DEUTSCHES ELEKTRONEN-SYNCHROTRON

DESY 93-198
December 1993

Stochastic Beam Dynamics in Storage Rings

A. Pauluhn
Fachbereich Physik, Universität Hamburg

ISSN 0418-9833

NOTKESTRASSE 85 - 22603 HAMBURG

To be sure that your preprints are promptly included in the
HIGH ENERGY PHYSICS INDEX,
send them to (if possible by air mail):

DESY **Bibliothek** **Notkestraße 85** **22603 Hamburg** **Germany**	**DESY-IfH** **Bibliothek** **Platanenallee 6** **15738 Zeuthen** **Germany**

DESY 93-198
December 1993

ISSN 0418-9833

Stochastic Beam Dynamics in Storage Rings

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Anuschka Pauluhn
aus Wilhelmshaven

Hamburg
1993

Gutachter der Dissertation:	Prof. Dr. R.D. Kohaupt Prof. Dr. P. Schmüser
Gutachter der Disputation:	Prof. Dr. R.D. Kohaupt Prof. Dr. Dr.h.c. G.-A. Voss
Datum der Disputation:	16.12.1993
Sprecher des Fachbereichs Physik und Vorsitzender des Promotionsausschusses:	Prof. Dr. E. Lohrmann

Übersicht

In dieser Arbeit werden verschiedene Modelle für die Beschreibung stochastischer Einflüsse auf die Speicherringdynamik untersucht. Stochastische Effekte werden beispielsweise verursacht durch Bodenbewegungen, Schwankungen der Netzgeräte, Fluktuationen durch die Quantenprozesse der Photoemission und Rauschen im Hochfrequenzsystem.
Im ersten Teil wird die Theorie von stochastischen Differentialgleichungen und Fokker-Planck Gleichungen benutzt, mit deren Hilfe sich Markov Prozesse beschreiben lassen. Obwohl für diese Prozesse ein wohlbekannter mathematischer Formalismus existiert, sind Lösungen nur in wenigen Fällen einfach zu erhalten. Vielfach ist man auf numerische Behandlung der Bewegungsgleichungen angewiesen. Es wird deshalb besonders auf numerische Verfahren eingegangen. An analytisch lösbaren Beispielen werden numerische Integrationsalgorithmen getestet und auf spezielle Beispiele aus der Beschleunigerphysik angewendet. Ausführlich wird die stochastisch angeregte Synchrotronbewegung für das Spezialbeispiel eines Doppel HF-Systems mit verschiedenen Methoden untersucht. Hierbei werden verschiedene analytische störungstheoretische Methoden zur Herleitung der zugehörigen Fokker-Planck Gleichung in der Wirkungsvariablen benutzt und mit numerischen Verfahren überprüft.

Der zweite Teil der Arbeit beschäftigt sich mit der Bewegung von Elektronen in einem Speicherring, die aufgrund der Synchrotronstrahlung sowohl Dämpfung (Strahlungsrückwirkungskraft) als auch stochastischer Anregung durch die Quantenfluktuationen in der Photoemission ausgesetzt sind. Es wird ein Algorithmus zur Berechnung der Wahrscheinlichkeitsdichtefunktion im Phasenraum für ein solches dissipatives, stochastisch angeregtes System vorgestellt.
Die Dichtefunktion ist von fundamentaler Bedeutung für den gesamten Prozeß, da aus ihr sämtliche Information über die mittleren Strahlgrößen und die Lebensdauer des Strahls erhalten werden kann. Der Algorithmus besteht aus der Berechnung eines Zeitpropagators für die Dichtefunktion auf dem Phasenraum, mit dessen Hilfe dann die Zeitentwicklung der Dichtefunktion sehr rechenzeitsparend modelliert wird. Das Verfahren wird auf einfache Modelle der Strahl-Strahl Wechselwirkung (eindimensional, runde Strahlen) angewandt, und die Ergebnisse werden mit Berechnungen der Dichtefunktion mittels Vielteilchen-Bahnverfolgungsrechnungen verglichen. Desweiteren werden Möglichkeiten zur Abwandlung des Verfahrens vorgestellt, die Verbesserungen in bezug auf Rechenzeit und Speicherplatzbedarf bieten. Schließlich wird der Algorithmus für zweidimensionale Strahl-Strahl Systeme erweitert.

Abstract

In this thesis several approaches to stochastic dynamics in storage rings are investigated. In the first part the theory of stochastic differential equations and Fokker-Planck equations is used to describe the processes which have been assumed to be Markov processes. The mathematical theory of Markov processes is well known. Nevertheless, analytical solutions can be found only in special cases and numerical algorithms are required. Several numerical integration schemes for stochastic differential equations will therefore be tested in analytical solvable examples and then applied to examples from accelerator physics. In particular the stochastically perturbed synchrotron motion is treated. For the special case of a double rf system several perturbation theoretical methods for deriving the Fokker-Planck equation in the action variable are used and compared with numerical results.

The second part is concerned with the dynamics of electron storage rings. Due to the synchrotron radiation the electron motion is influenced by damping and exciting forces. An algorithm for the computation of the density function in the phase space of such a dissipative stochastically excited system is introduced.
The density function contains all information of a process, e.g. it determines the beam dimensions and the lifetime of a stored electron beam. The new algorithm consists in calculating a time propagator for the density function. By means of this propagator the time evolution of the density is modelled very computing time efficient. The method is applied to simple models of the beam-beam interaction (one-dimensional, round beams) and the results of the density calculations are compared with results obtained from multiparticle tracking. Furthermore some modifications of the algorithm are introduced to improve its efficiency concerning computing time and storage requirements. Finally, extensions to two-dimensional beam-beam models are described.

Contents

1 Introduction

The motion of particles in storage rings is influenced by various kinds of perturbations. Apart from the classical theory of particle motion in electromagnetic fields, stochastic phenomena such as ground motion, power supply ripple, noise caused by the quantum emission of synchrotron radiation and noise in the radiofrequency (rf) system have to be taken into consideration. Therefore it is important to include the stochasticity caused by these effects in the calculation of the beam dynamics. Stochastic dynamics in storage rings has been investigated by various authors (see e.g.[1],[2],[3],[4],[5]). In the most general description of accelerator dynamics (still neglecting the spin of the particles) one has to solve a nonlinear system of three coupled differential equations of second order, two for the transverse motion (betatron oscillations) and one for the longitudinal motion (synchrotron oscillations). Additionally a complete description has to include the previously mentioned stochastic fluctuations. A typical stored particle beam consists of bunches of 10^{10} to 10^{11} particles which is about 10 orders of magnitude less than the numbers of particles usually treated in statistical physics. Consequently most of the methods common in this field of physics cannot be applied to accelerator physics straightforwardly and different mathematical concepts to model these stochastic effects are required.
The aim of the present work is to introduce and test some of these methods. A short introduction to the basic mathematics is given in chapter 2, where some results of probability theory are resumed and definitions and expressions used further on are mentioned. Particularly the fundamental concept of Markov processes is introduced. These processes are the appropriate model for systems whose "memory time" is short compared to their other typical time scales.
In the second chapter methods for handling stochastic systems which are assumed to have the Markov property are studied. The appropriate tools for such systems are stochastic differential equations (SDEs) in the system variables and also equations for the corresponding probability density functions (Fokker-Planck equations, FPEs). Methods for solution are shown and applied to various examples. Although the theory of Markov processes is well known [6],[7],[8],[9], explicit analytic solutions are rare, especially for nonlinear and higher-dimensional systems. Numerical solutions become necessary and in this work special emphasis has been put on numerical algorithms. Several methods to model the influence of external fluctuations on the beam dynamics and in particular integration algorithms for stochastic differential equations are investigated. The validity of numerical integration schemes for stochastic differential equations is tested in solvable examples and then these schemes are applied to problems relevant for accelerator dynamics.
As a special example for stochastic differential equations in accelerator physics the longitudinal motion of particles in a storage ring has been chosen. By adding a second rf system whose frequency is a higher harmonic of the main frequency one is able to control the spread in synchrotron frequency. For suitable choice of the parameters of the second system the synchrotron oscillation frequency becomes strongly dependent on the oscillation amplitude. Different analytical and numerical approaches to a double rf system under the influence of phase noise will be studied. Double rf systems have been previously investigated by Hofmann and Myers [10] and Wei [11].
Perturbation techniques for the derivation of the rf Fokker-Planck equation in the action variable using perturbation theory in the stochastic differential equations like Krinsky and Wang [12] or using averaging methods in the Fokker-Planck equation like Schonfeld [13] and Ripken and Karantzoulis [14] will be applied to this special problem and compared to numerical solutions.

The second part of this thesis is concerned with stochastic dynamics in electron storage rings. Passing the bendings of a circular accelerator an ultrarelativistic charged particle emits synchrotron radiation. The average effect of the radiation process results in a damping of all degrees of freedom of the particle motion (radiation damping), whereas the quantum fluctuations in the radiation process act as a random excitation (see e.g. [5],[15]). The particle dynamics is therefore determined by the interaction of these two effects. The equilibrium ("natural") beam sizes are defined by a balance of the dissipative and exciting forces. When investigating these kinds of systems both damping and stochastic forces have to be included.
In this work the density function in phase space is studied for a simple model electron storage ring with additional nonlinear forces. The density function is of fundamental importance because it determines all the statistical properties of a stochastic process. The beam dimensions are defined by the density function and its time evolution is a measure of the "lifetime" of a beam, see e.g. [16]. Based on the ideas of Gerasimov [17] a new algorithm for computing the density function has been developed and tested. This algorithm consists in calculating a time propagator for the density function. The nonlinear beam-beam interaction in a storage ring is treated in detail. In colliding beam operation the counterrotating bunches of particles repeatedly cross each other at the interaction points and thus experience a force due to the charge distribution of the opposite bunches at these points. The highly nonlinear structure of the beam-beam interaction makes analytical investigations quite complicated. The new algorithm has been used for qualitative studies of the time evolution of the density function as well as for lifetime calculations. Density calculations for one-dimensional systems are compared with results of multiparticle tracking calculations. The last chapters are concerned with models of the two-dimensional beam-beam interaction and introduce possible extensions and variations of the algorithm to make it more efficient with respect to computing time and storage requirements.

2 Mathematical Principles

This chapter introduces concepts from probability theory which are important for the following. The general definitions and notations are given in the appendix.

2.1 Stochastic Processes

The time development of a physical system is described in terms of time-dependent functions. When in addition random fluctuations act on the system its process variables are no longer deterministic but become stochastic or random variables. Such time-dependent random variables are called stochastic processes. The definition of a random variable is given via a random experiment. All stochastic processes can thus be understood as sequences of random experiments. The probability P of an event is defined as its relative frequency in a series of experiments. For a random experiment there exists a set of possible outcomes (sample space) and one can assign to an event out of this sample space a function or random variable X which maps the outcomes to elements of another space, called the state space. Generally this state space consists of the set of real numbers R or of vectors of real numbers R^d. The function X thus defines a kind of measuring device to have a more convenient characterization for the outcome of a random experiment. This function X is called random variable. The set of possible outcomes, i.e. the sample space, that will be considered in the following is the set of real numbers or real vectors. Therefore one makes the canonical choice for the random variables, that means the state space is identified with the sample space and the function X is the identity function. A graphical representation of the relations between random variables and stochastic processes is given in figure 1.
The most important quantity characterizing a random variable is the density function $\rho(x)$. It is defined as the probability of finding a value of the random variable in the interval $[x, x+dx]$:

$$\rho(x) = P(x \leq X \leq x + dx).$$

The event "(the random variable y_1 has a value in $[y_1, y_1 + dy_1]$) and (the random variable y_2 has a value in $[y_2, y_2 + dy_2]$)" can be specified by a vector $\vec{y} = (y_1, y_2)$. This introduces the concept of joint probabilities. The joint density is written as $\rho(\vec{y})$ or $\rho(y_1, y_2)$. For stochastic processes one measures values $\vec{x}_1$, $\vec{x}_2$,...etc. of $\vec{x}(t)$ at times t_1, t_2,...etc. and assumes that a set of joint probability densities of the form

$$\rho(\vec{x}_1, t_1; \vec{x}_2, t_2; ...)$$

exist which describe the process completely.

2.2 White Noise

A process which is fundamental for the general treatment of stochastic processes is the so-called white noise $\xi(t)$. This is a stationary random process with mean zero , $< \xi(t) >= 0$ and δ-correlation $< \xi(t)\xi(t') >= \delta(t - t')$.
Therefore the white noise is totally uncorrelated for all $t \neq t'$ and has infinite variance for $t = t'$. These are rather artificial properties.
The joint probability density factorizes

$$\rho(\xi_1, t_1; ...; \xi_n, t_n) = \prod_{i=1}^{n} \rho(\xi_i)$$

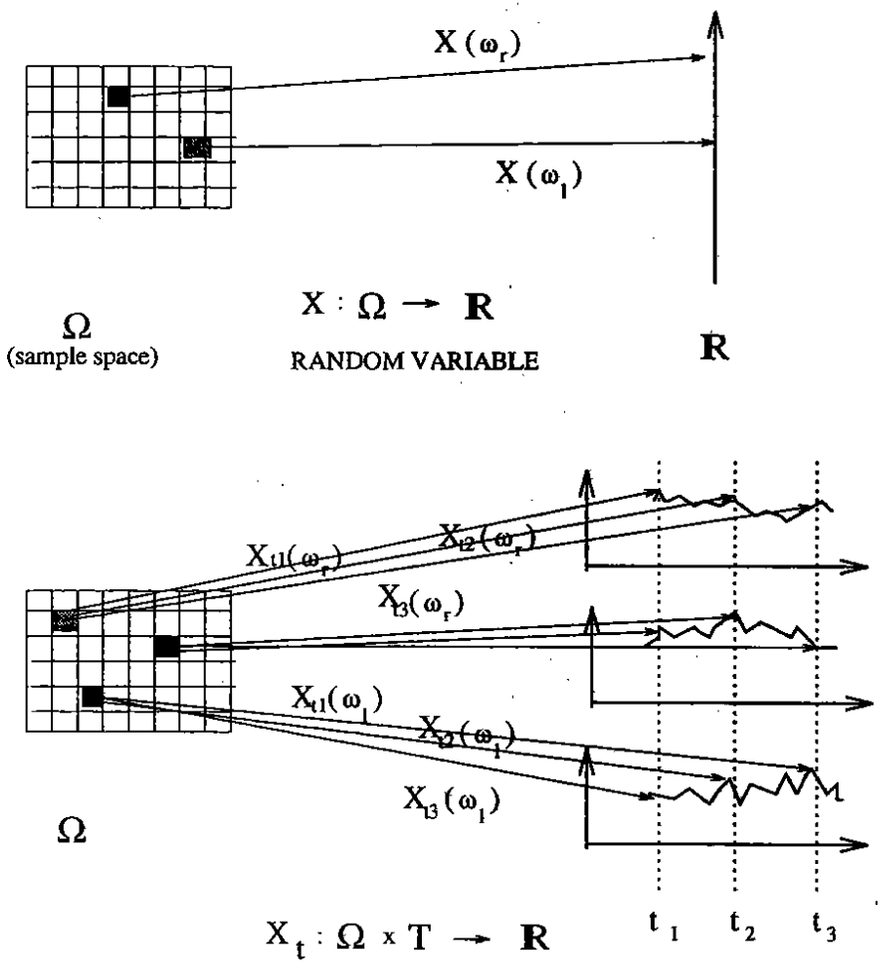

Figure 1: Random variables and stochastic processes.

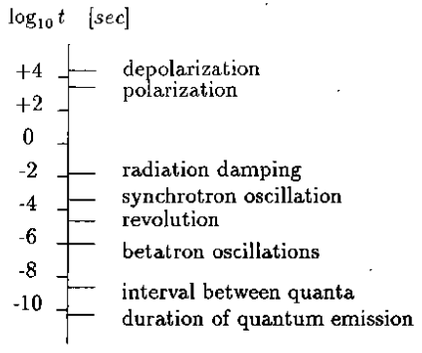

Figure 2: Characteristic time scales in a $25GeV$ electron storage ring, from [18].

because all single events are independent of each other. Moreover all single densities $\rho(\xi_i)$ are independent of time, i.e. the same probability law governs the process at all times. From the δ-correlation of this process follows that the spectrum is a constant in frequency space:

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}\delta(\tau)d\tau = \frac{1}{2\pi}.$$

All frequencies occur with the same weight, therefore the name "white noise" by analogy with "white light". If the probability distribution function of the white noise is the normal distribution, it is called Gaussian white noise. This is the most interesting case for physical applications and a consequence of the central limit theorem which states that the distribution of a large number of independent identically distributed random variables converges to a Gaussian distribution. In real physical systems the time scales of the (environmental) fluctuations and of the considered macroscopic system are in most cases totally different. This implies that the noise process consists of a large number of uncorrelated single random effects acting nearly at the same time on the system. (The memory time of the fluctuations is extremely short compared to the typical macroscopic time scales.) Thus one can assume that the conditions for applying the central limit theorem are fulfilled. An example for time scales in a noisy system is given in figure 2 where the time scales in an accelerator are represented [18]. These range from the time intervals for photon emission in the synchrotron radiation process to the times needed for the buildup of electron spin polarization and depolarizing effects.
White noise possesses such irregular features as having no continuous realizations and a total power $\int_{-\infty}^{\infty} S(\omega)d\omega$ that is infinite. Although processes with such extreme properties do not occur in nature, they are most useful as an approximative model for rapidly fluctuating systems. Another advantage of the white noise concept is the fact that the solutions of differential equations with white noise excitation are Markov processes. These can be treated with powerful mathematical tools.

2.3 Markov Processes

The next simple random processes are the Markov processes. They are characterized by the property that the conditional probability (see appendix) depends only on the immediately

preceding time step and is independent of the earlier history of the process (Markov property). For an ordered time sequence $t_n \geq t_{n-1} \geq \ldots \geq t_1$ call $\rho_n(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_1, t_1)dx_n$ the probability that a random variable (for example the position of a particle) is in the interval $(x_n, x_n + dx_n)$ at time t_n under the condition that it has been at x_{n-1} at the time t_{n-1} and at x_{n-2} at the time t_{n-2} ... etc.
Then the Markov property is

$$\rho_n(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = \rho_2(x_n, t_n|x_{n-1}, t_{n-1}). \tag{2.1}$$

The information relevant for the future of the process is contained in the present state and is independent of the past history, the system "has no memory".
Mathematically these processes arise as solutions of stochastic differential equations with Gaussian white noise:

$$\dot{\vec{X}}(t) = \vec{f}(\vec{X}(t)) + \underline{g}(\vec{X}(t))\vec{\xi}(t) \tag{2.2}$$

with $\vec{\xi}(t)$ describing a Gaussian white noise process and $\vec{f}$ and $\underline{g}$ being continuous functions where $\vec{f}$ is a vector valued function and $\underline{g}$ denotes a matrix function.
The Markov property implies the causality in a physical system.
Markovian processes are uniquely defined by $\rho_1(x,t)$ and $\rho_2(x_2,t_2|x_1,t_1)$ or $\rho_2(x_2,t_2;x_1,t_1)$, because all joint densities can be expressed by ρ_1 and ρ_2.
For $t_3 > t_2 > t_1$ this yields a Markov chain:

$$\begin{aligned}\rho_3(x_3,t_3;x_2,t_2;x_1,t_1) &= \rho_3(x_3,t_3|x_2,t_2;x_1,t_1)\rho_2(x_2,t_2;x_1,t_1)\\ &= \rho_2(x_3,t_3|x_2,t_2)\rho_2(x_2,t_2|x_1,t_1)\rho(x_1,t_1).\end{aligned} \tag{2.3}$$

Markov chains are a good approximation for processes for which the influence of the history becomes negligible in a time that is short compared to the other relevant time constants.
Using (2.3) one can write for the two-time density:

$$\begin{aligned}\rho_2(x_3,t_3;x_1,t_1) &= \int dx_2\rho_3(x_3,t_3;x_2,t_2;x_1,t_1)\\ &= \int dx_2\rho_2(x_3,t_3|x_2,t_2)\rho_2(x_2,t_2|x_1,t_1)\rho(x_1,t_1).\end{aligned} \tag{2.4}$$

Dividing by $\rho(x_1,t_1)$ this leads to

$$\rho_2(x_3,t_3|x_1,t_1) = \int dx_2\rho_2(x_3,t_3|x_2,t_2)\rho_2(x_2,t_2|x_1,t_1). \tag{2.5}$$

This functional relation is called Chapman-Kolmogorov equation. It states that the evolution in time from t_1 to t_2 and further from t_2 to t_3 is the same as from t_1 to t_3.
Likewise the density $\rho(x,t)$ for $t > t_0$ is calculated from an initial density $\rho(x,t_0)$:

$$\rho(x,t) = \int dx'\rho_2(x,t|x',t_0)\rho(x',t_0) \tag{2.6}$$

The two-time transition probability $\rho_2(x_2,t_2|x_1,t_1)$ which represents a time propagator for $t_2 > t_1$, is the characteristic quantity for the entire Markov process.
In a vanishing time interval there is no transition from one state to another:

$$\lim_{t_2\to t_3}\rho_2(x_3,t_3|x_2,t_2) = \delta(x_3 - x_2).$$

If at the time t_0 the value of the process variable is $x = x_0$ with certainty, i.e. $\rho(x_0, t_0) = \delta(x_0 - x)$, then one has using (2.6):

$$\rho(x,t) = \rho_2(x,t|x_0,t_0).$$

This means that the transition probability $\rho_2(x_2,t_2|x_1,t_1)$ is identical with the density function $\rho(x_2,t_2)$, if at $t_2 = t_1$ it is equal to $\delta(x_2 - x_1)$.
In the following the notation ρ will also be used for the two-time probability densities.

2.4 Markovian Diffusion Processes

Diffusion processes are characterized by the following properties:

1. For all $\epsilon > 0$ holds:
$$\lim_{\tau\to 0}\frac{1}{\tau}\int_{|\vec{x}_2-\vec{x}_1|>\epsilon}\rho(\vec{x}_2,t+\tau|\vec{x}_1,t)d\vec{x}_2 = 0.$$
This means that large changes within short time intervals are rare.

2. For all $\epsilon > 0$ exists a function $\vec{A}(\vec{x},t)$ called the drift term with
$$\lim_{\tau\to 0}\frac{1}{\tau}\int_{|\vec{x}_2-\vec{x}_1|\leq\epsilon}(\vec{x}_2-\vec{x}_1)_i\rho(\vec{x}_2,t+\tau|\vec{x}_1,t)d\vec{x}_2 = A_i(\vec{x}_1,t).$$

3. For all $\epsilon > 0$ there exists a function $\underline{D}(\vec{x},t)$ called the diffusion term with
$$\lim_{\tau\to 0}\frac{1}{\tau}\int_{|\vec{x}_2-\vec{x}_1|\leq\epsilon}(\vec{x}_2-\vec{x}_1)_i(\vec{x}_2-\vec{x}_1)_j\rho(\vec{x}_2,t+\tau|\vec{x}_1,t)d\vec{x}_2 = D_{ij}(\vec{x}_1,t).$$

To first order in the time difference τ a diffusion process is defined by the first two moments of the transition probability, all higher order moments vanish.
The evolution of probability is generated by a drift term which defines the deterministic part of the motion and a diffusion term that tends to "smear out" the distribution.

2.4.1 The Fokker-Planck Equation

The transition probability of a diffusion process fulfils the Fokker-Planck equation:

$$\frac{\partial}{\partial t}\rho(\vec{x},t|\vec{z},t') = -\frac{\partial}{\partial x_i}A_i(\vec{x},t)\rho(\vec{x},t|\vec{z},t') + \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_k}D_{ik}(\vec{x},t)\rho(\vec{x},t|\vec{z},t'). \qquad (2.7)$$

From the normalization condition $\int_{-\infty}^{\infty}\rho(\vec{x},t)d\vec{x} = 1$ for all $t \geq 0$ follows that ρ and the probability current $S_i = \frac{\partial}{\partial x_i}(A_i - \frac{1}{2}\frac{\partial}{\partial x_k}D_{ik})\rho$ disappear for $|x_i| \to \infty$ (natural boundary conditions, see [7]).
For indices appearing twice in an expression the summation convention has been adopted .
For the probability density

$$\rho(\vec{x},t) = \int d\vec{z}\rho(\vec{x},t|\vec{z},t')\rho(\vec{z},t')$$

the same equation holds.

2.5 Examples of Random Processes

Some fundamental random processes that play an important role in the modelling of stochastic processes will be introduced.

2.5.1 Wiener Process

One of the most important diffusion processes is the Wiener process $\vec{W}(t)$, named after N. Wiener who studied it extensively. It has for example been used as a mathematical description for the highly irregular Brownian motion, i.e. the motion of a small particle suspended in a fluid. All finite dimensional probability densities are Gaussian, i.e. they have the form:

$$\rho(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n\det(\underline{C})}}e^{-\frac{1}{2}(\vec{x}-\vec{m})^T\underline{C}^{-1}(\vec{x}-\vec{m})},$$

where $<\vec{x}>=\vec{m}$ and $\underline{C}$ is a positive definite $n\times n$ matrix defined by $<(\vec{x}-\vec{m})(\vec{x}-\vec{m})^T>=\underline{C}$.
The Wiener process has mean zero $<\vec{W}(t)>=0$ and its correlation function is $<W_i(t)W_j(t')>=\delta_{ij}\min(t,t')$. The variance is $<W_i(t)W_j(t)>=\delta_{ij}t$ which approaches infinity for $t\to\infty$. This shows that the sample paths of $\vec{W}(t)$ are very variable.
As $\vec{W}(t)$ is a diffusion process, its density obeys the Fokker-Planck equation

$$\frac{\partial}{\partial t}\rho(\vec{W},t|\vec{W}_0,t_0) = \frac{1}{2}\sum_i\frac{\partial^2}{\partial W_i^2}\rho(\vec{W},t|\vec{W}_0,t_0)$$

whose solution is

$$\rho(\vec{W},t|\vec{W}_0,t_0) = \frac{1}{\sqrt{(2\pi)^n t}}e^{-\frac{(\vec{W}-\vec{W}_0)^T(\vec{W}-\vec{W}_0)}{2t}}.$$

An initially sharp distribution spreads in time, see figure 3 for the one-dimensional case $W(t)$.
Like every diffusion process the Wiener process has continuous sample paths. However, these trajectories are nowhere differentiable. This can be seen by considering the difference quotient $\frac{W(t+h)-W(t)}{h}$. The probability P that this difference quotient is greater than an arbitrary integer k is given by:

$$P(|\frac{W(t+h)-W(t)}{h}| > k) = 2\int_{kh}^{\infty}dW\frac{1}{\sqrt{(2\pi)h}}e^{-\frac{W^2}{2h}}.$$

In the limit $h\to 0$ this is 1. Therefore, for arbitrary k, the differential quotient is almost certain greater than k, what means that the derivative at any point is almost certainly (i.e. with the probability 1) infinite. (So in this model the velocity of a Brownian particle is not defined, a more realistic model for this motion is given by the Ornstein-Uhlenbeck process, see below.)
An important property is the fact that the Wiener process has independent increments, i.e. for $t_1 < t_2 < ... < t_n$ the random variables
$W(t_1), W(t_2)-W(t_1), ..., W(t_n)-W(t_{n-1})$ are stochastically independent in the disjoint (non-overlapping) time intervals Δt_i. The Wiener process is the only homogeneous (time-independent coefficients in the Fokker-Planck equation) random process with independent increments and continuous realizations.
Although $W(t)$ is nowhere differentiable in an ordinary sense, one can consider it as a so-called generalized stochastic process, by analogy with the generalized functions or distributions in functional analysis and look at its derivative in this "generalized function sense".

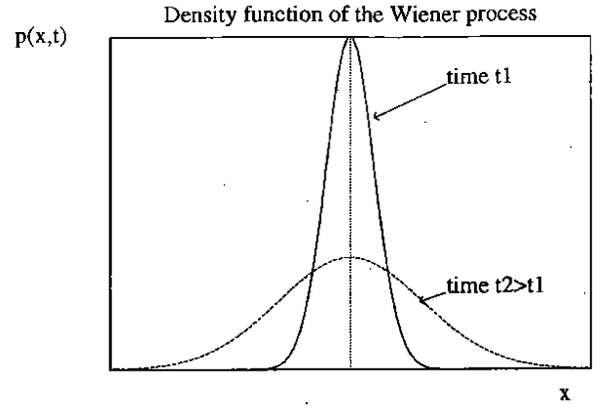

Figure 3: The density function $\rho(x,t) = \frac{1}{\sqrt{(2\pi)t}} e^{\frac{-x^2}{2t}}$ at different times.

Generalized functions $\Phi(\phi)$ are defined as functionals on the class of infinitely often differentiable functions with compact support, $C_K^\infty(\mathrm{R})$. (The latter means that the functions ϕ vanish outside some bounded interval of the t-axis.) An example for a generalized function is the Dirac δ-"function" which is formally written as an integral:

$$\Phi(\phi) = \int_{\mathrm{R}} \phi(t)\delta(t-s)dt = \phi(s).$$

Generalized functions are infinitely often differentiable and their derivatives are again generalized functions. These definitions are extended to stochastic processes.
For instance the derivative in this sense of the Wiener process is the Gaussian white noise.
This can be seen in the following way:
For the correlation function holds

$$< \dot{W}(t_1)W(t_2) >= \min(t_1,t_2) = \begin{cases} t_1 & t_1 < t_2 \\ t_2 & t_2 < t_1 \end{cases}. \tag{2.8}$$

The derivative of the Wiener process and its correlation are given by

$$< \dot{W}(t_1)\dot{W}(t_2) >= \frac{\partial^2}{\partial t_1 \partial t_2} < W(t_1)W(t_2) >= \frac{\partial^2}{\partial t_1 \partial t_2} \min(t_1,t_2). \tag{2.9}$$

Successively forming the derivatives leads to

$$\frac{\partial}{\partial t_1} H(t_1,t_2) \quad \text{with} \quad H(t_1,t_2) = \begin{cases} 0 & t_1 < t_2 \\ 1 & t_1 > t_2 \end{cases},$$

where $H(t_1,t_2)$ is the Heaviside step function in t_1 and its derivative with respect to t_1 is the δ-function $\delta(t_1-t_2)$.

This explains the fundamental relation for the stochastic integration theory

$$d\vec{W} = \int_t^{t+dt} \vec{\xi}(t')dt'.$$

2.5.2 Ornstein-Uhlenbeck Process

The Ornstein-Uhlenbeck process, named after G.E. Ornstein and L.S. Uhlenbeck, is the only Gaussian stationary process in one real variable. It has been used to model Brownian motion, taking the velocity of the particle as random process and calculating the particle position by integration.
The defining equation is:

$$\dot{v}(t) = -\gamma v(t) + \sqrt{d}\xi(t). \tag{2.10}$$

This linear equation can be solved via simple integration, leading to the expression:

$$v(t) = v(0)e^{-\gamma t} + \int_0^t \sqrt{d}\xi(s)e^{\gamma(s-t)}ds. \tag{2.11}$$

The mean value is calculated as

$$< v(t) >=< v(0) > e^{-\gamma t} \tag{2.12}$$

and the correlation function is

$$< v(t)v(t') >=< v(0)^2 > e^{-\gamma(t+t')} + \int_0^t \int_0^{t'} e^{\gamma(s-t)}e^{\gamma(s'-t')}d < \xi(s)\xi(s') > dsds'. \tag{2.13}$$

Of much more interest than the correlation function with a definite initial condition is the stationary correlation which is obtained in the limit $t \to \infty$:

$$< v(t)v(t') >_{st} = \frac{d}{2\gamma} e^{-\gamma|t-t'|}. \tag{2.14}$$

The position of the Brownian particle is obtained by integration:

$$x(t) = \int_0^t v(s)ds. \tag{2.15}$$

The Ornstein-Uhlenbeck Process as a Model for Coloured Noise

An important feature of the Ornstein-Uhlenbeck process is the applicability of its stationary solution in modelling slightly coloured noise processes, i.e. those noise processes which are not really "white" but have a finite though short correlation time τ_c.
The correlation function $G(\tau)$ of the Ornstein-Uhlenbeck process, τ denoting the time difference $t-t'$, is exponentially decreasing, its spectral density reads:

$$S(\omega) = \frac{1}{2\pi}\int_{\mathrm{R}} e^{-i\omega\tau}G(\tau)d\tau = \frac{1}{2\pi}\int_{\mathrm{R}} e^{-i\omega\tau}\frac{d}{2\gamma}e^{-\gamma|\tau|}d\tau = \frac{1}{2\pi}\frac{d}{\omega^2+\gamma^2}. \tag{2.16}$$

Equation (2.14) shows that $v(t)$ and $v(t')$ are only significantly correlated if $|t-t'|$ is of the order of γ^{-1}, that means the correlation time of this process is $\tau_c = \gamma^{-1}$.
Hence the "white noise limit" $\tau_c \to 0$ implies $\gamma \to \infty$. However, in this limit the mean square power of the oscillations of frequency ω vanishes.

So leaving the other parameters of the noise process constant, one has come to a completely noiseless limit, the random forces seem to have no more impact on the system at all. This stems from the fact that in this case the total input power $S = 2\int_0^\infty S(\omega)d\omega = \frac{d}{2\gamma}$ is spread uniformly over all frequencies.
The correct limit procedure from coloured to white noise with decreasing τ_c requires an appropriate increase in the noise strength. Equation (2.16) implies that the diffusion coefficient has to be modified from $\sqrt{d}$ to $\gamma\sqrt{d}$ in order to prevent $S(\omega)$ from vanishing in the limit $\tau_c \to 0$, leading to the equation

$$\dot{v}(t) = -\gamma v(t) + \gamma\sqrt{d}\xi(t).$$

The spectral density $S(\omega)$ is thus given by

$$S(\omega) = \frac{1}{2\pi}\int_{\mathrm{R}} e^{-i\omega\tau}G(\tau)d\tau = \frac{1}{2\pi}\int_{\mathrm{R}} e^{-i\omega\tau}\frac{d\gamma^2}{2\gamma}e^{-\gamma|\tau|}d\tau = \frac{1}{2\pi}\frac{d\gamma^2}{\omega^2+\gamma^2}. \tag{2.17}$$

In the limit $\tau_c \to 0$ this converges to $S(\omega) = \frac{d}{2\pi}$, i.e. a completely flat spectrum. The correlation function of the resulting process is $G(\tau) = d\delta(\tau)$ as is expected for a white noise process.

2.6 Stochastic Differential Equations

A first order stochastic differential equation has the form:

$$\dot{\vec{x}}(t) = \vec{A}(\vec{x},t) + \underline{B}(\vec{x},t)\vec{\xi}(t) \tag{2.18}$$

with $\vec{\xi}(t)$ describing a Gaussian white noise process, i.e.

$$<\vec{\xi}(t)>= 0 \quad , \quad <\xi_i(t)\xi_j(t')>= \delta_{ij}\delta(t-t').$$

If the diffusion matrix $\underline{B}$ is independent of the process variable $\vec{x}$, one speaks of additive noise, if $\underline{B}$ depends on $\vec{x}$, equation (2.18) is called a stochastic differential equation with multiplicative noise.
The equation (2.18) can be rewritten in the form of an integral equation:

$$\vec{x}(t) = \vec{x}_0 + \int_0^t \vec{A}(\vec{x},t')dt' + \int_0^t \underline{B}(\vec{x},t')\vec{\xi}(t')dt'. \tag{2.19}$$

At this point some particularities of the stochastic integration shall be explained.The stochastic differential equation (2.18) and the integral equation (2.19) cannot be interpreted unambiguously until one has given a "sensible" meaning to the stochastic integral.
The stochastic integral of a function is, like the deterministic integral, defined as the limit of its partial sums. But in contrast to the deterministic case these partial sums are not independent of the chosen partition of the intervals. In fact they depend on the particular intermediate points where the function is to be evaluated. This is a consequence of the highly irregular nature of the white noise process and its "δ-peak" structure. All definitions of a stochastic integral base on the relation between Gaussian white noise $\vec{\xi}(t)$ and the Wiener process:

$$d\vec{W}(t) = \vec{W}(t+dt) - \vec{W}(t) = \vec{\xi}(t)dt$$

and

$$\vec{W} = \int_0^t \vec{\xi}(t')dt'.$$

Further one defines the following integral

$$\int_0^t \underline{B}(\vec{x})\vec{\xi}(t')dt' = \int_0^t \underline{B}(\vec{x})d\vec{W}.$$

If one writes the stochastic equation (2.18) in differential form this yields

$$d\vec{x}(t) = \vec{A}(\vec{x},t)dt + \underline{B}(\vec{x},t)d\vec{W} \tag{2.20}$$

The differential $d\vec{x}(t)$ contains the differential of the Wiener process and thus the sample paths of $\vec{x}(t)$ will be at least continuous although they might be nowhere differentiable. The first integral in the integral equation (2.19) causes no problems. It can be understood as an ordinary Riemann integral for each realization:

$$\int_0^t A(\vec{x}(t'))dt' = \lim_{n\to\infty}\sum_{i=1}^n A(\vec{x}(\tau_i^{(n)}))(t_i^{(n)} - t_{i-1}^{(n)}),$$

where $\{t_0 = t_1^{(n)} < ... < t_n^{(n)} = t\}$ is a series of partitions of the interval $[t_0,t]$ and $\tau_i^{(n)}$ is a point in $[t_{i-1}^{(n)}, t_i^{(n)}]$.
Difficulties arise from the integral with respect to the Wiener process. Sample paths of this process are extremely irregular and so there appear two irregular quantities in the integral.

This can be seen from the example of the integral over the Wiener process itself $\int_{t_0}^t W(t')dW_{t'}$: The approximating sums S_n are

$$S_n = \sum_{i=1}^n W(\tau_i^{(n)})(W(t_i^{(n)}) - W(t_{i-1}^{(n)})).$$

The limit in the mean square is [8]:

$$\lim_{n\to\infty}[S_n - \sum_{i=1}^n(\tau_i^{(n)} - t_{i-1}^{(n)})] = \frac{1}{2}(W^2(t) - W^2(t_0)) - \frac{1}{2}(t-t_0)$$

or

$$\lim_{n\to\infty} S_n = \frac{1}{2}(W^2(t) - W^2(t_0)) - \frac{1}{2}(t-t_0) + \lim_{n\to\infty}\sum_{i=1}^n(\tau_i^{(n)} - t_{i-1}^{(n)})$$

and is thus dependent on the explicit choice of the $\tau_i^{(n)}$.
An unambiguous definition is achieved if the evaluation points $\tau_i^{(n)}$ are fixed once and forever. For instance one can take any point in the time interval, choosing

$$\tau_i^{(n)} = (1-\alpha)t_{i-1}^{(n)} + \alpha t_i^{(n)} \qquad 0 \le \alpha \le 1 \quad \text{fixed.}$$

Therefore one gets

$$\sum_{i=1}^n(\tau_i^{(n)} - t_{i-1}^{(n)}) = \alpha\sum_{i=1}^n(t_i^{(n)} - t_{i-1}^{(n)}) = \alpha(t-t_0).$$

For the mean square limit of the partial sums S_n this yields:

$$\lim_{n\to\infty} S_n = \frac{1}{2}(W^2(t) - W^2(t_0)) + (\alpha - \frac{1}{2})(t-t_0),$$

revealing an infinite number of possibilities for a definition of the integral.
Any choice of α leads to a mathematically correct definition of the integral.Two choices are particularly useful:

- $\alpha = \frac{1}{2}$, leading to the Stratonovich definition of the stochastic integral,
- $\alpha = 0$, leading to the Ito calculus.

If the external noise is described by a more realistic random process than the white noise, e.g. a slightly correlated Ornstein-Uhlenbeck process, the choice $\alpha = \frac{1}{2}$ is appropriate. If one chooses the middle points of the intervals $\tau_i = \frac{t_i - t_{i-1}}{2}$ as evaluation points, a slight correlation of the process is taken into consideration . Furthermore $\alpha = \frac{1}{2}$ leads to the results of ordinary calculus.
From a mathematical point of view $\alpha = 0$ is more appealing because it makes maximal use of the fact that the Wiener process has independent increments. With $\tau_i = t_{i-1}$, that means one takes the left border of the partition intervals as evaluation points, $W(\tau_i)$ and $(W(t_i) - W(t_{i-1}))$ are independent. Therefore the Ito integral has "nice" mathematical properties, it implies that there is no dependence between the random force $\vec{\xi}(t)$ and the considered process $\vec{x}(t)$ at the same instant of time t. The Ito interpretation of the integral is thus a reasonable choice in cases where noise is truly white.
However, if one arrives at a stochastic differential equation via a limiting procedure, namely starting from real fluctuations with a short but nonvanishing correlation time, the solution process $\vec{x}(t)$ depends on the stochastic force at the same time t. In a correct model this correlation should survive in the white noise limit. The Stratonovich integral definition takes into account this relation between the external noise and the system.
The approximating sums of the Stratonovich integral are "slightly anticipating", it is not sufficient to know the Wiener process up to time $t_s \in [t_{i-1}, t_i)$ to determine the value for $\underline{B}(\vec{x}(t_s))$ in the approximating sums, but the knowledge of the Wiener process up to t_i is needed since $\vec{x}(\tau_i) = \frac{\vec{x}(t_i) - \vec{x}(t_{i-1})}{2}$:

$$\int \underline{B}(\vec{x}_t) d\vec{W}_t = \lim \sum \underline{B}(\frac{\vec{x}_{t_i} - \vec{x}_{t_{i-1}}}{2}, t_{i-1})(\vec{W}_{t_i} - \vec{W}_{t_{i-1}}).$$

In the Ito calculus where the function is to be evaluated at the beginning of the infinitesimal intervals of the partial sums one gets:

$$\int \underline{B}(\vec{x}_t) d\vec{W}_t = \lim \sum \underline{B}(\vec{x}_{t_{i-1}}, t_{i-1})(\vec{W}_{t_i} - \vec{W}_{t_{i-1}}).$$

If one has chosen a certain "interpretation" of the integral, that is one special partition of the considered interval, and keeps in mind all the consequences of this choice, such as particular transformation rules when changing variables, then one has given a sensible meaning to the equations.
The Ito stochastic differential equation

$$d\vec{x}(t) = [\vec{A}(\vec{x}, t) + \frac{1}{2}\underline{B}(\vec{x}, t)' \underline{B}(\vec{x}, t)^T]dt + \underline{B}(\vec{x}, t) d\vec{W}, \tag{2.21}$$

where the i^{th} component of $\frac{1}{2}\underline{B}'\underline{B}^T$ reads:

$$\frac{1}{2} \sum_{i,j,k} (\frac{\partial}{\partial x_j} B_{ik}) B_{jk},$$

is equivalent to the Stratonovich stochastic differential equation

$$d\vec{x}(t) = \vec{A}(\vec{x}, t)dt + \underline{B}(\vec{x}, t) d\vec{W} \tag{2.22}$$

in the sense that they have one and the same solution.

Remarks:

- The Stratonovich integration calculus obeys the classical rules for coordinate transformations. (It is the only possible definition of a stochastic integral that fulfils the classical rules.)
- The Stratonovich version takes into consideration the stochastic dependence between the state of the system and the environmental fluctuations.
- The Stratonovich version of a stochastic differential equation describes a diffusion process with
  - drift term $\vec{A} + \frac{1}{2}\underline{B}'\underline{B}^T$,
  - diffusion term $\underline{B}\underline{B}^T$.
- It is always possible to change from the Stratonovich *calculus* to the Ito *calculus* by adding the term $\frac{1}{2}\underline{B}(\vec{x}, t)'\underline{B}(\vec{x}, t)^T$.

  And vice versa one can change from the Ito version of a stochastic differential equation to the Stratonovich version by subtracting $\frac{1}{2}\underline{B}(\vec{x}, t)'\underline{B}(\vec{x}, t)^T$.
- For additive noise there is no difference in the various calculi. This difference appears only for multiplicative noise where the influence of the random force depends on the state of the process. The correlation between the process $\vec{x}(t)$ and the random force $\vec{W}(t)$ is modelled implicitly in the Stratonovich definition and leads to a systematic contribution to the evolution of $\vec{x}(t)$, the noise induced drift $\frac{1}{2}\underline{B}'\underline{B}^T$.
- The Wong-Zakai theorem [8] states, that starting from a phenomenological equation containing real noise and passing to the white noise limit one arrives at a stochastic differential equation of the form:

  $$d\vec{x}(t) = \vec{A}(\vec{x}, t)dt + \underline{B}(\vec{x}, t) d\vec{W}$$

  which has to be interpreted as a Stratonovich equation.
- It is worth noting that due to the unproblematic transformation and the much easier stochastic relations in the Ito calculus, all results concerning the Stratonovich integral are proved using Ito's theory.

The Fokker-Planck equations in the two different calculi are the following:
For $\underline{D} = \underline{B}\underline{B}^T$ the Fokker-Planck equation corresponding to the stochastic differential equation (2.20) in the Ito calculus is

$$\frac{\partial}{\partial t}\rho(\vec{x}, t|\vec{z}, t') = -\frac{\partial}{\partial x_i} A_i(\vec{x}, t)\rho(\vec{x}, t|\vec{z}, t') + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_k} D_{ik}(\vec{x}, t)\rho(\vec{x}, t|\vec{z}, t'). \tag{2.23}$$

The Stratonovich interpretation of the stochastic differential equation (2.20) leads to the Fokker-Planck equation

$$\frac{\partial}{\partial t}\rho(\vec{x},t|\vec{z},t') = -\frac{\partial}{\partial x_i}\hat{A}_i(\vec{x},t)\rho(\vec{x},t|\vec{z},t') + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_k}D_{ik}(\vec{x},t)\rho(\vec{x},t|\vec{z},t'). \qquad (2.24)$$

with

$$\hat{A}_i = A_i + \frac{1}{2}(\frac{\partial}{\partial x_j}B_{ik})B_{jk},$$

that is:

$$\frac{\partial}{\partial t}\rho(\vec{x},t|\vec{z},t') = -\frac{\partial}{\partial x_i}[A_i(\vec{x},t)+\frac{1}{2}(\frac{\partial}{\partial x_j}B_{ik}(\vec{x},t))B_{jk}(\vec{x},t)]\rho(\vec{x},t|\vec{z},t')+\frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_k}D_{ik}(\vec{x},t)\rho(\vec{x},t|\vec{z},t'). \qquad (2.25)$$

Summarizing one can say that the Stratonovich interpretation is appropriate if the stochastic differential equation has been obtained as a white noise limit of a coloured noise process. In the following all stochastic equations will be interpreted in the Stratonovich way.

3 Methods for Solving Stochastic Differential Equations

In this chapter solution methods for stochastic differential equations are summarized and numerical integration algorithms are introduced. Results obtained in testing these integration schemes are presented and compared with analytically solvable models. The examples that will be treated are the Ornstein-Uhlenbeck process, the harmonic oscillator and a double well potential (Duffing oscillator).

3.1 The Linear Stochastic Differential Equation

The simplest case of a stochastic system is a linear system subject to external noisy forces. The linear stochastic differential equation

$$\dot{\vec{X}} = -\underline{A}\vec{X} + \underline{B}\vec{\xi}(t); \quad \vec{X}(0) = \vec{x}_0 \qquad (3.1)$$

has the solution

$$\vec{X}(t) = e^{-\underline{A}t}\vec{x}_0 + \int_0^t e^{-\underline{A}(t-t')}\underline{B}\vec{\xi}(t')dt'.$$

The mean value is computed as

$$< \vec{X}(t) >= e^{-\underline{A}t}\vec{x}_0,$$

where $< \vec{\xi}(t) >= 0$ was used.
This result is also found by solving the differential equations

$$< \dot{X}_i > (t) = -A_{ij}(t) < X_j >$$

with the initial condition

$$< X_i(0) >= x_{i0}.$$

The differential equations for the first moment have the same form as for the noiseless case. This simple relation holds only for linear systems.
For a linear system one finds the solution via the transfer matrix or Green's function

$$\Phi(t-t') = e^{-\underline{A}(t-t')}.$$

If the matrix $\underline{A}$ is independent of t and if all its eigenvalues are positive, the mean value $< \vec{X}(t) >$ vanishes for $t \to \infty$.
The equation for the second moments $< X_iX_j >$ is found to be

$$\frac{d}{dt} < X_iX_j >= -A_{il}(t) < X_lX_j > - < X_iX_l > A^T_{lj}(t) + D_{ij}$$

with $\underline{B}\underline{B}^T = \underline{D}$.
The variance matrix

$$\Sigma_{ij}(t) =< X_iX_j > - < X_i >< X_j >$$

satisfies

$$\begin{aligned}\frac{d}{dt}\Sigma_{ij}(t) &= -A_{ik}(t) < X_kX_j > - < X_iX_k > A^T_{kj}(t) + D_{ij} \\ &\quad +A_{ik}(t) < X_k >< X_j > + < X_i >< X_k > A^T_{kj}(t).\end{aligned}$$

Written in matrix form this equation reads:

$$\dot{\Sigma} = -[\underline{A}\Sigma + \Sigma\underline{A}^T] + \underline{D}. \tag{3.2}$$

It has the solution

$$\Sigma(t) = \Phi(t)\Sigma(0)\Phi^T(t) + \int_0^t \Phi(t)\Phi^{-1}(t')\underline{D}(\Phi^T)^{-1}(t')\Phi^T(t)dt'.$$

The corresponding Fokker-Planck equation for the system (3.1) with linear drift and additive noise is

$$\frac{\partial}{\partial t}\rho(\vec{x},t|\vec{x}_0,0) = -\frac{\partial}{\partial x_i}A_i(\vec{x},t)\rho(\vec{x},t|\vec{x}_0,0) + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_k}D_{ik}\rho(\vec{x},t|\vec{x}_0,0) \tag{3.3}$$

with the initial condition for the transition probability

$$\rho(\vec{x},t=0|\vec{x}_0,0) = \delta(\vec{x}-\vec{x}_0).$$

The density function is completely determined by the mean and the variance and therefore the explicit solution of this equation can be given:

$$\rho(\vec{x},t) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma(t))}} e^{-\frac{1}{2}(\vec{x}-<\vec{X}(t)>)_i(\Sigma^{-1}(t))_{ij}(\vec{x}-<\vec{X}(t)>)_j}. \tag{3.4}$$

Under the condition that $<\vec{X}(t)> \rightarrow 0$ for $t \rightarrow \infty$ the stationary density is

$$\rho_{st}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma(\infty))}} e^{-\frac{1}{2}x_i(\Sigma^{-1}(\infty))_{ij}x_j} \tag{3.5}$$

where $\Sigma(\infty)$ stands for the stationary variance matrix. It is calculated from (3.2) by setting the time derivative equal to zero:

$$\underline{D} = \underline{A}\Sigma(\infty) + \Sigma(\infty)\underline{A}^T. \tag{3.6}$$

Equilibrium solutions of the Fokker-Planck equation can also be found when both exciting and damping forces act on the system and fulfil certain conditions of balance [6] such that a stationary state can be reached.
The Kramers equation describes a damped system which has a stationary solution

$$m\ddot{x} + \alpha\dot{x} + U'(x) = \sqrt{d}\xi(t).$$

It can be rewritten as

$$\begin{aligned} \dot{x} &= \frac{1}{m}p \\ \dot{p} &= -\frac{\alpha}{m}p - U'(x) + \sqrt{d}\xi(t) \end{aligned}$$

with the corresponding Fokker-Planck equation

$$\frac{\partial}{\partial t}\rho(x,p,t) = -\frac{1}{m}\frac{\partial}{\partial x}(p\rho) + \frac{\partial}{\partial p}(\frac{\alpha}{m}p + U'(x))\rho + \frac{1}{2}d\frac{\partial^2}{\partial p^2}\rho.$$

Equations of this kind appear in accelerator dynamics in the description of "smeared out" multipoles, i.e. one neglects the s-dependence of the electromagnetic fields, s parametrizing the arc length along the accelerator (see appendix), and substitutes these fields by equivalent averages all around the ring.
The stationary solution is given by

$$\rho_{st}(x,p) = N\exp\left(-\frac{2\alpha}{d}(\frac{p^2}{2m} + U(x))\right).$$

The equilibrium state is determined by the damping parameter α and the diffusion strength d. In many problems in accelerator physics one has to consider s-dependent potential terms $U(x,s)$ with $U(x,s+L) = U(x,s)$. When only quadratic terms are included in the potential (corresponding to linear forces) and the motion is stable, there exists a periodic steady state solution for the density function of an electron storage ring [21]. The moments determining the beam sizes (beam emittance matrix) have for this case been given in [1] and [2]. It is therefore expected to find an L-periodic steady state solution for the motion in an electron storage ring with damping and noise effects even when small nonlinear forces are included in the beam dynamics. In the absence of damping forces, i.e. for Hamiltonian sytems, one cannot get a normalizable stationary solution for the density function when additional noise is acting on the system. Stochastic excitation will lead to permanent growth of the moments of the system variables. An example of this kind of motion is the synchrotron motion in an accelerator which can in a first approximation be treated independently of the transverse motion [14]. However, for systems with time-varying coefficients or undamped nonlinear systems no general theorem for the existence and calculation of solutions is available. It is therefore necessary to have reliable numerical methods to handle these cases. Some approaches will be introduced and investigated in this work.
At first a short description of the used numerical algorithms will be given.

3.2 Numerical Solution of Stochastic Differential Equations

In numerical approach to stochastic differential equations one performs the following steps:

- Taylor expansion of the approximate solution in the step width h
- modelling of the noise process
- simulating a sufficient number of realizations for averaging.

A difficulty lies in the modelling of the white noise or, if necessary, of higher order functionals of it.
Required is an approximative solution of the equation

$$\dot{\vec{x}} = \vec{f}(\vec{x}) + \underline{g}(\vec{x})\vec{\xi}(t); \qquad \vec{x}(0) = \vec{x}_0. \tag{3.7}$$

If the solution $\vec{x}(t)$ at a given time t is known, then for the solution $\vec{x}(t+h)$ at the next time step holds:

$$x_i(t+h) = x_i(t) + \int_t^{t+h} ds f_i(\vec{x}(s)) + \int_t^{t+h} ds g_{ij}(\vec{x}(s))\xi_j(s).$$

Iteratively substituting this expression into the terms inside the integrals yields

$$x_i(t+h) = x_i(t) + \int_t^{t+h} ds f_i[x_k(t) + \int_t^s du f_k(\vec{x}(u)) + \int_t^s du g_{kl}(\vec{x}(u))\xi_l(u)]$$
$$+ \int_t^{t+h} ds g_{ij}[x_k(t) + \int_t^s du f_k(\vec{x}(u)) + \int_t^s du g_{kl}(\vec{x}(u))\xi_l(u)]\xi_j(s).$$

Now one applies the Taylor expansion around $x_i(t)$. In this expansion functionals of the white noise term appear, e.g.:

$$W_i(h) = \int_t^{t+h} ds\xi_i(s) \quad , \quad F_i(h) = \int_t^{t+h} ds W_i(s).$$

$$C_{ij}(h) = \int_t^{t+h} ds W_i(s)\xi_j(s) \quad , \quad G_{ij}(h) = \int_t^{t+h} ds W_i(s) W_j(s).$$

A quantity y is said to be of the order of m in the step width h, $y \simeq h^m$, when there exists an $n \in \mathrm{N}, c \in \mathrm{R}$ with $< y^n > \simeq c \cdot h^{mn}$.
This definition gives for the order of the functionals appearing in the expansion:

$$W_i(h) \simeq h^{\frac{1}{2}}, \quad F_i(h) \simeq h^{\frac{3}{2}}, \quad C_{ij}(h) \simeq h, \quad G_{ij}(h) \simeq h^2.$$

Inserting these expressions, one gets with $f_i = f_i(\vec{x}(t))$:

$$x_i(t+h) = x_i(t) + g_{ij}W_j + f_i h$$
$$+(\frac{\partial}{\partial x_k} f_i) g_{kl} F_l(h) + (\frac{\partial}{\partial x_k} f_i) f_k \frac{1}{2} h^2 + \frac{1}{2}(\frac{\partial^2}{\partial x_k \partial x_l} f_i) g_{km} g_{ln} G_{mn}(h) + (\frac{\partial}{\partial x_k} f_i) f_k \frac{1}{2} h^2 + o(h^3)$$
$$+(\frac{\partial}{\partial x_k} g_{ij}) g_{kl} C_{jl}(h) + o(h^2, \underline{g}').$$

The last term is to be evaluated if one has a stochastic differential equation with multiplicative noise which means that $\underline{g}'$, the derivative of $\underline{g}$ with respect to $\vec{x}$, does not vanish.
These expansions include derivatives of the drift term f_i and of the diffusion term g_{ij} and also functionals of the white noise process.
This latter is a restriction for numerical approaches. In deterministic theory it is well known that higher than linear functionals of δ-functions make no sense. Here it is not clear how to model higher functionals of the white noise terms. In simulations these expressions have to be substituted by simple functions of the random vector $W_j(h)$ so that they yield the same moments up to a given order which is, compared to the gain in precision, often too computing time consuming. Therefore one is limited to algorithms of low order like the Euler and Heun schemes.
For more detailed treatment of the numerical approach to stochastic processes see for example [22],[23],[24].

The Euler Scheme

The Euler integration scheme includes terms up to the order of h for additive noise. For multiplicative noise in the simple Euler algorithm one term of order h is neglected, the term $C_{ij}(h)$,so one loses one order of convergence in this case and has to apply a slightly modified scheme which can, at least in the one-dimensional case, also be written down explicitly (Milsthein algorithm).

For additive noise g_{ij} is independent of x_i and therefore one gets

$$x_i(t+h) = x_i(t) + f_i(x_i(t)) \cdot h + g_{ij} \cdot \tilde{W}_j(h).$$

$\tilde{W}_j(h)$ is a random vector with the same first moments (mean and variance) as the normally distributed $W_j(h)$, that means: $< \tilde{W}_k >= 0, < \tilde{W}_k \tilde{W}_{k'} >= h\delta_{kk'}$.
It can be simulated by setting $\tilde{W}_j(h) = \sqrt{12h}(r_j - 0.5)$ with the r_j being independent uniformly distributed random numbers. It is not necessary to simulate Gaussian random numbers because the higher order moments do not appear in the approximate solution up to the desired order of convergence and one can take uniformly distributed random numbers with the same first moments.
Treating multiplicative noise the terms containing the derivative of the diffusion matrix g_{ij} have to be included, so in the (modified) Euler case the functional $C_{ij}(h)$ has to be simulated which leads to

$$x_i(t+h) = x_i(t) + f_i(x_i(t)) \cdot h + [\frac{\partial}{\partial x_k} g_{ij}(x_i(t))] g_{kl}(x_i(t)) \cdot \tilde{C}_{ij}(h) + g_{ij}(x_i(t)) \cdot \tilde{W}_j(h).$$

Now one has to decide in which calculus to work when evaluating $C_{ij}(h)$ and substituting it by suitable random numbers (denoted by a tilde) in the algorithm.
The Ito calculus gives for the one-dimensional case:

$$C(h) = \frac{1}{2}[(W^2(h) - W^2(0)) - (h - 0)] = \frac{1}{2}(W^2(h) - W^2(0)) - \frac{1}{2}h.$$

From the Stratonovich calculus follows:

$$C(h) = \frac{1}{2}(W^2(h) - W^2(0)),$$

and $W(h)$ is approximated as before.

The Heun Scheme

For the case of additive noise an algorithm of second order in the step width h is introduced, the Heun scheme. It gives a better convergence for additive noise but no improvement in the multiplicative case. The difference to the Euler scheme is an additional predictor step and one gets for the approximate solution:

$$x_i(t+h) = x_i(t) + \frac{1}{2}(f_i(x_i(t) + f_i(\bar{x}_i(h)) \cdot h + g_{ij}(x_i(t)) \cdot \tilde{W}_j(h),$$

with

$$\bar{x}_i(h) = x_i(t) + f_i(x_i(t)) \cdot h + g_{ij}(x_i(t)) \cdot \tilde{W}_j(h).$$

Especially for small step sizes these algorithms are very computing time consuming. Besides, as the statistical error decreases with $\frac{1}{\sqrt{N}}$, big numbers of realizations are necessary. The statistical error σ_{err} in measuring a random variable X is calculated as the standard deviation in a series of N independent realizations X_i of X and is given by:

$$\sigma_{err} = \sqrt{Var(\bar{X}_N)} \tag{3.8}$$

with

$$\begin{aligned} Var(\bar{X}_N) &= Var(\frac{1}{N}\sum_{i=1}^{N} X_i) \\ &= <(\frac{1}{N}\sum_{i=1}^{N} X_i - <\frac{1}{N}\sum_{i=1}^{N} X_i>)^2> \\ &= \frac{1}{N^2}(\sum_{i=1}^{N} <X_i^2> - <X_i>^2 + \sum_{i\neq j}^{N} <X_i X_j> - <X_i><X_j>) \\ &= \frac{1}{N^2}\sum_{i=1}^{N} <X_i^2> - <X_i>^2 \\ &= \frac{1}{N^2}\sum_{i=1}^{N} VarX_i \end{aligned}$$

The independence of the realizations leads to vanishing covariances, and as the error in the single measurements is supposed to be identical ($VarX_1 = ... = VarX_N$) this yields

$$\sigma_{err} = \sqrt{\frac{1}{N} VarX_1} = \frac{1}{\sqrt{N}}\sigma_1.$$

4 Study of Stochastic Processes in Accelerator Physics by Means of Stochastic Differential Equations

4.1 Test of the Methods for Analytically Solvable Problems

In this section exactly solvable models are investigated numerically and compared with the analytic solutions.

Ornstein-Uhlenbeck Process

An example of a linear one-dimensional stochastic differential equation is the Ornstein-Uhlenbeck process.

$$\dot{x} + \alpha x = \sqrt{d} \cdot \xi(t).$$

The analytic expressions for the moments are:

$$< x > (t) = x_0 e^{-\alpha t}$$

and

$$< x^2 > (t) = x_0^2 e^{-2\alpha t} + \frac{d}{2\alpha}(1 - e^{-2\alpha t}).$$

For the numerical calculation the Euler and the Heun schemes are used. The Heun scheme needs about 20% more computing time. For this simple example the two algorithms did not lead to different results. The parameters were $\alpha = 1.0, d = 1.0$, and $x_0 = 0$. Figure 4 shows the

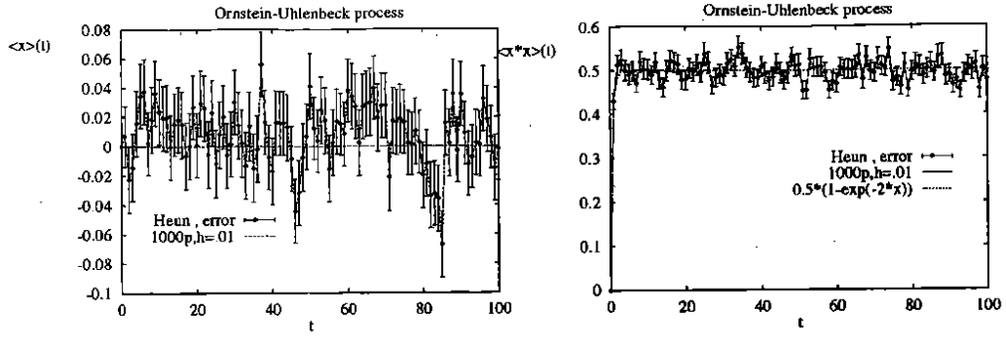

Figure 4: First moment $< x > (t)$ and second moment $< x^2 > (t)$ for initial value $x_0 = 0$, damping constant $\alpha = 1$ and noise parameter d=1, calculated from 1000 samples with a step width $h = 0.01$ using the Heun algorithm.

results for a simulation with the Heun algorithm for 1000 samples and a step width $h = 0.01$ in comparison with the analytic result. The agreement between theoretical and numerical solution is obvious. The statistical error has been calculated using (3.8). In some of the following figure the error bars have been omitted for clarity.

Harmonic Oscillator

The following equation describes a harmonic oscillator with frequency ω, damping coefficient α and stochastic excitation.

$$\ddot{x} + \alpha\dot{x} + \omega^2 x = \sqrt{d} \cdot \xi(t).$$

For the initial conditions $x_0 = 0$ and $\dot{x}_0 = 0$ the undamped and the damped case are investigated.

- undamped case, $\alpha = 0$:
  The Green's function for the undamped harmonic oscillator reads

  $$\Phi(t) = \begin{pmatrix} \cos(\omega t) & \frac{1}{\omega}\sin(\omega t) \\ -\omega\sin(\omega t) & \cos(\omega t) \end{pmatrix}.$$

  The mean value for these initial conditions is found to be

  $$< x > (t) = 0,$$

  and the second moment is calculated as

  $$< x^2 > (t) = \sigma^2(t) = \frac{d}{2}(t - \frac{1}{2\omega}\sin(2\omega t)).$$

  The variance increases continuously with time, no stationary solution exists.

  In figure 5 it can be seen that the deviations from the mean value $< x > (t) = 0$ are oscillating and grow with time. The second moment shows similar behaviour. From a

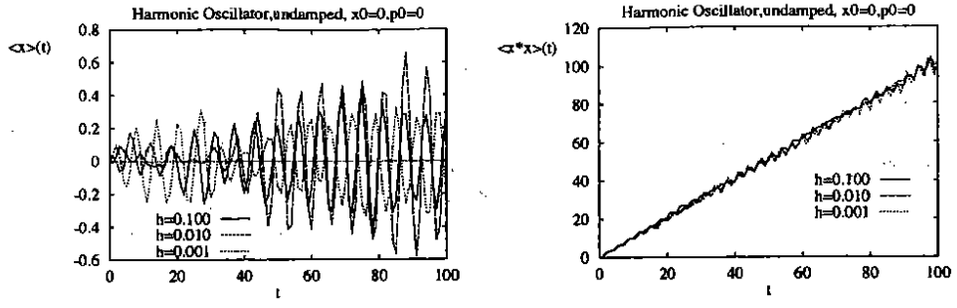

Figure 5: First moment $< x > (t)$ and second moment $< x^2 > (t)$ for initial values $x_0 = \dot{x}_0 = 0$, damping parameter $\alpha = 0$, noise strength $d = 2$, 1000 samples, different step sizes, calculated with the Heun algorithm.

certain time on the amplitude of the oscillations about the theoretical result starts to grow. Nevertheless the error in the second moment is much smaller than in the first moment. It can be concluded that for the undamped case the numerical calculations stay stable only for relatively short time periods.

- damped case, $\alpha > 0$:
  For the damped harmonic oscillator the stationary moments are computed from equation (3.6). They are found to be

$$\begin{aligned} < x >_{st} &= < \dot{x} >_{st} = 0, \\ < x^2 >_{st} &= \sigma_x^2 = \frac{d}{2\alpha\omega^2}, \\ < \dot{x}^2 >_{st} &= \sigma_p^2 = \frac{d}{2\alpha}. \end{aligned} \tag{4.1}$$

  The stationary solution for the density $\rho(x, \dot{x})$ is a Gaussian.

  In the damped case the solutions for the mean and the second moment stay stable, see figure 6. So these integration schemes can be applied for long term studies of damped systems.

  In figure 7 the time evolution of the mean total action I is given. This quantity is defined as $< I >= \frac{1}{2}(\omega < x^2 > + \frac{1}{\omega} < p_x^2 >)$.

  The left hand side of figure 7 corresponds to the mean total action in the undamped system. Although the result agrees with the theoretical curve it can be seen that the errors of the calculation grow with time. For the damped harmonic oscillator (right hand side of figure 7) a growth of the errors cannot be observed.

Double well potential

A stochastic differential equation with a (nonlinear) time-independent potential $U(x)$ is called Kramers equation. An example for a Kramers equation is the stochastically excited Duffing

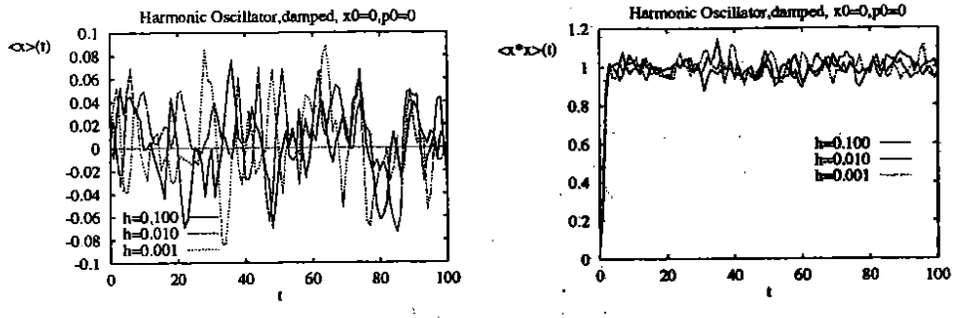

Figure 6: First moment $< x > (t)$ and second moment $< x^2 > (t)$ for $x_0 = \dot{x}_0 = 0, \alpha = 1, d = 2$, 1000 samples, different step sizes, calculated with the Heun algorithm.

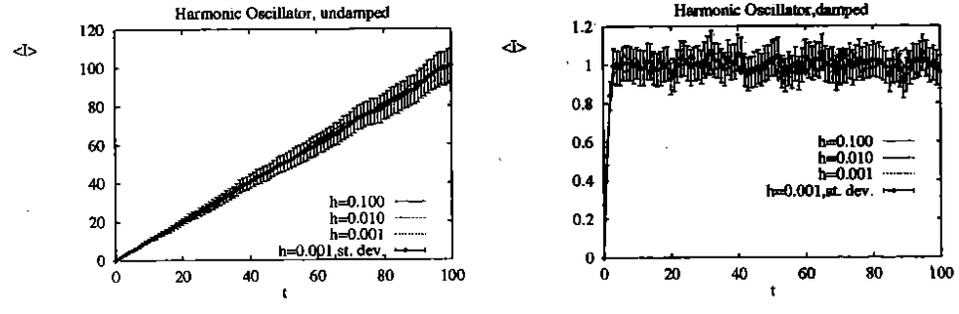

Figure 7: Mean total action $< I > (t) = \frac{1}{2}(\omega < x^2 > + \frac{1}{\omega} < p_x^2 >)$ as a function of time for the initial conditions $x_0 = \dot{x}_0 = 0$, in the undamped case $\alpha = 0$ (left) and damped case $\alpha = 1$ (right), calculated with the Heun algorithm.

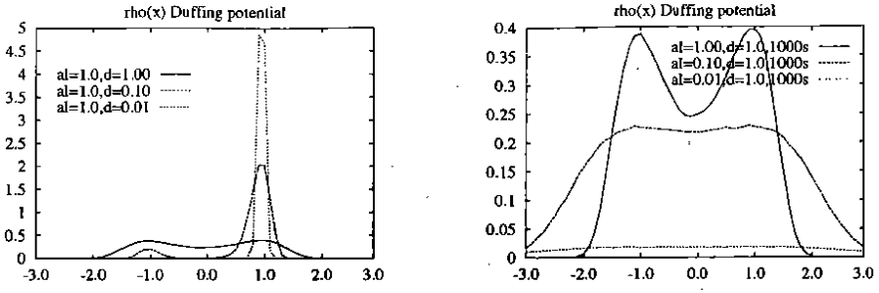

Figure 8: Stationary density $\rho_{st}(x)$ for the Duffing oscillator.Shown is the projection onto the x-plane calculated with the Heun algorithm.
a) different noise strengths, same damping parameter
b) different damping parameters, same noise strength.

oscillator characterized by the double well potential $U(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2$. This system is named after G. Duffing who originally studied the particle motion a double well potential under the additional influence of a periodic driving force.

$$\ddot{x} + \alpha\dot{x} + x^3 - x = \sqrt{d} \cdot \xi(t).$$

Potential functions of this kind are e.g. "smeared-out" multipoles along a storage ring.
The stationary solution is

$$\rho_{st}(x, \dot{x}) = Ne^{-\frac{2\alpha}{d}(\frac{1}{2}\dot{x}^2 + (\frac{1}{4}x^4 - \frac{1}{2}x^2))}.$$

The marginal density of (=projection onto) the $\dot{x}$-variable (=p) is a Gaussian, whereas the x-axis projection of the density function is determined by the nonlinear potential term.
The maxima of the density function are found to be at the stable fixed points of the underlying Hamiltonian system which are $+1.0$ and -1.0. When noise is switched on transition of the particles between the maxima around the fixed points can be observed. This transition behaviour depends strongly on the noise strength and the damping (see figure 8). If the noise strength is reduced, the system approaches the dynamics of the deterministic case, see figure 8 a), where the density is given for different noise strengths and constant damping term.
For vanishing damping and constant noise the stochastic excitation tends to smear out the structure, see figure 8 b).
The figures 9 to 11 show the stationary density $\rho(x, p_x)$ for three different damping parameters α. For both strong damping and strong noise intensity (figure 9) the density in p_x-direction is clearly visible to be Gaussian. The density in x-direction is determined by the double well potential. If the damping is reduced while the noise excitation is kept constant the structure is smeared out, see figures 10 and 11.
The calculations were made with the Heun algorithm and the step size $h = 0.1$ for 1000 samples and an iteration time of 200 steps. The initial conditions have been $x_0 = 0.0$ and $p_0 = 1.0$.

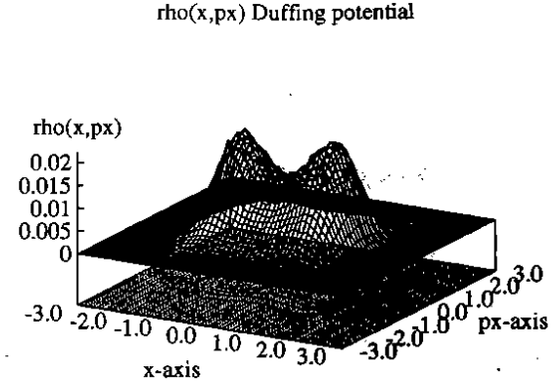

Figure 9: Stationary density $\rho_{st}(x, p)$ for the large damping parameter $\alpha = 1.0$ and the strong noise excitation $d = 1$. The Gaussian form in p-direction and the double peak shape in x-direction can easily be recognized.

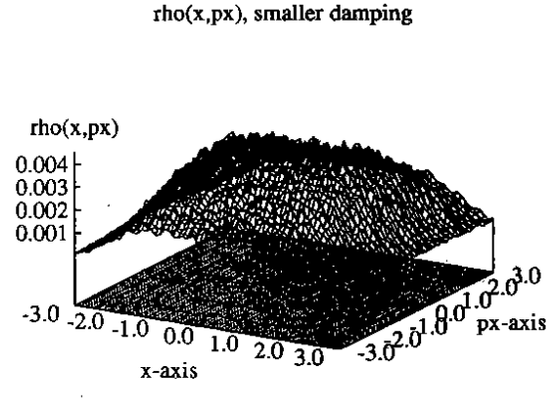

Figure 10: Stationary density $\rho_{st}(x, p)$ for the damping parameter $\alpha = 0.1$ and the strong noise excitation $d = 1$. It can be seen that the distinct contours visible in figure 9 have nearly disappeared.

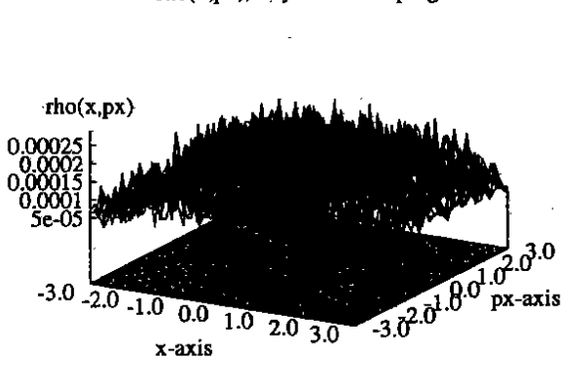

Figure 11: Stationary density $\rho_{st}(x,p)$ for the damping parameter $\alpha = 0.01$ and the same large noise parameter as before. The motion is dominated by the stochastic excitation. The structure of maxima and minima of the density has completely vanished.

4.2 Beam-Beam Interaction

In colliding beam facilities where two counterrotating beams of ultrarelativistic charged particles repeatedly cross each other, the additional fields at the interaction area of the two charge distributions have a strong perturbing effect on the particle dynamics. Here the effect of a Gaussian charge distribution (strong beam) on a single particle (weak beam) is considered. This approximation is called the weak-strong model of the beam-beam effect.
The potential of such a charge distribution has the form

$$U(x,z,s) = \frac{N_b r_e}{\gamma}\int_0^\infty \frac{1-\exp(\frac{-x^2}{2\sigma_x^2+q}+\frac{-z^2}{2\sigma_z^2+q})}{(2\sigma_x^2+q)^{\frac{1}{2}}(2\sigma_z^2+q)^{\frac{1}{2}}}\,dq \quad \delta_p(s-s_0).$$

The δ-function represents the s-periodicity for the successive crossings of the two beams. The periodic δ-function is given by

$$\delta_p(s) = \sum_{n=-\infty}^{\infty}\delta(s-nL).$$

In the particular case of round beams this yields for the kicks in the momenta of the particle

$$\begin{aligned}\Delta x' &= -2\frac{N_b r_e}{\gamma}\frac{x}{r^2}(1-\exp(-\frac{r^2}{2\sigma^2})) \\ \Delta z' &= -2\frac{N_b r_e}{\gamma}\frac{z}{r^2}(1-\exp(-\frac{r^2}{2\sigma^2})) \qquad r^2 = x^2+z^2.\end{aligned} \tag{4.2}$$

A one-dimensional model of this interaction is described by the stochastic equation:

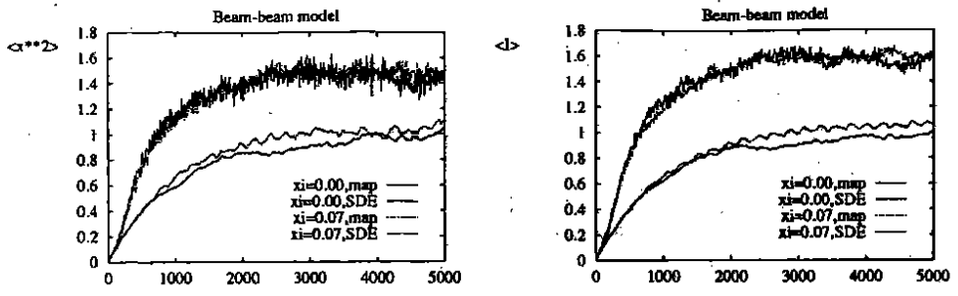

Figure 12: Time evolution of $< x^2 >$ and $< I >$ in the linear case (lower curves) and with beam-beam kick force (upper curves). Shown are the results of calculations with a stochastic differential equation (Heun scheme) and of a simulation with an equivalent transfer map.

$$\ddot{x} + \alpha\dot{x} + \omega^2 x + f(x,s) = \sqrt{d}\cdot\xi(s),$$

with

$$f(x,s) = -8\pi\xi_{bb}(\frac{1-e^{-\frac{1}{2}x^2}}{x})\delta_p(s),$$

where

$$\sqrt{d} = \sigma\omega\sqrt{2\alpha}$$

and the quantity

$$\xi_{bb} = \frac{N_b r_e}{\gamma}\frac{\beta^*}{2\pi\sigma^2}$$

is called the beam-beam parameter. It determines the maximum tune shift in linear approximation (valid up to 1 σ) where the effect of the beam-beam potential can be approximated by a quadrupole kick.
For the following studies a simple model ring of length $23.25m$ with a Q-value of 3.7 and a strong beam-beam parameter of $\xi_{bb} = 0.07$ was used, the damping time was taken to be 1000 turns. The weight of the noise has been adjusted to achieve an equilibrium value of $< x^2 > = 1$ for the corresponding harmonic oscillator representing the linear system without beam-beam force. The parameters have been chosen in a way to make the effect of the periodic kick force easily visible. For the stochastic differential equation simulation the Heun algorithm has been used with 500 samples for the averaging over the noise process.
Figure 12 shows the time development of the second moment of x, $< x^2 >$, that measures the bunch size in x-direction and of the mean total action $< I >$ which determines the bunch dimensions in phase space.
In both cases, unperturbed (= noisy harmonic oscillator, linear system, $f(x,s) = 0$) and perturbed (= noisy harmonic oscillator plus beam-beam kick force), an equilibrium value is reached.

For the beam-beam system comparative computations with the equivalent transfer map for the

harmonic oscillator have been performed. In this case the integrated damping and noise has been applied at one interaction point.
The results are in agreement within the statistical errors but the mapping scheme is by far less computing time intensive.

4.3 Stochasticity in the Synchrotron Motion

The accelerating system consists of radiofrequency (rf) cavities with longitudinal electric fields on the s-axis that provide the energy needed for acceleration and (for electron beams) substitute the energy radiated away in the bending magnets.
For protons the energy losses due to the synchrotron radiation are negligible because of their higher mass. After acceleration the rf system is used to keep the proton beam "bunched", that means the 10^{10} to 10^{11} particles in a bunch are performing longitudinal (phase) oscillations around some reference particle.
This reference or synchronous particle is taken to be a particle with the nominal momentum and it is required that on each of its passages through the cavity the rf phase should have the same value ϕ_s so that this particle always receives the same energy gain. This condition is fulfilled if the frequency of the accelerating system ω_{rf} is an integer multiple of the revolution frequency ω_0, $\omega_{rf} = h\omega_0$, with $\omega_0 = \frac{2\pi}{T_0}$. The integer h is called the harmonic number.
Noise in the amplitude and in the phase of the voltage of an rf cavity introduces a diffusion of particles in the longitudinal phase space. In the absence of damping mechanisms this will lead to a permanent growth of the bunch length and hence to a reduction of the beam lifetime because particles leaving the stable phase space region, the rf bucket, are lost from the beam. The effect of cavity noise on the particle motion in proton storage rings has been treated previously with different approaches, see [12],[14],[19],[20],[26],[27],[28]. In the following a proton rf system with additional noise perturbation will be studied. At first some general features of such rf systems are summarized. To this end the variables $q(t)$ as the deviation of the rf phase of a proton from the synchronous value and $p(t) = \dot{q}(t)$, its conjugate momentum, corresponding to a deviation from the nominal energy are introduced. The quantity that characterizes the beam size in phase space is the (single particle) emittance ϵ which is related to the canonical action $J = \int p dq$ via $J = \pi\epsilon$.
This system is represented by the Hamiltonian

$$H = \frac{1}{2}p^2 + V(q),$$

with $V(q)$ being an arbitrary rf potential.
In general $V(q)$ has the form

$$V(q) = \sum_{n=1}^{\infty} a_n \sin(nq) + b_n \cos(nq).$$

The Hamilton equations of motion read:

$$\begin{aligned} \dot{q} &= \frac{\partial H}{\partial p} = p \\ \dot{p} &= -\frac{\partial H}{\partial q} = -\frac{dV}{dq}(q). \end{aligned} \tag{4.3}$$

Any deviation of the rf voltage from its nominal value can be decomposed into a phase error and an amplitude error of the voltage.

If now noise is included in the description (4.3), this yields

$$H = \frac{1}{2}p^2 + V(q) + h(q)\xi(t), \tag{4.4}$$

with $\xi(t)$ describing a random process with the properties $< \xi(t) >= 0$ and $< \xi(t)\xi(t') >= \delta(t - t')$,
and where one has to set

- for phase noise: $h(q) = -q$,
- for amplitude noise: $h(q) = V(q)$.

Phase noise was shown to have the more destructive effect [28] and the following studies are restricted to this case.

4.3.1 The Fokker-Planck Equation for Longitudinal Motion with Phase Noise

The form of the Fokker-Planck equation for the average synchrotron motion on a long time scale is [12],[20]

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial J}(A_1\rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(A_2\rho).$$

J stands for the action variable and $\rho(J,t)dJ$ defines the probability of finding a particle with action values between J and $J + dJ$ at the time t.
The coefficients A_1 and A_2 are given by

$$A_1 = << \frac{\Delta J}{\Delta t} >>_Q,$$

$$A_2 = << \frac{(\Delta J)^2}{\Delta t} >>_Q,$$

with ΔJ being the change in the action variable during the time interval Δt.
The double brackets indicate the two averaging processes involved in deriving the equation: an average over the noise process and additionally an average over the (initial) values of the angle variable Q which is the conjugate variable to the action J.
Boussard, Dôme and Graziani have shown in [20] that for particular (linear and sinusoidal) rf potentials the drift coefficient is related to the diffusion coefficient by

$$A_1 = \frac{1}{2}\frac{\partial}{\partial J}A_2,$$

and therefore the Fokker-Planck equation can be written in the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial J}(\frac{1}{2}A_2\frac{\partial}{\partial J}\rho). \tag{4.5}$$

Krinsky and Wang [12] proved this relation for the general case of an arbitrary rf potential. They also derived an explicit expression for the diffusion term A_2.
A short outline of their proof which is based on canonical perturbation theory for stochastic differential equations to second order in the noise term (and therefore to first order in time, see before), is given in the appendix.

4.4 Double Rf System with Phase Noise

In the following a double rf potential with additional phase noise is investigated using several numerical and analytical methods.
By adding a higher harmonic of the rf to the main system (therefore the name "double rf") one is able to control the synchrotron frequency, the synchrotron frequency spread and the bunch length [10],[11],[28]. Wei showed the advantages of a double rf system for stochastic cooling.
During this section the notation used in [11] will be followed.
The starting point is a short description of the Hamiltonian dynamics. Then different ways to calculate the Fokker-Planck equation will be applied, like the method of averaging [14] or perturbation theory for the stochastic differential equations [12]. Finally the system is investigated numerically with the methods described before.
Let now ϕ denote the phase deviation of a particle from the synchronous phase. The conjugate variable W is given by the deviation in energy divided by ω_{rf}: $W = \frac{\Delta E}{\omega_{rf}} = \frac{\Delta E}{h\omega_0}$.
A particle with the coordinates ϕ and W sees the voltage

$$V(\phi) = \hat{V}\sin(\phi + \phi_s) + k\hat{V}\sin(m\phi + m\phi_{2s}),$$

with
$\hat{V}$= peak voltage of the fundamental rf system,
$k\hat{V}$= peak voltage of the secondary rf system,
ϕ_s= stable phase angle relative to the fundamental wave form,
ϕ_{2s}= stable phase angle relative to the higher frequency wave form.

The first and the second derivative of $V(\phi)$ should vanish at the center of the bunch in order to maximize the bunch length and to avoid other regions of stability close by.
These two conditions give $mk\cos(m\phi_s) = -\cos(\phi_s)$ and $m^2k\sin(m\phi_s) = -\sin(\phi_s)$.
If the particle beam is stored without acceleration one has $\hat{V}\sin(\phi_s) + k\hat{V}\sin(m\phi_s) = 0$.
All this implies $mk = 1$, and $\phi_s = 0, m\phi_{2s} = \pi$ or $\phi_s = \pi, m\phi_{2s} = 0$.
Assuming $\phi_s = 0$ and $m\phi_{2s} = \pi$ this yields the following Hamiltonian:

$$H(\phi, W) = C_W W^2 + \frac{1}{2}C_\phi[(1 - \cos\phi) - \frac{1}{m^2}(1 - \cos m\phi)] - \phi\sqrt{d}\xi(t), \qquad (4.6)$$

where
$C_W = \frac{h^2\omega_0^2\eta}{2E\beta^2}$, $\quad C_\phi = \frac{qe\hat{V}}{\pi h}$,
qe=electric charge of the particle,
h= harmonic number,
γ_t= transition energy,
$\eta = \frac{1}{\gamma_t^2} - \frac{1}{\gamma^2}$,
$\alpha = \frac{1}{\gamma_t^2}$ = momentum compaction factor,
ω_0=revolution frequency of the synchronous particle,
βc=velocity of the synchronous particle,
$E = m_0c^2\gamma$ = synchronous energy,
$\sqrt{d}$= scaling diffusion parameter for the noise term,
$\xi(t)$= noise process (properties see before).
The potential $V(\phi)$ and its derivative $V'(\phi) = \frac{\partial V(\phi)}{\partial\phi}$ are shown in figure 13.

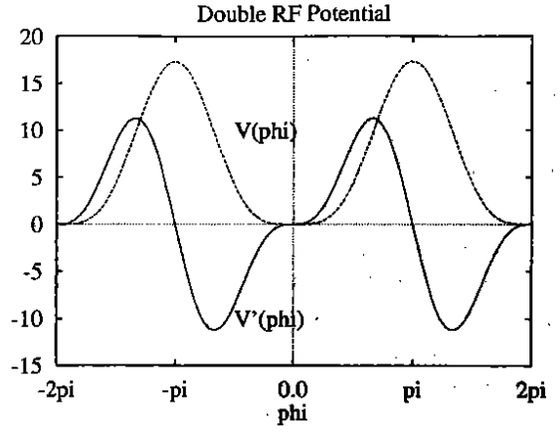

Figure 13: $V(\phi)$ and $V'(\phi) = \frac{\partial V(\phi)}{\partial\phi}$ for the double rf system.

Let α_H be the value of the Hamiltonian (4.6) in the deterministic case ($\sqrt{d} = 0$). For α_H holds $0 \le \alpha_H \le C_\phi$.
The Hamiltonian equations are

$$\begin{aligned} \dot{\phi} &= \frac{\partial H}{\partial W} = 2C_W W \qquad (4.7) \\ \dot{W} &= -\frac{\partial H}{\partial\phi} = -\frac{1}{2}C_\phi\sin(\phi) + \frac{1}{2m}C_\phi\sin(m\phi) + \sqrt{d}\xi(t). \end{aligned}$$

This system of stochastic differential equations will be used later on for the direct numerical simulations in the original phase space variables ϕ and W.
The quantities of interest are the mean bunch length $\sigma(t)$ and the mean action $< J(t) >$.
The second moment of ϕ is related to the bunch length σ by

$$\sigma = \frac{L}{\pi h}\sqrt{<\phi^2>} = \frac{2c}{\omega_0 h}\sqrt{<\phi^2>},$$

where $L = cT_0 = \frac{c2\pi}{\omega_0}$ is the circumference of the accelerator.

4.4.1 The Deterministic System

As a starting point the deterministic dynamics of the system will be considered, that means equation (4.6) without noise, $d = 0$.
The Hamiltonian is then independent of time and completely integrable. Therefore the action $J = \oint W d\phi$ is a constant of motion.
The solution of this system is found in the usual way using Hamilton-Jacobi theory [25], see appendix.

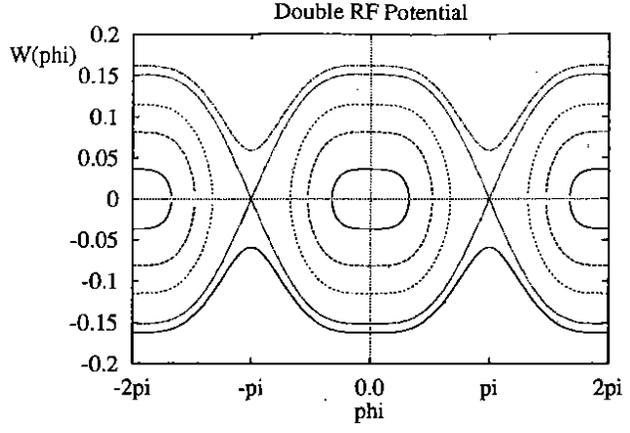

Figure 14: Phase space trajectories $W(\phi)$ for different energy values.

In the small amplitude limit $m\phi << 1$ the potential

$$V(\phi) = \frac{1}{2}C_\phi((1-\cos\phi) - \frac{1}{m^2}(1-\cos m\phi))$$

can be approximated with the ϕ^4-potential

$$V(\phi) = \frac{1}{2}C_\phi(\frac{m^2-1}{4!})\phi^4.$$

The action variable J can be expressed as a function of the value of the Hamiltonian α_H which is the energy of the system.

$$J = \underbrace{\frac{8\sqrt{2}K(\frac{1}{\sqrt{2}})}{3C_W^{\frac{1}{2}}C_\phi^{\frac{1}{4}}}(\frac{3}{m^2-1})^{\frac{1}{4}}}_{=const=:\tilde{c}}\alpha_H^{\frac{3}{4}} = J(\alpha_H), \qquad \alpha_H << C_\phi, \tag{4.8}$$

where $K(\frac{1}{\sqrt{2}}) = 1.8541$ is the complete elliptic integral of modulus $k = \frac{1}{\sqrt{2}}$ (see [29]).
This relation is used in the numerical calculations to simulate the growth of the action variable in the noise perturbed system:

$$< J >= \tilde{c} < (C_W W^2 + \frac{1}{2}C_\phi[(1-\cos(\phi)) - \frac{1}{4}(1-\cos(2\phi))])^{\frac{3}{4}} > . \tag{4.9}$$

For $\tilde{\mathcal{H}}(J) = \alpha_H(J)$ this yields:

$$\alpha_H(J) = \tilde{c}^{-\frac{4}{3}}J^{\frac{4}{3}} = (\frac{3}{8\sqrt{2}})^{\frac{4}{3}}\frac{C_W^{\frac{2}{3}}C_\phi^{\frac{1}{3}}}{K^{\frac{4}{3}}(\frac{1}{\sqrt{2}})}(\frac{m^2-1}{3})^{\frac{1}{3}}J^{\frac{4}{3}}. \tag{4.10}$$

The Hamilton equations of motion are:

$$\begin{aligned}\frac{\partial\tilde{\mathcal{H}}}{\partial J} &= \dot{Q} = \frac{4}{3}\tilde{c}^{-\frac{4}{3}}J^{\frac{1}{3}} = \frac{4}{3}(\frac{3}{8\sqrt{2}})^{\frac{4}{3}}\frac{C_W^{\frac{2}{3}}C_\phi^{\frac{1}{3}}}{K(\frac{1}{\sqrt{2}})}(\frac{m^2-1}{3})^{\frac{1}{3}}J^{\frac{1}{3}}\\ -\frac{\partial\tilde{\mathcal{H}}}{\partial Q} &= \dot{J} = 0.\end{aligned}$$

The synchrotron frequency Ω_s is given by the time derivative of the angle variable Q as $\Omega_s = 2\pi\dot{Q}$ and $\dot{Q} = \frac{\partial\tilde{\mathcal{H}}}{\partial J}$. (For this choice of variables the angle variable Q takes values in the range from 0 to 1.)

$$\Omega_s = 2\pi\dot{Q} = \frac{3^{\frac{1}{3}}\pi\Omega_0}{2^{\frac{2}{3}}K(\frac{1}{\sqrt{2}})}(\frac{m^2-1}{3})^{\frac{1}{3}}(\frac{J}{\hat{J}_0})^{\frac{1}{3}}, \qquad J << \hat{J}_0, \tag{4.11}$$

where $\hat{J}_0 = 8\sqrt{\frac{C_\phi}{C_W}}$ and $\Omega_0 = \sqrt{C_W C_\phi}$ are the bucket area and the zero amplitude synchrotron oscillation frequency for the single rf system.
From the $J^{\frac{1}{3}}$-dependence of Ω_s in equation (4.11) it becomes obvious how a secondary rf system modifies the synchrotron oscillation spread of the particles of different oscillation amplitudes.
For the further calculations the undisturbed rf phase $\phi(t)$ is needed as a function of the action variable J (or of the energy variable α_H) and the angle variable Q.
The solution for a ϕ^4-potential can be found via integration (see appendix) and reads

$$\phi(t) = \hat{\phi} \quad \mathrm{cn}(4K(\frac{1}{\sqrt{2}})Q), \tag{4.12}$$

where cn means the Jacobian elliptic cosine function.

4.4.2 Direct Simulations for the Rf Stochastic Equation

The rf system with phase noise described by the Hamiltonian (4.6) and the corresponding system of stochastic differential equations (4.7) has been treated with the numerical algorithms described before.
For this rf system the first and second moments of the phase and energy deviation from the reference particle, the action and the bunch length have been computed.
The parameters are: $q = 1$, $\hat{V} = 60kV$, $h = 1100$, $m = 2$, $\eta = 5.75 \cdot 10^{-4}$, $\omega_0 = 47kHz \cdot 2\pi$, $\beta = 1.0$, $E = 40GeV$ which have been chosen to be similar to the HERA-p injection optic.
For the given parameters and $\alpha_H = 7.58 \cdot 10^{-4}eV$, the theoretical value of Ω_s is $11.18s^{-1}$.
The numerical value using the Euler scheme is $\Omega_{s,num} = 11.16s^{-1}$ and one sees that in this case even the simple Euler algorithm delivers precise results.
In the figures the results of the calculations for initial conditions corresponding to energy values between $\alpha_H = 7.58 \cdot 10^{-4}eV$ and $\alpha_H = 17.3625eV$ are shown, the latter corresponding to the separatrix, and for two different noise levels, a large one of $\sqrt{d} = 0.1$ and a small one of $\sqrt{d} = 0.0001$.
Figure 15 presents the time development of the second moment of the energy deviation with respect to the synchronous particle $< W^2 > (t)$. For the large noise level the values grow with time, while the motion under influence of the small noise stays unaffected up to 10000 turns. (The different oscillation frequencies for different initial values result from the relation (4.11).) Similar behaviour holds for the bunch length, see figure 16. Whereas for the large

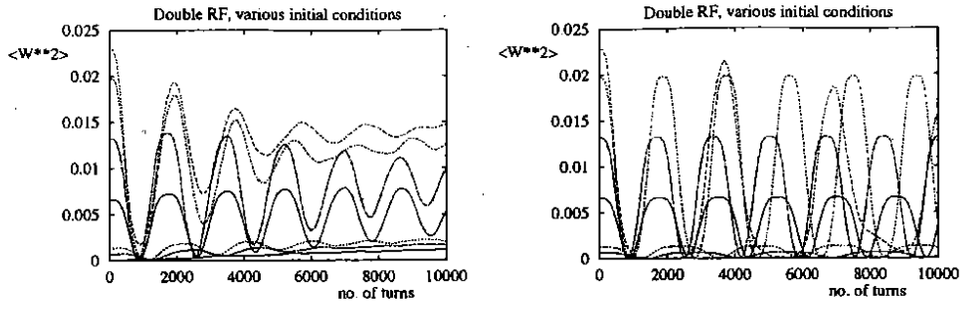

Figure 15: Second moment of the energy deviation $< W^2 >$ for several initial conditions reaching from the centre of the bucket to the separatrix for two different noise levels, a large one $\sqrt{d} = 0.1$ (left) and a small one $\sqrt{d} = 0.0001$ (right).

noise parameter the bunch length $\sigma(t)$ grows significantly within 10000 turns there is almost no growth for the small noise level. (The upper curve represents the motion for initial values on the separatrix and therefore the particle starting from there is likely to leave the rf bucket.) The following table 1 shows a comparison between the Euler and the Heun schemes for the stochastic differential equation with additive noise, simulated in the original variables ϕ and W. (The calculations were performed on an HP9000-730 workstation.)

	Euler	Heun
no. of samples	500	500
no. of turns	1000	1000
CPU time	10'40"	36'18"
no. of samples	1000	1000
no. of turns	1000	1000
CPU time	23'56"	57'56"
no. of samples	500	—
no. of turns	10^5	—
CPU time	18h24'14"	—

Table 1: Comparison of the integration schemes with respect to computing time and sample sizes. These calculations refer to the system of stochastic differential equations with additive noise in the variables ϕ and W.

4.4.3 Averaging the Fokker-Planck Equation

The stochastically driven motion in an rf system is described by the Fokker-Planck equation for the density function in the action variable J. The first method to derive the Fokker-Planck equation in the J-variable is based on an averaging method for this equation in action angle representation. It has been applied to linear problems for example in [14],[26].
The first step is to transform the Hamiltonian (4.6) to the action angle variables:

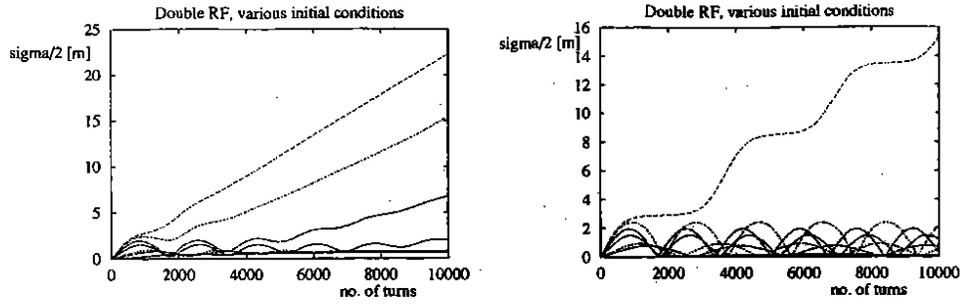

Figure 16: Time development of the (half) bunch length $\frac{1}{2}\sigma(t)$ for several initial conditions reaching from the centre of the rf bucket to the separatrix for a large noise level of $\sqrt{d} = 0.1$ (left) and a small one of $\sqrt{d} = 0.0001$ (right). The upper curves give the motion for initial values on the separatrix.

$$\mathcal{H}(\phi, W) \rightarrow \tilde{\mathcal{H}}(Q, J).$$

Making use of the relations (4.4) and (C.7) this yields with $\tilde{h} = -\phi(Q, J)$:

$$\begin{aligned} \tilde{\mathcal{H}}(Q,J) &= \alpha_H(J) + \tilde{h}(Q, \alpha_H(J)) \quad \xi(t) \\ &= c \cdot J^{\frac{4}{3}} - \sqrt{d} \quad \hat{\phi} \quad \mathrm{cn}(4KQ) \quad \xi(t), \end{aligned} \tag{4.13}$$

with sn, cn, dn meaning the Jacobi elliptic functions, $K = K(\frac{1}{\sqrt{2}})$ denoting the complete elliptic integral of modulus $\frac{1}{\sqrt{2}}$, see [30], and $c = (\frac{3C_W^{\frac{1}{2}} C_\phi^{\frac{1}{4}}}{8\sqrt{2}K(\frac{1}{\sqrt{2}})})^{\frac{4}{3}} (\frac{m^2-1}{3})^{\frac{1}{3}}$.
The Hamiltonian in Q and J representation is then:

$$\tilde{\mathcal{H}}(Q,J) = c \cdot J^{\frac{4}{3}} - \sqrt{d} 2 (\frac{3}{m^2-1})^{\frac{1}{4}} C_\phi^{-\frac{1}{4}} c^{\frac{1}{4}} J^{\frac{1}{3}} \mathrm{cn}(4KQ) \xi(t). \tag{4.14}$$

The differential equations in the new canonical conjugate coordinates read

$$\begin{aligned} \dot{Q} = \quad \frac{\partial \tilde{\mathcal{H}}}{\partial J} &= c \cdot \frac{4}{3} J^{\frac{1}{3}} - \sqrt{d} \frac{2}{3} (\frac{3}{m^2-1})^{\frac{1}{4}} C_\phi^{-\frac{1}{4}} c^{\frac{1}{4}} J^{-\frac{2}{3}} \mathrm{cn}(4KQ) \xi(t) \\ \dot{J} = -\frac{\partial \tilde{\mathcal{H}}}{\partial Q} &= -\sqrt{d} 8K (\frac{3}{m^2-1})^{\frac{1}{4}} C_\phi^{-\frac{1}{4}} c^{\frac{1}{4}} J^{\frac{1}{3}} \mathrm{sn}(4KQ) \mathrm{dn}(4KQ) \xi(t). \end{aligned} \tag{4.15}$$

This is a system of stochastic differential equations with multiplicative noise.
Rewritten in matrix form this yields:

$$\begin{aligned} \begin{pmatrix} \dot{Q} \\ \dot{J} \end{pmatrix} &= \begin{pmatrix} c \cdot \frac{4}{3} J^{\frac{1}{3}} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{d} \frac{2}{3} (\frac{3}{m^2-1})^{\frac{1}{4}} C_\phi^{-\frac{1}{4}} c^{\frac{1}{4}} J^{-\frac{2}{3}} \mathrm{cn}(4KQ) \\ 0 & -\sqrt{d} 8K (\frac{3}{m^2-1})^{\frac{1}{4}} C_\phi^{-\frac{1}{4}} c^{\frac{1}{4}} J^{\frac{1}{3}} \mathrm{sn}(4KQ) \mathrm{dn}(4KQ) \end{pmatrix} \cdot \begin{pmatrix} \xi_1(t) \\ \xi_2(t) \end{pmatrix} \\ &= \vec{A}(Q,J) + \underline{B}(Q,J) \vec{\xi}(t). \end{aligned} \tag{4.16}$$

The diffusion matrix $\underline{D}$ is computed as

$$\underline{D} = \underline{B}\underline{B}^T = \begin{pmatrix} D_{QQ} & D_{QJ} \\ D_{JQ} & D_{JJ} \end{pmatrix}$$

with

$$D_{QQ} = \frac{4}{9} d (\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{cn}^2(4KQ),$$

$$D_{QJ} = D_{JQ} = -\frac{16}{3} K d (\frac{3}{m^2-1})^{\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{sn}(4KQ)\mathrm{cn}(4KQ)\mathrm{dn}(4KQ),$$

$$D_{JJ} = 64 K^2 d (\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{\frac{2}{3}} \mathrm{sn}^2(4KQ)\mathrm{dn}^2(4KQ).$$

The Fokker-Planck equation in Q and J is, thus

$$\frac{\partial \rho}{\partial t}(Q,J,t) = -\frac{\partial}{\partial Q}(A_Q \rho) - \frac{\partial}{\partial J}(A_J \rho) + \sum_{i,k} \partial_i \partial_k D_{ik} \rho, \tag{4.17}$$

where the indices i, k denote Q and J, and

$$A_Q = a_Q + \frac{1}{2} \sum_{i,k} (\partial_i B_{Qk}) B_{ik},$$

$$A_J = a_J + \frac{1}{2} \sum_{i,k} (\partial_i B_{Jk}) B_{ik}.$$

The a_i stand for the deterministic part of the drift term and the remaining expressions give the noise induced drift.
Taking into account that some of the coefficients are zero this yields

$$\begin{aligned} \frac{\partial \rho}{\partial t} = &-\frac{\partial}{\partial Q}(c \cdot \frac{4}{3} J^{\frac{1}{3}} \rho) - \frac{1}{2}\frac{\partial}{\partial Q}[(\frac{\partial}{\partial Q} B_{QJ}) B_{QJ} \rho] - \frac{1}{2}\frac{\partial}{\partial J}[(\frac{\partial}{\partial Q} B_{JJ}) B_{QJ} \rho] \\ &-\frac{1}{2}\frac{\partial}{\partial Q}[(\frac{\partial}{\partial J} B_{QJ}) B_{JJ} \rho] - \frac{1}{2}\frac{\partial}{\partial J}[(\frac{\partial}{\partial J} B_{JJ}) B_{JJ} \rho] \\ &+\frac{1}{2}[\frac{\partial^2}{\partial Q^2} D_{QQ} \rho + 2\frac{\partial^2}{\partial Q \partial J} D_{QJ} \rho + \frac{1}{2}\frac{\partial^2}{\partial J^2} D_{JJ} \rho]. \end{aligned} \tag{4.18}$$

Inserting the corresponding expressions finally leads to

$$\begin{aligned} \frac{\partial \rho}{\partial t} = &-\frac{\partial}{\partial Q}(c \cdot \frac{4}{3} J^{\frac{1}{3}} \rho) - \frac{1}{2}\frac{\partial}{\partial Q}[-d\frac{32}{9}(\frac{3}{m^2-1})^{\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{sn}(4KQ)\mathrm{cn}(4KQ)\mathrm{dn}(4KQ)\rho] \\ &-\frac{1}{2}\frac{\partial}{\partial Q}[-d\frac{16}{9}(\frac{3}{m^2-1})^{\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{sn}(4KQ)\mathrm{cn}(4KQ)\mathrm{dn}(4KQ)\rho] \\ &-\frac{1}{2}\frac{\partial}{\partial J}[\frac{64}{3} K^2 d(\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{sn}^2(4KQ)\mathrm{dn}^2(4KQ)\rho] \\ &-\frac{1}{2}\frac{\partial}{\partial J}[\frac{32}{3} K^2 d(\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \{\mathrm{cn}^2(4KQ)\mathrm{dn}^2(4KQ) - \mathrm{sn}^2(4KQ)\mathrm{cn}^2(4KQ)\}\rho] \\ &+\frac{1}{2}[\frac{\partial^2}{\partial Q^2}\frac{4}{9} d(\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{cn}^2(4KQ)\rho] \end{aligned} \tag{4.19}$$

$$\begin{aligned} &+\frac{\partial^2}{\partial Q \partial J}[-\frac{16}{3} K d(\frac{3}{m^2-1})^{\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{-\frac{1}{3}} \mathrm{sn}(4KQ)\mathrm{cn}(4KQ)\mathrm{dn}(4KQ)\rho] \\ &+\frac{1}{2}\frac{\partial^2}{\partial J^2}[64 K^2 d(\frac{3}{m^2-1})^{-\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} J^{\frac{2}{3}} \mathrm{sn}^2(4KQ)\mathrm{dn}^2(4KQ)\rho]. \end{aligned}$$

It is assumed that the density ρ is equally distributed in the angle variable Q [13],[14] , that means the ansatz $\rho = \rho(J,t)$ is made and an averaging over the angle variable Q is performed. The Q-dependent terms in the equation are then replaced by their mean value $< \ldots >_Q = \int_0^1 \ldots dQ$. As already mentioned the angle variable Q has been normalized to values in [0,1]. Using the $4K$-periodicity of the Jacobian elliptic functions [30] one gets the averaged Fokker-Planck equation in the action variable J:

$$\frac{\partial \rho}{\partial t} = \frac{32}{3} K^2 d (\frac{3}{m^2-1})^{\frac{1}{2}} C_\phi^{-\frac{1}{2}} c^{\frac{1}{2}} \{-\frac{2}{3}\frac{\partial}{\partial J}(J^{-\frac{1}{3}} \rho) + \frac{\partial^2}{\partial J^2}(J^{\frac{2}{3}} \rho)\}. \tag{4.20}$$

Another method for deriving the rf Fokker-Planck equation, suggested in [12], is based on a perturbation theoretical treatment of the rf stochastic differential equations. A short outline of this approach is given in the appendix.

4.4.4 Perturbation Methods in the Rf Stochastic Equation

The long term motion in synchrotron phase space can be described by the following Fokker-Planck equation in J:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial J}(A_1 \rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(A_2 \rho) \tag{4.21}$$

with

$$A_1 = \frac{1}{2}\frac{\partial}{\partial J} A_2,$$

leading to the diffusion equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial J}(\frac{1}{2} A_2 \frac{\partial}{\partial J}\rho). \tag{4.22}$$

Moreover an explicit expression for the diffusion coefficient A_2 is derived via the perturbation theoretical approach:

$$A_2 = \lambda T_0 \oint p dq (h'(q))^2.$$

For the double rf system one has

$$A_2(J) = T(J) \oint p dq (h'(q))^2 = \frac{3}{4} \tilde{c}^{\frac{4}{3}} J^{-\frac{1}{3}} J 2 C_W d = \frac{3}{4} \tilde{c}^{\frac{4}{3}} 2 C_W d \quad J^{\frac{2}{3}}.$$

This yields a Fokker-Planck equation of the form

$$\frac{\partial \rho}{\partial t} = \frac{3}{8} d 2 C_W \tilde{c}^{\frac{4}{3}} \frac{\partial}{\partial J}(J^{\frac{2}{3}} \frac{\partial}{\partial J}\rho(J)). \tag{4.23}$$

Inserting the correct constants this agrees with equation (4.20). In both cases the resulting Fokker-Planck equation can be written as:

$$\frac{\partial \rho}{\partial t} = dK^{\frac{1}{3}} (\frac{3}{m^2-1})^{\frac{1}{3}} (\frac{C_W}{C_\phi})^{\frac{1}{3}} (\frac{2^8}{3})^{\frac{1}{3}} \{(-\frac{\partial}{\partial J}\frac{2}{3} J^{-\frac{1}{3}} \rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(2J^{\frac{2}{3}} \rho)\}. \tag{4.24}$$

Thus it was shown that both ways of treating the rf system, namely the perturbation theory for the system of differential equations as well as the averaging method in the diffusion equation, lead to the same Fokker-Planck equation in the action variable J.
The validity of this result has been tested with the numerical methods.

4.4.5 The Rf Stochastic Equation in the Action Variable

From the one-dimensional Fokker-Planck equation in the action J the corresponding stochastic differential equation in J is derived and the growth of the action is simulated numerically.
The Fokker-Planck equation can also be written as

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\frac{\partial}{\partial J}(\frac{2}{3}BJ^{-\frac{1}{3}}\rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(2BJ^{\frac{2}{3}}\rho) \qquad (4.25)\\ &= -\frac{\partial}{\partial J}([\frac{1}{3}BJ^{-\frac{1}{3}} + \frac{1}{2}\{\frac{\partial}{\partial J}\sqrt{2B}J^{\frac{1}{3}}\}\sqrt{2B}J^{\frac{1}{3}}]\rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(2BJ^{\frac{2}{3}}\rho),\end{aligned}$$

with

$$B = \frac{3}{8}d2C_W\tilde{c}^{\frac{4}{3}}.$$

Comparing this representation (4.25) with equation (4.21) one can easily identify the drift and diffusion coefficients A_1 and A_2. In the second expression the drift term has been divided into a "deterministic" and a "noise-induced" part.
As the diffusion coefficient is dependent on J this system involves multiplicative noise and so one has to be aware of the difference between the two calculi (Ito - Stratonovich) when setting up the stochastic differential equation or writing down the corresponding numerical approximation.
In the Stratonovich calculus one has to take into account that in the first coefficient of the Fokker-Planck equation already the noise induced drift is included. This can be seen from the second expression in equation (4.25).
The (Stratonovich-)stochastic differential equation reads:

$$\dot{J} = \frac{1}{3}BJ^{-\frac{1}{3}} + \sqrt{2B}J^{\frac{1}{3}}\xi(t). \qquad (4.26)$$

The corresponding Ito-stochastic differential equation for this system reads:

$$\dot{J} = \frac{2}{3}BJ^{-\frac{1}{3}} + \sqrt{2B}J^{\frac{1}{3}}\xi(t),$$

and for this equation the term $C(h)$ of the numerical approximation has to be calculated in the Ito way to lead to the same results.
The behaviour of $< J >$ and $< J^2 >$ has been studied using the two possibilities for calculating the moments:

- Simulate the stochastic differential equations in the original variables ϕ and W and then compute J from those by using (4.9)
- or perform simulations in the action variable J by using (4.26).

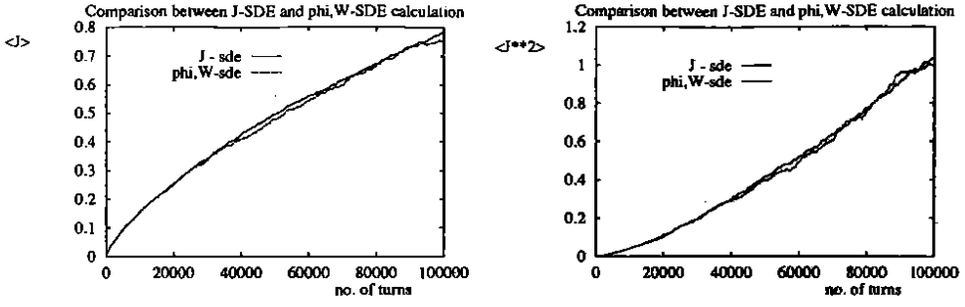

Figure 17: Synchrotron motion.First moment $< J(t) >$ and second second moment $< J^2(t) >$ vs. number of turns for $d = 10^{-2}$, 500 samples. (Error bars have been suppressed for clarity.) Long term comparison of the results from both stochastic differential equations, the two-dimensional system of equations in ϕ and W with additive noise term, and the one-dimensional multiplicative stochastic differential equation in the action variable J for the initial conditions $\phi_0 = 0.1deg$ and $W_0 = 1.0 \cdot 10^{-3}eVs$.

The parameters are the same as before, again in one case the noise level is chosen extremely large, $d = 0.01, \sqrt{d} = 0.1$, to make the effect of the growth of the action variable easily visible. For comparison a small noise parameter of $d = 10^{-8}, \sqrt{d} = 10^{-4}$ is used.
Figure 17 shows the results for the growth of the first moment $< J(t) >$ and for the second moment $< J^2(t) >$ for $d = 10^{-2}$. In these plots the ϕ, W-simulations of the stochastic differential equation are compared with the J-equation calculations (500 particles).
For the multiplicative stochastic differential equation in the action variable the modified Euler algorithm (Milsthein scheme) is used. The initial conditions have been $\phi_0 = 0.1deg$ and $W_0 = 1.0 \cdot 10^{-3}eVs$.
Both simulation results agree very well even on a long term scale of 10^5 turns. The simulation of the stochastic differential equation with multiplicative noise term (the equation in the action variable J) is by far more computing time consuming, see tables 1 and 2.

no. of samples	500	1000	500
no. of turns	1000	1000	10^5
CPU time	31'24"	1h02'38"	50h10'38"

Table 2: Computing time and sample sizes for the one-dimensional multiplicative stochastic differential equation in the action variable J.

5 The Beam-Beam Interaction

The counting rate R of events for a physical reaction in a collider is given by

$$R = \mathcal{L} \cdot \sigma. \tag{5.1}$$

The factor $\mathcal{L}$ is called luminosity. For head-on collisions $\mathcal{L}$ is given by

$$\mathcal{L} = f N_1 N_2 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} \rho_1(x,z,s-ct)\rho_2(x,z,s-ct)2cdt. \tag{5.2}$$

N_1, N_2 are the numbers of particles in bunch no. 1, resp. bunch no. 2,
ρ_i are the normalized bunch densities,
c is the velocity of light,
f designates the collision frequency of two bunches at the interaction point.
The factor 2 is a result of the fact that the bunches have a relative velocity of $2c$. For bunches with identical Gaussian transverse distributions characterized by σ_x^* and σ_z^*, equal numbers of particles in the bunches and bunch lengths short compared to the β-functions at the interaction points the formula for the luminosity reduces to

$$\mathcal{L} = \frac{N^2 f}{4\pi\sigma_x^*\sigma_z^*}, \tag{5.3}$$

where $\sigma_{x,z}^*$ are given by $\sigma_y^* = \sqrt{\epsilon_{y0}\beta_y^* + (\eta_y^*\sigma_\delta)^2}$, ϵ_{y0} is the emittance, η^* means the dispersion at the interaction point, and σ_δ is the energy spread.
The "natural" beam sizes are determined by the optics, mainly the linear focusing, and the effects of synchrotron radiation. As the number of particles within the beams is increased above a certain level, the transverse cross sections begin to enlarge beyond their "natural" sizes. The transverse growth of the beams implies that the luminosity grows only proportional to N rather than to N^2. Moreover the beam lifetimes are reduced below their values for Gaussian transverse distributions. These limitations are mainly caused by the beam-beam interaction.
Here the weak-strong model of the beam-beam effect is considered which describes the interaction of a single particle (weak beam) with the electromagnetic field of a charge distribution (strong beam). If this charge distribution is assumed to be Gaussian in all three spatial dimensions and in addition the bunch length is assumed to be short compared to the betafunctions at the interaction points, the corresponding potential is

$$U(x,z,s) = \frac{N_b r_e}{\gamma} \int_0^\infty \frac{1 - \exp(\frac{-x^2}{2\sigma_x^2+q} + \frac{-z^2}{2\sigma_z^2+q})}{(2\sigma_x^2+q)^{\frac{1}{2}}(2\sigma_z^2+q)^{\frac{1}{2}}}\ dq \quad \delta_p(s-s_0), \tag{5.4}$$

and the impulse like kicks felt by a particle crossing the counterrotating bunch are given by

$$\begin{aligned}
\Delta x' &= -\frac{\partial U}{\partial x} = -2\frac{N_b r_e}{\gamma} x \int_0^\infty \frac{\exp(\frac{-x^2}{2\sigma_x^2+q} + \frac{-z^2}{2\sigma_z^2+q})}{(2\sigma_x^2+q)^{\frac{3}{2}}(2\sigma_z^2+q)^{\frac{1}{2}}}\ dq \\
\Delta z' &= -\frac{\partial U}{\partial z} = -2\frac{N_b r_e}{\gamma} z \int_0^\infty \frac{\exp(\frac{-x^2}{2\sigma_x^2+q} + \frac{-z^2}{2\sigma_z^2+q})}{(2\sigma_x^2+q)^{\frac{1}{2}}(2\sigma_z^2+q)^{\frac{3}{2}}}\ dq.
\end{aligned} \tag{5.5}$$

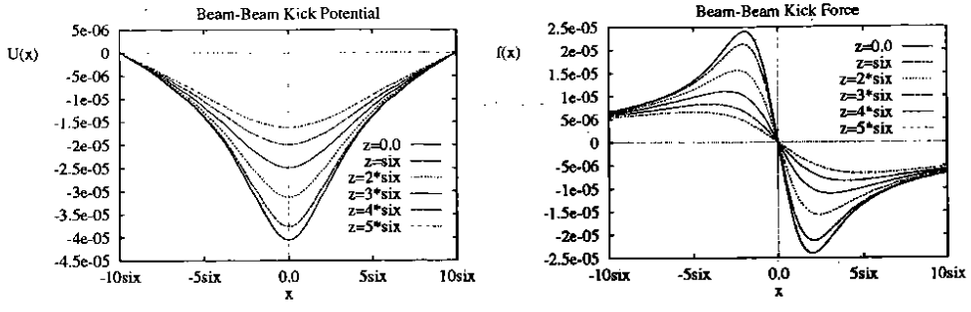

Figure 18: Beam-beam kick potential and force as a function of displacement.

In the central region of the beam (up to about 1 σ) the kick is similar to a quadrupole kick, but outside the central region the deflection is highly nonlinear and in addition couples horizontal and vertical motion.
These features are the major reason for the growth of the transverse beam sizes.
In the following the case of round beams, $\sigma_x = \sigma_z = \sigma$, will be studied, so the kicks are given by

$$\begin{aligned}
\Delta x' &= -2\frac{N_b r_e}{\gamma}\frac{x}{r^2}(1 - \exp(-\frac{r^2}{2\sigma^2})) \\
\Delta z' &= -2\frac{N_b r_e}{\gamma}\frac{z}{r^2}(1 - \exp(-\frac{r^2}{2\sigma^2})) \\
r^2 &= x^2 + y^2.
\end{aligned} \tag{5.6}$$

The quantity

$$\xi_y = \frac{N_b r_e}{\gamma}\frac{\beta_{0y}^*}{2\pi\sigma_y(\sigma_x+\sigma_z)}$$

is called the beam-beam parameter and characterizes the strength of the interaction. In linear approximation ξ_y is the maximum tune shift due to the perturbing beam-beam kick, $\xi_y = \Delta Q_{y,lin}$.
The following chapters are concerned with the description of electron dynamics including damping, noise and the beam-beam interaction.

6 Discrete Models

6.1 Introduction

In the preceding sections "continuous" algorithms have been applied to the direct numerical simulation of stochastic differential equations. Now discrete models will be used for the calculation of stochastic beam dynamics. Instead of considering functions of the form $\dot{x}(t) = f(x,t)$ one rather uses maps to describe the motion, i.e. transformations of the form $x_{n+1} = Mx_n$.
This formulation is most appropriate for accelerator physics because if once the transfer map M for the whole periodic ring structure has been computed, the motion of a particle can be followed by iteratively applying the operator M to an initial value x_0 and the successively generated vectors x_n. Thus one computes the particle coordinates at the discrete time steps $t_0 + kT$, $k = 0, 1, 2, ...$ with T being the revolution time. If the parametrization is changed from time t to arc length s, these times correspond to successive passings of a reference point s_0 in the ring, $s_0 + kL$, where L is the storage ring circumference. The transfer map M can be computed from the ring optics. First one sets up the transfer maps for each ring element and obtains the transformation through the whole ring as a concatenation of the single maps. In the linear case (when only drift spaces, dipoles and quadrupoles are considered) or having applied certain approximations, this reduces to matrix multiplication of the transfer matrices of the single elements or structures [2],[31],[32].
A quantity of fundamental interest is the phase space density function $\rho(\vec{x}, t)$ of an electron storage ring and particularly its time evolution. This function contains all information about the size and the lifetime of a stored electron beam. The usual method to compute density functions and their dynamics is to track a large number of particles and to follow their motion for a large number of turns. This makes numerical calculations for particle densities quite computing time consuming.
A new kind of algorithm will be described which uses a map for the particle tracking and moreover is based on the generation of a stochastic mapping operator for the density function $\rho(\vec{x}, t)$. The main ideas of this approach have been suggested by A. Gerasimov [17].
The density $\rho(\vec{x}, t)$ at the discrete successive time steps t_n is calculated by repeatedly applying a time propagator A:

$$\rho(t_n) = A\rho(t_{n-1})$$
$$\rho_n = A\rho_{n-1} = ... = A^n\rho_0.$$

Thus the time evolution of the density function of a dissipative, stochastically excited nonlinear system is modelled by approximating it by a homogeneous discrete Markov process (i.e. the propagator A is taken to be time independent). Applying the model of discrete Markov processes means that one has performed a second discretization, namely a spatial one, partitioning the phase space into a set of discrete states. The calculations are done on a grid, where each state in the phase space is attached to a cell of the grid.
The method will be demonstrated for simple examples of the one-dimensional beam-beam interaction. Later on more complicated systems will be considered, e.g. two-dimensional beam-beam simulations. The calculations are performed with a map describing a model electron storage ring with damping, noise excitation and a beam-beam kick at one interaction point in the ring.
The linear part of the map reads

$$\begin{pmatrix} x_{s_{ip}+L-\ell} \\ p_{s_{ip}+L-\ell} \end{pmatrix} = \begin{pmatrix} \cos 2\pi Q & \beta^* \sin 2\pi Q \\ -\frac{1}{\beta^*}\sin 2\pi Q & \cos 2\pi Q \end{pmatrix} \begin{pmatrix} x_{s_{ip}+\ell} \\ p_{s_{ip}+\ell} \end{pmatrix}, \tag{6.1}$$

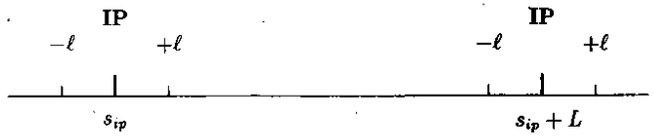

Figure 19: Model structure of the map given by the equations (6.1)and (6.2). The transformation between the interaction points (IP) is linear. At the interaction points the beam-beam kick, damping and noise are applied.

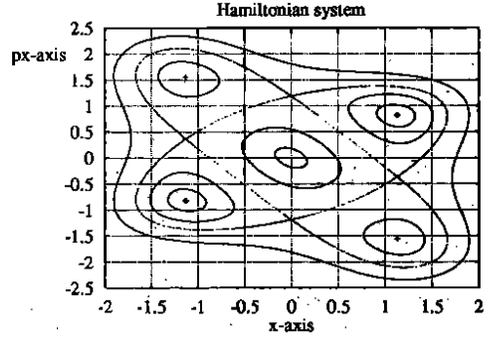

Figure 20: Phase space structure of the Hamiltonian system with beam-beam kick of $\xi = 0.07$.

and the transformation around the interaction point is

$$\begin{aligned} x_{s_{ip}+\ell} &= x_{s_{ip}-\ell} \\ p_{s_{ip}+\ell} &= p_{s_{ip}-\ell} - \alpha p_{s_{ip}-\ell} - 8\pi\xi\frac{\sigma^2}{\beta^*}\left(\frac{1-e^{(-x^2_{s_{ip}-\ell}/2)}}{x_{s_{ip}-\ell}}\right) + \sigma\frac{1}{\beta^*}\sqrt{2\alpha}\eta. \end{aligned} \tag{6.2}$$

α means the damping parameter which is the inverse of the damping time τ_D in number of turns and η denotes here a random number with mean zero and variance 1.
For the parameter values $Q = 3.7, \sigma = \beta^* = 1.0$ and $\xi = 0.07$ the phase space trajectories of the corresponding deterministic and Hamiltonian system ($\alpha = 0$, no dissipation and no external noise forces) are shown in figure 20.
The resonance structure of the system is clearly visible in the picture. One recognizes four stable and four unstable fixed points.
The motion on the separatrix curve is highly irregular.
In general for beam-beam interaction without crossing angle and no offsets between the counterrotating bunches only even order resonances appear because of the symmetry of the potential function. If one admits offsets at the interaction points also resonances of odd order occur, see e.g. [34],[35].

6.2 The Theory of Stochastic Matrices

To compute the time propagator for the density on phase space, the phase space is divided into a set of discrete states $I = \{1, ..., N^2\}$. This means that the $x - p_x$-space is covered with a grid, each cell representing a state that can be reached by a particle during its motion. The

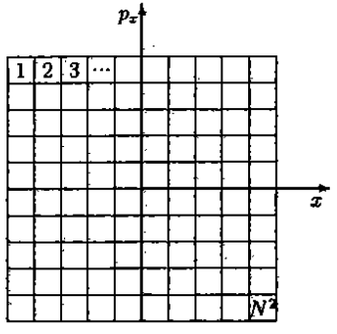

Figure 21: The phase space grid.

fluctuating forces are assumed to be totally uncorrelated, therefore the resulting process has the Markov property.
Consequently its time evolution is fully determined by an initial density ρ_0 and a two-time transition probability $\rho_{\Delta t} = \rho(x_{n+1} = j | x_n = i)$ which means the probability of a particle for being in state j at time $n+1$ after having been in state i at time n. If the transition probabilities are independent of n (i.e. of time), the process is called a homogeneous Markov chain. For arbitrary states $i_0, i_1, ..., i_n$ one gets for the joint probability of the process variables

$$\rho(x_n = i_n, x_{n-1} = i_{n-1}, ..., x_0 = i_0) = \rho_{i_{n-1} i_n} \cdot ... \cdot \rho_{i_0 i_1} \rho_{i_0}. \tag{6.3}$$

To compute a path in time one simply multiplies the single transition probabilities.
From the definition of the ρ_{ij} as transition probabilities follows

$$\begin{aligned} \rho_{ij} &\geq 0 \quad \forall i, j \in I \\ \sum_{j \in I} \rho_{ij} &= 1 \quad \forall i \in I. \end{aligned} \tag{6.4}$$

For a transition from state i to state k in n steps $\rho_{ik}^{(n)} = \rho(x_n = k | x_0 = i)$ the following relation holds

$$\rho_{ik}^{(n)} = \sum_{j \in I} \rho_{jk}^{(n-1)} \rho_{ij} = \sum_{j \in I} \rho_{jk} \rho_{ij}^{(n-1)}. \tag{6.5}$$

Let the particle start in state i. After one step it will be in the state j with the probability ρ_{ij}, and, starting from there, in $n-1$ steps reach state k, with the probability $\rho_{jk}^{(n-1)}$.
Following the "formula for the total probability" (see e.g.[36]) one has to sum over all possible states in between these two states.

For $n = 2$ this yields

$$\rho_{ik}^{(2)} = \sum_{j \in I} \rho_{jk} \rho_{ij}. \tag{6.6}$$

Consequently one arranges the probabilities for all possible transitions, ρ_{ij}, as a matrix A and gets $A_{ij} = \rho_{ij}$. This quadratic $|I| \times |I|$ matrix is called stochastic matrix. The elements of this matrix fulfil the relation (6.4).
By induction it can be shown that the n-fold transition probabilities $\rho_{ij}^{(n)}$ are just the elements of the n^{th} power of A.

6.3 Algorithm for the Density Calculation

The central task of the algorithm is the calculation of the stochastic matrix A which forms the time propagator for the density function ρ. This matrix is computed as the matrix of relative frequencies for all possible transitions between any two cells of the phase space grid. The first step of the algorithm consists in generating the $N \times N$ grid on the phase space. Each point (x, p_x) is assigned an integer number from the set $I = \{1, 2, ..., N^2\}$ denoting its state.

The Tracking Section of the Algorithm

In each cell a certain number of particles is started and their path is followed by iterative application of the one turn map. The trajectory of every particle is registered in an array and an initial cell and a corresponding final cell are recorded. A particle trajectory is thus defined by the sequence of cells on the phase space grid. The tracking is performed in the $x - p_x$-phase space. After each step the state i corresponding to the $x - p_x$-coordinates is calculated. All (x, p_x)-values inside a cell are assigned to the same state i. The time parameter n_t for the tracking, that is the number of turns for which the particles are followed, depends on the damping time of the system. A good choice for this parameter is half of the damping time ($n_t = \frac{1}{2}\tau_D$). A condition for the cell size can be derived from the diffusion time along the grid. To have sufficiently good resolution the cell width Δx has to be smaller than the rms diffusion length σ.
The central part of the program is the determination of the propagator matrix. During the tracking the "transition pairs" of initial state i_{in} and final state i_{fin} are found, and for each such transition its relative frequency is noted. The frequencies for the transition pairs are computed after the particle has already been circulating for a certain number of turns ($= n_{tm} = \frac{1}{2}n_t$).
This means that the particles are followed for n_t turns and for every particle which has been around for more than n_{tm} turns its final state i_{fin} and the corresponding initial state i_{in} which has been recorded in an array, are noted.
In this way one performs an averaging over these turns $n_t - n_{tm}$, see figure 22.
The time parameter n_{tm} over which the corresponding final and initial cells are computed, is identical with the time step Δt for the time evolution of the density function: $\Delta t = n_{tm}$.

Computation of the Transition Matrix

Each time a state i_{in} appears as initial cell, a counter is incremented, i.e. the number of particles having actually started there is stored. This counter is used for the normalization. All the grid cells that are reached from this particular cell i_{in} as final cells i_{fin} are recorded.
So for every cell i_{in} there exists a list of corresponding final cells i_{fin}. Consequently, when a new transition pair is found this list will be searched and if the actual final cell i_{fin} is found

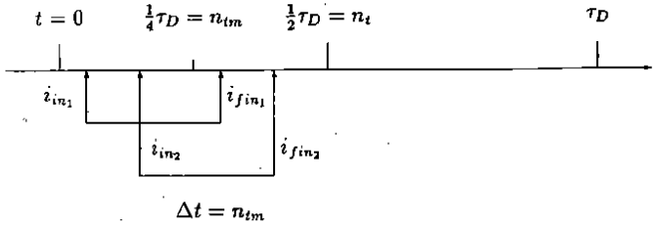

Figure 22: Time steps for calculating the transition pairs and performing the averaging.

there, a counter $a_{i_{in},i_{fin}}$ is incremented, otherwise the element i_{fin} is added to the existing list. This second counter $a_{i_{in},i_{fin}}$ records the occurrence of the event "transition between states $i_{in} \rightarrow i_{fin}$" and counts its frequency. Dividing a_{ij} by the number of particles starting in state i, the element A_{ij} of the propagator matrix is computed as the relative frequency of the transitions $i \rightarrow j$. By covering the whole phase space grid in the algorithm, the $N^2 \times N^2$ matrix A_{ij} is constructed. Due to the fact that the majority of cells is not connected by transitions, this matrix contains many zeros, i.e. it is a sparse matrix. This fortunate property is taken advance of in the way the matrix elements are stored and addressed. The addressing is done indirectly via index arrays and most of the vanishing elements are not registered in the arrays. This sparseness of the matrix A is used in the second part of the program where the time evolution of the density function ρ is obtained via multiplication with the transition matrix A. This multiplication consequently can be done very computing time efficient. Unfortunately, Fortran77 is not adapted to variable array sizes and therefore one has to know the maximum array dimensions before starting the program or insert for safety the maximum system sizes as array bounds. This leads to strong limitations especially for higher-dimensional cases. By using a C code equivalent to the Fortran code which allocates the necessary storage at run time of the program this restriction can be avoided.

Time Scales for the Averaging

Important parameters of the algorithm are the time steps for calculating the transitions $i \rightarrow j$.

- The time interval Δt should be not too large, otherwise the tails of the distribution will be lost.
- On the other hand, if Δt is too small, the dynamics of the system, especially the evolution towards a "relaxed distribution function" [37] corresponding to a stationary steady state for systems in equilibrium, cannot be represented.

The relevant parameter for the computation of the transition matrix A is the damping time τ_D. Damping and the nonlinear perturbations are small in the considered systems, and thus the time parameters can be estimated by considering the linear case. As a picture of the stochastic motion on the phase space grid one can have in mind an initial δ-peak distribution

starting in a particular cell on the grid which evolves in time under the influence of drift (deterministic part of the motion) and diffusion (spreading and widening of the distribution due to fluctuating forces), see figure 23. For an estimate of the time parameters the evolution of the mean value is approximated by a simple one-dimensional linear process. The condition is imposed that starting from an arbitrary point x_0 on the grid, after the tracking time t_{tr} the initial δ-distribution shall not have become wider than about one σ:

$$< x_0 > - < x(t_{tr}) > \simeq \frac{1}{2}\sigma(t_{tr}).$$

The temporal evolution of the mean is given by the expression

$$< x(t) > = < x_0 > e^{-\alpha t}$$

and the spreading of σ is approximated by $\sigma(t) \simeq \sqrt{dt}$ where $d = 2\alpha$.
For small damping and $t_{tr} \leq \tau_D$ this can be written as

$$\begin{aligned} < x_0 > (1 - e^{-\alpha t_{tr}}) &\simeq \frac{1}{2}\sqrt{dt_{tr}} \\ < x_0 > (1 - \{1 - \alpha t_{tr}\}) &\simeq \frac{1}{2}\sqrt{dt_{tr}} \end{aligned} \tag{6.7}$$

This leads to an approximate expression for the tracking time:

$$t_{tr} \simeq \frac{1}{< x_0 >^2}\frac{1}{2}\frac{1}{\alpha}.$$

Neglecting the dependence on the initial value one gets the following estimate

$$t_{tr} \simeq \frac{1}{2}\tau_D. \tag{6.8}$$

Calculation of the Density Function

The time propagator matrix A as computed in the first step of the program is then successively applied to an initial density ρ_0 and the temporal evolution of ρ is modelled:

$$\begin{aligned} \rho_1 &= A\rho_0 \\ &\vdots \\ \rho_n &= A^n\rho_0 = A\rho_{n-1}. \end{aligned} \tag{6.9}$$

In this way one has constructed a "numerical Markov chain" for modelling the dynamics of the phase space density $\rho(x, p_x, t)$.
The time step between the successive density functions $\rho_n = \rho(t)$ and $\rho_{n+1} = \rho(t + \Delta t)$ equals the time step Δt which has been used to compute the transition matrix A.
The substantial assumption in this model is that processes with s-dependent coefficients can be approximated by a homogeneous Markov chain, i.e. the time propagator is taken to be time independent and constant after once it has been calculated. These approximations are justified since for the investigated systems the perturbing force is weak (small beam-beam parameter). Moreover the time parameters for the calculation of the transition matrix have to be chosen to represent the dynamics of the system as well as possible.
In the next sections some simple example systems are used to demonstrate the advantages and limitations of the algorithm.

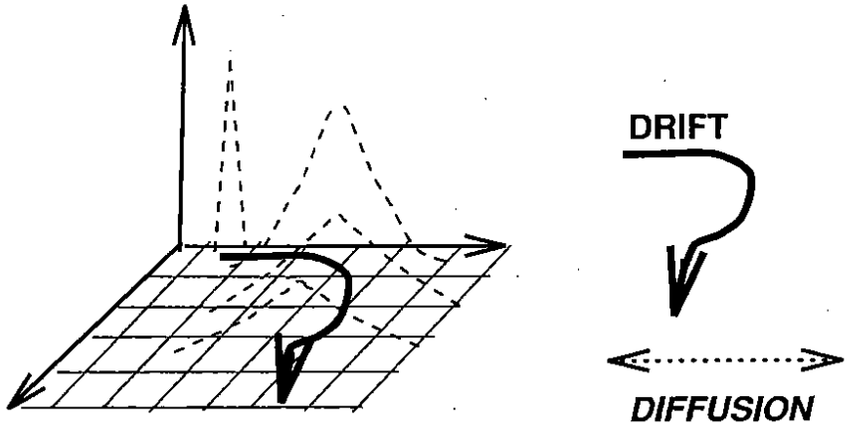

Figure 23: Schematic view of the evolution of a δ-localized initial distribution undergoing deterministing motion due to the drift terms and diffusion resulting from the stochastic excitation.

6.4 One-dimensional Calculations with the Mapping Algorithm

In the following certain properties of the matrix algorithm for several example systems will be demonstrated. The phase space has been partitioned into an $N \times N$ grid with $N = 30$. It is shown how this method can be used for fast lifetime calculations and consequently how frequency scans for this purpose can be done very quickly. The density mapping method is compared to direct multiparticle tracking with the same parameters and the same grid structure, computing and following particle trajectories exactly the given number of turns. The density is given as the relative frequency of finding the particle in a particular state i_0 after the prescribed number of turns.

Simple Beam-Beam Model

Density calculations for the one-dimensional beam-beam interaction which has already been discussed in chapter 4 are presented as a first example. The fractional part of the Q-value is $Q = 0.7$, the frequency of the corresponding unperturbed linear system (harmonic oscillator) is normalized to 1, the damping constant α has the value $\alpha = 0.001$, corresponding to a damping time τ_D of 1000 turns and the noise strength is chosen to be $d = 2\alpha$.
The particle coordinate and momentum are normalized to the unperturbed equilibrium beam sizes σ_x and σ_p. These can be calculated by means of the equations (4.1).
Strong aperture limitations are imposed: absorbing boundaries at $\pm 3\sigma$ in either variable. Absorbing boundary conditions mean that particles exceeding the aperture limit are lost, the density function vanishes at the boundaries. This of course leads to rapid decrease of the density with time, see figure 24 .
A relatively large beam-beam parameter of $\xi = 0.07$ is taken to make the influence of the perturbing nonlinearity clearly visible. The time parameters for the matrix method are: $n_t = 750$ turns, $n_{tm} = 300$ turns, the number of starting particles per grid cell is 200 for the direct tracking as well as for the mapping method. The projection of the function $\rho(x, p_x)$ on the

x-axis is shown in figure 25 . Both curves, one generated by the mapping method and the other one computed via direct tracking, agree very well. In the first picture of figure 25 the initial particle density is shown for comparison. The computing time needed for these calculations (on an HP9000-730) was about 1 minute for the mapping algorithm whereas the direct tracking took 2 hours and 17 minutes which means that the direct tracking was slower by a factor of 137.

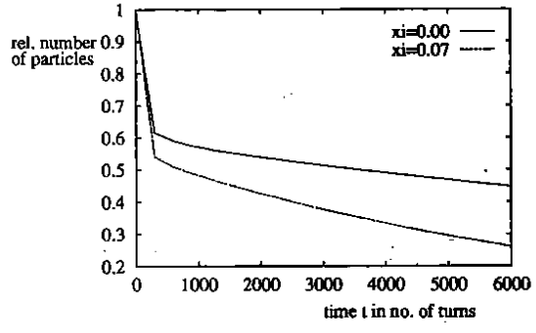

Figure 24: Strong decrease of the density function due to the restrictive aperture limitations. Shown is the relative number of particles which stay within the boundary for $\xi = 0.00$ (upper curve) and $\xi = 0.07$ (lower curve).

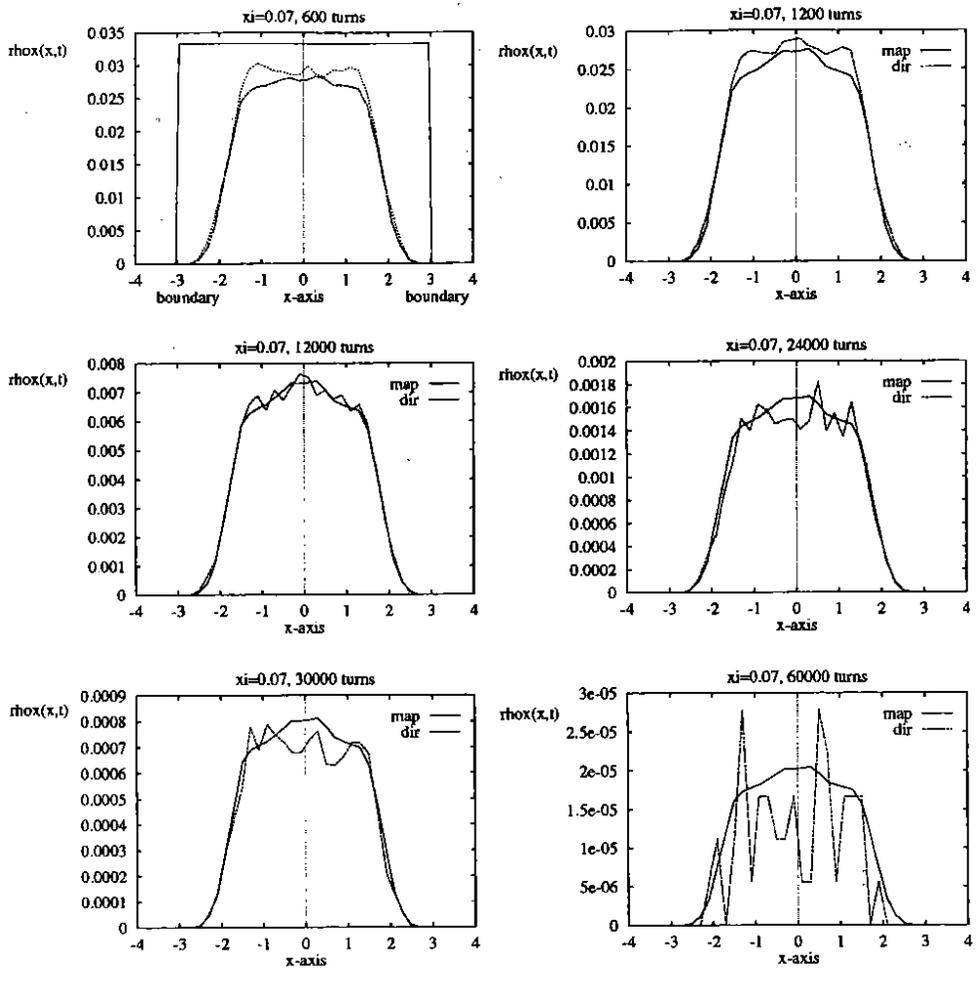

Figure 25: Time development of the density function (x-axis projection) of the first example system with $Q = .7$ and $\xi = 0.07$, calculated with the density mapping algorithm and direct tracking. The last picture clearly shows how the density has been reduced due to the restrictive absorbing boundary conditions. The loss of particles has a strong affect on the direct tracking statistics because in this technique the density is calculated by counting the particle number at the particular time steps.

For comparison the density of the unperturbed system $\xi = 0.00$ is also calculated, again with "direct tracking" and the matrix formalism. In this case the form of the resulting "relaxed" density function is known to be Gaussian. Nevertheless, due to the absorbing boundary conditions, no stationary solution exists, there is a constant loss of particles from the distribution. The density cannot preserve its normalization to unity because of this permanent outwards flux of particles. As the initial distribution is homogeneous, i.e. the particles are distributed uniformly all over the phase space at the beginning of the calculations, the normalization is destroyed soon by the loss of the particles near the aperture limit. This influence of the "scraping" at the boundary is clearly seen in figure 24 which depicts the loss of particles from inside the aperture for the first 6000 turns around the storage ring. After the density has been reduced near the boundaries and also concentrated towards the center because of the focusing forces and the damping, the initially high loss rate is seen to decrease strongly and stay constant then. However, due to the stochastic diffusion term there is still a nonvanishing flux of particles outwards, leading to a reduction of the density function. The high loss rate during the first turns is strongly dependent on the initial density ρ_0 at $t = 0$, whereas the loss rate after the "relaxation" is independent of the initial distribution.

Figure 26 shows the marginal densities for the unperturbed systems. In the first plots a Gaussian density which has been reduced to 12% of its initial normalization is shown for comparison. The first two pictures again contain also the initial density distribution. The expected Gaussian form of the density of the unperturbed system is clearly visible, also the good agreement between direct tracking and the matrix algorithm. For these calculations the same time parameters and particle numbers as before have been chosen.

The computations took 30 seconds of computing time for the mapping and 9 hours and 38 minutes of CPU time for the direct tracking method, the latter one being slower this time by a factor of 1155.

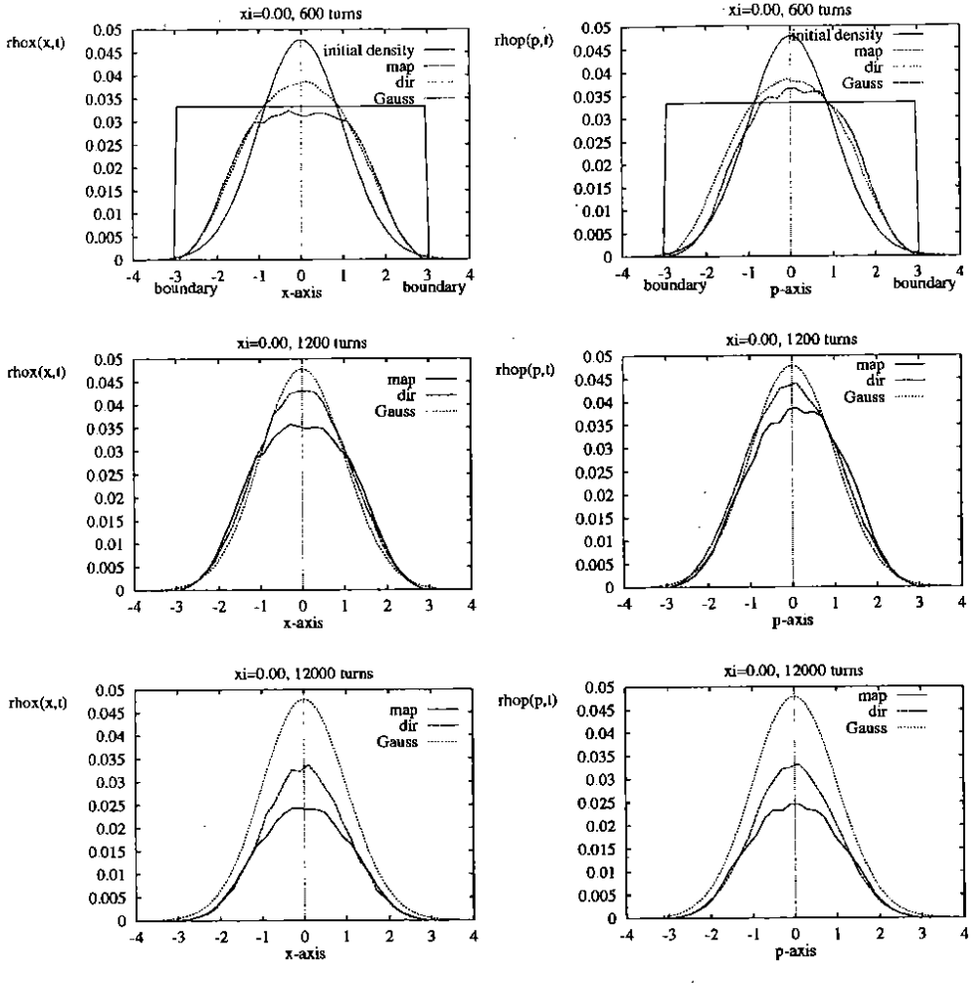

Figure 26: Time development of the density function of the unperturbed (linear) example system with $Q = .7$ and $\xi = 0.00$, calculated with the density mapping algorithm and direct tracking. On the left hand side the x-axis projection of $\rho(x,p)$ is given and on the right hand side the p-axis projection. For comparison a Gaussian density that has been reduced to 12% of its initial normalization is shown. The first two pictures also contain the initial homogeneous density.

Determination of Lifetimes

Particles exceeding the given limiting coordinates or momenta are lost rapidly. The density function vanishes at the absorbing boundaries which are for example given by the walls of the vacuum chamber of the beam pipe or some scraping devices in the beam line. The diffusion process due to the stochastic forces continually replenishes the regions near the boundary and thus there is a constant flux of particles out of the core of the distribution (see figure 24).

The lifetime τ_{life} is defined as the inverse of the exponential decay rate of the density, i.e. the time in which the total number of particles in the beam has dropped by a factor of $\frac{1}{e}$.
In the approximation of small damping the lifetime τ_{life} can be estimated from the Kramers formula [16]:

$$\tau_{life} = \frac{\epsilon}{\alpha y_{max}^2} e^{\frac{y_{max}^2}{2\epsilon}}, \tag{6.10}$$

where here ϵ is given by the mean value $\epsilon = \frac{<x^2>+<p_x^2>}{2}$ (the mean of the total action of the system) and y_{max} is determined by the physical aperture A of the vacuum chamber (in multiples of σ). For the investigated example system the aperture is calculated to $y_{max}^2 = (\frac{A^2}{\beta})_{min} = 9$.
In the used normalization the value for the unperturbed system is $\epsilon = 1$, leading to an estimated lifetime of about 10^4 turns.
According to chapter 4 the mean of the total action of the system with beam-beam kick of the strength $\xi = 0.07$ has the value $\epsilon = 1.6$. This gives an approximate lifetime of some 10^3 turns.
The lifetime calculations with the density mapping method have been performed in the following way. A measure for the total number of particles is obtained by summing the density of all cells of the phase space grid. The loss rate consequently is the decrease of this summed density ρ_{sum}. With the ansatz

$$\rho_{sum}(t) = \rho_{sum}(0)e^{-\lambda t} \tag{6.11}$$

the lifetime is $\tau_{life} = \frac{1}{\lambda}$.

The lifetimes in the direct multiparticle tracking are computed by counting the number of residual particles within the boundary at the particular time steps. The decrease of this number determines the loss rate and the lifetime is defined as before.
For $\xi = 0.00$ one obtains:

- with direct tracking: $\tau_{life} = 3 \cdot 10^4$ turns
- with the mapping technique: $\tau_{life} = 2.2 \cdot 10^4$ turns,

and for $\xi = 0.07$:

- with direct tracking: $\tau_{life} = 7.5 \cdot 10^3$ turns
- with the mapping technique: $\tau_{life} = 8 \cdot 10^3$ turns.

In figure 27 the negative logarithm of the summed densities is shown for the investigated system in the perturbed and unperturbed case, calculated with the mapping algorithm and direct tracking. The inverse value of the slope of the curves gives the lifetimes. The results agree with the analytical computation of τ_{life} with the formula (6.10).

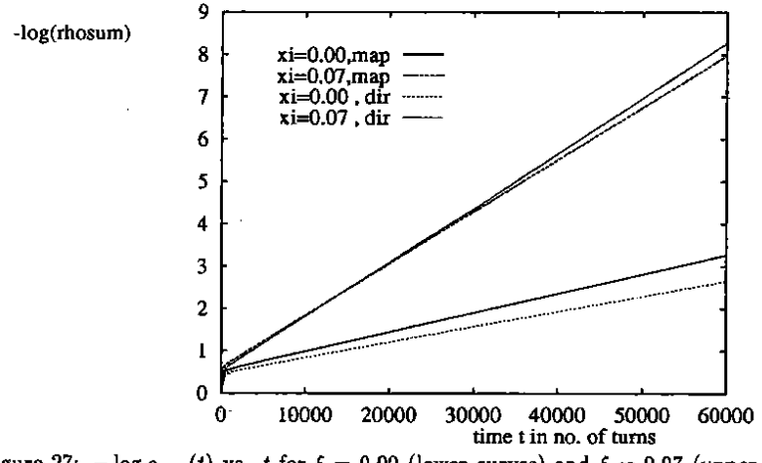

Figure 27: $-\log \rho_{sum}(t)$ vs. t for $\xi = 0.00$ (lower curves) and $\xi = 0.07$ (upper curves). The inverse values of the slopes of the curves define the lifetime. Shown is a comparison between direct multiparticle tracking and the mapping algorithm for the example system with $Q = .7$.

Beam-Beam Model with HERA-like Parameters

Here a one-dimensional model is investigated with parameters similar to the HERA-e ring. The damping time τ_D is taken to be 400 turns, corresponding to a damping constant of $\alpha = 0.0025$. The unperturbed beam size of the weak beam is $\sigma = 2 \cdot 10^{-4} m$, and the size of the counterrotating beam is $\sigma_{cr} = 2.5 \cdot 10^{-4} m$. The beam-beam parameter is chosen to be $\xi = 0.017$. (These values do not correspond to the actual HERA optic but have been taken from an earlier description of a HERA lattice [38].) Absorbing boundary conditions are imposed at $\pm 5\sigma_x$ for the x-coordinate and at $\pm 5\sigma_p$ for the p_x-coordinate. This is again an artificially strong restriction.
The parameters for the mapping algorithm are the following:
The phase space is partitioned into a 30×30 grid, the number of particles starting per cell is 200, the tracking is performed over $n_t = 200$ turns and the time for the averaging, i.e. the time over which the transition pairs are calculated, is $n_{tm} = 100$ turns.
Consequently, 100 applications of the propagator matrix in the algorithm correspond to tracking of 10^4 turns.
In figure 28 the time evolution of the density function is shown, computed for a Q-value of $Q = 63.12$. The shape of the density clearly represents the resonance structure of the system. In the figure the 8-resonance can be recognized.

Increasing the boundary limits from $\pm 5\sigma$ to $\pm 6\sigma$ leads to a clear increase in τ_{life}.
The results of a betatron frequency scan from $Q = 63.10$ to $Q = 63.25$ are presented in table 3.

Q-value:	63.10	63.11	63.12	63.13	63.14	63.15	63.16	63.17
$\pm 5\sigma$	$6 \cdot 10^6$	$9.8 \cdot 10^5$	$7.5 \cdot 10^5$	$8 \cdot 10^6$	$7.6 \cdot 10^6$	$5.5 \cdot 10^6$	$3 \cdot 10^5$	$3.3 \cdot 10^6$
$\pm 6\sigma$	$5.4 \cdot 10^8$	$4.6 \cdot 10^5$	$1.1 \cdot 10^8$	$1.6 \cdot 10^8$	$4.4 \cdot 10^8$	$5.3 \cdot 10^8$	$1.1 \cdot 10^7$	$4.4 \cdot 10^8$

Q-value:	63.18	63.19	63.20	63.21	63.22	63.23	63.24	63.25
$\pm 5\sigma$	$9.6 \cdot 10^6$	$7.8 \cdot 10^6$	$2.5 \cdot 10^6$	$1.5 \cdot 10^7$	$9 \cdot 10^7$	$1.8 \cdot 10^7$	$1.1 \cdot 10^7$	$2 \cdot 10^5$
$\pm 6\sigma$	$1.7 \cdot 10^8$	$6.3 \cdot 10^8$	$3.3 \cdot 10^8$	$1.2 \cdot 10^8$	$1.8 \cdot 10^8$	$5.8 \cdot 10^8$	$1.4 \cdot 10^8$	$5.8 \cdot 10^8$

Table 3: Lifetimes (in number of turns) dependent on the Q-value for two different boundaries for a one-dimensional beam-beam model with $\xi = 0.017$.

The computing time needed for these calculations has been about 14 to 25 CPU seconds.

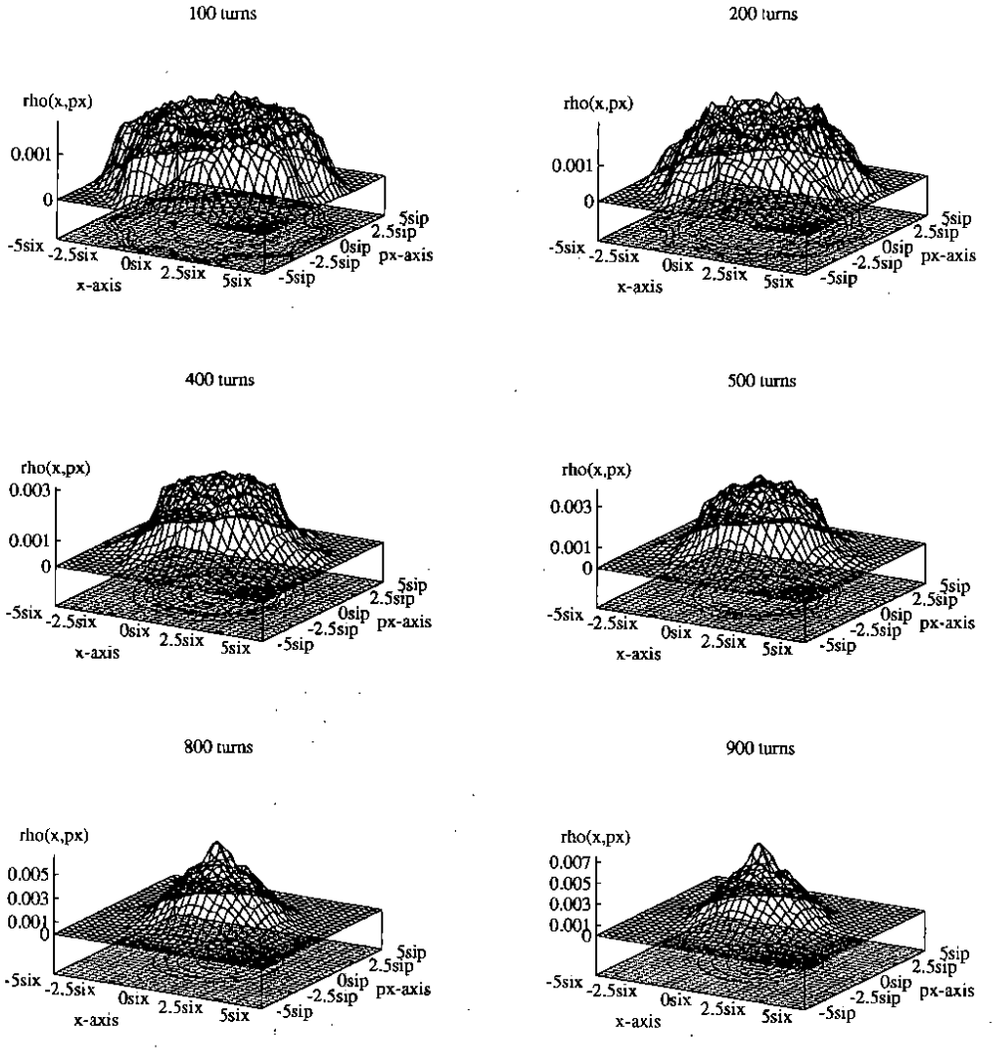

Figure 28: Time development of the density function $\rho(x,p)$ for $Q = .12$ and $\xi = 0.017$ up to 1000 turns. The initial density has been chosen to be homogeneous on the $x-p$-phase space grid.

6.5 The Macrostate Technique

The computing time and the storage requirements can be reduced by performing the calculations on larger structures of the phase space which A. Gerasimov called "macrostates" [17].
Computing the elements a_{ij} of the transition matrix between any two elements of a large phase space grid one very soon hits the limits imposed by the storage capacity and CPU time. For a two-dimensional simulation the transition probabilities on an $N \times N \times N \times N$ grid have to be calculated which means a transition matrix of $N^4 \times N^4$ elements.
The time propagator for larger units of the phase space, the "macrostates", can be computed with much reduced effort.
These macrostates have to be chosen in such a way as to represent the whole phase space adequately.

Partition of the Phase Space

In accelerator physics one is generally concerned with weakly damped, low noise systems. Damping times of the order of 10^2 to 10^3 turns are common. For the first few revolutions around the storage ring the influence of noise and damping is small. Therefore the behaviour of the system is nearly that of the corresponding Hamiltonian one. This feature is made use of when generating the partition of the phase space that represents its dynamic structure suitably.
The macrostate algorithm begins like the mapping algorithm with the generation of a fine grid of "microstates" on the phase space. In each single state or cell of this grid a test particle is started and its trajectory is followed for a certain number of turns nt_{part}.
nt_{part} is a time parameter which is considerably shorter than the damping time τ_D ($nt_{part} \simeq 0.03 - 0.10 \cdot \tau_D$ for one-dimensional systems). The larger the tracking time nt_{part} is chosen, the coarser the grid structure of macrostates will be which is generated in the algorithm described below. Again there have to be considered restrictions given by the storage capacity, especially for higher-dimensional systems.
A "trajectory" is defined as the sequence of cells a particle traverses. Starting a particle e.g. in state (=cell) number i its trajectory (=sequence of cells) is stored in an array and all the cells of this sequence are joined together in the macrostate ms_1. Then the next particle is started in the neighbouring cell, microstate number $i+1$, and again its path is followed and stored.
Two cases are now distinguished: In case the new path and all previously calculated trajectories are disjoint the new trajectory is defined as next macrostate ms_2.
However, as soon as the new trajectory begins to overlap with one of the earlier recorded macrostates it is assigned to that macrostate and the tracking of this test particle is terminated. That means all the cells of the fine grid that the test particle has traversed so far are also added to the macrostate. The algorithm continues with the start of a particle in the next cell and so on until the whole phase space is partitioned into such units. Owing to their construction these macrostates resemble "quasi-Hamiltonian" orbits and thus represent the basic dynamics of the system. A schematic view of a possible phase space partition is presented in figure 29.
After the phase space has been disjointly covered with macrostates one computes the transition probabilities and thereby the time propagator as the matrix of relative frequencies of all possible transitions for these larger units in the same way as before. In this case the sample particles are started in each of the macrostates.
To collect enough statistics the number of test particles per macrostate has to be sufficiently large, usually a factor of 5 to 10 larger than in the fine grid algorithm, and the starting cells of the particular macrostate are chosen randomly.

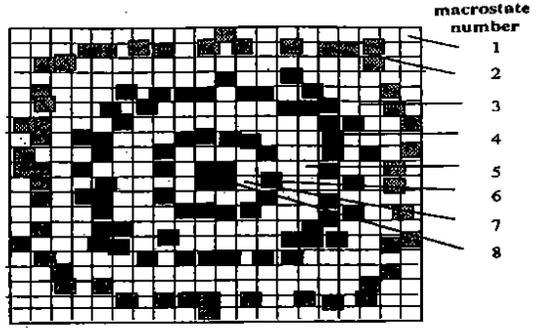

Figure 29: Schematic view of a partition of the phase space of a harmonic oscillator into macrostates.

Moreover the time parameters have to be larger to resemble the coarser structure. The best choice for the number of turns for the tracking of particles n_t is about $0.75 - 1.0 \cdot \tau_D$, and n_{tm} should be about $0.5 \cdot \tau_D$.
The transformation from the macrostate-density back to the fine grid-density is done by dividing the former quantity by the number of cells in the particular macrostates.
In the following the results of the density calculations with the macrostate method are compared to those gained with the fine grid mapping algorithm and to direct tracking.
The figures 30 and 31 illustrate the evolution of the phase space density $\rho(x, p_x)$ for the beam-beam model with $Q = 63.21$ and $\xi = 0.07$. The maxima of the density function correspond to the fixed points of the underlying Hamiltonian system and the resonance structure of the 4-resonance can easily be recognized.
In the density calculation with the macrostate technique the density maximum at the origin does not reach the the same value as in the fine grid case and the structure around this main density peak is broadened. This is due to the relatively large number of cells collected together in the macrostate containing the origin.
A comparison of the density calculations with all three methods can be found in figure 32. Shown is the projection of the density onto the x-axis.

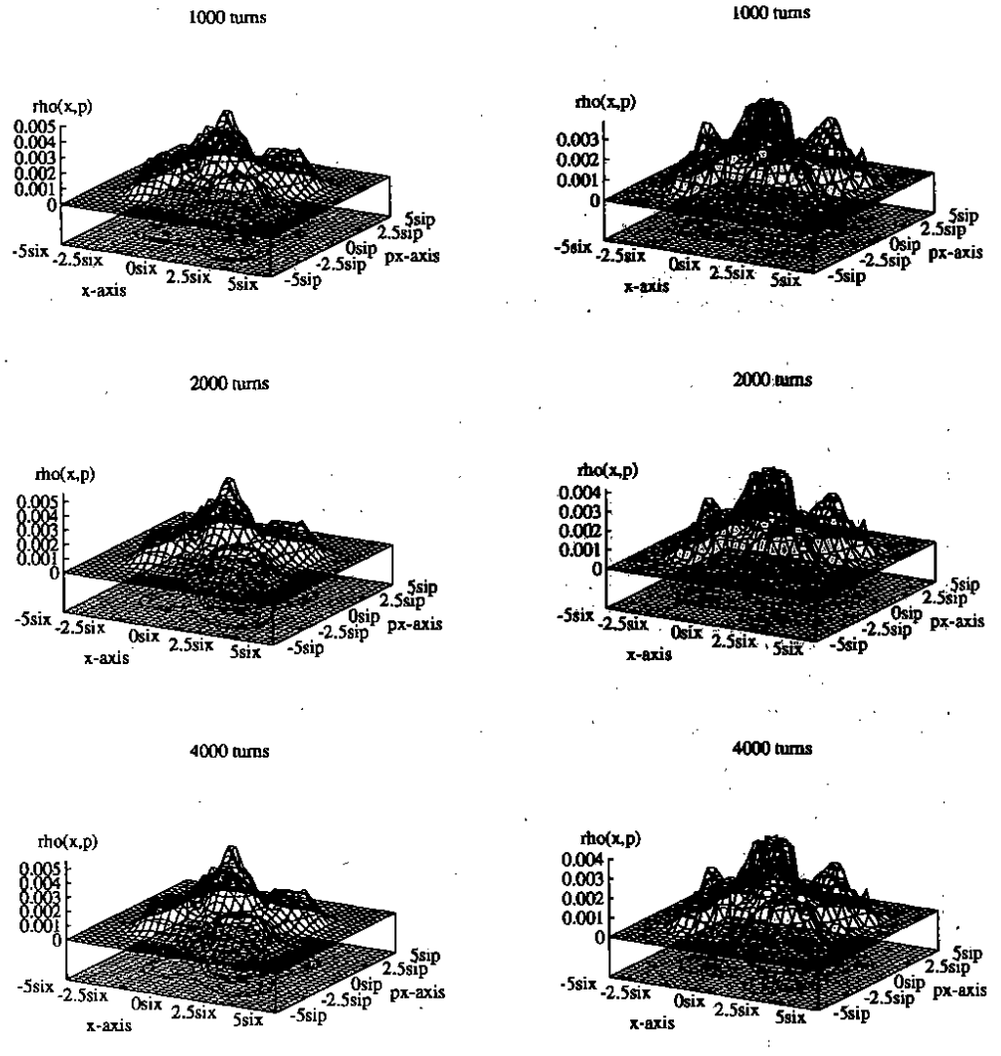

Figure 30: Evolution of the density function calculated via the fine grid mapping (left hand side) and the macrostate technique (right hand side), for $Q = .21$ and $\xi = 0.07$.
Shown are $\rho(1000)$, $\rho(2000)$ and $\rho(4000)$.

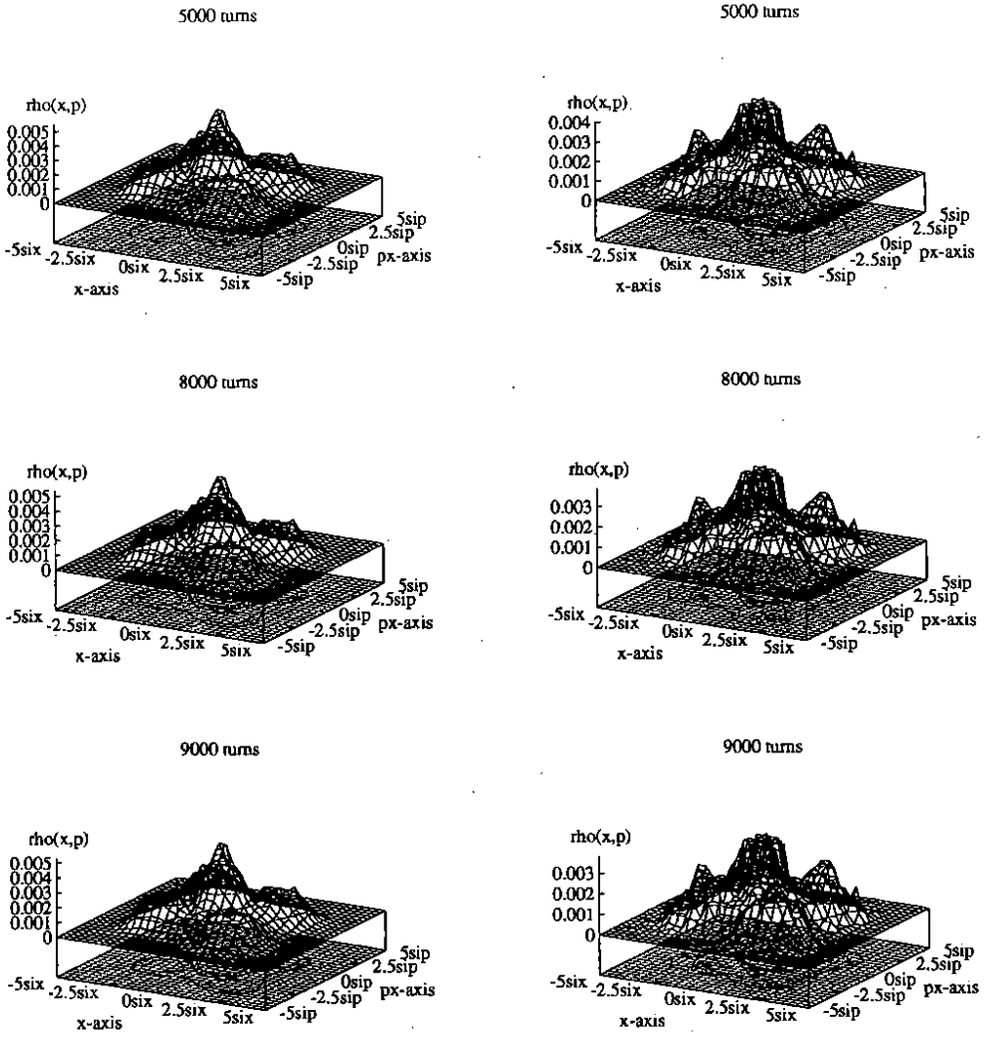

Figure 31: Evolution of the density function calculated via the fine grid mapping (left hand side) and the macrostate technique (right hand side), $Q = 63.21$, $\xi = 0.07$.
Shown are $\rho(5000)$, $\rho(8000)$ and $\rho(9000)$.

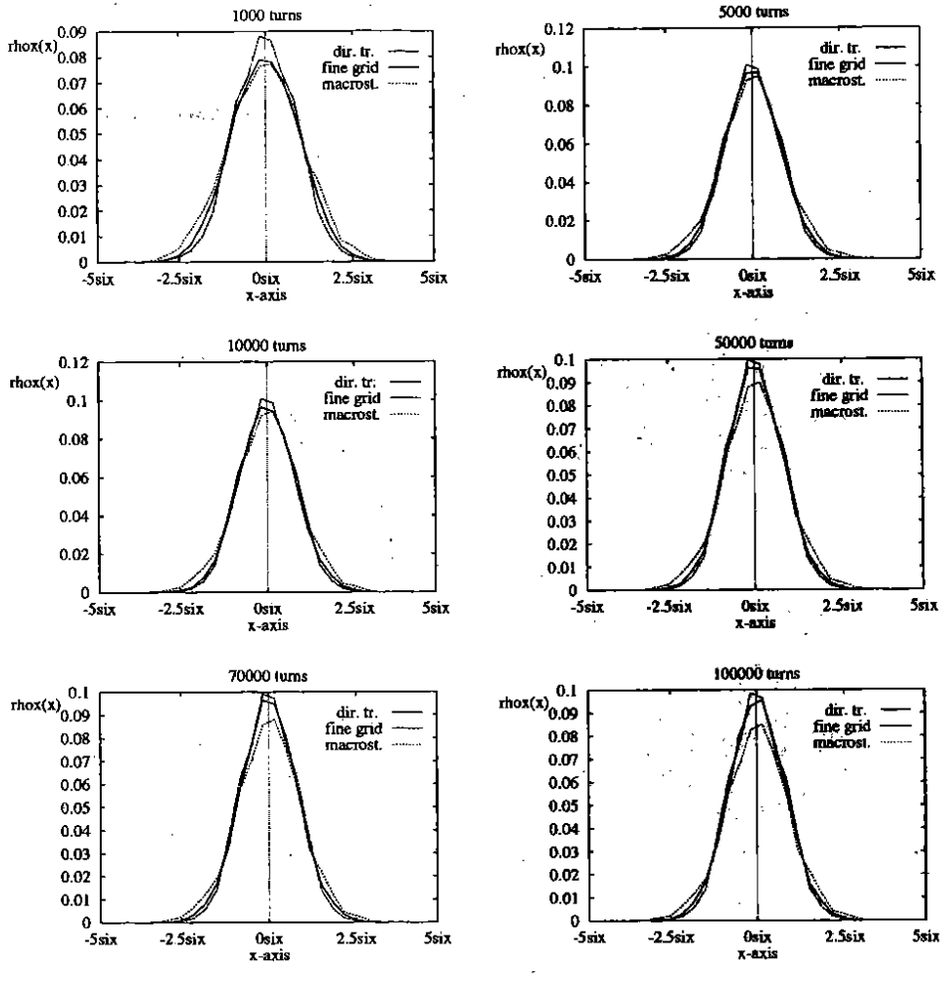

Figure 32: Comparison between the three different methods for the density calculation.
Shown are the x-projections of the density $\rho(x, p)$ for $Q = .12$ and $\xi = 0.07$.

The macrostate technique leads to a considerable reduction of computing time and storage requirements of the program. For the long term tracking studies up to 10^5 turns of figure 32 the consumed computing time of the programs was 13 hours for the direct tracking, 2 minutes for the fine grid version of the matrix algorithm, and about 15 CPU seconds for the macrostate technique.
Up to a large number (10^5) of considered turns the agreement of the methods is quite good. Deviations appear especially in the tails which are slightly overestimated by the macrostate technique. As the number of particles in the tails of the distribution determines the lifetime of a stored beam the macrostate method here yields results for τ_{life} that are too pessimistic. The deviations could possibly be reduced by further adjusting the time scales for the computation of the transition matrix for the macrostates.
In table 4 we compare the results of lifetime calculations with the three methods for absorbing boundaries at $\pm 5\sigma_x$ in the x-coordinate and at $\pm 5\sigma_p$ in the momentum.
The direct tracking was performed with a sample size of 50 particles per bin, the number of turns was 10^4.

Q-value:	63.10		63.11		63.12		63.13		63.14		63.15	
method	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU
dir.tracking	$1.5 \cdot 10^6$	55'	$4.5 \cdot 10^6$	57'	$6.8 \cdot 10^6$	1h01'	$1.9 \cdot 10^6$	1h02'	$1.2 \cdot 10^5$	56'	$1.2 \cdot 10^5$	38'
map,fine gr.	$4.9 \cdot 10^5$	29"	$1.1 \cdot 10^6$	30"	$4.2 \cdot 10^5$	31"	$5.7 \cdot 10^5$	33"	$1 \cdot 10^5$	32"	$6.6 \cdot 10^4$	25"
macrostates	$2 \cdot 10^5$	26"	$3 \cdot 10^5$	18"	$9 \cdot 10^5$	16"	$2 \cdot 10^5$	19"	$3.4 \cdot 10^4$	18"	$2.2 \cdot 10^4$	13"

Q-value:	63.16		63.17		63.18		63.19		63.20		63.21	
method	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU
dir.tracking	$1.6 \cdot 10^6$	1h00'	$4.5 \cdot 10^6$	1h05'	$5.1 \cdot 10^6$	1h07'	$5.1 \cdot 10^6$	1h11'	$1.6 \cdot 10^6$	1h35'	$2.3 \cdot 10^5$	1h35'
map,fine gr.	$9 \cdot 10^5$	25"	$1.6 \cdot 10^6$	26"	$1.4 \cdot 10^6$	28"	$1.7 \cdot 10^6$	27"	$5.6 \cdot 10^5$	38"	$1.1 \cdot 10^5$	38"
macrostates	$3.2 \cdot 10^5$	13"	$1.1 \cdot 10^5$	16"	$7.7 \cdot 10^5$	14"	$8 \cdot 10^5$	14"	$3.1 \cdot 10^5$	14"	$8.4 \cdot 10^4$	11"

Table 4: Lifetimes (in number of turns) and computing time, dependent on the Q-value for the various methods for the one-dimensional beam-beam model with $\xi = 0.07$.

In figure 33 the computed lifetime for this system in dependence on the Q-value for the three different algorithms is shown. The agreement of the direct tracking and the fine grid mapping algorithm is satisfactory. The macrostate method yields results that are generally too low. This is due to the fact that the density in the tails is still overestimated for the chosen time parameters in the tracking and averaging procedure. This difficulty should be overcome by adjusting the time scales adequately.

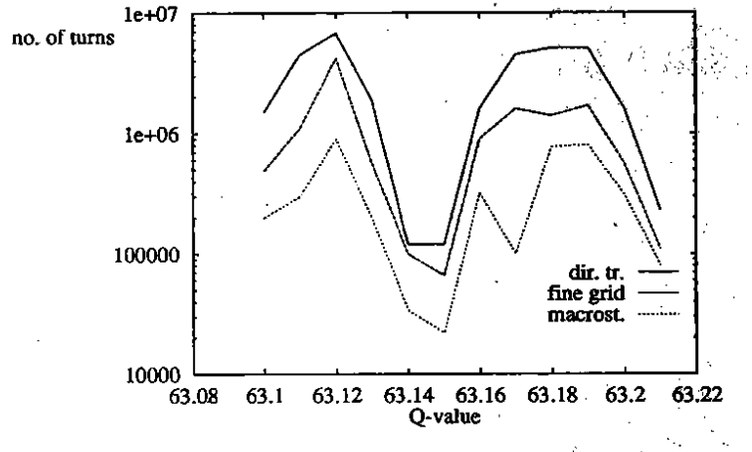

Figure 33: Lifetime in dependence on the Q-value for the one-dimensional beam-beam model with $\xi = 0.07$ calculated with the three different methods. The agreement of the direct tracking and the mapping algorithm on the fine grid is good. The deviations of the macrostate method are relatively large, nevertheless the general behaviour is similar.

Beam-Beam Model with DORIS-like Parameters

The next beam-beam model is a one-dimensional model of a storage ring with DORIS-like parameters, such as a damping time of 1850 turns, corresponding to $\alpha = 5.4 \cdot 10^{-4}$, a beam-beam parameter of $\xi = 0.029$ and beam sizes of $\sigma = 7 \cdot 10^{-4} m = \sigma_{cr}$ and $\beta^* = 0.9m$. The range of Q-values is taken from $Q = 5.10$ to $Q = 5.25$. Boundaries are imposed at $\pm 6\sigma$ in both coordinates. The parameters used in the mapping algorithm are the following: The particles are tracked for $n_t = 925 = \frac{1}{2}\tau_D$ turns, the transitions between the states are calculated after $n_{tm} = 463 = \frac{1}{4}\tau_D$ turns and the number of sample particles is 200. The density is computed up to 25 damping times, corresponding to 46300 turns or 100 applications of the transition matrix. Figure 34 shows the time evolution of the density for the Q-value $Q = 5.16$. One clearly recognizes the 6-resonance. The table 5 shows the comparison of the three methods for this system. The macrostate calculations were performed with the parameter values $n_t = 1388 = \frac{3}{4}\tau_D$ and $n_{tm} = 925 \doteq \frac{1}{2}\tau_D$ for the tracking times and a sample size of 1000 particles.

Q-value:	5.10		5.11		5.12		5.13		5.14		5.15	
method	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU
dir.tracking	$5 \cdot 10^5$	4h32'	$2.6 \cdot 10^6$	4h54'	$7.3 \cdot 10^5$	4h45'	$2.5 \cdot 10^6$	4h44'	$5.3 \cdot 10^5$	4h54'	$2.7 \cdot 10^6$	4h59'
map,fine gr.	$1.2 \cdot 10^6$	25"	$2.4 \cdot 10^6$	23"	$8.7 \cdot 10^5$	22"	$1.2 \cdot 10^6$	26"	$2.8 \cdot 10^7$	27"	$7 \cdot 10^7$	25'
macrostates	$1.2 \cdot 10^6$	21"	$3 \cdot 10^6$	20"	$1.3 \cdot 10^5$	17"	$6.4 \cdot 10^7$	17"	$7.4 \cdot 10^7$	28"	$7.2 \cdot 10^6$	22"

Q-value:	5.16		5.17		5.18		5.19		5.20		5.21	
method	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU	τ_{life}	CPU
dir.tracking	$3 \cdot 10^5$	4h08'	$8 \cdot 10^7$	5h52'	$1.3 \cdot 10^6$	5h10'	$5.4 \cdot 10^6$	4h56'	$2.7 \cdot 10^6$	6h33'	$5.6 \cdot 10^6$	6h43'
map,fine gr.	$2.2 \cdot 10^5$	29"	$1.4 \cdot 10^6$	24"	$3.5 \cdot 10^5$	23"	$3.3 \cdot 10^6$	25"	$6 \cdot 10^6$	25"	$1.3 \cdot 10^6$	26"
macrostates	$3.3 \cdot 10^5$	25"	$4.4 \cdot 10^6$	19"	$5 \cdot 10^5$	15"	$7 \cdot 10^6$	17"	$2 \cdot 10^6$	20"	$2.3 \cdot 10^7$	23"

Table 5: Lifetimes (in number of turns) and computing time, dependent on the Q-value for the various methods, $\xi = 0.029$.

The projected densities in the x-coordinate up to 25 damping times for this system are compared in figure 35.
The direct tracking has been performed with a sample size of 50 particles, the computing time that was required for this method was 4 hours and 8 minutes, the mapping algorithm on the fine grid needed 30 CPU seconds, and for the macrostate technique 16 seconds were necessary. The mapping algorithm on the fine grid and the direct multiparticle tracking are in excellent agreement. In the macrostate method the particle distribution in the tails again seems to be slightly overestimated resulting in a shorter lifetime.

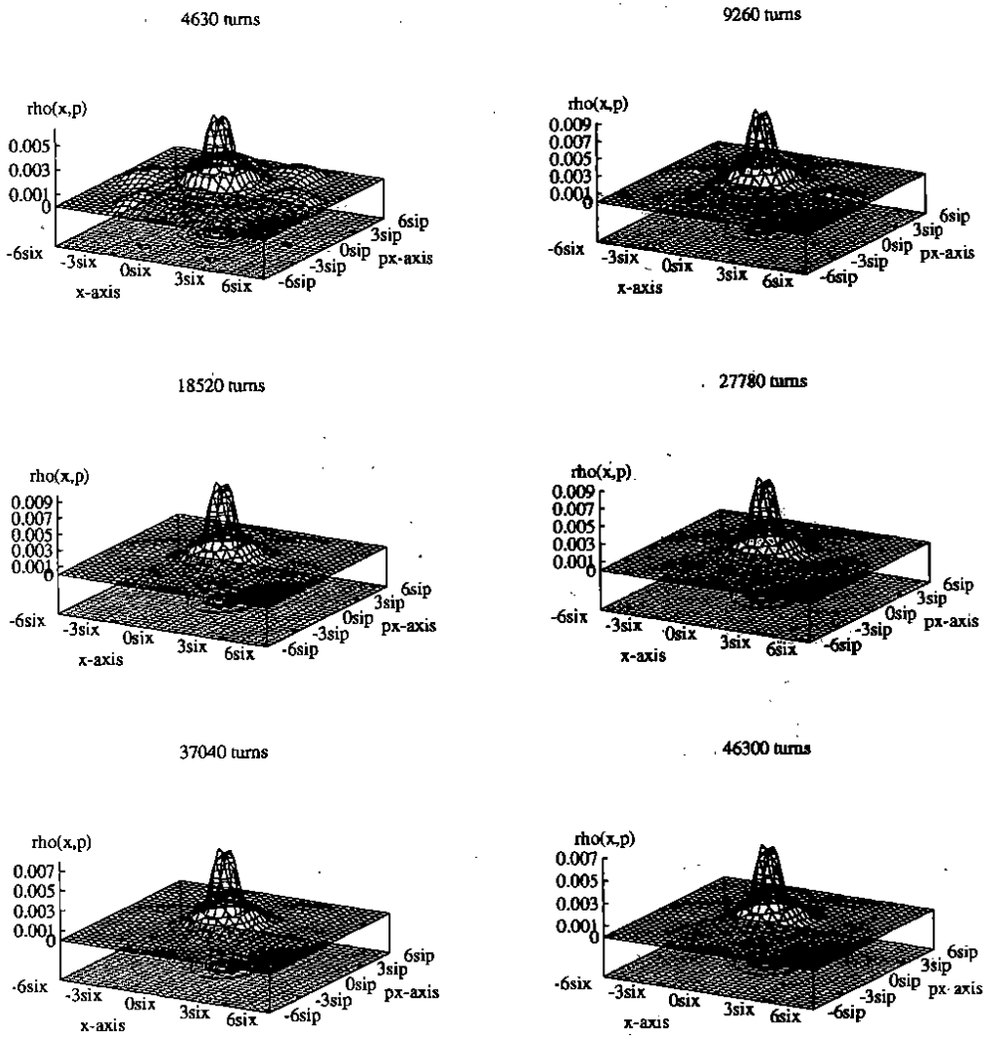

Figure 34: Evolution of the density function calculated via fine grid mapping for the example beam-beam model with $Q = .16$ and $\xi = 0.029$.
Shown are $\rho(4630)$, $\rho(9260)$, $\rho(18520)$, $\rho(27780)$, $\rho(37040)$ and $\rho(46300)$. In the first pictures the maxima of the density corresponding to the 6-resonance can be recognized.

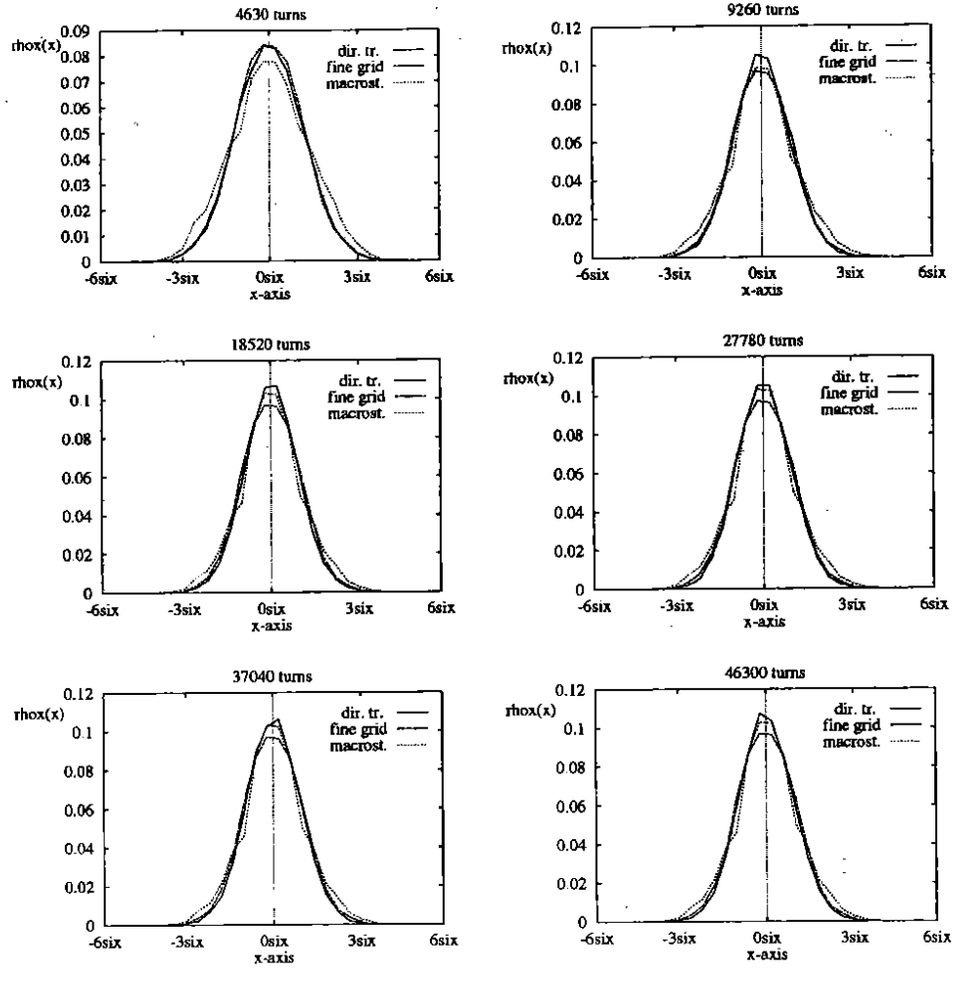

Figure 35: Comparison between the three different methods for the density calculation. The shown densities are $\rho_x(1000)$, $\rho_x(5000)$, $\rho_x(10000)$, $\rho_x(50000)$, $\rho_x(70000)$ and $\rho_x(100000)$, for $Q = 5.13$ and $\xi = 0.029$.

6.6 Higher-dimensional Calculations

For realistic applications it is necessary to extend the matrix algorithm to two-dimensional beam-beam systems. This extension is in principle possible but requires the generation of an $N \times N \times N \times N$ grid on the phase space and thus the storage for an $[N \times N \times N \times N]^2$ matrix as time propagator. For a 30×30 grid this would mean a need for storage in the range of several Gbyte.

In more than one dimension the structure of the phase space becomes much more complicated and there are effects that do not occur in the one-dimensional case. It is therefore necessary to approximate the complicated dynamics as close as possible in the matrix algorithm which is based on a lot of simplifications and averaging techniques. The main difficulties are the limitations in computing time and available storage capacity.

Several ways to treat these problems have been investigated and the results will be shown.

The first method is to make the grid partition as fine as possible. In order to make sure that there will be sufficient storage available for the propagator matrix, one has to know the maximum possible states that can appear as initial and final states in this matrix. This knowledge is needed before the calculation of the matrix starts and can be gained by making a run of a "just counting" version of the program which does not calculate any matrices but only counts the maximum numbers of possible initial states and corresponding final states that are needed for the indirect addressing.

The second way of computing the time propagator avoids this need for previous information by allocating the storage at run time of the program which can be done in a C version of the program. In this case the indirect addressing of the matrix elements is done via pointers and linked lists. The disadvantage of this second method is the additional amount of storage for the pointer structures.

The third method that has been tested for handling the two-dimensional beam-beam system is the macrostate technique. Here several difficulties appear, for example the number of macrostates has to be sufficiently low in order to make the matrices treatable and to yield acceptable results. The number of macrostates should not be larger than about 1000 to resemble the correct behaviour of the system. Although higher numbers of macrostates could in principle be handled, they did not lead to reasonable results. The time parameters for the first part of this method, namely the partition of the four dimensional grid into the macrostates, are larger than for the one-dimensional equivalent of the program. This is due to the fact that many more grid cells belong to one of the "quasi-Hamiltonian" orbits that represent the phase space structure. Tracking times of more than $0.20 \cdot \tau_D$ for finding the partition of the fine grid into macrostructures have shown to yield the optimal system representations. These tracking times depend sensitively on the phase space structure of the investigated systems, i.e. on the Q-values. For certain example systems featuring complicated dynamics like resonances and coupling, tracking times nt_{part} of about $0.80 \cdot \tau_D$ delivered the optimal partition.

Again models of simple storage rings are treated, consisting of linear elements like drift spaces, dipoles and quadrupoles and a beam-beam kick at the interaction area where also damping and noise terms are introduced.

The finest grid partition of the phase space that could be achieved within the limits of computing time and storage request was a $16 \times 16 \times 16 \times 16$ partition. For the density mapping technique on the fine grid 50 particles are started on each grid cell and tracked for half of the damping time to compute the density propagator matrix.

The parameters of the first example system are: $Q_x = .12$ and $Q_z = .31$ for the fractional parts of the Q-values, $\alpha = 0.0025$, corresponding to a damping time of 400 turns, the beam-

beam parameters are $\xi_x = 0.017$ and $\xi_z = 0.019$, the unperturbed beam sizes are taken to be $\sigma_x = 2 \cdot 10^{-4}m$ and $\sigma_z = 3 \cdot 10^{-5}m$, the transverse size of the counterrotating charge distribution is assumed $\sigma_{cr} = 2.5 \cdot 10^{-4}m$, and absorbing boundaries at $\pm 6\sigma_x$ in x-direction, at $\pm 6\sigma_{px}$ in p_x-direction, at $\pm 4\sigma_z$ in z-direction and at $\pm 4\sigma_{pz}$ in the corresponding momentum are considered.

In the presented figures the results of calculations with the fine grid mapping for the density are shown and compared to direct tracking.

The figures 36 to 38 give the projections of the four dimensional density function $\rho(x, p_x, z, p_z)$ on the $x - p_x$-plane, the $z - p_z$-plane, and the $x - z$-plane after 400, 1000, and 10000 turns respectively.

On the left hand sides the results of the mapping algorithm are given and on the right hand side the corresponding densities that have been calculated via direct tracking.

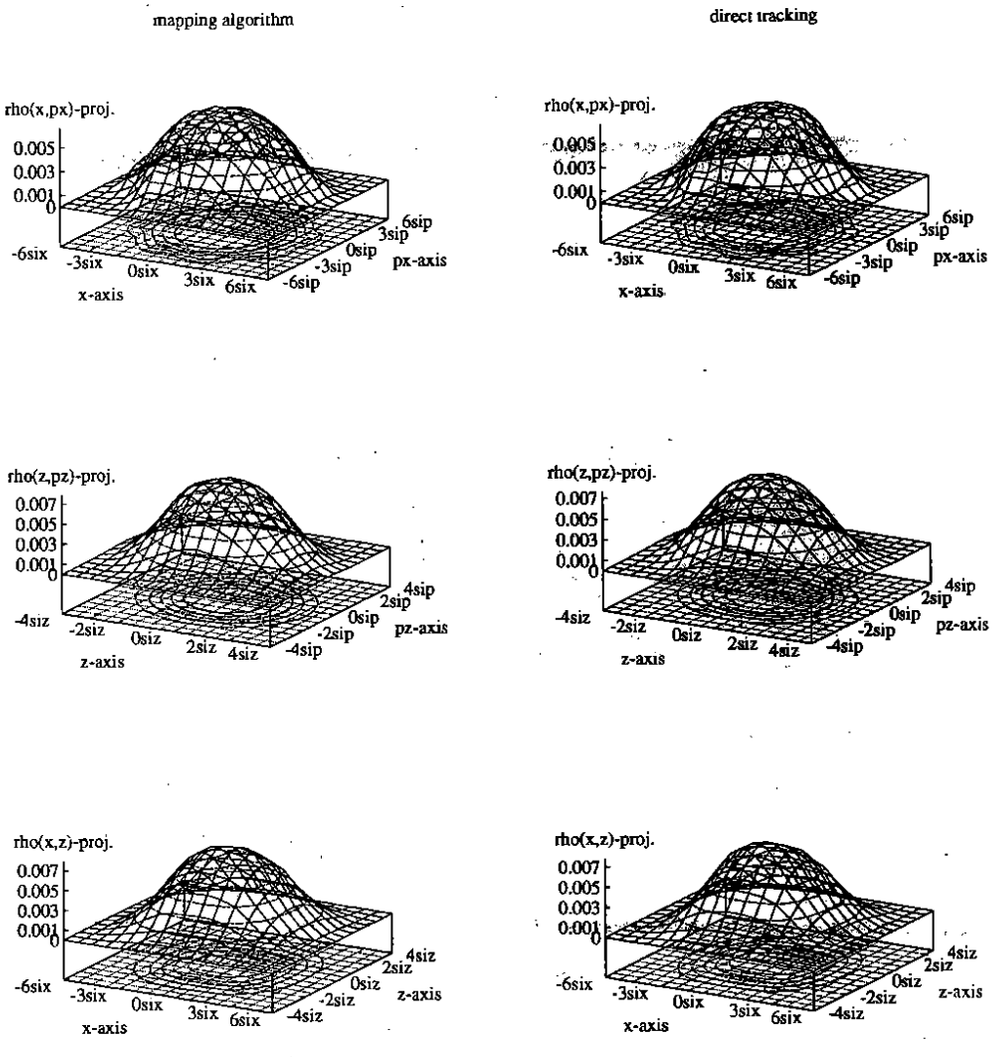

Figure 36: Phase space density for the two-dimensional beam-beam model with $Q_x = .12, Q_z = .31$ and $\xi_x = 0.017, \xi_z = 0.019$ after 400 turns. Shown are the projections $\rho(x, p_x)$, $\rho(z, p_z)$ and $\rho(x, z)$, calculated with the mapping scheme (left) and dir. tracking (right). The initial density has been homogeneous on the phase space.

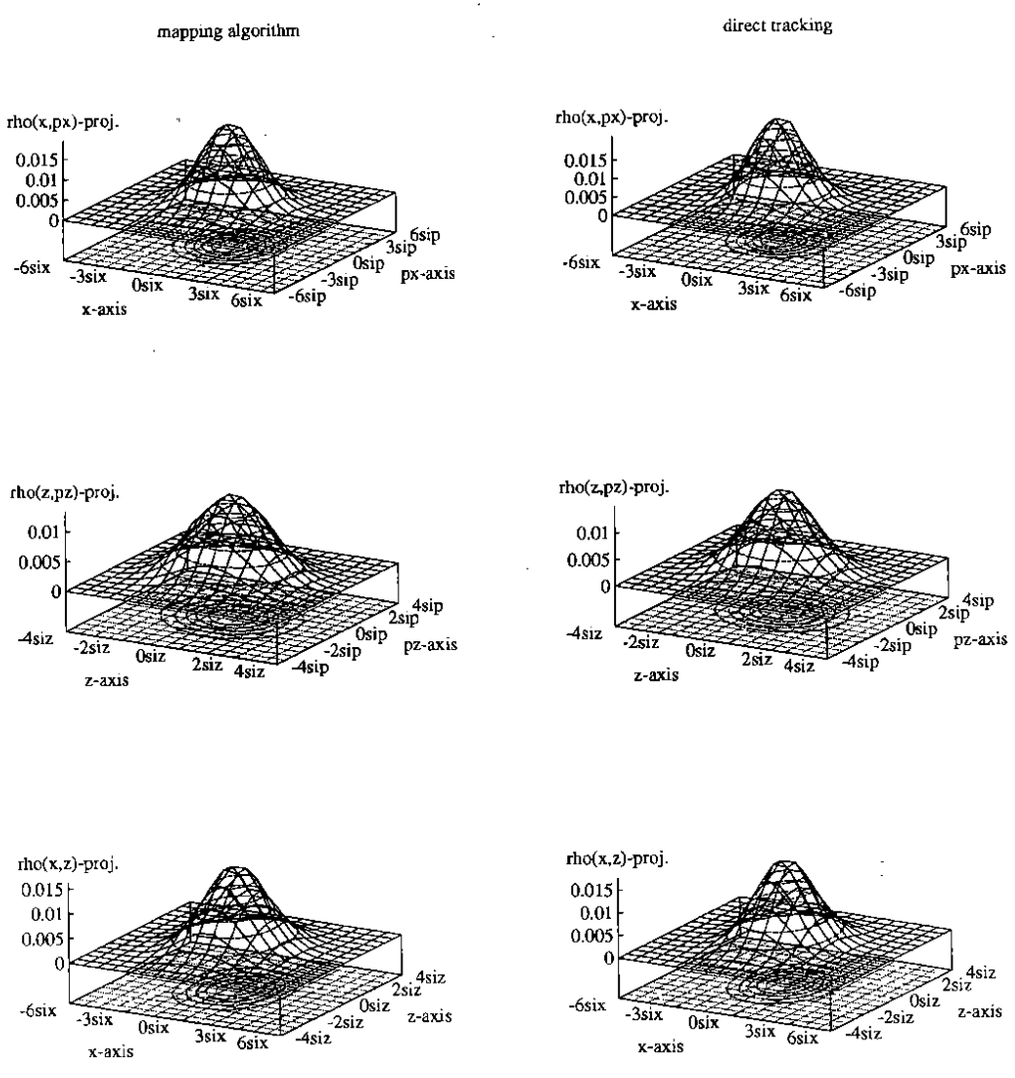

Figure 37: Phase space density for the two-dimensional beam-beam model with $Q_x = .12, Q_z = .31$ and $\xi_x = 0.017, \xi_z = 0.019$ after 1000 turns. Shown are the projections $\rho(x, p_x)$, $\rho(z, p_z)$ and $\rho(x, z)$, calculated with the mapping scheme (left) and dir. tracking (right).

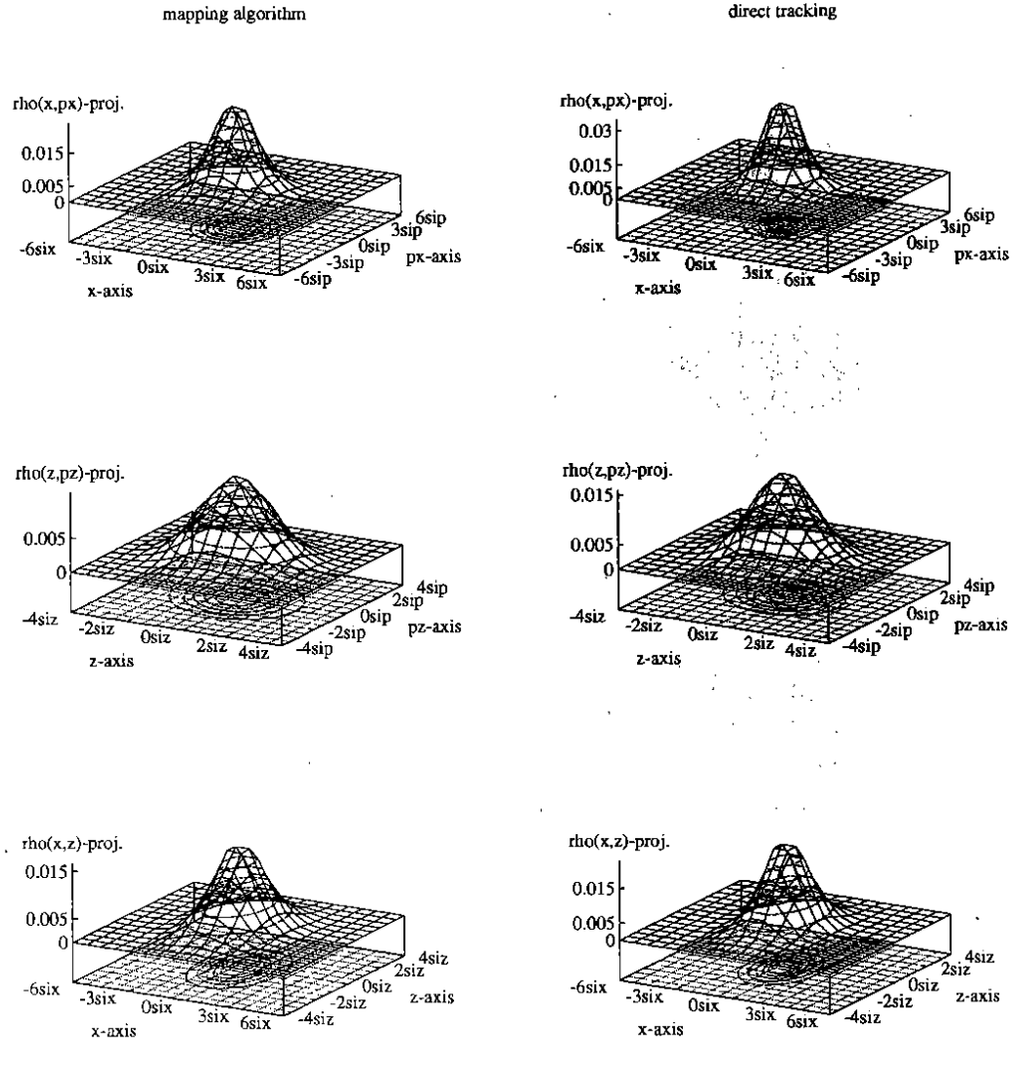

Figure 38: Phase space density for the two-dimensional beam-beam model with $Q_x = .12, Q_z = .31$ and $\xi_x = 0.017, \xi_z = 0.019$ after 10000 turns. Shown are the projections $\rho(x, p_x)$, $\rho(z, p_z)$ and $\rho(x, z)$, calculated with the mapping scheme (left) and dir. tracking (right). It can be seen that the density calculated with the mapping algorithm here is generally lower than the directly computed one.

Lifetime Calculations for Two-dimensional Systems

The dependence of the lifetime on the boundary conditions can be estimated from the Kramers formula which is valid for the linear case, at least for the vertical motion. In horizontal direction the off-energy orbit, that means the dispersion, has to be included in the lifetime calculation in general. Neglecting dispersion in the calculations, the Kramers formula can be applied for both transversal directions [16] and as an approximation one takes the minimum of the separately estimated lifetimes for both horizontal and vertical space.
Adding the nonlinear perturbation to the system these linear lifetime estimates can only give an idea of the order of magnitude of the density decay and of its dependence on the position of the absorbing boundaries.
Lifetime calculations yield $\tau_{life} = 6 \cdot 10^4$ turns with the mapping algorithm and $\tau_{life} = 1.3 \cdot 10^5$ turns in the direct tracking. These short lifetimes are a result of the small aperture limit, especially in the vertical coordinates.
For these calculations the particle motion was simulated for 10000 turns which took about 6 hours of computing time for the mapping program. The direct tracking program needed 68 CPU hours. In both cases the sample size has been taken to be 50 particles per grid cell. (Again all computations were done on HP9000-730 workstations.)
However, these preliminary results show that there still appear discrepancies between the mapping scheme results and the direct tracking ones. This is probably an effect caused by the way the transition matrix is normalized, taking also into consideration all the particle that are lost soon. This problem might be overcome by suitably adjusting the normalization of the matrix or the averaging time scales over which the transition matrix is calculated.

An extreme case is given if the fractional parts of the unperturbed tunes are $Q_x = Q_z = 0.25$. When adding a perturbation this system will therefore resemble a 4-resonance.

In figure 39 the projections of the density function after 200 turns for the two different systems, the first one having the fractional Q-values of $Q_x = .12$ and $Q_z = .31$ and the second one with $Q_x = Q_z = 0.25$, are presented.
Due to the 4-resonance the decay of the density function is very fast and the lifetime calculations of this example system is only a few thousand turns. The direct tracking result is τ_{life}=6000 turns and the mapping algorithm yields τ_{life}=3000 turns. Again following 50 particles per grid cell and computing the density for 10000 turns the consumed computing time was 12 hours for the direct tracking and 1 hour in the mapping case.

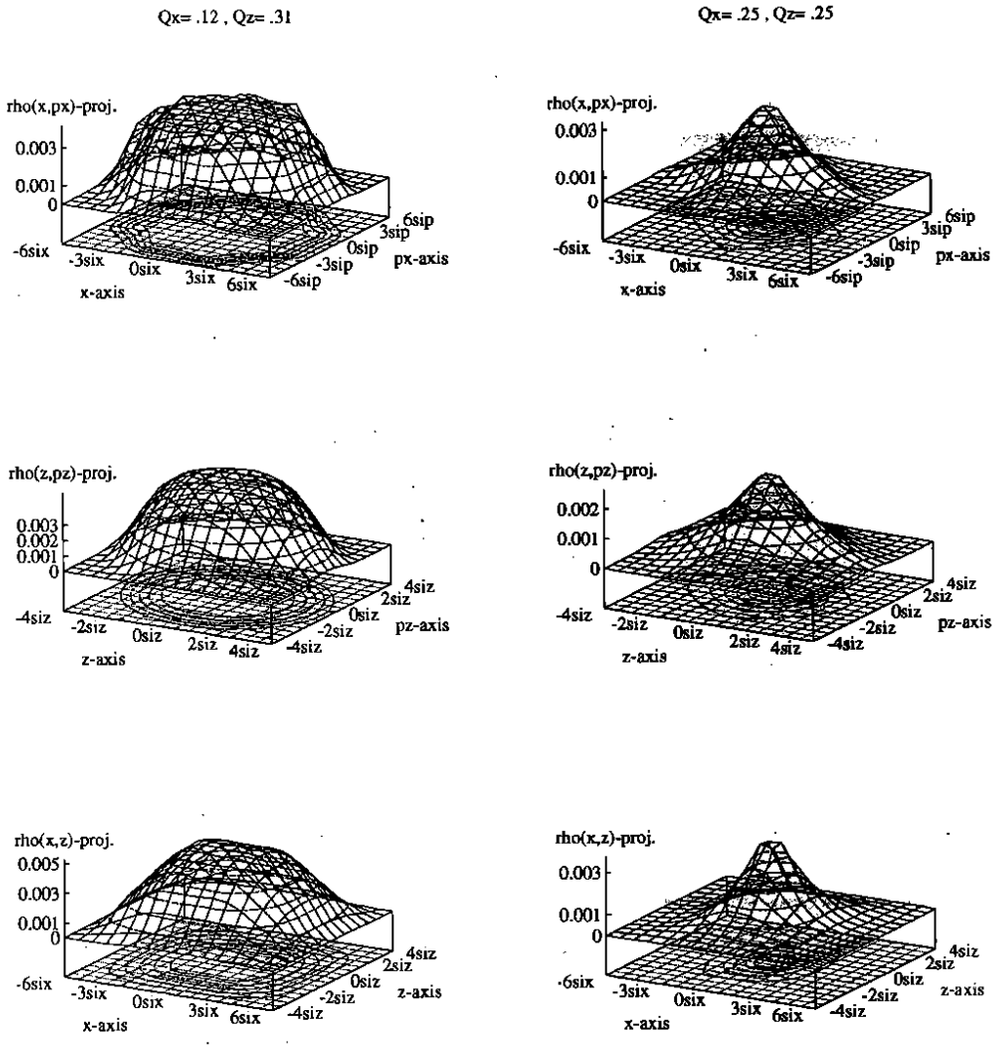

Figure 39: Phase space density for the two-dimensional beam-beam model after 200 turns. Shown are the projections $\rho(x, p_x)$, $\rho(z, p_z)$ and $\rho(x, z)$ calculated with the density mapping algorithm for $Q_x = 0.12, Q_z = 0.31$ (left) and $Q_x = 0.25, Q_z = 0.25$ (right) and $\xi_x = 0.017, \xi_z = 0.019$.

Macrostates for Two-dimensional Systems

In applying the macrostate method to calculate the temporal and spatial evolution of the density function several difficulties are encountered.
The time parameters of the tracking part of the partition algorithm where the macrostate structure is determined show a strong dependence on the detailed features (e.g. the Q-values) of the system. The tracking time for finding a partition of the phase space varied from $0.2-0.8\cdot\tau_D$ for different Q-values.
Moreover in the averaging over the more complicated phase space structures of the two-dimensional systems some information is possibly lost.
Here still some work has to be done to determine the correct time scales and averaging parameters that represent the behaviour of the system in the right way.

7 Summary

Several methods for the investigation of stochastic effects in accelerator physics are introduced. In the first part of this thesis the mathematical concept of stochastic differential equations is used for the description of simple models from accelerator dynamics with additional random forces. Special emphasis has been put on the installation and test of numerical algorithms for the solution of stochastic differential equations. These are applied to example systems for the computation of beam sizes or the emittance. Most of the calculations turn out to be very computing time intensive due to large numbers of samples needed for achieving sufficient statistics and to the small step sizes in the integration procedure. Thus in general these algorithms do not seem to be suitable for long term studies. In particular a double radiofrequency structure with phase noise has been investigated. Good agreement is observed between perturbation theoretical treatment of the stochastic differential equations or the corresponding Fokker-Planck equation and the numerical simulation.
In the second part a new algorithm for computing the temporal and spatial evolution of the phase space density function $\rho(\vec{x},\vec{p},t)$ for dissipative stochastically excited systems is introduced. Including nonlinear terms like for example the force generated by the beam-beam interaction of two counterrotating particle distributions, the density function can no longer be calculated analytically and numerical methods become indispensable. In general the numerical calculation of densities needs a lot of computing time. This new algorithm was developed to reduce the computing effort. By calculating a time propagator for the density function of a Markov process the method can model the dynamical evolution of the phase space density of an electron storage ring in the presence of damping (radiation damping) and fluctuating external forces (noise). The evolution of the density is modelled via successive application of this propagator to an initial density. The time propagator is generated by calculating the probabilities for all possible transitions between states in phase space within a certain time interval. These probabilities are computed as the relative frequencies for the transitions from one state to another. For this purpose the phase space is partitioned into a set of discrete cells which means that the calculations are performed on a grid. All the transition probabilities are stored in a matrix that represents the time propagator for the density. Using indirect addressing of the matrix elements the matrix multiplications can be performed very computing time efficient. The algorithm has been tested with simple beam-beam systems. Considering one degree of freedom (two-dimensional phase space) good agreement is observed with the far more computing time intensive direct multiparticle tracking calculations. In a next step the algorithm has

been applied to the two-dimensional beam-beam interaction. Although in this case the phase space grid could not be chosen as fine as in the one-dimensional case due to storage capacity limitations, the results were acceptably good compared to direct tracking.
For several simplified model systems of electron storage rings the phase space density and the electron beam lifetime have been evaluated.
A further reduction of the computing time and the storage requirements is provided by the macrostate technique. By joining suitable parts of the phase space into larger units, "macrostates", and calculating the transition matrix for these larger structures the density function computation can be speeded up by several orders of magnitude. This latter method has been tested and compared to the mapping algorithm on the fine grid and to the direct tracking. In the case of the one-dimensional systems the macrostate method yields quite good agreement with the other algorithms. The shape of the density is almost perfectly represented although an average over the large areas in phase space forming the macrostates is involved.
The application of the macrostate technique to higher-dimensional systems leads to severe problems and large differences to the other methods are observed.

8 Acknowledgements

I would like to thank Dr. H. Mais for continuous support and guidance and many helpful discussions.
Special thanks to Dr. A. Gerasimov for many valuable ideas and suggestions.
I want to express my gratitude to Prof. Dr. R.D. Kohaupt for continuous encouragement and interest, helpful discussions and valuable suggestions.
I am grateful to Prof. Dr. P. Schmüser for careful reading of the manuscript and many helpful and valuable suggestions.
Especially I want to thank my colleagues Dipl.Phys., resp. Dres., Michael Böge, Oliver Brüning, Winfried Decking and Klaus Flöttmann for many helpful and stimulating discussions and innumerable good and valuable suggestions and hints. In particular I am indebted to Michael's profound knowledge about everything around computers (and about the preparation of excellent coffee).
Finally, I want to express my gratitude to all colleagues at MPY for the nice and stimulating atmosphere, and last but not least I want to thank the DESY directorate for giving me the possibility to write a thesis in the field of accelerator physics and for the continuous support.

A The Storage Ring Coordinates

A detailed description of the motion of a relativistic charged particle in a storage ring is obtained using the Lorentz equation

$$\frac{d}{dt}\left(\frac{E}{c^2}\dot{\vec{r}}\right) = e\vec{\mathcal{E}} + \frac{e}{c}\left(\dot{\vec{r}} \times \vec{B}\right) + \vec{R} \qquad \text{(A.1)}$$

with
E = energy of the particle, $E = \frac{m_0c^2}{\sqrt{1-\frac{\dot{r}^2}{c^2}}} = \gamma m_0 c^2$,
c = velocity of light,
$\vec{B}$ = magnetic field,
$\vec{\mathcal{E}}$ = electric field and
$\vec{R}$ gives the radiative part of the force acting on the particle.
$\vec{R}$ consists of a continuous part and a stochastic one modelling the fluctuations.
$\vec{R}$ can be expressed by considering the radiation power $\mathcal{P}$ of an ultrarelativistic particle in a purely magnetic field [3],[5]. $\mathcal{P}$ is divided into a mean value and a fluctuating part $\mathcal{P} = \mathcal{P}^d + \delta\mathcal{P}$.
In a storage ring the magnetic fields are mainly transverse to the direction of motion and thus $\dot{\vec{r}} \cdot \vec{B} = 0$ holds. It is assumed that the radiation reaction force is collinear with $\dot{\vec{r}}$ which is fulfilled in good approximation at high energy because the photons are emitted in the direction of the momentum of the particle within an opening angle of $\frac{1}{\gamma} = \frac{m_0c^2}{E}$.
For the radiation power $\mathcal{P} = \vec{R} \cdot \dot{\vec{r}}$ one gets, see [3],[5]:

$$\mathcal{P}^d = E^2 \cdot \frac{2r_e}{3(m_0c)^3}\left(\frac{e}{c}|\vec{B}|\right)^2$$

and

$$< \delta\mathcal{P}(s)\delta\mathcal{P}(s') >= E^4 \cdot \frac{55 r_e \hbar}{24\sqrt{3}(m_0c)^6}\left(\frac{e}{c}|\vec{B}|\right)^3 \cdot \delta(s-s').$$

The position vector $\vec{r}$ refers to a fixed coordinate system. It is customary in accelerator physics to introduce a system of curvilinear coordinates x, z, s. These describe the motion of a particle with respect to an ideal closed design orbit which is defined as the path of a particle of constant design energy E_0. In addition it is required that the design orbit consists of piecewise flat curves either in vertical or horizontal direction. This reference orbit is denoted by $\vec{r}_0(s)$ where s is the arc length along the design orbit. An arbitrary particle trajectory $\vec{r}(s)$ is therefore described by the deviation $\delta\vec{r}(s)$ from the design orbit

$$\vec{r}(s) = \vec{r}_0(s) + \delta\vec{r}(s).$$

The vector $\delta\vec{r}(s)$ is represented in an orthogonal coordinate system accompanying the particles travelling along the design orbit.
This comoving frame is given by

- the unit tangent vector $\vec{e}_s(s) = \frac{d}{ds}\vec{r}_0(s)$,
- a unit vector perpendicular to $\vec{e}_s(s)$ in the horizontal plane, $\vec{e}_x(s)$
- and the unit vector $\vec{e}_z(s)$, perpendicular to $\vec{e}_s(s)$ and $\vec{e}_x(s)$, $\vec{e}_z(s) = \vec{e}_s(s) \times \vec{e}_x(s)$.

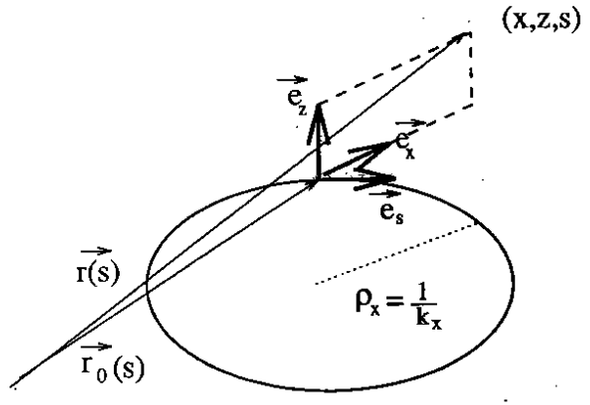

Figure 40: The storage ring coordinate system.

In these coordinates $\delta\vec{r}(s)$ is expressed as

$$\delta\vec{r}(s) = (\delta\vec{r}(s) \cdot \vec{e}_x)\vec{e}_x + (\delta\vec{r}(s) \cdot \vec{e}_z)\vec{e}_z.$$

Consequently the orbit vector $\vec{r}(s)$ can be written in the form

$$\vec{r}(s) = \vec{r}_0(s) + x(s)\vec{e}_x(s) + z(s)\vec{e}_z(s).$$

The Fresnet formulae for the accompanying coordinate system read:

$$\begin{aligned}
\frac{d}{ds}\vec{e}_x(s) &= +K_x(s)\vec{e}_s(s) \\
\frac{d}{ds}\vec{e}_z(s) &= +K_z(s)\vec{e}_s(s) \\
\frac{d}{ds}\vec{e}_s(s) &= -K_x(s)\vec{e}_x(s) - K_z(s)\vec{e}_z(s)
\end{aligned}$$

where $K_x(s)$ and $K_z(s)$ denote the curvatures in x- and in z-direction respectively and it is assumed that $K_x(s) \cdot K_z(s) = 0$, i.e. piecewise no torsion.
Using this coordinate system the storage ring motion is described by a 6-dimensional vector $\vec{y}$ with the two transverse coordinates x and z and the longitudinal coordinate ϕ (ϕ giving the time deviation from the reference particle in the bunch center) and the conjugate momenta. Thus the equations of motion get the form

$$\dot{\vec{y}} = \vec{f}(\vec{y}) + \underline{g}(\vec{y}) \cdot \vec{\xi}$$

where $\vec{\xi}$ denotes the stochastic part of the force.

B Definitions and Notations from Probability Theory

Probability theory deals with the calculation of new probabilities from given ones. In general these initial probabilities are the results of observations of frequencies in long experimental series.

B.1 The Sample Space

Let Ω be the set of all possible outcomes of a probability experiment. This set is called sample space. Its elements $\omega \in \Omega$ are referred to as elementary outcomes. Observable outcomes A are subsets of Ω, $A \subset \Omega$. But in general not all subsets of Ω are observable or interesting outcomes. Let $\mathcal{A}$ denote the set of all observable outcomes (= events) of a probability experiment.

$\mathcal{A}$ has the following properties:

1. $\Omega \in \mathcal{A}$ (the certain event),
2. $\emptyset \in \mathcal{A}$ (the impossible event),
3. $A \in \mathcal{A} \rightarrow A^c \in \mathcal{A}$ (the complementary event)
4. $A, B \in \mathcal{A} \rightarrow A \cup B \in \mathcal{A}$ and $A \cap B \in \mathcal{A}$
   From this follows that $\mathcal{A}$ is an algebra.
   If in addition this last property holds also for all countable unions
5. $A_n \in \mathcal{A} \rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$,
   $\mathcal{A}$ is said to be a σ-algebra.

The pair $(\Omega, \mathcal{A})$ is called measure space.

On the measure space one can define a real valued function P.

This function is called a probability measure if it fulfils the conditions

1. $0 \leq P(A) \leq \infty \quad \forall A \in \Omega$,
2. $P(\emptyset) = 0$,
3. $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n), \quad A_n \in \mathcal{A} \quad \forall n \geq 1$ and $A_n \cap A_m = \emptyset \quad \forall n \neq m$,
4. $P(\Omega) = 1$.

The triple $(\Omega, \mathcal{A}, P)$ is called probability triple.
$P(A)$ means the theoretical value of the relative frequency of A in a large number of realizations of a probability experiment which is characterized by the measure space $(\Omega, \mathcal{A})$.

B.2 Random Variables

Let $(\Omega, \mathcal{A}, P)$ denote a probability space and $(\Omega', \mathcal{A}')$ a measure space.
In the following Ω', also called "state space", will be the set of real numbers R or some subset thereof or more general R^{d}, that means the set of d-dimensional vectors whose components are real numbers.
A random variable X is a function from the sample space Ω into the state space Ω'.
Here one has $X : \Omega \rightarrow \mathrm{R}$. X represents a kind of measuring device and the state space R can be considered as a new sample space. The events in R are given by the intervals: $B = [x, y], x, y \in \mathrm{R}$. These intervals generate the Borel σ-algebra $\mathcal{B}$. (For generating this algebra it is sufficient to use only the intervals $(-\infty, x], \quad x \in \mathrm{R}$.)
The random variable $X : (\Omega, \mathcal{A}, P) \rightarrow (\mathrm{R}, \mathcal{B})$ induces a probability measure from the underlying sample space Ω into the state space R by

$$P_X(B) = P(X^{-1}(B)) = P(\omega | X(\omega) \in B) = P(X \in B) \quad \forall B \in \mathcal{B}.$$

P_X is called the distribution of the random variable. Due to the structure of the Borel σ-algebra $\mathcal{B}$ it is sufficient to consider events of the form $B_x = (-\infty, x]$ and P_X is completely defined by its distribution function $F(x)$:

$$F(x) = P(\omega | X(\omega) \leq x) = P(X \leq x), \quad x \in \mathrm{R}.$$

If $\vec{X}$ has values in R^{d} one has

$$F(\vec{x}) = F(x_1, ..., x_d) = P(\omega | X_1(\omega) \leq x_1, ..., X_d(\omega) \leq x_d) = P(\vec{X} \leq \vec{x}), \quad \vec{x} \in \mathrm{R}^{\mathrm{d}}.$$

The distribution function $F(\vec{x})$ has the properties:

- $F(-\infty) = 0$
- $F(+\infty) = 1$
- F is an increasing right continuous function.

Together with the distribution function the state space and its associated Borel σ-algebra $\mathcal{B}$ can be completed to a probability triple $(\mathrm{R}^{\mathrm{d}}, \mathcal{B}^d, P_X)$.
In case one is interested only in the values of the quantity X it suffices to consider $(\mathrm{R}^{\mathrm{d}}, \mathcal{B}^d, P_X)$ as a new sample space and take the random variable X as the identity.
In the following therefore simply P will be used instead of P_X.
For continuous random variables one can define a probability density ρ, that means there exists an integrable function $\rho(\vec{x}) \geq 0$ with

$$F(x_1, ..., x_d) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_d} \rho(y_1, ..., y_d) dy_1 ... dy_d.$$

Equivalently

$$\rho(\vec{x}) = \frac{\partial^d}{\partial_{x_1} ... \partial_{x_d}} F(\vec{x})$$

and $\rho(\vec{x}) d\vec{x}$ can be interpreted as the probability that $X(\omega)$ takes a value in the infinitesimal neighbourhood of $\vec{x}$.

$$\rho(\vec{x}) d\vec{x} = P(\omega | \vec{x} \leq X(\omega) \leq \vec{x} + d\vec{x})$$

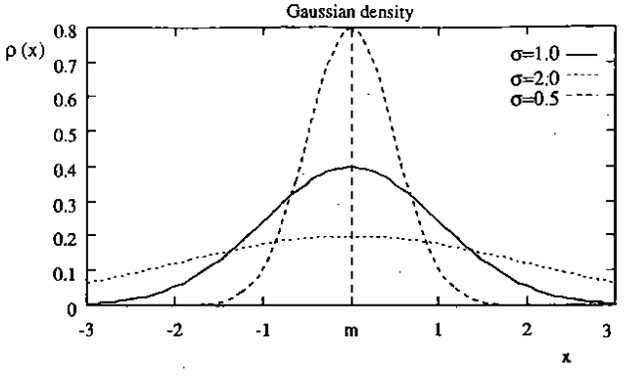

Figure 41: One-dimensional Gaussian density $\rho(x)$ for different σ-values, mean $m = 0$.

Example: Gaussian Distribution

The density function of the Gaussian or normal distribution $\mathcal{N}(\vec{m}, \underline{C})$ reads:

$$\rho(\vec{x}) = \frac{1}{\sqrt{(2\pi)^d \det \underline{C}}} e^{-\frac{1}{2}(\vec{x}-\vec{m})^T \underline{C}^{-1}(\vec{x}-\vec{m})}.$$

The importance of this particular distribution follows from the "central limit theorem" which will be explained below.

The following table shows the fraction of the area below the curve of the normalized Gaussian density which is obtained when integrating over the particular intervals.

interval	fraction of the area
$[m - \frac{\sigma}{\sqrt{2}}, m + \frac{\sigma}{\sqrt{2}}]$	52.22 %
$[m - \sigma, m + \sigma]$	68.27 %
$[m - \sqrt{2}\sigma, m + \sqrt{2}\sigma]$	84.14 %
$[m - 2\sigma, m + 2\sigma]$	95.45 %
$[m - 3\sigma, m + 3\sigma]$	99.73 %

B.3 Moments

The density function gives a complete characterization of the random variable $\vec{X}$. Often one is only interested in part of the information about the random variables, namely the moments: The mean or expectation value of $\vec{X}$ is defined as

$$< \vec{X} >= E(\vec{X}) = \int_{\mathrm{R^d}} \vec{x}\rho(\vec{x})d\vec{x} = \begin{pmatrix} \int_{\mathrm{R^d}} x_1\rho(x_1,...,x_d)dx_1...dx_d \\ \vdots \\ \int_{\mathrm{R^d}} x_d\rho(x_1,...,x_d)dx_1...dx_d \end{pmatrix}.$$

The k^{th} moment is $< X^k >$, the k^{th} central moment is $< (X - < X >)^k >$.

$$< X_1^{i_1} \cdot ... \cdot X_n^{i_n} >= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} x_1^{i_1} \cdot ... \cdot x_n^{i_n} \rho_{X_1,...X_n}(x_1,...x_n).$$

In the one-dimensional case $d = 1$ the quantity

$$VarX =< (X - < X >)^2 >= \sigma^2(X)$$

is called the variance of X
and

$$\sigma(X) = \sqrt{VarX}$$

is the standard deviation.
To characterize the relation between two or more random variables one defines the covariance

$$cov(X,Y) =< (X - < X >)(Y - < Y >) >$$

which is given by a matrix in the higher-dimensional case $d > 1$:

$$cov(X_i, Y_j).$$

The mean value of the expression $e^{i\vec{k}\vec{X}}$ is called the characteristic function:

$$\phi(\vec{k}) =< e^{i\vec{k}\vec{X}} >= \int_{\mathrm{R^d}} e^{i\vec{k}\vec{x}}\rho(\vec{x})d\vec{x}.$$

It is the Fourier transform of the density function.

For the normal distribution with mean $\vec{m}$ and covariance matrix $\underline{C}$ holds

$$\phi(\vec{k}) = e^{i\vec{k}\vec{m} - \frac{1}{2}\vec{k}^T \underline{C}\vec{k}}, \quad \vec{k} \in \mathrm{R^d}.$$

In one dimension the normal distribution $\mathcal{N}(m, \sigma^2)$ posesses the central moments:

$$< (X - m)^n >= \begin{cases} 0, & n \geq 1, \quad n \quad \text{odd} \\ 1 \cdot 3 \cdot 5 \cdot ... (n-1) \cdot \sigma^n, & n \geq 2, \quad n \quad \text{even} \end{cases}$$

B.4 Joint Probability

The joint probability for a pair of random variables $X < Y$ is given by

$$F_{XY}(x,y) = P(X \leq x, Y \leq y),$$

for its joint probability density holds:

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \rho_{XY}(x',y')dx'dy'$$

and

$$\rho_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y).$$

The random variables X, Y are called independent if their joint density is multiplicative:

$$\rho_{XY}(x,y) = \rho_X(x)\rho_Y(y).$$

B.5 Conditional Probability

The conditional probability of an event $A \in \mathcal{A}$ given the event $B \in \mathcal{A}$ with $P(B) \geq 0$ is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

This is the probability of an occurrence of A under the condition that the event B has already occurred. The expression in the random variables X and Y reads:

$$P(x_1 \leq X \leq x_2 | y_1 \leq Y \leq y_2) = \frac{\int_{y_1}^{y_2}\int_{x_1}^{x_2} \rho_{XY}(x,y)dxdy}{\int_{y_1}^{y_2}\int_{-\infty}^{\infty} \rho_{XY}(x,y)dxdy} = \frac{\int_{y_1}^{y_2}\int_{x_1}^{x_2} \rho_{XY}(x,y)dxdy}{\int_{y_1}^{y_2} \rho_Y(y)dy}$$

and the conditional distribution function of X given Y is defined by

$$F_{XY}(x|y) = P(X \leq x, Y = y) = \frac{\int_{-\infty}^{x} \rho_{XY}(x',y)dx'}{\rho_Y(y)}$$

with the conditional density function

$$\rho_{XY}(x|y) = \frac{\rho_{XY}(x,y)}{\rho_Y(y)} \quad \text{for} \quad \rho_Y(y) > 0.$$

Therefore, if $\rho_{XY}(x,y) = \rho_X(x)\rho_Y(y)$, i.e. X and Y are stochastically independent,

$$\rho_{XY}(x|y) = \rho_X(x).$$

In general the joint probability density is computed from the hierarchy of conditional density functions:

$$\rho_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \rho_{X_1,\ldots,X_n}(x_1|x_2,\ldots,x_n)\rho_{X_2,\ldots,X_n}(x_2|x_3,\ldots,x_n)\ldots\rho_{X_{n-1},X_n}(x_{n-1}|x_n)\rho_{X_n}(x_n)$$

B.6 Central Limit Theorem

The distribution of the sum of a large number of random variables is (under rather moderate conditions for these) a normal distribution:
Let $S_n = X_1 + \ldots + X_n$ be a sequence of independent identically distributed random variables with $< X_i >= m$ and $0 \leq < (X_i - < X_i >)^2 > = \sigma^2 \leq \infty$.
Then the following relation holds

$$\lim_{n\to\infty} P(\frac{S_n - mn}{\sigma\sqrt{n}} \leq x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}dy = \mathcal{N}(0,1)$$

or equivalently

$$\lim_{n\to\infty} P(\frac{S_n}{\sqrt{n}} \leq x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(y-m)^2}{2\sigma^2}}dy = \mathcal{N}(m,\sigma^2).$$

B.7 Stochastic Processes

The behaviour of a system that is subject to stochastic fluctuations is modelled by random variables which represent the state of the system at each time. This leads to random variables that have an additional parameter t indicating the time dependence.
A family $\{X_t; t \in T\}$ of random variables in $\mathrm{R^d}$ is called a stochastic or random process.
In the following the parameter t always will be interpreted as a time and the index set T as the real axis or a subset thereof.
The random process is characterized by the properties:

- For constant $t \in T$:
  $X_t(.)$ is a random variable with values in $\mathrm{R^d}$.
- For constant $\omega \in \Omega$:
  $X(\omega)$ is a function defined on T with values in $\mathrm{R^d}$.
  This function is called realization (or trajectory) of the stochastic process.

Summarizing one has for

- **random variables:**
  $$X : \Omega \to \mathrm{R^d}.$$
- **stochastic processes:**
  $$X_t : \Omega \times T \to \mathrm{R^d}.$$

A random variable is characterized by its distribution function, a pair of random variables by its joint distribution function and a random process is determined by the (infinite) hierarchy of its joint distribution functions:

$$\begin{aligned} F(x,t) &= P(X_t \leq x) \\ F(x_1,t_1;x_2,t_2) &= P(X_{t_1} \leq x_1; X_{t_2} \leq x_2) \\ &\vdots \\ F(x_1,t_1;\ldots;x_n,t_n) &= P(X_{t_1} \leq x_1;\ldots;X_{t_n} \leq x_n) \quad \ldots \end{aligned}$$

This system of the probability distributions has the properties:

1. Symmetry: F is invariant with respect to permutations of the indices.
2. Compatibility: All lower dimensional distributions are marginal distributions of the higher dimensional ones.

   For $m \leq n, \quad t_1, ..., t_m, t_{m+1}, ..., t_n \in T$ one has:

   $$F(x_1, t_1; ...; x_m, t_m; \infty, t_{m+1}, ..., \infty, t_n) = F(x_1, t_1; ...; x_m, t_m).$$

   In terms of the probability densities which are defined via

   $$F(x_1, t_1; ...; x_n, t_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} \rho(x_1', t_1; ...; x_n', t_n) dx_1' ... dx_n',$$

   this means integration over the variables $x_{m+1}, ..., x_n$:

$$\rho(x_1, t_1; ...; x_m, t_m) = \int_{\mathrm{R}} dx_{m+1} ... \int_{\mathrm{R}} dx_n \rho(x_1, t_1; ...; x_{m+1}, t_{m+1}; ...; x_n, t_n).$$

The fundamental theorem of Kolmogorov [8] states that the inverse is also true, i.e. for a hierarchy of distribution functions having these properties, there exist a probability triple $(\Omega, \mathcal{A}, P)$ and a stochastic process with the given distribution functions.
Therefore processes can be specified by their distribution functions and not by a probability triple $(\Omega, \mathcal{A}, P)$. For this probability triple the canonical choice can be made:
The sample space Ω is taken to be the set of all real-valued functions $\omega(t)$ defined on the interval T. In this choice any real-valued function is admitted as an element of the sample space Ω. Consequently these functions can be very irregular.
As an extension of the definition of random variables for random processes one has therefore: $X_t(\omega) = \omega(t)$ = value of the function $\omega(t)$ at time t. This canonical choice is characterized by the fact that the elementary outcomes correspond to the trajectories of the process.

B.8 The Spectrum

A measure for the underlying dynamics of the process is the autocorrelation function

$$< X_t X_{t+\tau} >=< X(t) X(t+\tau) >= G(\tau).$$

(In the following the notations X_t as well as $X(t)$ will be used for a stochastic process.)
For stationary processes (which have to be ergodic as an additional requirement) one can equivalently use the time average instead of the ensemble or sample average [6]:

$$G(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t) x(t+\tau) dt.$$

Ergodicity is not a strong condition and is in most cases fulfilled by processes whose correlation $< X_{t_1} X_{t_2} >$ converges to zero sufficiently fast for time differences $|t_2 - t_1|$ going to ∞.
The Fourier transform of the autocorrelation function of a stationary process is called the spectrum.

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} G(\tau) d\tau,$$

$$G(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) d\omega.$$

B.9 Stationarity

A stochastic process is called stationary if all its finite-dimensional probability densities are invariant with respect to time translations:

$$\rho(x_1, t_1; ...; x_n, t_n) = \rho(x_1, t_1 + \tau; ...; x_n, t_n + \tau).$$

This condition is for example fulfilled if the experimental conditions are kept constant after the initial oscillations have damped out.
As a consequence the one-dimensional probability density is constant in time:

$$\rho(x, t) = \rho_s(x).$$

The same applies for the mean of a stationary process:

$$< X_t >= \int_{\mathrm{R}} x \rho(x, t) dx = \int_{\mathrm{R}} x \rho_s(x) dx = m.$$

The two-dimensional probability density $\rho(x_1, t_1; x_2, t_2)$ depends in this case only on the time difference $t_2 - t_1$:

$$\rho(x_1, t_1; x_2, t_2) = \rho(x_1, x_2; t_2 - t_1)$$

and the same holds for the two-time covariance or correlation function $C_X(t_1, t_2)$:

$$< (X_{t_1} - m)(X_{t_2} - m) >= \int_{\mathrm{R}} \int_{\mathrm{R}} (x_{t_1} - m)(x_{t_2} - m) \rho(x_1, t_1; x_2, t_2) dx_1 dx_2.$$

C Solution of the Deterministic Double Rf System

The aim is to transform the Hamiltonian

$$H(\phi, W) = C_W W^2 + \frac{1}{2} C_\phi [(\frac{m^2 - 1}{4!}) \phi^4], \tag{C.1}$$

to action-angle variables Q and J. H is not explicit time dependent and therefore the Hamilton-Jacobi equation reduces to (see [25]):

$$H(\phi, \frac{\partial S_2}{\partial \phi}) = \alpha_H$$

where α_H is the energy of the system and the generating function S_2 is dependent on the old coordinate ϕ and the new momentum α_H.
The transformation equations for the coordinates read

$$\begin{aligned} Q = \frac{\partial S_2}{\partial \alpha_H} &= \frac{1}{\sqrt{C_W}} \frac{\partial}{\partial \alpha_H} \int_0^\phi d\phi' \sqrt{\alpha_H - V(\phi')} d\phi' \\ &= \frac{1}{\sqrt{2C_W}} \int_0^\phi \frac{d\phi'}{\sqrt{\alpha_H - V(\phi')}} \end{aligned} \tag{C.2}$$

$$W = \frac{\partial S_2}{\partial \phi} = \frac{1}{\sqrt{C_W}} \sqrt{\alpha_H - V(\phi)}. \tag{C.3}$$

The action variable of the ϕ^4-potential

$$V(\phi) = \frac{1}{2}C_\phi(\frac{m^2-1}{4!})\phi^4$$

is calculated as

$$\begin{aligned} J = \oint W d\phi = \oint \frac{\partial S_2}{\partial \phi} d\phi &= \oint \frac{1}{\sqrt{C_W}}\sqrt{\alpha_H - V(\phi)} \\ &= \frac{1}{\sqrt{C_W}} \oint \sqrt{\alpha_H - \frac{C_\phi}{2}(\frac{m^2-1}{4!})\phi^4} \\ &= 2 \cdot \sqrt{\frac{\alpha_H}{C_W}} \int_{-\hat{\phi}}^{\hat{\phi}} \sqrt{1 - \frac{C_\phi}{2\alpha_H}(\frac{m^2-1}{4!})\phi^4}. \end{aligned} \tag{C.4}$$

The maximal amplitude of oscillation in phase space $\hat{\phi}$ is given by

$$\hat{\phi} = (\frac{C_\phi}{2\alpha_H}(\frac{m^2-1}{4!}))^{-\frac{1}{4}}.$$

The solution for the action variable J in dependence on α_H in the limit of small amplitudes is the expression

$$J = \underbrace{\frac{8\sqrt{2}K(\frac{1}{\sqrt{2}})}{3C_W^{\frac{1}{2}}C_\phi^{\frac{1}{4}}}(\frac{3}{m^2-1})^{\frac{1}{4}}}_{=const=:\tilde{c}} \alpha_H^{\frac{3}{2}} = J(\alpha_H), \qquad \alpha_H << C_\phi,$$

where $K(\frac{1}{\sqrt{2}}) = 1.8541$ is the complete elliptic integral of modulus $k = \frac{1}{\sqrt{2}}$ (see [29]).

In the ϕ^4-potential approximation the rf phase $\phi(t)$ as a function of the action variable J (or of the energy variable α_H) and the angle variable Q can be found by integration

$$\dot{\phi} = \frac{\partial H}{\partial W} = 2C_W W \tag{C.5}$$

$$= 2\sqrt{C_W}\sqrt{\alpha_H - \frac{1}{2}C_\phi(\frac{m^2-1}{4!})\phi^4} \tag{C.6}$$

and by using the relations $\Omega_s = 2\pi\dot{Q}$ and $\dot{Q} = \frac{\partial \tilde{\mathcal{H}}}{\partial J}$.
The solution for ϕ is

$$\begin{aligned} \phi(t) &= \hat{\phi}\ \mathrm{cn}(\frac{2\sqrt{2\alpha_H C_W}}{\hat{\phi}}t) \\ &= \hat{\phi}\ \mathrm{cn}((\frac{m^2-1}{3})^{\frac{1}{4}}\sqrt{2}C_W^{\frac{1}{2}}C_\phi^{\frac{1}{4}}\alpha_H^{\frac{1}{4}}t) \\ &= \hat{\phi}\ \mathrm{cn}(\frac{2K(\frac{1}{\sqrt{2}})}{\pi}\Omega_s t) \\ &= \hat{\phi}\ \mathrm{cn}(4K(\frac{1}{\sqrt{2}})\dot{Q}t) \\ &= \hat{\phi}\ \mathrm{cn}(4K(\frac{1}{\sqrt{2}})Q), \end{aligned} \tag{C.7}$$

where cn means the Jacobian elliptic cosine function.

D Perturbation Theory for the Rf Stochastic Differential Equation

A perturbation theoretical approach to calculate the rf Fokker-Planck equation from the system of stochastic differential equations (4.7) was suggested in [12]. Here the basic ideas will be summarized.
The long time scale motion in synchrotron space can be described by the following Fokker-Planck equation in J:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial J}(A_1\rho) + \frac{1}{2}\frac{\partial^2}{\partial J^2}(A_2\rho),$$

with

$$A_1 = \frac{1}{2}\frac{\partial}{\partial J}A_2,$$

leading to the diffusion equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial J}(\frac{1}{2}A_2\frac{\partial}{\partial J}\rho). \tag{D.1}$$

As the noise term is considered weak, a perturbation series ansatz for the solution is made.
The starting point is the Hamiltonian describing the synchrotron dynamics in the dynamical variables q and p

$$\mathcal{H} = \frac{1}{2}p^2 + V(q) + h(q)\xi(t),$$

where $V(q)$ is an arbitrary rf potential and $\xi(t)$ is a white noise process.
The equation of motion reads

$$\ddot{q} + V'(q) + h'(q)\xi(t) = 0,$$

with phase noise corresponding to $h(q) = -q$ and amplitude noise corresponding to $h(q) = V(q)$.
One makes a canonical transformation to new phase space variables Q and P with the generating function $S_2(q,P)$ which is defined by

$$\begin{aligned} Q &= \frac{\partial S_2}{\partial P} = \int_0^q \frac{dq'}{\sqrt{2(P-V(q))}} \\ p &= \frac{\partial S_2}{\partial q} = \sqrt{2(P-V(q))}. \end{aligned}$$

The transformed Hamiltonian reads

$$\tilde{\mathcal{H}} = P + h(q(Q,P))\xi(t) \tag{D.2}$$

and the stochastic equations are now

$$\begin{aligned} \dot{Q} = \frac{\partial \tilde{\mathcal{H}}}{\partial P} &= 1 + \frac{\partial h(q(Q,P))}{\partial P}\xi(t) \\ \dot{P} = -\frac{\partial \tilde{\mathcal{H}}}{\partial Q} &= -\frac{\partial h(q(Q,P))}{\partial Q}\xi(t) \end{aligned} \tag{D.3}$$

which is a description where the noise term couples multiplicatively to the system.
Integrating equation (D.3) yields

$$\begin{aligned} Q(t) &= Q_0 + t + \int_0^t dt' \frac{\partial \tilde{h}(Q(t'), P(t'))}{\partial P} \xi(t') \\ P(t) &= P_0 - \int_0^t dt' \frac{\partial \tilde{h}(Q(t'), P(t'))}{\partial Q} \xi(t'). \end{aligned}$$

These terms are expanded into a perturbation series to second order in the noise term (and therefore to first order in t).

$$\begin{aligned} Q(t) &= Q_0 + t + Q_1(t) + Q_2(t) + ... \\ P(t) &= P_0 \quad + P_1(t) + P_2(t) + ... \end{aligned} \tag{D.4}$$

where the indices correspond to the order of the noise term ξ, such that
$Q_0 + t$ and P_0 are of 0^{th} order in the noise, $o(\xi^0)$;
Q_1 and P_1 are the terms of first order, $o(\xi^1)$, and so on.
The first order terms are

$$Q_1(t) = \int_0^t dt' \frac{\partial \tilde{h}(Q_0 + t', P_0)}{\partial P_0} \xi(t'),$$

$$P_1(t) = -\int_0^t dt' \frac{\partial \tilde{h}(Q_0 + t', P_0)}{\partial Q_0} \xi(t').$$

The action variable J is related to P by

$$J = \oint p dq = \oint \sqrt{2(P - V(q))} dq,$$

and the oscillation period $T(P)$ can be expressed by the partial derivative of J with respect to P:

$$\frac{\partial J(P)}{\partial P} = \oint \frac{dq}{\sqrt{2(P - V(q))}} = T(P). \tag{D.5}$$

The action $J(P)$ is expanded around the lowest order term of the P-perturbation series, P_0, again to second order in $\xi(t)$, using (D.4) and (D.5).

$$J(P) = J(P_0) + \frac{\partial J(P_0)}{\partial P}(P_1 + P_2) + \frac{1}{2}\frac{\partial^2 J(P_0)}{\partial P^2}(P_1 + P_2)^2 + ... \quad ,$$

yielding expressions for
$< \Delta J >=< J - J_0 >$ and $< (\Delta J)^2 >=< (J - J_0)^2 >$
where $J_0 = J(P_0)$ and $T_0 = T(P_0)$.
The averages are taken with respect to the noise samples. $\Delta J = J - J_0$ means the change in in the action variable during a time interval Δt which is long compared to the correlation time of the noise and short compared to the time scales within which J changes appreciably.

$$\begin{aligned} < \Delta J > &= < P_2 > T_0 + \frac{1}{2} < P_1^2 > \frac{\partial T_0}{\partial P_0}, \\ < (\Delta J)^2 > &= < P_1^2 > T_0^2, \qquad \text{using } < P_1 >= 0. \end{aligned} \tag{D.6}$$

The coefficients A_1 and A_2 of the Fokker-Planck equation are obtained by averaging over the lowest order term of the Q-series, Q_0:

$$A_1 = \frac{1}{T_0}\int_0^{T_0} dQ_0 < \frac{\Delta J}{\Delta t} >=<< \frac{\Delta J}{\Delta t} >>_Q$$

and

$$A_2 = \frac{1}{T_0}\int_0^{T_0} dQ_0 < \frac{(\Delta J)^2}{\Delta t} >=<< \frac{(\Delta J)^2}{\Delta t} >>_Q .$$

For white noise with $< \xi(t)\xi(t') >= \lambda\delta(t - t')$ this leads to an explicit expression for the diffusion coefficient A_2:

$$A_2 = \lambda T_0 \oint p dq (h'(q))^2,$$

and together with the relation

$$A_1 = \frac{1}{2}\frac{\partial}{\partial J} A_2$$

the Fokker-Planck equation is completely determined.

E Flow Chart Representation of the Mapping Algorithm

In this part of the appendix the principles of the matrix algorithm which is used for the density calculations are summarized in a flow chart form.

The first chart describes that part of the program where the tracking of the test particles is performed and the propagator matrix is computed by calculating the transition probabilities.

The second chart refers to the "matrix multiplication", i.e. the successive application of the propagator matrix to the phase space density which simulates the temporal evolution of the density function.

COMPUTING THE TRANSITION MATRIX

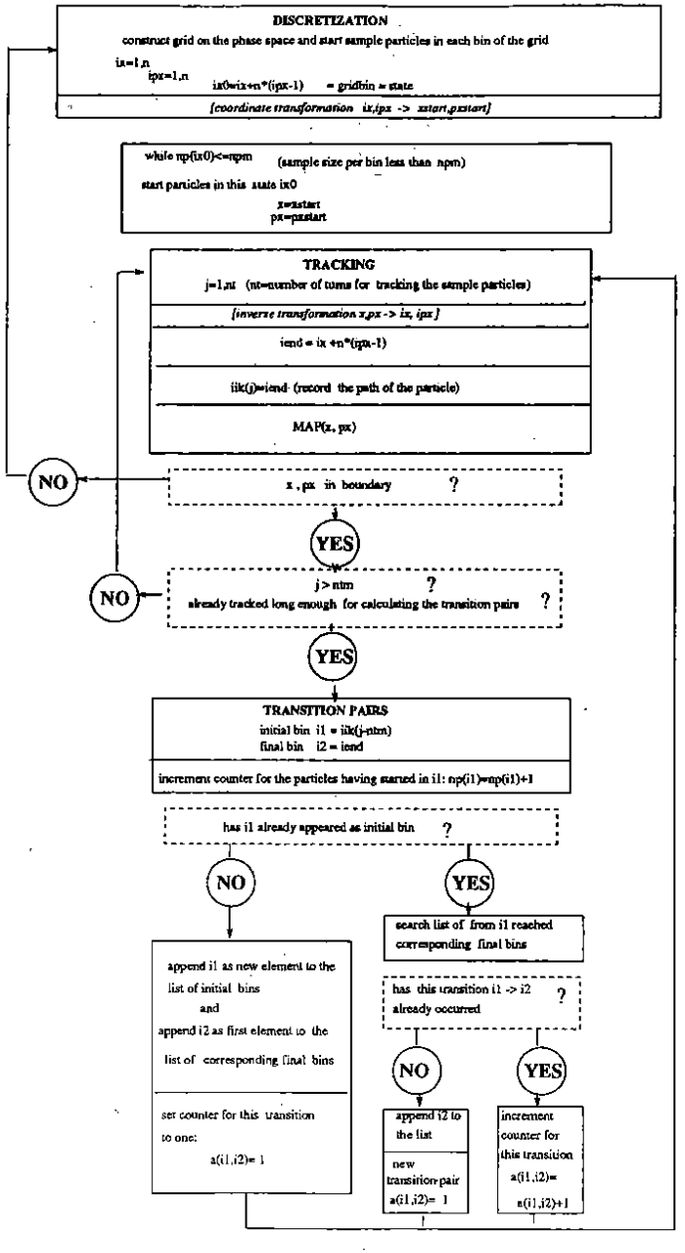

Figure 42: Outline of the first part of the algorithm

COMPUTING THE DENSITY FUNCTION

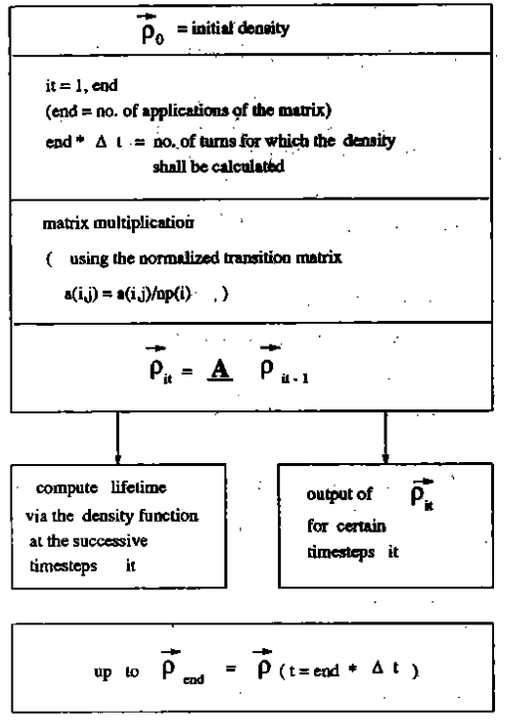

Figure 43: Outline of the second part of the algorithm

References

[1] Chao,A.W., Evaluation of Beam Distribution Parameters in an Electron Storage Ring, J.Appl.Phys.,50, (1979)

[2] Mais,H.,Ripken,G.,Influence of the Synchrotron Radiation on the Spin-Orbit Motion of a Particle in a Storage Ring, DESY M-82-20, (1982)

[3] Barber,D.P.,Heinemann,K.,Mais,H.,Ripken;G., A Fokker-Planck Treatment of Stochastic Particle Motion within the Framework of a Fully Coupled 6-dimensional Formalism for Electron-Positron Storage Rings including Classical Spin Motion in Linear Approximation, DESY 91-146, (1991)

[4] Brinkmann,R., A Simulation Study for the Beam-Beam Interaction of Protons with a Flat Beam in HERA, DESY HERA 89-24, (1989)

[5] Jowett,J.M., Introductory Statistical Mechanics for Electron Storage Rings, SLAC-PUB-4033, (1986)

[6] Gardiner,C.W., Handbook of Stochastic Methods, Springer, (1990)

[7] Risken,L., The Fokker-Planck Equation, Springer, (1984)

[8] Horsthemke,W., Lefever,R., Noise-Induced Transitions, Springer, (1984)

[9] Soong,T.T., Random Differential Equations in Science and Engineering, Mathematics in Science and Engineering, Vol. 103, (1979)

[10] Hofmann,A.,Myers,S., Beam Dynamics in a Double RF System, Proc.XI^{th} Internat. Conf. on High Energy Accelerators, CERN (1980),p.610

[11] Wei,J., Stochastic Cooling with a Double RF System, BNL-46792, AD/RHIC-105, Informal Report, (1991)

[12] Krinsky,S.,Wang,J.M., Bunch Diffusion due to RF Noise, Part. Acc., 12, (1982)

[13] Schonfeld,J.F., Statistical Mechanics of Colliding Beams, Ann.of Phys. 160, (1985)

[14] Ripken,G., Karantzoulis,E., A Fokker-Planck Treatment of Coupled Synchro-Betatron Motion in Proton Storage Rings under the Influence of Cavity Noise, DESY 90-135, (1990)

[15] Sands, M., The Physics of Electron Storage Rings: An Introduction, SLAC-121, (1970)

[16] Chin, Y.H., Quantum Lifetime, DESY 87-062, (1987)

[17] Gerasimov, A.L., private communication, (1990)

[18] Montague,B.W., Polarized Beams in High Energy Storage Rings, Phys. Rep.,113,Vol.1, (1984)

[19] Dôme, G., Theory of RF Acceleration, CAS, Oxford, (1985)

[20] Boussard,D., Dôme, G.,Graziani,C., Proc.XI^{th} Internat. Conf. on High Energy Accelerators, CERN, (1980), p.620

[21] Virichenko,Yu.P.,Grigor'ev,Yu.N., Equilibrium Distribution of Charged Particles in the Phase Space of a Cyclic Accelerator, Ann.Phys.,Vol. 209,No. 1, (1991)

[22] Honerkamp,J., Stochastische dynamische Systeme, VCH, Weinheim, (1990)

[23] Klauder,J.P.,Petersen,W.P., Numerical Integration of Multiplicative-Noise Stochastic Differential Equations, SIAM J.Num.Anal., Vol.22,No.6, (1985)

[24] Strittmatter,W., Numerische Behandlung von stochastischen dynamischen Systemen, Dissertation, Freiburg, (1988)

[25] Kuypers,F., Klassische Mechanik, VCH Weinheim, (1993)

[26] Pauluhn,A., Some Aspects of RF Noise in Storage Rings, DESY HERA 92-07,ed.:F.Willeke, p.421, (1992)

[27] Pauluhn,A., Stochastic Beam Dynamics with Special Applications to a Double RF System, DESY HERA 93-02, (1993)

[28] Leeman,B.T.,Forest,E.,Chattopadhyay,S., Simulation of Synchrotron Motion with RF Noise, Proc.$XIII^{th}$ Internat. Conf. on High Energy Accelerators, Novosibirsk, (1986)

[29] Gradshteyn, Ryshik, Tables of Integrals, Series and Products, Academic Press, New York, London, (1965)

[30] Byrd,P.F., Friedman,M.D., Handbook of Elliptic Integrals for Engineers and Scientists, Springer, Berlin, (1971)

[31] Courant,E.P.,Snyder,H.S., Theory of the Alternating-Gradient Synchrotron, Ann. of Phys. 3, (1958)

[32] Schmüser, P., Basic Course on Accelerator Optics, DESY HERA 87-02, (1982)

[33] Milton, S.V., The Beam-Beam Interaction in Electron Storage Rings: A Study of the Weak-Strong Case, Ph.D. thesis, Cornell, (1990)

[34] Zholents,A.A., Beam-Beam Effects in Electron-Positron Storage Rings, preprint 91-18, Novosibirsk, (1990)

[35] Brüning, O., Emittance Growth in Proton Storage Rings due to the Combined Effect of Tune Modulation and Beam-Beam Interaction for Round Beams, DESY HERA 93-05, (1993)

[36] Chung,K.L., Elementare Wahrscheinlichkeitstheorie und stochastische Prozesse, Springer, Berlin, (1978)

[37] Gerasimov, A.L., Phase Convection and Periodically Driven Brownian Motion, Physica D (41), (1989)

[38] Brinkmann,R., HERA, DESY HERA 88-03, (1988)