambda = \sigma_{xs}/\sigma_{x\beta}$. As we can see, the widths of resonances with $l \neq 0$ really fall down with the increase in λ.

A large number of bunches in the B-Factory leads to parasitic crossings (PC's). As a rule, there are two PC at a distance of about 2 meters from the main IP, all the next bunches are shielded by a vacuum chamber. In spite of considerable separation (20–40 beam sizes), the PC can essentially disturb the opposite beam due to high value of a beta-function (see Figure 13, where there are no PC's, and compare with Figures 14 and 15, where they are present). The direction of the separation is also very important. The technique was used to search for the minimum separation value which is allowed for both horizontal and vertical separations. The specific instability, which results in losing the particle when it achieves a certain threshold in the vertical amplitude, was discovered for the vertical separation (see Figure 14). This threshold can be several times smaller than the separation value! In case of the horizontal separation, the PC's have much less influence (see Figure 15). The lifetime versus the separation value is shown in Figure 16 for both cases (the vertical aperture is $30\sigma_z$, the horizontal aperture is $10\sigma_x$, all the particles were lost at the vertical aperture).

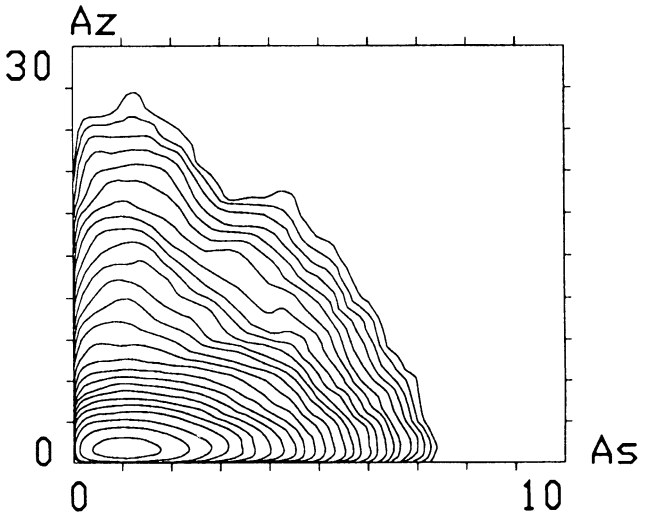

FIGURE 12: The same working point as in Figure 11, but monochromatization parameter $\lambda = 5$. The widths of resonances $l \cdot Q_x + m \cdot Q_z + n \cdot Q_s = k$ with $l \neq 0$ are decreasing here due to increasing of λ.

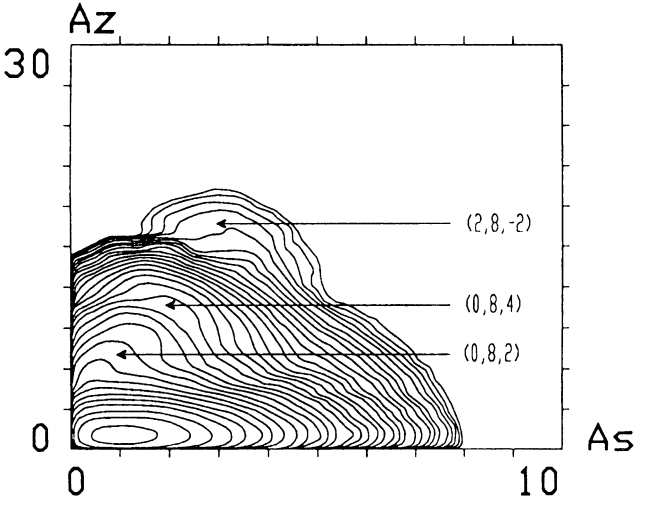

FIGURE 13: The B-Factory working point without PC's: $\{Q_x\} = 0.08$, $\{Q_z\} = 0.11$, $Q_s = 0.02$, $\xi_x = 0.01$, $\xi_z = 0.05$, $\sigma_x^*/\sigma_z^* = 300$, $\lambda = \sigma_{xs}^*/\sigma x\beta^* = 10$. Here asterisk (*) denotes the values at the main IP. Resonances $l \cdot Q_x + m \cdot Q_z + n \cdot Q_s = k$ are shown as (l, m, n).

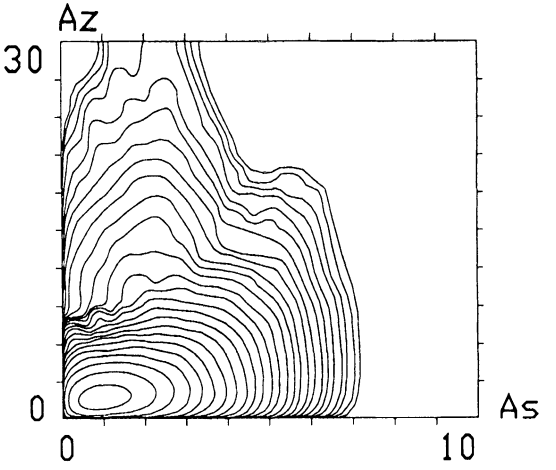

FIGURE 14: The same working point as in Figure 13 with account of two PC's (vertical separation on $50\sigma_z$). An instability arises, which results in short lifetime.

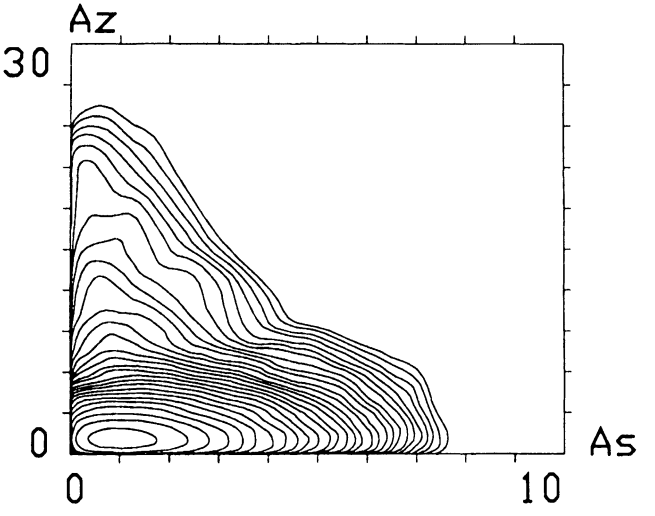

FIGURE 15: The same working point, but horizontal separation on $20\sigma_x$ instead of the vertical one. Pay attention that σ_x and σ_z becomes almost equal at the PC.

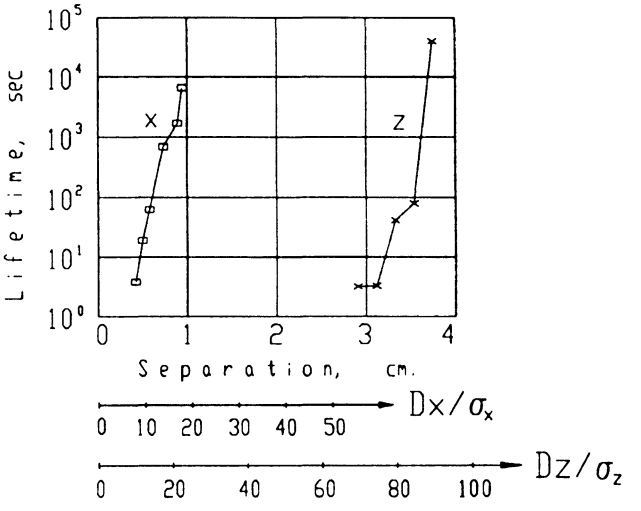

FIGURE 16: Lifetime versus the separation value (horizontal or vertical) at the PC.

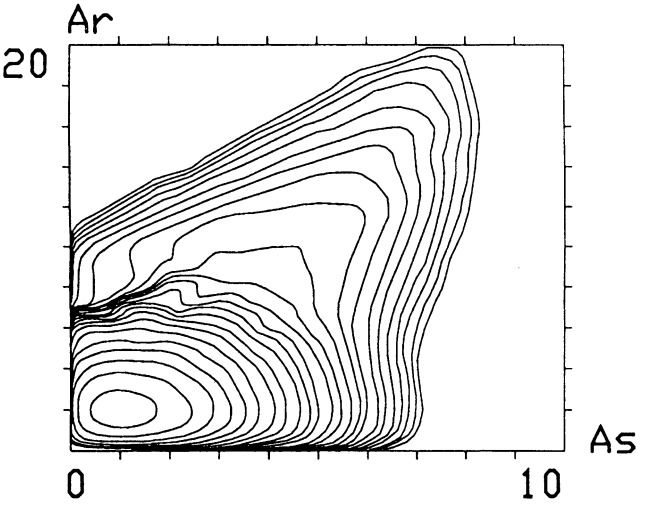

FIGURE 17: Equilibrium distribution for φ-Factory with round beams. $\{Q_{x,z}\} = 0.04$, $Q_s = 0.02$, $\xi_{x,z} = 0.2$. The strong flow to large amplitudes arises due to longitudinal beam-beam effects.

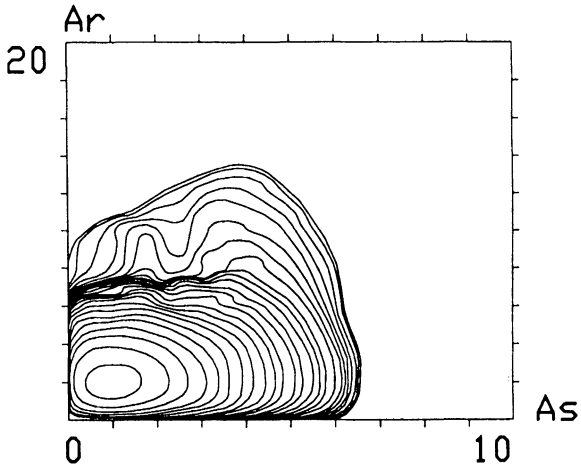

FIGURE 18: The same working point as in Figure 17, but momentum compaction factor is negative.

The technique was also used in simulations for the Novosibirsk φ-Factory Project with round beams. The high intensity and low energy of the beams led to strong longitudinal beam-beam effects.[5] Particularly, strong flows to high amplitudes (both betatron and synchrotron) could arise due to these effects (see Figure 17). In order to suppress such flows, an interesting idea of negative momentum compaction[6] was suggested by V.V. Danilov and E.A. Perevedentsev. The simulation results for the same working point as in Figure 17 and inverse sign of the momentum compaction factor are shown in Figure 18.

Moreover, some simulations results were obtained for the HERA electron beam (see Shatilov[8]). Any suggestions concerning new fields of application of the technique will be welcome.

Acknowledgements

I would like to acknowledge many useful discussions with A.A. Zholents. In addition, A.L. Gerasimov made a number of valuable remarks concerning the technique.

References

[1] J. Irwin, *Simulation of Tail Distributions in Electron-Positron Circular Colliders* (Third Adv. ICFA Beam Dynamics Workshop, Novosibirsk, 29 May – 3 June, 1989).

[2] J.L. Tennyson, *Resonance Transport in Near Integrable Systems with Many Degrees of Freedom*, Physica **D5** (1982) 123.

[3] A.L. Gerasimov *et al.*, *Beam-Beam Effects with Big Dispersion Function at the Interaction Point*, NIM **A305** (1991) 25.
[4] V.M. Aulchenko *et al.*, *Scattered Electrons Registration System of KEDR Detector for Two-Photons Process Investigation* (INP 91-49, Novosibirsk, 1991).
[5] V.V. Danilov *et al.*, *Longitudinal Beam-Beam Effects for an Ultra-High Luminosity Regime* (IEEE Part. Acc. Conf., San-Francisco, May 6–9, 1991).
[6] V.V. Danilov *et al.*, *Negative Momentum Compaction in the Longitudinal Beam-Beam Effects* (Int. Conf. on High Energy Acc., Hamburg, July 20–24, 1992).
[7] D.N. Shatilov, *Beam-Beam Simulations at Large Amplitudes and Lifetime Determination* (INP 92–79, Novosibirsk, 1992).
[8] D.N. Shatilov, *Beam-Beam Simulations for the HERA Electron Beam* (DESY HERA 93–12, Hamburg, 1993).
[9] T. Chen *et al.*, *Simulation of the Beam Halo from the Beam-Beam Interaction,* Phys. Rev. **E49** (1994) 2323.
[10] J. Tennyson, *Undocumented code "TRS"*, 1989.
[11] M. Furman *et al.*, *Comparisons of Beam-Beam Code Simulations* (CBP Tech Note-59, 1995).